A Search for New Resonant Phenomena in Dijet Final States With the ATLAS Detector at the Large Hadron Collider

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Clark, Brian. 2016. A Search for New Resonant Phenomena in Dijet Final States With the ATLAS Detector at the Large Hadron Collider. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493365">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493365</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
A search for new resonant phenomena in dijet final states with the ATLAS detector at the Large Hadron Collider

A DISSERTATION PRESENTED
BY
Brian Lee Clark
TO
The Department of Physics
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy
IN THE SUBJECT OF
Physics

Harvard University
Cambridge, Massachusetts
May 2016
A search for new resonant phenomena in dijet final states with the ATLAS detector at the Large Hadron Collider

Abstract

This dissertation describes a model-agnostic search for new resonant phenomena decaying to two jets (dijets) in 3.6 fb$^{-1}$ of proton-proton collisions at center-of-mass energy $\sqrt{s} = 13$ TeV. The collisions were produced by the Large Hadron Collider and recorded by the ATLAS detector in 2015. A dijet invariant mass spectrum is examined for local excesses above a data-driven background estimation accounting for Standard Model production. No statistically significant excess is observed. New physics models are excluded at 95% credibility-level for quantum black holes with production mass thresholds below 5.3 TeV, 8.1 TeV, and 8.3 TeV for three different production scenarios and for excited quarks of masses less than 5.2 TeV. Furthermore, limits are provided on generic Gaussian signal shapes for additional generalization of the 2015 dijet resonant analysis search results.
## Contents

0 Introduction ................................................................. 1

I Theory Overview ............................................................ 4

1 The Standard Model and Beyond ........................................ 5
   1.1 The Standard Model .................................................. 5
   1.2 Beyond the Standard Model ......................................... 8

2 Dijet Production at the LHC .............................................. 16
   2.1 Quantum Chromodynamics .......................................... 16
   2.2 Proton Structure ..................................................... 21
   2.3 Overview of a Collision ............................................. 22
   2.4 Dijet Production at Hadron Colliders ............................. 24

II Experimental Measurement .............................................. 30

3 The Large Hadron Collider ............................................... 31
   3.1 Experiments ............................................................ 31
   3.2 Injection Chain ....................................................... 32
   3.3 Luminosity Measurement ............................................ 33
   3.4 2015 Running Conditions ........................................... 34

4 The ATLAS Experiment .................................................... 36
   4.1 Coordinate System and Nomenclature ............................. 39
   4.2 Magnet System ....................................................... 41
   4.3 Inner Tracker .......................................................... 42
   4.4 Calorimeters ........................................................... 47
   4.5 Muon Spectrometer ................................................... 51
   4.6 Summary of Detector Resolution and Coverage ................. 53
   4.7 Detectors as a System ............................................... 53
   4.8 Trigger ................................................................. 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.9</td>
<td>2015 Data Collection</td>
<td>56</td>
</tr>
<tr>
<td>5</td>
<td>JETS AT ATLAS</td>
<td>57</td>
</tr>
<tr>
<td>5.1</td>
<td>Building Topological Clusters</td>
<td>57</td>
</tr>
<tr>
<td>5.2</td>
<td>Jet Reconstruction Algorithm</td>
<td>59</td>
</tr>
<tr>
<td>5.3</td>
<td>Jet Energy Calibration</td>
<td>61</td>
</tr>
<tr>
<td>5.4</td>
<td>Jet Cleaning</td>
<td>66</td>
</tr>
<tr>
<td>III</td>
<td>Dijet Resonance Analysis</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>Building a Dijet Spectrum</td>
<td>70</td>
</tr>
<tr>
<td>6.1</td>
<td>Data Sample</td>
<td>71</td>
</tr>
<tr>
<td>6.2</td>
<td>Monte Carlo Sample</td>
<td>71</td>
</tr>
<tr>
<td>6.3</td>
<td>Trigger Strategy</td>
<td>72</td>
</tr>
<tr>
<td>6.4</td>
<td>Event Selection</td>
<td>73</td>
</tr>
<tr>
<td>6.5</td>
<td>Analysis Selection Summary</td>
<td>77</td>
</tr>
<tr>
<td>6.6</td>
<td>Dijet Mass Spectrum</td>
<td>79</td>
</tr>
<tr>
<td>6.7</td>
<td>$m_{jj}$ Spectrum for 2015 Data</td>
<td>86</td>
</tr>
<tr>
<td>6.8</td>
<td>Additional Kinematic Distributions</td>
<td>86</td>
</tr>
<tr>
<td>6.9</td>
<td>Highest Mass Dijet Event</td>
<td>87</td>
</tr>
<tr>
<td>7</td>
<td>The Search Phase</td>
<td>89</td>
</tr>
<tr>
<td>7.1</td>
<td>Statistical Framework</td>
<td>90</td>
</tr>
<tr>
<td>7.2</td>
<td>Background Estimation</td>
<td>92</td>
</tr>
<tr>
<td>7.3</td>
<td>Uncertainty on Background Estimation</td>
<td>98</td>
</tr>
<tr>
<td>7.4</td>
<td>Search Results for 2015 Data</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>Limit Setting</td>
<td>101</td>
</tr>
<tr>
<td>8.1</td>
<td>Overview of Limit Setting Strategy</td>
<td>102</td>
</tr>
<tr>
<td>8.2</td>
<td>Benchmark Signal Models</td>
<td>103</td>
</tr>
<tr>
<td>8.3</td>
<td>Limits Using 2015 Dataset</td>
<td>104</td>
</tr>
<tr>
<td>IV</td>
<td>Looking Forward</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>Discovery Potential</td>
<td>110</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Example $q^*$ Signal Shapes</td>
<td>113</td>
</tr>
</tbody>
</table>
Appendix B  Kinematic Distributions  115

Appendix C  Preliminary Studies Using Wide Jets  125
  C.1  Simple Model .......................... 125
  C.2  Improved Sensitivity .................... 127

References  134
Listing of figures

1. Presented are cross-sections for various final-states at hadron colliders as a function of center-of-mass collision energy. The total proton-proton cross-section at the LHC is approximately 80 mb. Figure taken from [1].

1.1. SM particles are summarized. There are three generations of quarks (purple) and leptons (green) with increasing mass from left to right. Gauge bosons are shown in orange, while the Higgs boson is in the far right corner. Figure is taken from [22].

1.2. Feynman diagram for a $qg \rightarrow q^* \rightarrow qg$ process.

1.3. A QCD dijet invariant mass background is modeled for proton-proton collisions at a hypothesized $\sqrt{s} = 40$ TeV collider. Excited quark signals for various masses are drawn as Breit-Wigner resonances. Figure from [27].

1.4. QCD background for dijet production is a smooth, monotonically falling background. The analysis presented in this dissertation is sensitive to resonant signals, or bumps, in the dijet mass distribution, such as the one drawn in (a). If new physics is slightly beyond the reach of our dijet mass distribution it may be modeled as a contact interaction, where the compositeness scale is assumed to be greater than achieved center-of-mass collision energies. Such contact interactions would present as a rising tail in the dijet mass spectrum. A companion analysis, the dijet angular analysis, is designed to search for contact interactions.

2.1. Feynman diagram interaction vertices in QCD. Figures 2.1b and 2.1c corresponds gluon-gluon self-interaction.

2.2. Presented is a summary of experimental measurements of $\alpha_s$ as a function of the energy scale $Q$. QCD perturbation theory is used to extract a value for the strong coupling constant, and the respective order of perturbation theory is indicated in parentheses. This figure is taken from [22].

2.3. A proton is composed of two $u$ quarks and one $d$ quark, called valence quarks, bound together by gluons. There are also sea quarks and additional gluons that arise from pQCD loop corrections.

2.4. Shown are two PDFs at (a) low-energy scales and (b) high-energy scales [1].
2.5 Presented is a simulated $pp \to ttH$ event produced by an event generator. Incoming partons are shown in blue. The hard scattering process is represented by the large red filled circle and the resulting parton shower is shown in red. A secondary interaction, or underlying event, is represented in purple. Light green blobs represent the initial hadronization and dark green blobs correspond to hadron decay. Note that the incoming protons are represented in green as they are hadrons. Additionally, yellow lines correspond to photon emission. Figure is taken from [39].

2.6 Feynman digram for the $gg \to ttH$ process simulated in Figure 2.5.

2.7 Cartoon illustrating a $1 + 2 \to 3 + 4$ scattering process. The lab frame (left) is boosted to the center-of-mass frame (right) where the dijets are back-to-back.

2.8 A subset of leading-order Feynman diagrams for QCD scattering are shown. This figure is taken from [40] and is an adaptation of similar figure appearing in [41].

2.9 Monte Carlo simulation showing dijet mass curves for processes with initial $gg$, $qg$, and $qq$ states.

3.1 Shown is the CERN accelerator complex. For proton-proton collisions, protons will progressively be accelerated though the ordered chain: LINAC 2, BOOSTER, PS, SPS, LHC. Figure is provided by CERN [47].

3.2 Shown are 2015 peak instantaneous luminosity per proton fill at the LHC (a) and cumulative luminosity delivered to the ATLAS experiment (b) as a function of time. Figures taken from [48].

4.1 Shown is an overview of the ATLAS detector. Individual detector components are labeled. Figure taken from [43].

4.2 Coordinate system for the ATLAS detector is shown. A right-handed Cartesian coordinate system is shown in (a) with the origin at the interaction point, the $x$-axis pointing towards the LHC ring center, the $y$-axis pointing upwards to the surface, and the $z$-axis pointing along the beamline. A more typical cylindrical system is shown in (b) and measures pseudorapidity $\eta$, radius $r$, and azimuthal angle $\phi$.

4.3 (a) The ATLAS magnet system, shown in orange, consists of one 2 T solenoid servicing the inner detector and large toroidal magnet system servicing the muon spectrometer. The colorful cylindrical layers encasing the solenoid represent the calorimeter system. (b) Shown is the approximate field integral for perfectly straight tracks traversing the toroidal system in one octant. The red curve maps the field for $\phi = 0$ and the black curve maps the field integral for $\phi = \pi/8$. Figures are taken from [43].

4.4 Cut-away of the ATLAS inner detector. The three independent sub-detectors are the pixel detector, semiconductor tracker, and transition radiation tracker. Figure is taken from [43].
4.5 Drawn are the sensors and structural components of the innermost barrel components of the ATLAS detector. The red line corresponds to the simulated track of a \( p_T = 10 \) GeV charged particle at \( \eta = 0.3 \). The particle crosses the beryllium beam pipe, four silicon-pixel layers (including the newly-installed insertable b-layer), four paired layers of silicon microstrips, and approximately 36 transition radiation tube straws contained within their supporting medium. Figure taken from [43].

4.6 Displayed is an accounting of material in the inner detector as measured in radiation \( (X_0) \) and nuclear interaction \( (\lambda) \) lengths as a function of \( |\eta| \) and averaged over \( \phi \). The distributions account for external services in red and the individual sub-detectors, where the services internal to the sub-detector value are included. The large increase in material around \( |\eta| \approx 1 \) corresponds to service cables and structural supports for the SCT and and TRT, which enter between the barrel and end-cap. The increase in material around \( |\eta| \approx 2.7 \) indicates the concentration of service material for the pixel detector, which enters beyond the SCT end-cap. Figure is taken from [43].

4.7 Overview of the ATLAS calorimeter. Figure is taken from [43].

4.8 Drawn is an EM calorimeter barrel module where the different layers and their geometrical properties are evident. Figure taken from [43].

4.9 The cumulative amount of detector material, in nuclear interaction lengths, is shown as a function of \( |\eta| \). The unlabeled contribution in beige represents all material in front of the EM calorimeter. The calorimeter layers are labeled with text. The light blue shading represents all detector material before the first active layer of the muon spectrometer. The shape of this figure will play a noticeable role in \( \eta \) dependent jet calibrations. Figure is pulled from [43].

4.10 Muon sub-detector components are highlighted. The light blue corresponds to muon detection, while the toroidal magnet system is in light orange. Image pulled from [43].

4.11 Shown are various particle types traversing the ATLAS detector. Combining various detectors allows for general identification of particles. Image taken from [54].

4.12 Various L1 (a) and HLT (b) trigger rates are shown. The total trigger rate is the sum of all constituent trigger rates. Plots are taken from [57].

4.13 Cumulative 2015 luminosity delivered (green), recorded (yellow), and made available for physics analyses (blue) [48].

5.1 A hard-scatter event may manifest as jets in the ATLAS detector. Careful measurement of these jets gives insight into the original partons emitted after the hard-scatter. Figure taken from [58].

5.2 ATLAS uses a 4/2/0 clustering scheme to build topoclusters from calorimeter cells. As seen in the drawing, neighboring cells included in the clustering scheme need not be in the same calorimeter layer. Figure taken from [59].
5.3 Four different jet clustering algorithms offer four different interpretations of the same physics event. Notice the distinct shape differences in the formed jet areas. Figure taken from [61].

5.4 Energy response is shown for jet reconstruction and truth jets, as simulated in Monte Carlo. Large dips in the response can be directly contributed to transition regions in the calorimeter. Figure taken from [62].

5.5 Jet transverse momentum response is shown as a function of the number of muon segments associated with the jet. Figure taken from [62].

5.6 Punch through corrections are applied for any jet with greater than 20 associated muon segments. Notice in (a) that punch through is most likely at calorimeter gaps or transition regions.

5.7 Jet energy resolution uncertainties estimated for 2015 data as a function of jet transverse momentum (a) and pseudorapidity (b) are shown for 25 ns bunch spacing. Figure taken from [62].

5.8 Jet energy scale uncertainties estimated for 2015 data as a function of jet transverse momentum are shown for 25 ns running conditions (a) and 50 ns running conditions (b). The difference in bunch spacing will most directly affect pile-up correction. Figure taken from [62].

6.1 (a) The $y^*$ cut value is varied and significance curves are calculated and plotted. (b) Each significance curve, pertaining to a specific $q^*$ mass, is normalized to unit area for comparison of all signal optimal values. The optimal centrality cut is approximately $|y^*| < 0.6$.

6.2 (a) The $y_B^*$ cut value is varied, where events are required to have $|y_B^*| < y_B^{cut}$, and significance curves are calculated and plotted. (b) Each significance curve, pertaining to a specific $q^*$ mass, is normalized to unit area for comparison of all signal optimal values. No optimal cut value was identified, therefore a $y_B^*$ was not implemented in the resonance analysis.

6.3 Figure shows the detector and $y^*$ cuts drawn in detector $\eta$. Events with leading and sub-leading jets appearing in red and blue shaded regions of the detector will pass $\eta$ and $y^*$ analysis cuts.

6.4 The dijet mass resolution is shown.

6.5 Binned MC dijet mass spectrum is plotted in bins of reconstructed $m_{jj}$ versus truth $m_{jj}$.

6.6 Efficiency (a) and purity (b) for each of our derived $m_{jj}$ bins. A smooth, monotonic efficiency curve is expected for a smooth binning distribution.

6.7 A $M_{q^*} = 4$ TeV resonant signal is simulated and the peak width is conservatively estimated.
Displayed is the $m_{jj}$ spectrum for 2015 data (black), next-to-leading order Standard Model predictions (blue), and jet energy scale uncertainty on Monte Carlo predictions (light blue shading). The binning reflects the derived dijet binning and the analysis cuts have been applied, including $m_{jj} > 1.1$ TeV, leading jet $p_T > 440$ GeV and $|y^*| < 0.6$.  

Leading (a) and sub-leading (b) jet $p_T$ distributions with analysis cuts applied.  

Highest 2015 recorded $m_{jj}$ event satisfying the resonant dijet analysis cuts. The $m_{jj} = 6.9$ TeV event consists of two central jets with $p_T \approx 3.2$ TeV.  

Highest 2015 recorded $m_{jj}$ event satisfying the resonant dijet analysis cuts. The $m_{jj} = 6.9$ TeV event consists of two central jets with $p_T \approx 3.2$ TeV.  

Each horizontal line represents a BumpHunter mass window. The most discrepant region from the background estimation is at approximately 1.6 TeV.  

Many pseudo-experiments are generated to quantify the probability of randomly observing a spectrum at least as bumpy as the one observed. The observed BumpHunter test statistic is represented by the red arrow. In this case, approximately 670 of the 1000 pseudo-experiments resulted in a spectrum at least as bumpy as the one observed in data.  

Shown is the reconstructed $m_{jj}$ spectrum with 2015 data (filled dots) fitted with the dijet fit function described by Equation 7.7 (solid line). The blue vertical lines indicate the region of greatest excess. Bin-by-bin significance of the data-fit difference is shown in the bottom panel, considering statistical uncertainties only.  

Background only estimation is shown in red. The statistical uncertainty on the fit and the choice of fit parameterization are shown.  

The reconstructed dijet mass for data passing event selection criteria is plotted as filled dots. The solid red line is the background estimation. The most discrepant region is identified by vertical blue lines. Two resonant models, $q^*$ and QBH (BlackMax generator), are shown above the fit using a predicted cross-section. The $q^*$ cross-section has been increased by a factor of three for visualization of the signal shape. Bin-by-bin comparisons of the data to the background fit are shown in the middle panel. The bottom panel shows NLO MC background predictions with JES uncertainty bands.  

The 95% credibility-level upper limits for $\sigma \times A$ are shown for quantum black hole signals as a function of mass. ADD models are denoted by QBH (BM) and QBH (QBH) where $n = 6$ extra dimensions are assumed. A RS model with $n = 1$ extra dimension is denoted by QBH (RS). The observed 95% credibility-level lower limit on QBH mass threshold in a given model is identified as the intersection of the theory and observed curves. Result presented in [20].
8.2 The 95% credibility-level upper limits for $\sigma \times A$ are shown for $q^*$ signals as a function of mass. The observed 95% credibility-level lower limit on $q^*$ mass is identified as the intersection of the theory and observed curves. Result presented in [20]. 107

8.3 The 95% credibility-level upper limits for $\sigma \times A \times BR$ are presented for a Gaussian-shaped signal width $\sigma_G$ and mass $m_G$, where the two values refer to the Gaussian standard deviation and mean, respectively. These results are useful for extrapolating analysis results to signal models that were not explicitly considered. Result presented in [20]. 108

9.1 The ratios (13 TeV/8 TeV) of LHC parton luminosities are shown as a function of mass. This increase in parton luminosity for massive object production allows further reach into the high-mass region of a dijet mass spectrum, which increases discovery potential. Figure taken from [81]. 111

9.2 Discovery potential for $q^*$ and QBH benchmark signals with a dijet mass resonance search with the ATLAS detector at $\sqrt{s}$ TeV. Discovery is defined as a 5$\sigma$ excess identified by the search procedure described in Chapter 7. Figures taken from [82]. 112

9.3 Limit setting potential for $q^*$ and QBH benchmark signals with a dijet mass resonance search with the ATLAS detector at $\sqrt{s}$ TeV. Expected upper limits are set using Bayesian methods described in Chapter 8 and correspond to a 95% credibility level. Figures taken from [82]. 112

A.1 Six $q^*$ signals generated in MC are presented. Peak fits were accomplished with an iterative Gaussian fitting technique to give conservative estimates of the signal width. Analysis cuts were applied and the number of events assumes $\approx 1$ fb$^{-1}$ of data. 114

B.1 Shown are kinematic distributions for the dijet system. 117
B.2 Shown are kinematic distributions showing angular and transverse momentum balance for the dijet system. 118
B.3 Shown are kinematic distributions for the leading-$p_T$ jet. 119
B.4 Shown are kinematic distributions for the subleading-$p_T$ jet. 120
B.5 Shown are kinematic distributions for the sub-subleading-$p_T$ jet. 121
B.6 Shown are kinematic distributions displaying the correlation of average dijet invariant mass as a function of leading and sub-leading kinematic variables. 122
B.7 Shown is the average interactions per crossing, $<\mu>$, and the correlation of several quantities with $<\mu>$. 123
B.8 Shown are the angular correlations between the identified dijets and the third jet. 124

C.1 Cone size is optimized for a simple wide jet model applied to $q^*$ resonances and MC QCD background. Sensitivity increase is plotted with respect to narrow jets. 126
C.2 $q^*$ resonances for wide and narrow jets are compared. 127
### Listing of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Relative branching ratios for decay of $q^<em>$. $f_s = f = f' = 1$, $M_{q^</em>} = \Lambda$, and $\alpha_s = 0.11$. Table adapted from [27].</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Properties of quarks are summarized for reference. All of the SM particles, including quarks, were given in Figure 1.1.</td>
<td>17</td>
</tr>
<tr>
<td>2.2</td>
<td>QCD dijet cross-sections are presented and evaluated at $\theta^* = \pi / 2$. Table recreated from [36, 41].</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>A subset of values parameterizing the operation of the LHC in 2015 and its design operation [42, 48, 49].</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>General design resolution and coverage of the various components of the ATLAS detector. Units for $p_T$ and $E$ are GeV. Detector components are approximately uniform in $\phi$. Table is reconstructed from [43].</td>
<td>53</td>
</tr>
<tr>
<td>6.1</td>
<td>Trigger plateau values, as defined by 99.5% trigger efficiency, and off-line cut values. The $m_{jj}$ cut further ensures kinematic trigger bias is removed.</td>
<td>74</td>
</tr>
<tr>
<td>6.2</td>
<td>Cut-flow for 2015 data events with resonance analysis cuts applied.</td>
<td>78</td>
</tr>
<tr>
<td>6.3</td>
<td>Cut-flow for MC-generated events with resonance cuts applied.</td>
<td>78</td>
</tr>
<tr>
<td>6.4</td>
<td>Dijet mass binning used to bin 2015 dataset.</td>
<td>82</td>
</tr>
<tr>
<td>6.5</td>
<td>Presented is the signal width for select simulated $q^<em>$ signal masses. Also, signal width divided by detector resolution is provided. Notice that the $q^</em>$ signals are multiples of detector resolution and will be distributed over multiple $m_{jj}$ bins. Signal fits are provided in Appendix A.</td>
<td>84</td>
</tr>
<tr>
<td>8.1</td>
<td>New and previous ATLAS dijet resonance 95% credibility-level lower limit for QBH and $q^*$ model masses are summarized.</td>
<td>105</td>
</tr>
</tbody>
</table>
To the teachers who embolden us to ask questions, seek answers, and THINK BIG.
Acknowledgments

As one chapter of my life flows rapidly onto the next, I pause to acknowledge the depth of my gratitude for the students and teachers, colleagues and mentors, friends and family, who have guided my journey.

The Harvard Physics Department and ATLAS group have been tremendously supportive and intellectually challenging. Thank you to John Huth (my PhD supervisor) for your guidance and patience. You have entertained my physics questions with zeal and encouraged me to challenge assumptions. Thank you for the conversations – both about physics and the world. Thank you David López Mateos for introducing me to the dijet analysis and the broader ATLAS collaboration. Doug Finkbeiner and Matt Reece, thank you for serving on my dissertation committee. Melissa Franklin and Masahiro Morii, thank you for the helpful professorial advice along the way. Nathan Felt and John Oliver, thank you for your guidance during my time designing a micromegas trigger algorithm. Chris Rogan, thank you for discussions about physics when I was preparing for my qualification exam. And thank you to my fellow graduate students of the Harvard ATLAS group. Emma Tolley, thank you for walking me through jet cleaning. Stephen Chan, thank you for the great conversations, programming help, and reading this dissertation. Tomo Lazovich, thank you for being a great colleague, neighbor, friend, and Geneva roommate; I am honored to be graduating alongside you in two short weeks.

Thank you to Gabriel Facini for leading the dijet resonance search team. I am grateful for your supervision and friendship through the analysis. James Frost, thank you for leading the final leg of the resonance analysis. Thank you to Lydia Beresford, Jeff Dandoy, Meghan Frate, Antonio Boveia, Katherine Pachal, Lene Bryngemark, and the other members of the dijet analysis team. You are a fun, brilliant, and inspiring group of colleagues and friends.

Thank you to those organizations that have supported me through this journey. Thank you to the National Science Foundation for funding part of my graduate education. Thank you to the Marshall Scholars program for allowing me to study in the United Kingdom and for tremendous friendships born from the experience. To the Caldwell Fellows program, thank you for the daily inspiration to “think big.” (And thank you Tomás Carbonell; it may seem like a small gesture, but you introduced me to the Caldwell program... and the Marshall program... and it is no hyperbole to say that both have profoundly impacted my life.) Dr. O., thank you for approving a stipend experience to travel to England ten years ago... which brings me to Todd Huffman. As a college sophomore I emailed you out of the blue and said I would be traveling through Oxford. Thank you for responding, chatting, touring me around the university, and ultimately bringing me back to Oxford for an incredible summer of research that would lay the foundation for my continued study.

Thank you Stephen Reynolds, Albert Young, and Paul Huffman for teaching me undergraduate physics
at North Carolina State University. Thank you to Jennie LaMonte for your continued encouragement and support.

Ben and Sharon Dotger, thank you for being inspiring teachers, mentors, and friends. Ben, I have never forgotten our conversation from freshman year when you kept me after class and explicitly told me I could, and should, consider attending college. You both were an inspiration as the first people I knew to pursue PhDs.

Dave, Gina, and John Wesley Goff, thank you for supporting Allison and me on this journey. John, thank you for letting me use your computer remotely to make plots when I was working under a deadline last August; as promised, here is the public acknowledgment of your moment in experimental particle physics.

Thank you to my Kirkland family for making Cambridge a home. The spirit of this community has defined my time at Harvard. I am eternally grateful to my students, Tom and Verena Conley, Kate Cavell, and my Tutor colleagues for your friendship and support.

Thank you to all the friends who have made this journey possible and enjoyable.

Thank you to my family, who are my inspiration. Your support and encouragement have given me a valuable thing in life: choice. Alicia, thank you for your support and kind heart. Mom, thank you for reminding me to call home more often. Dad, you inspired me to get the education – and then some – that you did not have a chance to pursue. I am still striving to make half the leap you made.

Thank you to my grandmother. She passed away as I was completing the last chapters of this dissertation, and my heart sinks to know that she will not see me in my cap and gown. Grandma, thank you for raising me, for always supporting me, for never doubting in my future, and for being the epicenter of our family.

The final thank you is reserved for my partner, my dearest, Allison. There’s a million things we haven’t done... just you wait – but this one we have accomplished together. Thank you for every day and every moment – iloveyou. (And I love that you will recognize two Hamilton references inserted to mark the year!)
Introduction

In 2015, the Large Hadron Collider (LHC) at CERN produced the first laboratory-based particle collisions at center-of-mass energy $\sqrt{s} = 13$ TeV, thereby opening a new energy regime to experimental inquiry. The most abundant high-energy objects produced at hadron colliders are jets, which are highly collimated sprays of hadrons that offer experimental insight to hard-scatter parton processes of high energy collisions. This abundance is shown in Figure 1 where the cross-section for jet production far exceeds the cross-section for other processes. At leading-order in perturbative quantum chromodynamics, this background is produced in $2 \rightarrow 2$ scattering processes, where the resulting two high-energy jets are collectively referred to as dijets. Standard Model dijet production results in a smoothly falling dijet invariant mass spectrum. Many new massive particle states, including excited quarks and quantum black holes, are theorized to decay with high branching ratio to two strongly charged final-state partons. Such a decay would appear as resonance, or localized “bump,” in an otherwise smooth dijet invariant mass spectrum. This dissertation describes a search for high-mass resonances produced at the LHC and measured with the ATLAS detector. New lim-
Figure 1: Presented are cross-sections for various final-states at hadron colliders as a function of center-of-mass collision energy. The total proton-proton cross-section at the LHC is approximately 80 mb. Figure taken from [1].

Its are set on Beyond the Standard Model physics processes resulting in two high-energy jets. These limits exceed those offered in previous searches at the SppS [2–4], the Tevatron [5, 6], and the LHC [7–19]. The 2015 dijet resonance search results presented in this dissertation, and those produced in a companion analysis that measured the angular production of dijets, were recently published in Physics Letters B [20]. Additional details beyond the scope of this dissertation are discussed in our internal ATLAS document [21].

This dissertation is organized into nine chapters, which are grouped into three parts: Theory Overview, Experimental Measurement, and Dijet Resonance Analysis. Chapter 1 offers a brief overview of the Standard Model of particle physics and introduces two theoretical models that propose further understanding of the observed universe. Chapter 2 discusses the properties of strong interactions at the LHC, as described quantum chromodynamics, and the production of dijets. The LHC and ATLAS detector are described in
Chapters 3 and 4, respectively. Jet measurement and energy calibration at ATLAS is discussed in Chapter 5. The next three chapters discuss details of the 2015 dijet resonance analysis. Chapter 6 covers data collection and event selection protocols to build a dijet invariant mass spectrum, which is fitted and interrogated for evidence of physics beyond the Standard Model in Chapter 7. The resonance search results are used to produce limits on new physics processes, which are presented in Chapter 8. Finally, Chapter 9 seeks to motivate future dijet searches by reviewing the discovery potential of excited quark and quantum black hole benchmark resonant signals.

A Note on Collaboration

ATLAS succeeds as a dynamic international collaboration of dedicated scientists, engineers, and staff. Even small analysis teams devoted to one type of measurement are exceptionally collaborative in their operation – the dijet analysis team was no exception. I was fortunate to work as part of a phenomenal team of brilliant graduate students, postdocs, and faculty. Due to the collaborative nature of our work, explicit divisions of labor are not outlined. This dissertation provides an overview of our 2015 dijet resonance search analysis and presents selected results. Additional results taken from outside the dijet analysis team, including those throughout the larger ATLAS collaboration, are explicitly cited.
Part I

Theory Overview
1

The Standard Model and Beyond

This chapter offers a brief overview of the Standard Model of particle physics and introduces the fundamental particles and forces of nature. But the Standard Model is not the ultimate theory of everything. Many puzzling mysteries of the universe – both big and small in scale – are left conspicuously unanswered, which is the subject of the second half of this chapter.

1.1 The Standard Model

The Standard Model (SM) of particle physics describes the fundamental particles of nature and their electromagnetic, chromodynamic, and weak interactions. Elementary particles of the SM come in two basic types, fermions and bosons, which differ in their intrinsic angular momentum, or spin, quantized in reduced Planck units ($\hbar$). Fermions are half-spin particles and obey Fermi-Dirac statistics, thereby adhering to the Pauli Exclusion Principle; bosons are integer-spin particles and obey Bose-Einstein statistics. In addition to spin, bosons and fermions are also categorized by their charge. The SM particle are summarized
in Figure 1.1. Two categories of fundamental fermions appear in the SM: leptons and quarks. There are twelve different fermions and their distinct antiparticle counterparts, in which an antiparticle is identical to the particle with the exception of possessing mirror charge. For instance, the electron and positron constitute a particle-antiparticle pair and are defined by the same quantum numbers with the sole exception that an electron has $-1e$ electromagnetic charge and a positron has $+1e$ charge. An interesting mystery remains unsolved in defining the antiparticle of neutrally-charged neutrinos – are they their own antiparticle? The twelve elementary particles are often described as belonging to three generations, with each successive generation representing higher mass. These generations are evident in the left-to-right fermion arrangement of Figure 1.1. There are six quark flavors ($u, d, c, s, t, b$), three charged leptons ($e, \mu, \tau$), and three uncharged neutrinos ($\nu_e, \nu_\mu, \nu_\tau$). Quarks interact via strong, electromagnetic, and weak interactions. Charged leptons interact via electromagnetic and weak interactions, while neutrinos only interact weakly.

![Figure 1.1: SM particles are summarized. There are three generations of quarks (purple) and leptons (green) with increasing mass from left to right. Gauge bosons are shown in orange, while the Higgs boson is in the far right corner. Figure is taken from [22].](image)

Gauge bosons are spin-1 mediators of SM interactions. The electromagnetic force governs the motion of electrically charged particles, and it is mediated through the exchange of photons ($\gamma$). A particle’s cou-
pling strength to the electromagnetic force is proportional to its charge. The range of electromagnetic interactions is infinite due to the massless nature of a photon.

Gluons are the mediators of the strong force, which is responsible for binding quarks together to give composite particles, such as protons. Coupling to the strong force is indicated by color charge of which there are three types: red, blue, and green. Each quark will carry a unit of color charge. Eight massless gluons exist as color-anticolor states. While the gluon is massless, additional complexity in the theory of strong interactions limits it to a finite range; the effective range of the strong force is approximately $10^{-15}$m, or the diameter of a proton. This complexity leads to confinement and asymptotic freedom, and is further discussed in Chapter 2. Leptons do not couple to the strong force as they do not carry color charge.

The weak force, which plays an important role in nuclear fission, is mediated by massive $Z$ and $W^\pm$ bosons. Massive mediators limit the range of the weak force to be approximately the size of a nucleus. All left-handed fermions (and right-handed anti-fermions) experience the weak force, which maximally violates parity conservation resulting in only left-handed neutrinos and right-handed antineutrinos being observed in nature. A particle’s coupling strength is given by a quantity known as weak isospin. An intriguing property of the weak force is its unique role as the only force that does not preserve particle flavor. For instance, an up quark can decay to a down quark via the weak force ($d \rightarrow u + e + \bar{\nu}_e$); this is the fundamental process that allows for beta decay where a neutron decays to a proton. This lack of flavor conservation is also visible in the lepton sector; for instance, a muon can decay to an electron via the weak force ($\mu \rightarrow e + \bar{\nu}_e + \nu_\mu$).

An additional boson, the Higgs boson, serves a different role from the gauge bosons described above. The Higgs boson is the only scalar (spin-0) boson in the SM, and it is not a force carrier like the aforementioned bosons. It is a quantum excitation of the Higgs field, which gives rise to masses of the heavy gauge bosons and fermions in nature via the Higgs Mechanism reviewed in [23]. Discovery of the Higgs boson in 2012 [24, 25] constituted the last predicted building block of the SM. (But there is more to be discovered beyond the SM!)

These three interactions are described by quantum field theories. Electromagnetism is described by
quantum electrodynamics (QED); the strong force is described by quantum chromodynamics (QCD); and the weak force is described by electroweak theory, which is a unified description of electromagnetic and weak interactions. The relativistic quantum field theory description of particle physics belongs to the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge group.* A detailed description of the theory can be found in [23].

1.2 Beyond the Standard Model

The experimental frontier of high-energy physics is now firmly into the TeV scale and the SM continues to offer an accurate description of our high-energy universe; however, many questions remain unanswered. Why are there at least two drastically different scales of nature: the $\approx 1$ TeV electroweak scale governing SM interactions and the $\approx 10^{16}$ TeV Planck scale governing gravitational interactions? Might there be extra dimensions of space? More than 95% of the universe is of unknown composition and does not interact electromagnetically – what is the nature of this dark matter and dark energy? Why are there three distinct mass generations of quarks and leptons – are our elementary fermions truly elementary? These questions – and many, many others – constitute the realm of theoretical physics known as Beyond the Standard Model (BSM).

The analysis presented in this dissertation is a search for new phenomena produced at the Large Hadron Collider. Exotic particles and interactions manifesting at the LHC will necessarily couple to SM partons, which means that they can also decay to SM objects, such as jets. While there are many models of BSM physics decaying to two-jet final states – and this analysis takes a model-agnostic approach in its search technique – two common benchmark models are often used in dijet experimental searches: quark compositeness and quantum black holes.

1.2.1 Excited Quark Model

The history of particle physics can be viewed as a reductionist pursuit to describing nature. Each successive expedition to higher energy, or smaller distances, has led to new definitions of “fundamental.” We now know atoms are comprised of nuclei and electrons; nuclei are composed of protons and neutrons; protons

*Subscripts refer to color charge ($C$), left-handed fermions ($L$), and hyper-charge $Y$ related to the $U(1)$ group.
and neutrons are composed of quarks – but are quarks truly elementary objects as the SM assumes? Or are quarks composed of smaller particles, typically referred to as *preons*?

Emission of photons from excited atomic states offers insight to the structure and composition of the atom. Similarly, quark compositeness would be directly observable via excited states, which would present as high mass objects at collider experiments. Consider an excited quark model, as presented by Baur, *et al.* in [26, 27], where spin and weak isospin are set to $\frac{1}{2}$. The Lagrangian description of excited quark interactions with known quarks is given by:

$$\mathcal{L} = \frac{1}{2\Lambda} \bar{q}_R \sigma^{\mu\nu} \left[ g_s f_s \frac{\lambda^a}{2} G^a_{\mu\nu} + g f \frac{\tau^2}{2} \tilde{W}_{\mu\nu} + g' f' \frac{Y}{2} B_{\mu\nu} \right] q_L + h.c. \quad (1.1)$$

where $G^a_{\mu\nu}$, $\tilde{W}_{\mu\nu}$, and $B_{\mu\nu}$ are the SM field strength tensors for the $SU(3)$, $SU(2)$, and $U(1)$ gauge fields. SM gauge structure ($\lambda^a$, $\tau$, and $Y$) and coupling constants ($g_s$, $g$, and $g'$) are assumed. The three free parameters, $f_s$, $f$, and $f'$, are related to the dynamics of quark compositeness and are generally assumed to be equal to one ($f_s = f = f' = 1$). The compositeness scale ($\Lambda$) is set to be the excited quark mass ($M_{q^*}$), which gives a resonant production behavior. (Similar excited quark models are motivated in [28, 29].)

An excited quark can decay to a typical quark through the emission of a $\gamma$, $W$, or $Z$, but the dominant decay process will be through the emission of a gluon, where the decay width is given by:

$$\Gamma(q^* \rightarrow gg) = \frac{1}{3} \alpha_s f^2 \frac{M^3_{q^*}}{\Lambda^2}. \quad (1.2)$$

For first-generation excited quarks ($u^*$ and $d^*$) and assuming $\alpha_s = 0.11$, one finds that the relative branching ratio of this decay path is approximately 85% [27]. The branching ratios for different decay paths are given in Table 1.1. Furthermore, Equation 1.2 gives an intrinsic decay width of approximately:

$$\Gamma \approx 0.04 M_{q^*}, \quad (1.3)$$

where free parameters are assumed to be approximately one. Note that changing these structure constants from unity has a proportional affect on signal width. This resonance width is of particular interest for
excited quarks of a few TeV in mass, as we will see in Chapter 6, because it is wide enough to be resolved in a dijet search at the ATLAS detector.

<table>
<thead>
<tr>
<th>Decay channel</th>
<th>Branching Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u^* \rightarrow u g$</td>
<td>0.85</td>
</tr>
<tr>
<td>$u^* \rightarrow u \gamma$</td>
<td>0.02</td>
</tr>
<tr>
<td>$u^* \rightarrow u Z$</td>
<td>0.03</td>
</tr>
<tr>
<td>$u^* \rightarrow u W$</td>
<td>0.10</td>
</tr>
<tr>
<td>$d^* \rightarrow d g$</td>
<td>0.85</td>
</tr>
<tr>
<td>$d^* \rightarrow d \gamma$</td>
<td>0.005</td>
</tr>
<tr>
<td>$d^* \rightarrow d Z$</td>
<td>0.05</td>
</tr>
<tr>
<td>$d^* \rightarrow d W$</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 1.1: Relative branching ratios for decay of $q^*$. $f_s = f = f' = 1$, $M_{q^*} = \Lambda$, and $\alpha_s = 0.11$. Table adapted from [27].

A $q^*$ process of general interest at modern hadron colliders is $qg \rightarrow q^* \rightarrow qg$. The cross-section for this scattering process is given by [30]:

$$\frac{d\sigma}{dt} = \frac{2\pi\alpha_s}{9M_{q^*}} \frac{\hat{s}}{(\hat{s} - M_{q^*}^2) + \Gamma_{q^*}^2 M_{q^*}^2},$$  \hspace{1cm} (1.4)$$

and corresponds to the Feynman diagram drawn in Figure 1.2. The variable $\hat{s}$ is the square of center-of-

![Figure 1.2: Feynman diagram for a $qg \rightarrow q^* \rightarrow qg$ process.](image)
mass energy for the parton-parton hard-scattering process. Notice the Breit-Wigner resonance form; such resonances would appear as bumps in a dijet mass spectrum as seen in Figure 1.3. The displayed signals are for a hadron collider with center-of-mass energy $\sqrt{s} = 40$ TeV, which is much higher than collisions considered in this dissertation. However, the general signal and background shapes are relevant as they conform to the presented model for quark compositeness and SM physics, respectively.

![Figure 1.3](image)

Figure 1.3: A QCD dijet invariant mass background is modeled for proton-proton collisions at a hypothesized $\sqrt{s} = 40$ TeV collider. Excited quark signals for various masses are drawn as Breit-Wigner resonances. Figure from [27].

**Contact Interactions**

The resonance structure of decay width $\Gamma$ described above assumes the compositeness scale $\Lambda$ is equal to the $q^*$ mass. However, excited quarks may also couple to ordinary quarks through contact interactions resulting from unknown preon interactions mediated by massive particles. Such an interaction can be modeled by a four-fermion effective field theory given by [20, 27]:

$$\mathcal{L} = \frac{2\pi}{\Lambda^2} \left[ \eta_{LL} (\bar{q}_L \gamma^\mu q_L)(\bar{q}_L \gamma_\mu q_L) + \eta_{RR} (\bar{q}_R \gamma^\mu q_R)(\bar{q}_R \gamma_\mu q_R) + 2\eta_{RL} (\bar{q}_R \gamma^\mu q_R)(\bar{q}_L \gamma_\mu q_L) \right] \quad (1.5)$$
where $L$ and $R$ refer to quark chirality and the coefficients $\eta_{LL}$, $\eta_{RR}$, and $\eta_{RL}$ are set to $\pm 1$ or 0 to account for modeling of interference with QCD. The cross-section for a pure contact interaction will scale as $1/\Lambda^4$ and will result in an excess in the high-mass tail of a dijet invariant mass spectrum, where $\sqrt{s} \ll \Lambda$. Figure 1.4 provides a cartoon depiction of resonant and contact interaction signals on a dijet mass distribution. The analysis presented in this dissertation is not sensitive to contact interactions as only resonant-like features are examined. However, the full dijet analysis team worked collectively to produce a companion angular analysis result, which is sensitive to contact interactions. The full analysis is described in [20].

![Figure 1.4](image.png)

**Figure 1.4:** QCD background for dijet production is a smooth, monotonically falling background. The analysis presented in this dissertation is sensitive to resonant signals, or bumps, in the dijet mass distribution, such as the one drawn in (a). If new physics is slightly beyond the reach of our dijet mass distribution it may be modeled as a contact interaction, where the compositeness scale is assumed to be greater than achieved center-of-mass collision energies. Such contact interactions would present as a rising tail in the dijet mass spectrum. A companion analysis, the dijet angular analysis, is designed to search for contact interactions.

### 1.2.2 Extra Dimensions and Quantum Black Holes

The electroweak scale of SM physics and the Planck scale of gravitational interactions are separated by 16 orders of magnitude. This drastic difference in energy scale limits attempts to unify SM and gravitational physics into one quantum field theory framework. A popular solution to this hierarchy problem is the introduction of extra spatial dimensions, which are accessible to gravitational fields but not SM fields; thus, the force of gravity is effectively diluted in our familiar $3 + 1$ universe. The proposed topology and number of these extra dimensions is the subject of many BSM models. This analysis considers two benchmark models: Arkani-Hamed, Dimopoulos, and Dvali (ADD) and Randall-Sundrum (RS) scenarios. ADD
models consider multiple flat extra dimensions and RS models require a single warped extra dimension. The ADD model is motivated below and black hole production at the LHC is introduced as an observable consequence of extra dimensions. Many of the same arguments concerning black hole production in RS models will apply, but motivation of this model is left entirely to other sources, such as [31].

ADD Model Overview

ADD models attempt to lower the true Planck scale to LHC-observable electroweak scale of a few TeV, thereby offering one fundamental scale of nature [32]. Consider a $D = (4 + n)$ dimensional universe with $3 + 1$ familiar dimensions and $n$ extra compact spatial dimensions of radius $R$. Assume that only gravity propagates in the extra dimensions, thereby preserving SM physics. By Gauss' law in $D$ dimensions, two point-like masses $m_1$ and $m_2$ separated by a distance $r << R$ will experience a gravitational potential given by:

$$V(r) \approx \frac{m_1 m_2}{M_D^{n+2} r^{n+1}}, \ (r << R). \quad (1.6)$$

However, if the two masses are separated by a distance $r >> R$, then the gravitational potential is seemingly unaffected by the additional dimensions, and the typical $1/r$ potential is recovered as:

$$V(r) \approx \frac{m_1 m_2}{M_D^{n+2} R^n r}, \ (r >> R). \quad (1.7)$$

The observed Planck scale in 4 dimensions can be identified as $M_{Pl}^2 \approx M_D^{n+2} R^n$. If we assume there is one fundamental scale in nature, the electroweak scale ($M_{EW}$), then with $n \geq 2$ extra dimensions we are probing sub-millimeter distances of gravity.

Black Hole Production in ADD Models

In our familiar $3 + 1$ dimensional universe, the quantum-gravity scale ($M_D$) for microscopic black hole production is far beyond the reach of any conceivable particle accelerator. However, if the universe contains many extra dimensions of space, high-energy particle colliders may be able to produce black holes of a few TeV mass. Following the general argument presented above, the quantum-gravity scale is given by
With the true Planck mass lowered, a black hole in \((4 + n)\) dimensions will have a Schwarzschild radius \(R_S\) given by \([34]\):

\[
R_S = \frac{1}{\sqrt{\pi M_D}} \left[ \frac{M_{BH}}{M_D} \left( \frac{8\Gamma \left( \frac{n+3}{2} \right)}{n + 2} \right) \right]^{\frac{1}{n+1}},
\]

(1.9)

where the extra dimensions are assumed to be large \((R \gg R_S)\). For black hole production to occur, semi-classical arguments require the impact parameter of two colliding partons to be smaller than the Schwarzschild radius. Two colliding partons, \(a\) and \(b\), with center-of-mass energy \(\sqrt{s} = E\) have a cross-section for black hole production given by \([34]\):

\[
\sigma(ab \to BH) \approx \pi R_S^2(E) = \frac{1}{\sqrt{\pi M_D}} \left[ \frac{E}{M_D} \left( \frac{8\Gamma \left( \frac{n+3}{2} \right)}{n + 2} \right) \right]^{\frac{1}{n+1}},
\]

(1.10)

where \(\Gamma\) is a Gamma function. This production considers a simple, non-rotating scenario, where the black hole mass is taken to be the center of mass collision energy; however, the collision energy would certainly need to be much higher than the produced black hole as energy will be radiated as gravitational waves during the collapse process. Therefore, semi-classical black hole production, which requires \(M_{BH} \gg M_D\), is highly unlikely to be observed at the LHC.

If microscopic black holes were produced near the adjusted Planck-scale \((M_{BH} \approx M_D)\) they would exhibit quantum rather than classical behavior. A notable difference between semi-classical black holes and quantum black holes (QBHs) is the latter may carry color charge, which will necessarily be accounted for during production and decay processes. Semi-classical decay models often predict isotropic decay to high-multiplicity particle states via Hawking radiation. Using the methods of \([33, 35]\) one can show that the expected particle multiplicity for a QBH of mass just above the mass threshold for production, where mass threshold is taken to be the Planck mass \((M_{th} = M_D)\), is on the order of unity. For \(M_{BH} < 1.5M_D\), the average particle multiplicity is approximately two. Furthermore, since QBHs can carry color charge and would arise from the interaction of color-charged partons at the LHC, the dominant decay mode is two strongly interacting partons, which would be detected as jets at the LHC.
Notice that the cross-section present in Equation 1.10 is not suppressed by any small coupling factors and grows rapidly with center-of-mass energy. If the LHC successfully probes energy scales greater than $M_D$, then QBHs will be produced at enormous rates. However, parton distribution function effects diminish QBH production cross-sections as black hole mass increases because the probability of an individual parton carrying sufficient energy decreases rapidly at high-energy. The resulting behavior is a resonant-like signal at the tail of a dijet invariant mass spectrum.
2

Dijet Production at the LHC

Quantum chromodynamics (QCD) is the theory of strong interactions. Processes mediated by QCD are dominant at hadron colliders. Section 2.1 offers a brief introduction to QCD, where the notation and approach is a fusion of [22, 23, 36, 37]. Section 2.2 discusses the structure of a proton. Section 2.3 introduces jet production at hadron colliders. Section 2.4 discusses the kinematics and cross-sections of dijet production.

2.1 Quantum Chromodynamics

Strongly interacting particles, called hadrons, are composed of quarks bound together by gluons. The quark model is rooted in the $SU(3)$ flavor symmetry identified by Gell-Mann and George Zweig in the 1960s while classifying low-mass hadrons such as pions and kaons. A subset of quark properties are summarized in Figure 1.1 and also in Table 2.1. Flavor refers to the type of quark: up, down, strange, charm, bottom, top. However, only the first three – and least-massive – quarks were necessary for Gell-Mann’s
<table>
<thead>
<tr>
<th>Flavor</th>
<th>Spin</th>
<th>Charge [e]</th>
<th>Mass [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>up ((u))</td>
<td>(\frac{1}{2})</td>
<td>(\frac{2}{3})</td>
<td>(2.3 \times 10^{-3})</td>
</tr>
<tr>
<td>down ((d))</td>
<td>(\frac{1}{2})</td>
<td>(-\frac{1}{3})</td>
<td>(4.8 \times 10^{-3})</td>
</tr>
<tr>
<td>charm ((c))</td>
<td>(\frac{1}{2})</td>
<td>(\frac{2}{3})</td>
<td>(1.28)</td>
</tr>
<tr>
<td>strange ((s))</td>
<td>(\frac{1}{2})</td>
<td>(-\frac{1}{3})</td>
<td>(9.5 \times 10^{-2})</td>
</tr>
<tr>
<td>top ((t))</td>
<td>(\frac{1}{2})</td>
<td>(\frac{2}{3})</td>
<td>(173)</td>
</tr>
<tr>
<td>bottom ((b))</td>
<td>(\frac{1}{2})</td>
<td>(-\frac{1}{3})</td>
<td>(4.18)</td>
</tr>
</tbody>
</table>

Table 2.1: Properties of quarks are summarized for reference. All of the SM particles, including quarks, were given in Figure 1.1.

initial description of known low-mass bound states; the additional quarks were discovered in the second-half of the 20th century, as particle accelerators reached higher center-of-mass collision energies. “Color” charge, corresponding to the \(SU(3)_C\) gauge group, was introduced to describe observed hadrons of spin-\(\frac{3}{2}\) within the quark model of spin-\(\frac{1}{2}\) partons. For instance, a \(\Delta^{++}\) is composed of three up quarks \((uuu)\) and is a spin-\(\frac{3}{2}\) particle; therefore, all three quarks must be in a spin-up state making the total wave function symmetric, thereby violating Fermi-Dirac spin statistics; but if each quark carries a different color charge, then the total wave function is antisymmetric as expected for a spin-\(\frac{3}{2}\) state. Quarks carry one of three color charges (red, blue, green), while gluons carry one color charge and one anticolor charge.

Direct searches for quarks were conducted by deep inelastic scattering experiments of electrons scattering off protons. These experiments conclusively showed that there are point-like partons comprising the proton. Furthermore, they appear to be loosely-bound, or free, within the proton; that is, quarks exhibit asymptotic freedom at high-energy, or small distance, scales. This property was puzzling as quarks are never observed to be free particles in nature, but are always “confined” within hadrons – a property called confinement.
2.1.1 Lagrangian Formulation of QCD

These observed elements are expressed in the classical QCD Lagrangian density* given by [22]

\[ \mathcal{L}^{\text{classical}}_{\text{QCD}} = \sum_f \bar{q}_{f,a} (i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t^C_{ab} A^C_\mu - m_f \delta_{ab}) q_{f,b} - \frac{1}{4} F^{A}_{\mu\nu} F^{\mu\nu A}, \quad (2.1) \]

where the gluon field tensor \( F^{A}_{\mu\nu} \) is defined by

\[ F^{A}_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g_s f_{ABC} A^B_\mu A^C_\mu. \quad (2.2) \]

Here, repeated indices are summed over and \( \gamma^\mu \) are the Dirac \( \gamma \)-matrices. Quark spinor fields are represented by \( \bar{q}_{f,a} \), have flavor \( f \) and color-index \( a \in [1, ..., 3] \), corresponding to three colors in QCD. Each quark spinor field flavor has a corresponding mass \( m_f \), which is measured by experimental observation. The \( A^A_\mu \) are gluon fields, where \( A \in [1, ..., 8] \) specifies one of the eight types of gluons. (Gluons transform under the adjoint representation of the SU(3) color group, implying that there are \( N^2_{\text{colors}} - 1 = 8 \) gluons from the \( N_{\text{colors}} = 3 \) colors in QCD.) The \( t^C_{ab} \) correspond to eight \( 3 \times 3 \) matrices and serve as the \( SU(3) \) color group generators. Totally anti-symmetric structure constants \( f_{ABC} \) are given by

\[ [t^A, t^B] = i f_{ABC} t^C. \quad (2.3) \]

Finally, \( g_s \) is the gauge coupling constant and is the only free parameter in QCD. This quantity is closely related to the strong coupling constant, which will be discussed later in this section.

Upon expanding Equation 2.1 three types of interaction vertices can be identified in QCD: \( q\bar{q}g \) \( (g_s) \), \( ggg \) \( (g_s) \), and \( gggg \) \( (g_s^2) \), where the coupling strengths are given in parentheses. The corresponding Feynman vertices are drawn in Figure 2.1. Gluon self-interactions result from the last term in Equation 2.2 which makes QCD a non-abelian field theory. For reference, QED is an abelian field theory, thus photons do not self-interact and there is only one interaction vertex to consider similar to the form of Figure 2.1a. As

*Quantizing the Lagrangian and developing Feynman rules for perturbative calculations requires the introduction of two additional terms: a gauge-fixing term and a ghost term. Therefore, the full QCD Lagrangian density is given by \( \mathcal{L}_{\text{QCD}} = \mathcal{L}^{\text{classical}} + \mathcal{L}^{\text{gauge-fixing}} + \mathcal{L}^{\text{ghost}} \).
discussed in the next section, gluon self-interaction gives rise to confinement and asymptotic freedom in QCD.

![Feynman diagram interaction vertices in QCD. Figures 2.1b and 2.1c corresponds gluon-gluon self-interaction.](image)

**Figure 2.1**: Feynman diagram interaction vertices in QCD. Figures 2.1b and 2.1c correspond to gluon-gluon self-interaction.

Individual quarks and gluons are not observed as free particles in nature. Instead, we observe hadrons as color-singlet, or “colorless,” combinations of quarks, antiquarks, and gluons. Color singlets can be formed by combining two quarks – color and its anticolor – or by combining three quarks – each of the three colors or each of three anticolors. Observable quark combinations are classified as mesons ($qar{q}$) and baryons ($qqq$ or $qar{q}ar{q}$).\(^1\) This is equivalent to stating that mesons and baryons are invariant under $SU(3)$ transformations while the individual quarks are not invariant and, hence, not observable in nature.

### 2.1.2 Asymptotic Freedom and Confinement

Asymptotic freedom arises from the quantization of QCD, where perturbative expansion of Feynman diagrams introduces virtual loops of arbitrary energy. Thus, a simple two-point propagator from an initial to final state will include corrections with one-loop, two-loops, etc. The infinite number of loop-corrected diagrams results in a divergent sum, which is circumvented in quantum field theories by a mathematical tool called renormalization. In renormalization of QCD, an arbitrary scale is introduced where a perturbative calculation (pQCD) may be performed if the four-momentum transfer between two particles is greater than the assumed scale of QCD. However, physics must not depend on an arbitrarily chosen scale; therefore, the strong coupling constant has a running energy dependence at leading-order perturbation.

\(^1\)Additional quark combinations are allowed for forming colorless objects. Examples of textit{quark-quasiquartet} ($qqqq$) and textit{pentaqark} ($qqqqq$ or $qar{q}ar{q}ar{q}ar{q}$) combinations have been observed at LHCb [38]. However, the associated particles are exceptionally rare as more stable combinations of mesons and baryons are favored to the exotic, high-quark content particles.
theory given by:

$$\alpha_s(Q^2) = \frac{12\pi}{(11n_c - 2n_f) \ln \frac{Q^2}{\Lambda_{QCD}^2}}$$  \hspace{1cm} (2.4)$$

where $\alpha_s = g_s^2/4\pi$. $n_c$ is the number of colors, $n_f$ is the number of quark flavors, $Q$ is the four-momentum transfer characterizing the energy scale of interest, and $\Lambda_{QCD}$ is the scale of QCD. $\Lambda_{QCD}$ is not a well-defined value, but one can assume it is approximately given by the energy scale characterizing hadrons, which is approximately 200 MeV.

There are three color charges and six flavors in QCD resulting in $\alpha_s$ rapidly decreasing as higher energy scales, or smaller distances, are probed. Figure 2.2 shows experimentally measured values of $\alpha_s$ as a function of $Q$ confirming this behavior. When two quarks are separated by a small distance, such as a fraction of a proton diameter, the QCD coupling constant is weak enough allow for perturbative calculations; this property is called asymptotic freedom. Conversely, as the two partons are pulled apart to greater distances of separation, the coupling strength rapidly increases such that pQCD is no longer applicable. General models, often involving color flux lines connecting the two partons, argue that the potential energy between the two partons grows until it is energetically favorable to create a quark-antiquark pair from the vacuum. The result is one bound state of two quarks becomes two isolated color-neutral bound states of two quarks each. This observed behavior is called confinement.

![Figure 2.2: Presented is a summary of experimental measurements of $\alpha_s$ as a function of the energy scale $Q$. QCD perturbation theory is used to extract a value for the strong coupling constant, and the respective order of perturbation theory is indicated in parentheses. This figure is taken from [22].](image-url)
2.2 Proton Structure

The LHC collides protons, therefore it is important to consider the structure of a proton as visualized in Figure 2.3. Protons are composed of two $u$ quarks and one $d$ quark, referred to as *valence quarks*. Collectively, valence quarks sum to yield the quantum numbers of the proton, including $+1e$ electric charge. The valence quarks are bound into a colorless hadronic state by the exchange of gluons. These gluons will self-interact to produce more gluons or pair-produce additional quarks referred to as *sea quarks*. The proton momentum is distributed among all of these constituents: valence quarks, sea quarks, and gluons.

In collisions at high-energy colliders, it is important to characterize the momentum participatory partons in the hard scattering process. Parton distribution functions (PDFs) give the probability of finding a parton with a given fraction $x_i$ of the total proton momentum $P$ (i.e., $p_i = x_i P$) at an energy scale $Q$. At low-energy scales, or distances comparable to the proton diameter, valence quarks share most significantly in the proton’s total momentum. At high-energy scales, or short distance scales, a rich complexity of sea quarks manifests. These sea quarks will carry low momentum fractions and the probability of carrying a specific momentum fraction will increase with $Q$. Valence quarks will typically carry high momentum fractions with the probability of carrying a specific fraction will decreasing with increasing $Q$. These behaviors are evident in two example PDFs presented in Figure 2.4.

![Figure 2.3: A proton is composed of two $u$ quarks and one $d$ quark, called valence quarks, bound together by gluons. There are also sea quarks and additional gluons that arise from pQCD loop corrections.](image-url)
2.3 **Overview of a Collision**

High-energy collisions of two protons are complex events often resulting in showers of low-energy hadrons, which are measured by particle detectors. A collision, such as the one simulated in Figure 2.5, can generally be summarized into four components:

- **Partons approach**: Two protons approach the collision point. Each proton is comprised of partons that each carry a fraction of the total proton momentum. Before the hard-scattering process occurs, *initial state radiation* may be emitted by the incoming partons.

- **Hard scattering**: One parton from each incoming proton interacts to produce high-energy final state partons. These partons may radiate additional partons to create a *parton shower*. This process in calculable using perturbative QCD.

- **Hadronization**: Non-perturbative QCD effects dominate the evolution of final state partons at large distances. As the partons fly apart, colorless states (i.e., hadrons) will form from the by-products of the hard-scatter event. Quark-antiquark pairs are generated from the vacuum as bind-

---

**Figure 2.4**: Shown are two PDFs at (a) low-energy scales and (b) high-energy scales [1].
ing energy increases with distance between two outgoing partons; these pairs provide the additional color-charged partons to form many colorless states from the original outgoing hard-scatter partons. This formation of hadrons, mostly pions and kaons, results in highly collimated sprays of particles, called jets. Careful energy deposition measurements in a detector allow for the calculation of a jet’s four-vector, which corresponds to the associated outgoing parton from the hard interaction point.

- **Underlying Event**: Hard-scatter processes generally involve two partons; however, the additional partons, called *spectator partons*, will undergo soft (non-perturbative) interactions. These interactions are included in the final event cross-section and constitute a source of background for the hard-scatter process of interest.

![Figure 2.5: Present is a simulated $pp \rightarrow ttH$ event produced by an event generator. Incoming partons are shown in blue. The hard scattering process is represented by the large red filled circle and the resulting parton shower is shown in red. A secondary interaction, or underlying event, is represented in purple. Light green blobs represent the initial hadronization and dark green blobs correspond to hadron decay. Note that the incoming protons are represented in green as they are hadrons. Additionally, yellow lines correspond to photon emission. Figure is taken from [39].](image)
Detector level physics objects, such as jets, are used to reconstruct the hard-scatter events, which may contain interesting physics. In the case of Figure 2.5, the hard-scatter process of interest is $pp \rightarrow ttH$, for which a production method is diagrammed in Figure 2.6.

![Feynman diagram for the $gg \rightarrow ttH$ process simulated in Figure 2.5.](image)

**Figure 2.6:** Feynman diagram for the $gg \rightarrow ttH$ process simulated in Figure 2.5.

### 2.4 Dijet Production at Hadron Colliders

A common event at high-energy hadron colliders is *dijet* production where parton-parton hard-scatter produces two high-$p_T$ partons. These partons will then undergo parton showering and hadronization to produce two observable jets, referred to as a dijet system. Momentum conservation requires these jets to be equal-and-opposite in the center-of-mass frame of the scattering process. In the lab frame, the longitudinal momentum of each incoming parton is rarely balanced, but both partons will carry approximately zero transverse momentum. Therefore, the two jets are the back-to-back in $\phi$ and balanced in $p_T$. Of course, this represents the ideal dijet scenario. In practice there will be initial and final state radiation, detector reconstruction efficiencies, underlying event contamination, pile-up, and a host of other issues that complicate the balance of a dijet event.

#### 2.4.1 Kinematics

Consider two incoming protons of momentum $P$ where each has a parton participating in a hard scattering event. Parton momentums will be a fraction of the total proton momentums given by $x_iP$, where $0 \leq x_i \leq 1$.
$x_i \leq 1$. The parton four-momenta are given by

\begin{align*}
p_1^\mu &= (x_1 P, 0, 0, x_1 P) \tag{2.5} \\
p_2^\mu &= (x_2 P, 0, 0, -x_2 P). \tag{2.6}
\end{align*}

The parton-parton center-of-mass energy is given by

\begin{align*}
\hat{s} &\equiv (p_1 + p_2)^2 \tag{2.7} \\
&= 4x_1 x_2 P^2 \tag{2.8} \\
&= x_1 x_2 s, \tag{2.9}
\end{align*}

where $s = 4P^2$. The rapidity of the system is given

\begin{align*}
y &\equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{2.10} \\
&= \frac{1}{2} \ln \frac{x_1}{x_2}. \tag{2.11}
\end{align*}

Rapidities are additive under longitudinal Lorentz boosts. Therefore, differences in rapidities are Lorentz invariant under longitudinal boosts. This property is useful in translating observable quantities between the center-of-mass and the laboratory frames. Consider the $1 + 2 \rightarrow 3 + 4$ scattering process shown in Figure 2.7. A balanced dijet event will have back-to-back jets, thus the two jets will have equal and opposite rapidities, $y_3^{\text{CM}} = y^*$ and $y_4^{\text{CM}} = -y^*$. These rapidities can be translated to the lab frame by providing a longitudinal Lorentz boost $y_B$, where $y_B = \frac{1}{2} \ln \frac{x_1}{x_2}$ by Equation 2.11. The lab rapidities are then given by:

\begin{align*}
y_3 &= y^* + y_B \tag{2.12} \\
y_4 &= -y^* + y_B. \tag{2.13}
\end{align*}
Figure 2.7: Cartoon illustrating a $1 + 2 \rightarrow 3 + 4$ scattering process. The lab frame (left) is boosted to the center-of-mass frame (right) where the dijets are back-to-back.

Rearranging these equations, the center-of-mass rapidities and necessary boost to translate between the two frames can be extracted as

\[
y_B = \frac{y_3 + y_4}{2}, \quad y^* = \frac{y_3 - y_4}{2}.
\]

Furthermore, it is worth noting that $y^*$ is related to the center-of-mass scattering angle $\theta^*$ by

\[
y^* = \frac{1}{2} \left( \frac{1 + \cos \theta^*}{1 - \cos \theta^*} \right).
\]

The invariant mass of the dijet system is equal to the center-of-mass energy of the parton-parton collision (Equation 2.9). Using the variables $y^*$ and $y_B$, the longitudinal momentum fractions of our incoming partons can be expressed as

\[
x_1 = x_T \cosh y^* e^{y_B},
\]

\[
x_2 = x_T \cosh y^* e^{-y_B},
\]
where \( x_T \equiv 2p_T/\sqrt{s} \). Therefore, using Equation 2.9 the invariant mass of the dijet system can be expressed as

\[
m_{jj} = 2p_T \cosh y^*,
\]

(2.19)

where the partons were assumed to be massless, i.e. \(|p| \gg m\). Thus, three measurable quantities in the lab frame \((p_T, y_3, y_4)\) are sufficient to calculate a dijet invariant mass in this simple scenario.

### 2.4.2 Dijet Cross-section

When discussing \(1 + 2 \rightarrow 3 + 4\) scattering processes Mandelstam variables are conventionally introduced. The variables are defined as

\[
\hat{s} \equiv (p_1 + p_2)^2,
\]

(2.20)

\[
\hat{t} \equiv (p_1 - p_3)^2 = -\frac{1}{2} \hat{s}(1 - \cos \theta^*),
\]

(2.21)

\[
\hat{u} \equiv (p_2 - p_3)^2 = -\frac{1}{2} \hat{s}(1 + \cos \theta^*).
\]

(2.22)

The cross-section for a proton-proton collision producing two partons is given by [36]

\[
\frac{d^3\sigma}{dp_T^2dy_3dy_4} = \frac{1}{16\pi^2s^2} \sum_{abcd} f_a(x_1, Q^2)f_b(x_2, Q^2) |\mathcal{M}(ab \rightarrow cd)|^2 \frac{1}{1+\delta_{cd}}
\]

(2.23)

where matrix elements \(|\mathcal{M}|\) are averaged and summed over initial and final color and spin states. The term \(\frac{1}{1+\delta_{cd}}\) is included to account identical of final-state parton processes. A full cross-section calculation will include all \(2 \rightarrow 2\) processes. A subset of leading-order diagrams are drawn in Figure 2.8. Matrix elements are calculated in Table 2.2 and evaluated at \(\theta^* = \pi/2\). Notice that \(gg \rightarrow gg\) scattering is the most likely QCD process at leading order. As seen in Figure 2.4, gluons will provide the dominant cross-section contribution for low-\(p_T\) dijet events, while quarks will be the most likely participants in high-\(p_T\) dijet events. Figure 2.9 exhibits this behavior; notice that the highest \(m_{jj}\) events primarily originate from events including at least one incoming quark.
Figure 2.8: A subset of leading-order Feynman diagrams for QCD scattering are shown. This figure is taken from [40] and is an adaptation of similar figure appearing in [41].

| Process | \( \sum |M^2| / g_s^4 \) | Value at \( \theta^* = \pi/2 \) |
|---------|-----------------|-----------------|
| \( qq' \rightarrow qq' \) | \( \frac{4}{9} \frac{s^2 + \hat{u}^2}{t^2} \) | 2.22 |
| \( q\bar{q}' \rightarrow q\bar{q}' \) | \( \frac{4}{9} \frac{s^2 + \hat{u}^2}{t^2} \) | 2.22 |
| \( qq \rightarrow qq \) | \( \frac{4}{9} \left( \frac{s^2 + \hat{u}^2}{t^2} + \frac{s^2 + \hat{t}^2}{\hat{u}^2} \right) - \frac{8}{27} \frac{s^2}{\hat{u}^2} \) | 3.26 |
| \( q\bar{q} \rightarrow q'\bar{q}' \) | \( \frac{4}{9} \frac{s^2 + \hat{u}^2}{s^2} \) | 0.22 |
| \( q\bar{q} \rightarrow q\bar{q} \) | \( \frac{4}{9} \left( \frac{s^2 + \hat{u}^2}{t^2} + \frac{t^2 + \hat{u}^2}{\hat{u}^2} \right) - \frac{8}{27} \frac{s^2}{\hat{u}^2} \) | 2.59 |
| \( q\bar{q} \rightarrow gg \) | \( \frac{32}{27} \frac{t^2 + \hat{u}^2}{t \hat{u}} - \frac{8}{3} \frac{t^2 + \hat{u}^2}{s^2} \) | 1.04 |
| \( gg \rightarrow q\bar{q} \) | \( \frac{1}{6} \frac{t^2 + \hat{u}^2}{t \hat{u}} - \frac{3}{8} \frac{t^2 + \hat{u}^2}{s^2} \) | 0.15 |
| \( gg \rightarrow gg \) | \( -\frac{4}{9} \frac{s^2 + \hat{u}^2}{s \hat{u}} + \frac{\hat{u}^2 + \hat{t}^2}{t^2} \) | 6.11 |
| \( gg \rightarrow gg \) | \( \frac{9}{2} \left( 3 - \frac{t \hat{u}}{s^2} - \frac{\hat{u} \hat{t}}{t^2} - \frac{s \hat{t}}{\hat{u}^2} \right) \) | 30.4 |

Table 2.2: QCD dijet cross-sections are presented and evaluated at \( \theta^* = \pi/2 \). Table recreated from [36, 41].
2.4.3 Signal from Background

Many resonant signal models, including the benchmark excited quark model presented in Chapter 1, involve $s$-channel processes. Recalling Equation 2.22, $s$-channel processes decay isotropically in $\theta^*$. For example, Equation 1.4 shows the excited quark benchmark signal process $qg \rightarrow q^* \rightarrow gq$ has an associated differential cross-section with no dependence on $\theta^*$. However, SM QCD dijet production is dominated by $t$-channel processes, as seen in Figure 2.8 and Table 2.2; therefore, QCD production will peak in the forward direction for high-mass events due to the pole in $t$-channel production. Exploiting the difference angular production of signal models and QCD, dijet resonant mass searches maximize signal-to-background ratio by isolating central decay angle events. This is accomplished by requiring a $y^* \propto \frac{1}{1-\cos \theta^*}$ event selection cut in the dijet resonance analysis. Further detail for the optimization of a $y^*$ cut to reject $t$-channel QCD while admitting $s$-channel production is presented in Chapter 6.
Part II

Experimental Measurement
The Large Hadron Collider (LHC) and its associated experiments are machines of immense complexity requiring the dedicated work of thousands of international collaborators. Located at the European Organization Nuclear Research (CERN) in Geneva, Switzerland, the LHC is a 26.7 km particle accelerator designed to circulate two 7 TeV proton beams, one clockwise and the other counterclockwise; the beams are allowed to intersect at four specific locations to produce proton-proton collisions. From high-energy collisions, scientists explore the fundamental structure of the universe.

The material presented in this chapter will offer a brief overview of the LHC, which is described in detail by [42].

3.1 Experiments

Large particle detectors, or experiments, are constructed at each of the LHC’s four beam-crossing points. The four primary experiments are: ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid),
LHCb (Large Hadron Collider beauty), and ALICE (A Large Ion Collider Experiment). ATLAS [43] and CMS [44] are large, general-purpose detectors built to study SM and proposed BSM physics processes by examining proton-proton and lead-lead collisions. LHCb is designed to study CP-violation in interactions of b-hadrons [45]. ALICE is optimized to study quark-gluon plasmas created in lead-lead nuclei collisions [46].

3.2 Injection Chain

The LHC is the ultimate accelerator in the multi-tiered CERN accelerator complex shown in Figure 3.1. Protons originate from hydrogen gas that is passed through an electric field of sufficient strength to strip the atoms of their electrons, thus leaving only protons. The protons are accelerated to 50 MeV by a linear series of radio frequency cavities that compose the LINAC 2. They are then boosted to 1.4 GeV by the Proton Synchrotron Booster (PSB), which is composed of four concentric synchrotron rings. The proton beam then feeds into the Proton Synchrotron (PS) for additional acceleration to approximately 25 GeV.† Protons are then injected into the Super Proton Synchrotron (SPS), which is a 7 km-circumference accelerator used to accelerate protons to 450 GeV.‡ From the SPS, protons are injected into the LHC for further acceleration to a design energy of 7 TeV per beam.§

Initially, the two LHC proton beams exhibit a structure corresponding to the injection frequency of the SPS. At this point, the transverse and longitudinal spread of protons in the beams is large and the collision rate is small due to low proton density. These low-density beams at the LHC are locally compressed into a bunch structure. The LHC is designed to operate with 2808 of 3564 available bunch regions filled, while the unfilled bunches are useful for further beam injection and beam abort operations.

*Three smaller experiments share the interaction points with the aforementioned experiments: TOTEM (TOTal Elastic and diffractive cross-section Measurement), LHCf (Large Hadron Collider forward), and MoEDAL (Monopole and Exotic Detector At the LHC).
†When the PS was commissioned in 1959 it was the highest-energy particle accelerator in the world.
‡As operated in 1983, the SPS accelerated protons and antiprotons for collision at the UA1 and UA2 experiments, which made the first observations of W and Z bosons.
§The LHC resides in the tunnel previously occupied by the Large Electron-Positron Collider (LEP) and is, therefore, 26.7 km in circumference. LEP’s physical size was motivated by physics goals as a Z factory and accelerator technology available in the late-1980s. This circumference, and the capability of superconducting magnet technology, ultimately set the design collision energy of the LHC.
3.3 Luminosity Measurement

Collision rates at the LHC depend on proton beam properties and the cross-section of proton interactions. The beam properties impacting collision rates at the LHC are collectively expressed as instantaneous luminosity $\mathcal{L}$. The rate of collisions $\mathcal{R}$ can then be expressed as $\mathcal{R} = \mathcal{L}\sigma$, where $\sigma$ is the cross-section for proton-proton collisions. The instantaneous luminosity is explicitly given by:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_t}{4\pi \epsilon_{\beta^*}} F,$$  \hspace{1cm} (3.1)

where $n_b$ corresponds to the number of proton bunches; $N_b^2$ is the number of protons per bunch; $f_{\text{rev}}$ is the revolution frequency of the accelerator; and $\gamma_t$ is the relativistic Lorentz factor. The normalized beam emittance $\epsilon_{\beta^*}$ serves as a measure of proton spread in position and momentum phase space. $\beta^*$ is the beta
function value for the beam at the interaction point. $\beta$ and $\epsilon$ give the transverse beam size $\sigma = \sqrt{\epsilon\beta}$, where $\sigma$ is a Gaussian width describing the beam size. $F$ is a geometric factor related to the beam crossing angle.

The total luminosity can be calculated as the time integral of the instantaneous luminosity. Notice that the units of luminosity are $1/\text{(Area)}$. The amount of data provided by the LHC is often reported in units of inverse femtobarns ($\text{fb}^{-1}$), where the reader will recall that 1 barn $= 10^{-24}$ cm$^2$. In 2015, the LHC’s peak instantaneous luminosity was $5.0 \times 10^{33}$ cm$^{-2}$ s$^{-1}$ and a total luminosity of 4.2 fb$^{-1}$ was delivered to the ATLAS experiment, as shown in Figure 3.2.

---

Figure 3.2: Shown are 2015 peak instantaneous luminosity per proton fill at the LHC (a) and cumulative luminosity delivered to the ATLAS experiment (b) as a function of time. Figures taken from [48].

### 3.4 2015 Running Conditions

The LHC collided 6.5 TeV beams for the first time in 2015 for a center-of-mass collision energy of $\sqrt{s} = 13$ TeV. The machine was primarily operated with a 25 ns bunch spacing and the peak instantaneous luminosity was $5 \times 10^{33}$. Additional 2015 and design operational parameters are presented in Table 3.1.

---

*A 50 ns bunch spacing was implemented during the summer of 2015; the associated D5 run period constituted less than 1% of the total luminosity used in this dissertation.*
<table>
<thead>
<tr>
<th>Parameter</th>
<th>2015</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam energy [TeV]</td>
<td>6.5</td>
<td>7</td>
</tr>
<tr>
<td>Max. instantaneous luminosity [cm$^{-2}$s$^{-1}$]</td>
<td>$5 \times 10^{33}$</td>
<td>$10^{34}$</td>
</tr>
<tr>
<td>Number of bunches</td>
<td>1825</td>
<td>2808</td>
</tr>
<tr>
<td>Bunch spacing [ns]</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Max. protons per bunch</td>
<td>$1.2 \times 10^{11}$</td>
<td>$1.15 \times 10^{11}$</td>
</tr>
<tr>
<td>$\beta^*$ [cm]</td>
<td>80</td>
<td>55</td>
</tr>
<tr>
<td>Pileup ($\langle \mu \rangle$)</td>
<td>13.7</td>
<td>–</td>
</tr>
</tbody>
</table>

**Table 3.1:** A subset of values parameterizing the operation of the LHC in 2015 and its design operation [42, 48, 49].
The ATLAS Experiment

The ATLAS (A Toroidal LHC ApparatuS) experiment is a 7-kiloton general purpose detector designed to study proton-proton and nuclei-nuclei collisions at the LHC. The detector is approximately symmetric about the interaction point, where the two notable symmetries are forward-and-backward along the beamline and rotationally about the beam axis. The detector was constructed to further explore the Standard Model of physics, including the highly anticipated discovery of the Higgs boson in 2012, and explore exotic models of new physics that may be revealed at the TeV scale. Figure 4.1 diagrams a full overview of the ATLAS detector and its various specialized detector systems.

Three general principles guide detector construction at a collider experiment: hermetic coverage, lateral granularity, and full energy deposition. A hermetic detector, often referred to as a 4π detector, attempts to fully cover the area around the interaction point so as to capture all scattering and decay angles. There are many physical limitations in designing a hermetic detector; most notably, there are required gaps for beam pipes and electrical, gas, and cryogenic services to the detector systems. The second guiding principle,
lateral granularity, refers to a detector’s ability to track energy deposits from individual particles over the ideal $4\pi$ coverage. Finally, while hermetic coverage generally refers to tracking every decay path, detector depth is important to fully contain the energy of a collision’s by-products. In an ideal detector nothing will escape full containment. Analyzing and understanding collisions at the LHC generally relies upon these three detector design principles to reconstruct physics events through detailed measurement of particle trajectories and energy deposition.

ATLAS is divided into detector components that specialize in a general measurement type, such as tracking or calorimetry. The inner detector is situated closest to the beamline and designed to track particles emitted from the interaction point. High resolution tracking of particles in the inner detector allows for vertex identification of decay points immediately following the collision and precise measurements of particle momentum. This detector is composed of three components: the pixel detector, semiconductor tracker, and transition radiation tracker. Each of these detector types is described in Section 4.3. Enclosing the inner detector is the electromagnetic calorimeter, which is optimized for measuring electromagnetic
showers, and the hadronic calorimeter, which is optimized for measuring hadronic showers. Described in Section 4.4, these calorimeters are the most crucial detector components for this analysis as they provide the primary measurement inputs for jet reconstruction. The final detection system is the muon spectrometer which encloses the full detector. The muon spectrometer reconstructs tracks and measures the momentum of muons, which are generally the only interacting particles that will survive the inner detector layers. This final sub-detector is described in Section 4.5.

Additionally, a large magnet system is deployed to allow momentum measurements in the inner detector and muon spectrometer. Tracking within magnetic fields also allows for determination of particle charge and aids in particle identification. The inner detector is surrounded by a 2 T solenoid detector, which provides a longitudinal magnetic field to bend particle trajectories around the beamline. Beyond the calorimeters, three toroidal magnet systems, from which ATLAS derives its name, bend particle trajectories along the beam axis. The magnet system is described in Section 4.2.

Each detector is generally divided into two components: barrel and end-cap. Barrel components are cylindrical with the symmetry axis aligned along the beam pipe. End-cap components, of which there is one for each end of the barrel, cover the plane transverse to the beamline.

The material presented in this chapter is a summary of technical information published in [43].
4.1 Coordinate System and Nomenclature

The coordinate system implemented by ATLAS is a right-handed Cartesian coordinate system, shown in Figure 4.2a, with origin at the interaction point of the two proton beams. The $z$-axis is aligned along the beamline with the positive direction towards Point 8 of the LHC ring, i.e. towards the city of Geneva. The $y$-axis is directed upwards towards the surface. The $x$-axis points to the center of the LHC ring.

![Coordinate System](image)

**Figure 4.2:** Coordinate system for the ATLAS detector is shown. A right-handed Cartesian coordinate system is shown in (a) with the origin at the interaction point, the $x$-axis pointing towards the LHC ring center, the $y$-axis pointing upwards to the surface, and the $z$-axis pointing along the beamline. A more typical cylindrical system is shown in (b) and measures pseudorapidity $\eta$, radius $r$, and azimuthal angle $\phi$.

Common practice throughout ATLAS is to implement a cylindrical transformation of the above coordinate system. The cylindrical system, shown in Figure 4.2b, is a natural choice given the near symmetry of the detector about the $z$-axis. Furthermore, there is a forward-backward symmetry about the interaction point. The azimuthal angle $\phi$ is measurement from the $+x$-axis and is taken to be within $[-\pi, +\pi]$, where the positive semi-cylinder is the top-half of the detector. Radii are measured in the plane transverse to the beamline and given by $r = \sqrt{x^2 + y^2}$. The polar angle $\theta$ is measured from the beam axis, but it is rarely used; instead, ATLAS communications rely on pseudorapidity $\eta$, which is given by:

$$\eta = -\ln \tan \frac{\theta}{2}. \quad (4.1)$$

This variable has a few interesting qualities. Pseudorapidity runs from 0 in the $xy$-plane to $\pm \infty$ along the
beamline. Therefore, uniform steps in $\eta$ represent decreasingly sized steps in geometrically-familiar units of $\theta$ as one moves from the detector $xy$-plane to the beamline. This is an important feature to note as detector segmentation will often be quoted in $\eta$ and $\phi$. A detector component uniformly segmented in $\eta$ will have increasingly finer segmentation in $\theta$ as one approaches the beamline. Another interesting feature of pseudorapidity for hadron colliders is that pion production from inelastic collisions is approximately uniform in $\eta$ for the coverage range of the ATLAS detector.

Recall from Chapter 2 that rapidity $y$ is often used to describe the trajectory of massive objects, such as jets. An object of energy $E$ and longitudinal momentum $p_z$ has rapidity:

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

Differences between rapidities are Lorentz invariant under longitudinal boosts; therefore, the difference between two jet rapidities, as measured in the lab frame, can be used as an indicator of how central a dijet event is in its rest frame. In the massless limit, an object’s rapidity is given approximately by its pseudorapidity ($y \approx \eta$). Therefore, pseudorapidity is often used as a substitute for rapidity for reconstruction of massive objects, such as jets, due to ease of measurement. The massless limit approximation is suitable for this analysis as the search space is limited to massive dijet objects reconstructed from two high-$p_T$ jets, where $p_T^\text{jet} \gg m^\text{jet}$.

Other variables, introduced elsewhere in this dissertation but offered here for completion, are a distance measure between tracks, $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$, and transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$. As mentioned in Chapter 2, quantities defined transverse to the beam are often introduced because initial parton momentums along the beamline $p_z$ are unknown, but transverse momentums are approximately zero before collision. Careful measurements of transverse momentum (energy) can be used to indicate any missing momentum (energy), which could be attributed to neutrinos or exotic particles that escape detection in our detector.
4.2 Magnet System

The ATLAS detector relies on two primary magnet systems, drawn in Figure 4.3a, to bend the paths of charged particles as they traverse the magnetic field region. This bending allows charge identification and particle momentum measurements. The inner detector is served by a 2 T NbTi magnet operating at 4.5 K. The solenoid axis is aligned with the +z detector axis such that charged particle trajectories are bent in $\phi$, and the field returns through steel structure supports of the hadronic calorimeter.

The muon spectrometer relies upon a system of three air-core toroidal magnets. Each of the three toroids consists of eight coils aligned radially in a symmetric pattern about the beam axis. End-cap toroidal assemblies are rotationally offset from the barrel toroidal assembly by $\pi/8$ radians, which serves to optimize bending power at the interface of the barrel and end-cap assemblies. Figure 4.3b shows the available bending power in the toroidal magnet system.

![Image of ATLAS magnet system](image)

**Figure 4.3:** (a) The ATLAS magnet system, shown in orange, consists of one 2 T solenoid servicing the inner detector and large toroidal magnet system servicing the muon spectrometer. The colorful cylindrical layers encasing the solenoid represent the calorimeter system. (b) Shown is the approximate field integral for perfectly straight tracks traversing the toroidal system in one octant. The red curve maps the field for $\phi = 0$ and the black curve maps the field integral for $\phi = \pi/8$. Figures are taken from [43].
4.3 Inner Tracker

The inner detector (ID) is a tracking detector designed to reconstruct particle trajectories and identify primary and secondary event vertices. The ID offers coverage for $|\eta| < 2.5$ and a transverse momentum resolution of $\sigma_{p_T}/p_T = 0.05\% p_T \text{GeV} \oplus 1\%$. Additionally, the ID offers a transverse impact parameter resolution down to 11 $\mu$m for central, high-$p_T$ particles. Designed with three independent and distinctive components, a typical particle traversing the ID will leave at least four hits in the pixel detector, eight hits in the semiconductor tracker, and approximately 36 hits in the transition radiation tracker. An overview of the ID is shown in Figures 4.4 and 4.5.

![Figure 4.4: Cut-away of the ATLAS inner detector. The three independent sub-detectors are the pixel detector, semiconductor tracker, and transition radiation tracker. Figure is taken from [43].](image)

4.3.1 Pixel Detector

Particles emerging from the interaction point first traverse the pixel detector [43, 50]. The pixel detector consists of three layers of semiconductors used for high-resolution tracking and vertex identification of charged particles. Detector layers are composed of pixel modules of size $19 \times 63$ mm$^2$, each containing 47232 silicon pixels of nominal detection area $50 \times 400 \mu$m$^2$ that are bump-bonded to a readout system. Sensors are approximately 250 $\mu$m thick. There are 1744 pixel modules spread over the three layers.
of pixel detection resulting in approximately 80.4-million individual readout channels. The intrinsic position resolution is 10 μm in r-φ and 115 μm in z (r) for the barrel (end-caps).

The three barrel layers of the pixel detector are shown in Figure 4.5. Barrel layers are located at radii 50.5μm, 88.5μm, and 122.5μm. For each layer, the pixel modules are rotated 20° with respect to the tangent of the barrel cylinder, thereby allowing for full φ coverage. End-cap layers are placed ±495 mm, ±580 mm, and ±650 mm from the interaction point. Each of the end-cap disks has two closely spaced layers of pixel modules that are offset by a 3.75° rotation around the beam axis, thereby allowing for full φ coverage.

The entire system is cooled to −7°C to reduce the effect of radiation damage. The operating voltage is

---

*Approximately 10% of pixels have a larger detection area of 50 × 600μm². Additionally, there are 47232 pixels on each module, but only 46080 are readout channels because the first four pixels of each column on the readout chip are grouped together to adhere to space constraints.*
between 150V and 600V, where higher voltages are used to compensate for radiation damage. To further protect the pixel detector from radiation damage, it is initially left inactive until the LHC achieves fully stable beams.

**Insertable B-Layer**

Due to its close proximity to the beamline, the pixel detector operates in an exceptionally high-radiation environment. Following Run I, approximately 5% of the pixel modules were inactive [51]. This operation depreciation was expected in the initial ID design [43] that called for the inner pixel layer, referred to as the b-layer, to be replaced after three years of operation. Therefore, a fourth and innermost layer, called the insertable b-layer (IBL) [52], was commissioned and installed during the long shutdown between Runs I and II. The upgrade included replacing the existing beam pipe, which occupied the region $29 < r < 36$ mm, with a thinner beam pipe, which now occupies the region $25 < r < 29$ mm. Decreasing the beam pipe footprint allowed enough space for the IBL, which has an effective sensitivity radius of 33.25 mm. The IBL preserves designed tracking performance and extends the lifetime of the pixel detector. Additionally, the close proximity of the IBL to the beamline allows for improved impact parameter resolution, which directly benefits vertex identification and b-tagging performance.

### 4.3.2 Silicon Microstrip Detector

The next detector system traversed by a particle emitted for the interaction point is the semiconductor tracker (SCT), which provides at least eight precision position measurements. While additional layers of silicon pixels would be ideal for position resolution, the increased cost and complexity of a larger pixel detector limits design. Instead, the SCT uses silicon microstrips placed in small-angle stereo. The detection elements of the SCT are silicon strip sensors of area $6.36 \times 6.40$ cm$^2$, each containing 768 readout strips with a pitch of 80 $\mu$m. Operation voltages are set to 150 V, but can be adjusted up to 350 V to compensate for radiation damage over the a 10-year lifespan of the module. The full SCT implements 6.2 million readout channels. Resolutions are approximately 17 $\mu$m in $r$-$\phi$ and 580 $\mu$m in $z$ ($r$) for the barrel (end-cap) [43].
4.3.3 Transition Radiation Tracker

The transition radiation tracker (TRT) provides approximately 36 additional hits to be used in mapping a particle’s trajectory from the interaction point. The TRT is composed of 4 mm polyamide drift tubes, often referred to in the TRT as straws, filled with a gas mixture of 70% Xe, 27% CO₂, and 3% O₂. The space between the tubes is filled with a mesh of polypropylene foil and flooded with CO₂. Straw tubes are oriented parallel to the beam axis in the barrel and radially in the end-cap region. This geometry offers $r$-$\phi$ information with an intrinsic accuracy of 130 $\mu$m per straw. Hits in the TRT contribute significantly to ID momentum measurements because the sheer number of points measured exceedingly compensates for less resolution than other components [43].

The TRT is also used to distinguish between electrons and charged pions. When relativistic charged particles traverse media with different dielectric constants, electromagnetic radiation is emitted at the boundary zone. The emitted radiation, typically keV photons, is proportional to the Lorentz factor of the particle. (For reference, $\gamma \approx E/m$.) Therefore, an electron will emit approximately 280 times more energy than a pion of similar incident energy. The emitted x-rays are absorbed by Xe gas in the straw tubes, which in turn emits electrons to further ionize the tube and be collected at a central wire anode.

4.3.4 Material Accounting for the Inner Detector

Tracking detectors measure a charged particle’s trajectory while interfering with the particle as little as possible. A method to quantify the likelihood of significant interference is accounting for the amount of material traversed by a particle in the tracker volume. Figure 4.6 shows the amount of material comprising the inner detector as traversed by a particle emitted from the interaction point. Material thickness is measured in electromagnetic and hadronic interaction lengths, respectively labeled $X_0$ and $\lambda$. 
Figure 4.6: Displayed is an accounting of material in the inner detector as measured in radiation ($X_0$) and nuclear interaction ($\lambda$) lengths as a function of $|\eta|$ and averaged over $\phi$. The distributions account for external services in red and the individual sub-detectors, where the services internal to the sub-detector value are included. The large increase in material around $|\eta| \approx 1$ corresponds to service cables and structural supports for the SCT and and TRT, which enter between the barrel and end-cap. The increase in material around $|\eta| \approx 2.7$ indicates the concentration of service material for the pixel detector, which enters beyond the SCT end-cap. Figure is taken from [43].
4.4 Calorimeters

Enveloping the inner detector is the calorimeter. ATLAS implements sampling calorimeters, which operate with alternating layers of passive absorber and active detection material. The choice of absorber and active materials is optimized for incident particle type and detector constraints, such as geometry and cost. Two primary types of calorimeters are deployed by ATLAS: a fine granularity electromagnetic calorimeter followed by a coarser granularity hadronic calorimeter. Total coverage of the calorimeter is up to $|\eta| < 4.9$. Figure 4.7 shows an overview of the various calorimeter components [43].

![Figure 4.7: Overview of the ATLAS calorimeter. Figure is taken from [43].](image)

In general, a calorimeter’s resolution can be expressed as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} + \frac{b}{E} + c$$ (4.3)

where $E$ is energy of the incident particle. The first term is referred to as the sampling term and accounts...
for the stochastic fluctuations in energy deposition of showering particles, which generally follow Poisson statistics. The second term represents electronic noise in the active detector readout chain. The final term is referred to as the constant term which accounts for non-uniformities and calibration uncertainties in the calorimeter. For additional information and a concise introduction to calorimetry in particle physics, see [53].

The general purpose of the calorimeter system is to fully contain and measure the energy deposition of all interacting particles originating from the interaction point. (Muons are the exception to this goal as a typical muon will traverse the full extent of the detector; hence, the muon spectrometer represents the final layer of detection at ATLAS.) Therefore, calorimeters should be of sufficient depth to fully stop incident particles of interest. For EM calorimeters, thickness is often reported as radiation lengths, $X_0$. For hadronic calorimeters, thickness is measured in nuclear interaction lengths, $\lambda$.

4.4.1 Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter is used to identify and measure electrons and photons through electromagnetic showering. This showering process is largely accomplished through bremsstrahlung radiation and pair production. The EM calorimeter is composed of active layers of LAr and passive absorber layers of lead. Incoming electrons and photons produce EM showers in the lead layers, which will then ionize the LAr where kapton electrodes collect the free electrons. To provide full coverage in $\phi$, the EM calorimeter implements an accordion geometry as seen in Figure 4.8. The total $\eta$ coverage is out to $|\eta| < 3.2$, as achieved by barrel and end-cap components. The barrel is composed of two identical half-barrels offering coverage out to $|\eta| < 1.475$. Each end-cap is mechanically separated as two disks: the outer disk covering $1.375 < |\eta| < 2.5$ and the inner disk covering $2.5 < |\eta| < 3.2$. The EM calorimeter thickness is $> 22 X_0$ in the barrel and $> 24 X_0$ in the end-caps [43]. (Note that the majority of the calorimeter stopping power is in the lead absorber, which has $X_0 \approx 0.5$cm.)
\[ \Delta \phi = 0.0245 \]
\[ \Delta \eta = 0.025 \]
\[ 37.5 \text{mm} / 8 = 4.69 \text{ mm} \]
\[ \Delta \phi = 0.0031 \]
\[ \Delta \eta = 0.0245 \times 4 \]
\[ 36.8 \text{mm} \times 4 = 147.3 \text{ mm} \]

\[ \text{Figure 4.8: Drawn is an EM calorimeter barrel module where the different layers and their geometrical properties are evident. Figure taken from [43].} \]

4.4.2 Hadronic Calorimeter

The hadronic calorimeter is designed to measure hadrons, mostly pions and kaons, through hadronic showering processes. Three sub-detectors comprise the hadronic calorimeter: tile calorimeter, hadronic end-cap calorimeter (HEC), and forward calorimeter (FCal). Enveloping the EM barrel calorimeter is the tile calorimeter. The tile barrel covers the region \( |\eta| < 1.0 \), while extended barrels cover \( 0.8 < |\eta| < 1.7 \). Steel is used as the absorber material and plastic scintillator tiles are used as the active medium. Beyond the EM end-caps is the HEC. Copper plates comprise the passive material and LAr is used as the active material. The HEC covers the detector region \( 1.5 < |\eta| < 3.2 \). The FCal is integrated into the high-\( \eta \) region of the end-cap and recessed approximately 1.2 m from the inner face of the EM calorimeter. Covering the region \( 3.1 < |\eta| < 4.9 \), the FCal is segmented into three distinct modules. The innermost module implements copper as its absorber medium and is designed for electromagnetic measurements, while the two remaining modules use tungsten to measure hadronic showers. All three FCal layers consist of a metallic matrix of passive material with LAr as the active medium. The total coverage of the hadronic calorimeter is \( |\eta| < 4.9 \) [43].
4.4.3 Summary of Calorimeters

The full material budget of the calorimeter and its different components is summarized in Figure 4.9. The irregularities in $\eta$ coverage from having distinct calorimeter components and necessary gaps for services have noticeable effects on missing energy measurements and jet calibration in event reconstruction, as seen in Chapter 5.

![Diagram of detector material accumulation](image)

**Figure 4.9:** The cumulative amount of detector material, in nuclear interaction lengths, is shown as a function of $|\eta|$. The unlabeled contribution in beige represents all material in front of the EM calorimeter. The calorimeter layers are labeled with text. The light blue shading represents all detector material before the first active layer of the muon spectrometer. The shape of this figure will play a noticeable role in $\eta$ dependent jet calibrations. Figure is pulled from [43].
4.5 Muon Spectrometer

The calorimeter will generally contain all interacting particles with the exception of muons. Muons are minimally ionizing particles and will traverse the full detector depositing only a few GeV (≈ 3 GeV for a 1 TeV muon) in the calorimeter. The outermost detection components of ATLAS are devoted to detection of these muons. The muon spectrometer (MS) is a precision tracking detector, which operates within the toroidal magnet system to provide precision momentum measurements. The MS covers the region $|\eta| < 2.7$ and triggers on the region $|\eta| < 2.4$. Design standards require that the MS be able to measure the momentum of a 1 TeV muon with resolution greater than 10%, independent of the inner detector. There are four sub-detector types: muon drift tubes (MDTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs), and thin gap chambers (TGCs).

Figure 4.10: Muon sub-detector components are highlighted. The light blue corresponds to muon detection, while the toroidal magnet system is in light orange. Image pulled from [43].
4.5.1 Monitored Drift Tubes

Monitored drift tubes (MDTs) provide precision spatial measurements of muon track bending over most \( \eta \) range covered by the MS. There are 1088 MDT chambers, which consist of three to eight layers of drift tubes containing 93% Ar and 7% CO\(_2\) gas at a pressure of three bar. When a muon traverses the tube it will ionize the Ar gas, and the freed electrons will drift with a known velocity towards an anode wire running along the symmetry axis of the cylindrical tube. By measuring the electron drift time relative to a trigger signal, a position resolution can be achieved that far exceeds the precision implied by the physical tube size. While each MDT has a diameter of 3 cm, the average resolution is approximately 80 \( \mu \)m. Tubes are oriented to give precision measurements in \( \eta \) \([43]\).

4.5.2 Cathode Strip Chambers

Cathode strip chambers (CSCs) are used in high-\( \eta \) regions for muon tracking due to the increased rate of muon production and background tracks. The covered region is \( 2.0 < |\eta| < 2.7 \). The CSCs are multwire proportional chambers with anode wires aligned radially outward from the beam axis. There are two cathodes: one with strips segmented perpendicular to the wires to provide a measurement in \( \eta \) and another set running parallel to provide a measurement in \( \phi \). The resolution is approximately 60\( \mu \)m in \( \eta \) and 5mm in \( \phi \). The gas mixture is 80% Ar and 20% CO\(_2\) gas\([43]\).

4.5.3 Resistive Plate Chambers

Resistive plate chambers (RPCs) provide the trigger signal and azimuthal coordinate measurement in the barrel region of the MS. The region covered is \( |\eta| < 1.05 \). An RPC consists of two resistive plates separated by 2 mm with a voltage difference of approximately 4 kV. The large electric field between the two plates results in a quick response time (\( \approx 1.5 \) ns time resolution) when a muon traverses an RPC, making them useful for triggering purposes\([43]\).
4.5.4 Thin Gap Chambers

Thin gap chambers (TGCs) are multwire proportional chambers covering the region $1.05 < |\eta| < 2.4$. TGCs provide end-cap triggers for the MS, with approximately 4 ns time resolution, and provide an azimuthal coordinate measure.

4.6 Summary of Detector Resolution and Coverage

The full ATLAS detector resolution and coverage is summarized in Table 4.1.

<table>
<thead>
<tr>
<th>Component</th>
<th>Resolution</th>
<th>Measurement Coverage</th>
<th>Trigger Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking</td>
<td>$\sigma_{p_T}/p_T = 0.05% p_T \oplus 1%$</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>EM calorimetry</td>
<td>$\sigma_{E}/E = 10% / \sqrt{E} \oplus 0.7%$</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>Hadronic calorimetry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrel and end-cap</td>
<td>$\sigma_{E}/E = 50% / \sqrt{E} \oplus 3%$</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>Forward</td>
<td>$\sigma_{E}/E = 100% / \sqrt{E} \oplus 10%$</td>
<td>$3.1 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>Muon spectrometer</td>
<td>$\sigma_{p_T}/p_T = 10%$ at $p_T = 1$ TeV</td>
<td>$</td>
<td>\eta</td>
</tr>
</tbody>
</table>

Table 4.1: General design resolution and coverage of the various components of the ATLAS detector. Units for $p_T$ and $E$ are GeV. Detector components are approximately uniform in $\phi$. Table is reconstructed from [43].

4.7 Detectors as a System

Different sub-detectors are optimized to interact with different particle types, which can be exploited for basic particle identification, as indicated by Figure 4.11. For instance, both photons and electrons will shower in the electromagnetic calorimeter, but only electrons will leave a track in the inner detector while photons escape the inner detector unseen. A similar behavior is noted for protons and neutrons with respect to the hadronic calorimeter. Muons will deposit very little energy in the inner detector and calorimeters, and are generally the only detectable particles traversing the outer layer of detection, thus named the muon detector. Neutrinos escape the detector entirely without depositing any energy; therefore, neutrinos are only indirectly measured through careful accounting of all detectable particle tracks and energy deposits.
Figure 4.11: Shown are various particle types traversing the ATLAS detector. Combining various detectors allows for general identification of particles. Image taken from [54].

The analysis presented in this dissertation relies heavily on the calorimeter for jet reconstruction. Using a jet clustering algorithm, as described in Chapter 5, carefully measured energy depositions in the calorimeter can be used to reconstruct jet transverse momentum and pseudorapidity, which are the ingredients for calculating a dijet invariant mass. Other sub-detectors are implemented in support of developing an accurate dijet mass measurement. For example:

- A high-momentum jet is capable of punching through the hadronic calorimeter and depositing energy in the muon spectrometer. Therefore, the number of segments activated in the muon spectrometer matched to a jet trajectory is used as an input for jet energy scale correction to account for energy escaping the calorimeter.

- The inner detector is used to ensure the two highest momentum jets comprising a dijet event correspond to the same primary vertex. This is especially useful in developing jet cleaning criteria.
4.8 Trigger

A tremendous amount of data is generated every second at the LHC. (Almost one billion events per second!) However, most events correspond to well-known SM background processes, and ATLAS data-accumulation is inherently limited by the time necessary to process, store, and analyze events at the LHC. Therefore, ATLAS implements a two-tier trigger system to select events of interest [55, 56]. The first tier is referred to as a Level-1 (L1) trigger. L1 triggers are hardware-based and handle a subset of detector information to reduce the event rate to approximately 100 kHz. The second tier is a software-based trigger referred to as the High-Level Trigger (HLT). The HLT has full access to event information in the detector and operates many algorithms, such as jet clustering, that are similar to those used off-line. The HLT further reduces the rate of events to 1 kHz for data recording. The trigger chain implemented in this analysis involves a L1 trigger requiring a single-jet of $p_T > 100$ GeV, which seeds a HLT trigger requiring a single jet of $p_T > 360$ GeV. Example L1 and HLT trigger rates are presented in Figure 4.12.

![Trigger Rates](image)

**Figure 4.12:** Various L1 (a) and HLT (b) trigger rates are shown. The total trigger rate is the sum of all constituent trigger rates. Plots are taken from [57].
4.9 2015 Data Collection

ATLAS recorded approximately 3.9 fb$^{-1}$ of data, of which 3.2 fb$^{-1}$ was declared of sufficient quality to be used in physics analyses [48]. Data is declared “all good for physics” when all reconstructed physics objects are of good quality. However, the high-$p_T$ dijet analysis discussed in this analysis does not require the full detector to reconstruct objects, thereby allowing the inclusion of additional data on top of the collaboration-wide approved dataset. Most notably, 0.2 fb$^{-1}$ of data was not declared to be of general physics quality because the IBL was turned off; however, the IBL has a negligible impact on high-$p_T$ jet reconstruction, thus allowing this analysis to include the additional data. The total dataset used for this analysis is approximately 3.6 fb$^{-1}$.

![Graph showing total integrated luminosity delivered, recorded, and made available for physics analyses.]

Figure 4.13: Cumulative 2015 luminosity delivered (green), recorded (yellow), and made available for physics analyses (blue) [48].
5

Jets at ATLAS

Parton-parton collisions at the LHC producing quark and gluon final states are not directly observable in the ATLAS detector. Instead, ATLAS observes collimated sprays of hadrons, called jets, that collectively correspond to the original parton produced in the collision. (Jet production was discussed in Chapter 2.) Energy deposits measured in the calorimeter are used as inputs for jet reconstruction algorithms to interpret parton-level events. This process is diagrammed by a cartoon in Figure 5.1, whereby a hard-scatter event results in parton showering, followed by hadronization to form jets composed of hadrons, which proceed to deposit their energy in the calorimeter.

5.1 Building Topological Clusters

The calorimeter is comprised of approximately 188000 cells of varying size. Hadrons traversing the calorimeter will deposit energy in these cells, which constitute the input objects for jet reconstruction. The first step within jet reconstruction is to group cells into topological clusters, called topoclusters. The clustering algo-
A hard-scatter event may manifest as jets in the ATLAS detector. Careful measurement of these jets gives insight into the original partons emitted after the hard-scatter. Figure taken from [58].

The algorithm used by ATLAS is often referred to as a 4/2/0 clustering scheme and is illustrated in Figure 5.2. Each cell with reported energy $E_{\text{cell}}$ is compared to the expected noise for that cell $N_{\text{cell}}$. Seed cells are identified as those cells with energy deposits greater than four times the expected noise ($E_{\text{cell}} > 4N_{\text{cell}}$). Neighboring cells with reported energy $E_{\text{cell}} > 2N_{\text{cell}}$ are then added to the cluster. As a final step, a single layer of adjacent cells with energy $E_{\text{cell}} > 0$ are added to encourage the inclusion of all deposited energy within the topocluster. If a topocluster has more than one local signal maximum, then the large cluster is divided along signal minima. Each topocluster is assumed to be a massless particle with four-momentum defined by its total energy $\sum E_{\text{cell}}$ and projective direction from the detector origin, as given by energy-weighted center of the topocluster.

![Figure 5.1](image1.png)

**Figure 5.1:** A hard-scatter event may manifest as jets in the ATLAS detector. Careful measurement of these jets gives insight into the original partons emitted after the hard-scatter. Figure taken from [58].

![Figure 5.2](image2.png)

**Figure 5.2:** ATLAS uses a 4/2/0 clustering scheme to build topoclusters from calorimeter cells. As seen in the drawing, neighboring cells included in the clustering scheme need not be in the same calorimeter layer. Figure taken from [59].
5.2 Jet Reconstruction Algorithm

Jets are not uniquely identified objects in nature. Event reconstruction algorithms consisting of deterministic processes for combining collections of input particles into single objects define jets for experimental and theoretical calculations. Different algorithmic approaches to reconstructing jets can yield significantly different interpretations of an event. The effect of jet algorithm choice on event interpretation is illustrated in Figure 5.3. A parton-level event with underlying soft radiation particles was reconstructed by four different algorithms: $k_T$, anti-$k_T$, Cambridge/Aachen, and SISConE, where the first three are sequential recombination algorithms and the last is a cone algorithm. Notice that jet shape and, consequently, event reconstruction are highly dependent on the choice of algorithm. These algorithms are infrared and collinear safe. Collinear safety ensures the same hard jets are identified if the inputs have collinear splittings. Infrared safety ensures the same hard jets are identified if the inputs are modified by the addition of soft particle emissions. Both properties aid in providing a jet definition that yields finite cross-sections at any order in perturbation theory and is largely insensitive to the randomness of hadronization.*

![Figure 5.3](image_url)

*Reference [60] offers a review of jet algorithms and additional details on properties introduced in this section.
Sequential recombination algorithms are most commonly used in modern high-energy physics experiments. Such algorithms use distance measures that are invariant under longitudinal Lorentz boosts and generally given by [61]:

\[ d_{ij} = \min \left( p_{T_i}^2, p_{T_j}^2 \right) \frac{\Delta R_{ij}^2}{R^2}, \]  
\[ d_{iB} = p_{T_i}^2, \]  
\[ \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2. \]  

Here, \( p_{T_i}, y_i, \) and \( \phi_i \) are respectively the transverse momentum, rapidity, and azimuthal angle of particle \( i. \) Jet size is defined by \( R \) in coordinates \( y \) and \( \phi. \) \( d_{iB} \) is a distance measure of particle \( i \) relative to the beam-axis \( B. \) Inclusive clustering algorithms begin with a list of all initial particles detected in the calorimeter. All possible distances \( d_{ij} \) and \( d_{iB} \) are calculated and the smallest distance is identified. If the smallest distance is \( d_{ij}, \) then objects \( i \) and \( j \) are combined to form one object. If the smallest distance is \( d_{iB}, \) then object \( i \) is called a jet and removed from the list of entries. The process is then repeated with the augmented list, and continues until all inputs have been clustered into a jet.

The parameter \( p \) in Equation 5.3 dictates the type of algorithm within a general class of clustering algorithms. \( p = 1 \) corresponds to the \( k_T \) algorithm; \( p = 0 \) corresponds to the Cambridge/Aachen algorithm; \( p = -1 \) corresponds to the anti-\( k_T \) algorithm. An overview of these algorithms and many others is provided in [60]. While the conventional values of \( p \) are 0 or ±1, it should be noted that the distinctive features of each algorithm hold, respectively, for \( p = 0, p > 0, \) or \( p < 0, \) where the value \( p \) most crucially affects the ordering of particles added to the jet cluster. The analysis discussed in this dissertation implements an anti-\( k_T \) algorithm with \( R = 0.4, \) which is the most common choice of jet algorithms used on ATLAS.

An advantage of the anti-\( k_T \) algorithm is the construction of approximately conical jets as evident in Figure 5.3. The anti-\( k_T \) algorithm proceeds by clustering neighboring hard particles first, then adding soft jet contributions. Therefore, jets are well established objects before the inclusion of increasingly soft particles, which results in a conical jet shape. A traditional shortcoming of this approach for some analyses
is the loss of jet substructure information. In reconstruction of a QCD process, one would expect to work backward from radiative branching in such a way that soft particles are combined, then added to the closest hard particle, before all neighboring soft-radiation-corrected hard particles are combined. This is the general approach of the $k_T$ algorithm, which leads to irregular jet boundaries but preserves greater insight to jet substructure.

The inputs to jet clustering algorithms on ATLAS are assumed to be massless particles with four-momentum $k_i$. A jet with four momentum $p$ is constructed by summing the constituent input four-vectors. This recombination scheme, which conserves energy and momentum, results in a massive jet.

5.3 Jet Energy Calibration

ATLAS calorimeters are calibrated at the “EM-scale” where the energy deposited in topoclusters reflects the energy deposited by showering electrons or photons. Furthermore, the calorimeters are non-compensating; the energy response for electrons and photons is greater than for hadronic particles depositing the same amount of energy. By applying calibrations to the measured jets based on local signal depositions, which will consist of electromagnetic and hadronic shower components, energy resolution of the jets can be improved. Additional corrections are applied to account for pile-up, detector geometry and inefficiencies, and jet reconstruction biases. These corrections are addressed in a process known as jet calibration.

Jet calibration techniques used for 2015 ATLAS data analyses involving jets are described in detail by [62] and partially summarized here, following the referenced ATLAS note. There are five sequential steps to jet calibration:

1. Jet four-vectors are corrected to point pack to a hard-scatter vertex identified by the inner detector.
2. Pile-up effects are removed using a subtraction procedure based on the jet area and median energy density of the event in the $\eta \times \phi$ plane.
3. Jet energy corrections dependent upon $p_T$ and $\eta$ are derived from Monte Carlo and applied.
4. Corrections are applied to reduce the dependence of jet energy on longitudinal and transverse properties of jet structure.
5. Data jet calibrations derived in Monte Carlo are corrected using in situ studies of 2012 and early-2015 data.

5.3.1 ORIGIN CORRECTION

Each topocluster has an associated four-vector pointing from the detector origin to the energy-weighted center of the topocluster. However, jets will actually originate at the hard-scat ter vertex, which can differ from the detector origin by as much as a few centimeters. The initial step in jet calibration corrects the four-vector direction of each jet to originate at the hard-scat ter primary vertex identified by the inner detector.

5.3.2 PILE-UP CORRECTION

Pile-up can increase a jet energy measurement value diminishing the true correspondence between jet energy and the associated parton level process originating from the hard-scat ter event of interest. An area-based subtraction method is implemented in which the median energy density of the event \( \rho \times \phi \) plane is removed from the estimated area of the jet in that plane \( \langle A \rangle \). Additional residual effects dependent on the number of primary vertices \( (N_{PV}) \) and average number of interactions per bunch crossing \( \langle \mu \rangle \) are also removed. The pile-up corrected jet transverse momentum, \( p_T^{\text{corr}} \), is therefore given by:

\[
p_T^{\text{corr}} = p_T^{\text{EM}} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \langle \mu \rangle,
\]

where \( p_T^{\text{EM}} \) is the measured EM-scale jet \( p_T \) with origin correction applied.

5.3.3 JET ENERGY SCALE AND \( \eta \) CORRECTIONS

Detector energy response is defined as \( R = E_{\text{jet}}^{\text{recon}} / E_{\text{jet}}^{\text{truth}} \), where \( E_{\text{recon}} \) is the reconstructed jet energy and \( E_{\text{jet}} \) is the truth jet energy. Energy response is inhomogeneous throughout the detector due to variations in calorimeter calibration, coverage, and depth. This response variance is most notable when plotting energy response as a function of \( \eta \), as seen in Figure 5.4. Notice that detector gaps and transition regions clearly have a degrading effect on energy response. After correcting jet energy as a function of \( \eta \),
there is also a recognizable bias in the reconstruction of $\eta$, which must also be corrected. This correction is shown in Figure 5.4b and also shows a clear dependence on detector gaps and transition regions. These corrections are generated using Monte Carlo simulations.

**Figure 5.4:** Energy response is shown for jet reconstruction and truth jets, as simulated in Monte Carlo. Large dips in the response can be directly contributed to transition regions in the calorimeter. Figure taken from [62].

### 5.3.4 Global Sequential Correction

Additional corrections are applied to remove jet energy reconstruction dependence on longitudinal and transverse jet structure. Corrections are applied sequentially for dependence on the following: fraction of jet energy deposited in the first tile layer; fraction of jet energy deposited in the third electromagnetic calorimeter layer; number of charged tracks in the inner detector associated with the jet; the $p_T$-weighted width of inner detector tracks associated with the jet; and the number of muon segments associated with the jet. These corrections are generated in Monte Carlo simulations.

**Punch Through**

The fifth sequential global correction is of particular interest for dijet resonance searches, which rely on high-$p_T$ jets. Muon segments in the muon spectrometer are used as a measure of jet punch through. While most jets are well-contained within the hadronic calorimeter, high-$p_T$ jets may traverse the full calorimeter depth and deposit energy in the muon spectrometer. Punch through corrections are applied to any jet with greater than 20 associated muon segments ($N_{\text{segments}}$). This $p_T$ correction is shown as a function of
$N_{\text{segments}}$ in Figure 5.5 and a 10\% uncertainty is applied to the muon momentum [62]. Figure 5.6a shows average number of muon segments associated with an analysis jet as a function of $\eta$, which clearly shows a dependence on the amount of calorimeter material before the muon spectrometer. The fraction of events with at least one punch through corrected jet is shown in Figure 5.6b; notice that fewer than 6\% of events contain a punch through corrected jet for $1 \text{ TeV} < m_{jj} < 6 \text{ TeV}$.

![Figure 5.5: Jet transverse momentum response is shown as a function of the number of muon segments associated with the jet. Figure taken from [62].](image)

### 5.3.5 In situ Corrections

The aforementioned corrections are all derived using Monte Carlo; a natural final-step correction accounts for the differences between Monte Carlo and data jet reconstruction. These differences may arise from many sources, including limited or inaccurate modeling of the following: underlying event processes, jet formation physics, pile-up, detector interactions, detector operation and material composition. Using $\gamma/Z +$ jet and multijet processes, the discrepancy between data and Monte Carlo can be quantified. $\gamma/Z +$ jet and early-2015 data are used in making the data-to-simulation comparison. Note that little has changed in
Figure 5.6: Punchthrough corrections are applied for any jet with greater than 20 associated muon segments. Notice in (a) that punchthrough is most likely at calorimeter gaps or transition regions.

Detector composition since 2012, which provides confidence in using this dataset to derive jet corrections. A notable exception is the introduction of the IBL; however, this thin detector layer adds little to the overall pre-calorimeter material budget and is expected to have a negligible effect on jet calibration. The general correction is applied following the formula:

\[
p_{\text{T}} \text{corrected} = f_{2012}(p_T, \eta) C_{2012}(p_{\text{T}}^{\text{EM}}, \eta, N_{\text{PV}}, \mu, \{g_i\}) p_{\text{T}}^{\text{EM}}
\]

where \( f_{2012} = \frac{r_{2012}}{r_{\text{data}}} \) and \( r \) is the ratio between jet transverse momentum and the transverse momentum of the reconstructed reference object from the selected calibration process. \( \{g_i\} \) corresponds to the global sequence corrections. \( C_{2012} \) accounts for all of the corrections described in previous jet calibration steps and their observable dependencies.

5.3.6 JES Uncertainty

Jet energy scale (JES) and resolution (JER) uncertainties often constitute the largest uncertainty in jet-based analyses. The use of corrections derived in situ with 2012 data to calibrate jets in 2015 data introduces uncertainty in the modeling of detector response and operating conditions used to translate data corrections. Changes from Run I to Run II include: beam conditions, installation of the IBL, LAr calorimeter energy reconstruction inputs, cell noise level calibration for topocluster reconstruction, changes in muon segment
reconstruction, and new detector simulation models [62]. Uncertainties related to the in situ studies include: jet $p_T$ and $\eta$ dependence on JES and JER calibrations, modeling of pileup conditions, jet flavor composition, global sequential corrections, and limited certainty in a new high-$p_T$ regime. Total uncertainties and selected contributions are summarized for JER in Figure 5.7 and for JES in Figure 5.8. Note that JER uncertainty is less than 0.5% for jets of $p_T > 300$ GeV and JES uncertainty is approximately 1% for jets of $200$ GeV $< p_T < 1800$ GeV. JES uncertainty cannot be calculated from in situ studies because multijet techniques do not have sufficient statistics beyond $p_T > 1.8$ TeV for 2012 data; therefore, uncertainty is estimated from response to single hadrons as described in [63] in addition to punch through correction uncertainty.

![Figure 5.7: Jet energy resolution uncertainties estimated for 2015 data as a function of jet transverse momentum (a) and pseudorapidity (b) are shown for 25 ns bunch spacing. Figure taken from [62].](image)

### 5.4 Jet Cleaning

In addition to reconstructing and calibrating jets, it is important to identify and remove those jets that are inconsistent with a hard-scatter event originating from the interaction point. There are three primary sources of background jets at the LHC: beam-induced background from proton interaction upstream of the interaction point; cosmic-ray showers produced in the upper atmosphere; and large-scale calorimeter noise or an isolated malfunctioning cell. ATLAS design and modeling largely accounts for these background sources, but additional jet-level cleaning is necessary. Jet cleaning attempts to efficiently remove
Figure 5.8: Jet energy scale uncertainties estimated for 2015 data as a function of jet transverse momentum are shown for 25 ns running conditions (a) and 50 ns running conditions (b). The difference in bunch spacing will most directly affect pile-up correction. Figure taken from [62].

Background jets, called “fake jets,” while not discarding jets originating from collisions at the interaction point, or “good jets.” Variables are identified to efficiently discriminate fake and good jets. These variables consider signal pulse shape in the LAr calorimeters, ratio of energy in different layers of the calorimeter, and ratio of track-based momentum measurements to calorimeter-based jet momentum measurements. A detailed discussion of the motivation and determination for each jet cleaning variable is provided in [64]. The relevant variables are defined as:

- \( \langle Q \rangle \): Measure of jet quality given by the weighted average of the pulse quality for each calorimeter cell in the jet, where pulse quality refers to the difference between expected and observed signal shapes.

- \( f_{EM} \): Fraction of energy deposited in the electromagnetic calorimeter.

- \( f_{LAr} \): Fraction of energy deposited in the LAr calorimeter.

- \( f_{HEC} \): Fraction of energy deposited in the hadronic calorimeter.

- \( f_{QHEC} \): Fraction of energy deposited in the hadronic calorimeter for which the signal shape was poor.

- \( |E_{neg}| \): Sum of all cells in the jet with negative energy values.
• $f_{\text{max}}$: Ratio of energy in calorimeter layer with the largest energy deposit to the overall energy measurement.

• $f_{\text{ch}}$: Jet charged fraction calculated as the ratio of $\sum p_T^{\text{tracks}} / p_T^{\text{jet}}$.

Using the above variables, two jet cleaning criteria are used to identify fake jets for exclusion. These criteria are called BadLoose and BadTight. A BadLoose jet will satisfy one of the following:

• $f_{HEC} > 0.5$ and $|f_Q^{HEC}| > 0.5$ and $\langle Q \rangle > 0.8$

• $|E_{\text{neg}}| > 60 \text{ GeV}$

• $f_{EM} > 0.95$ and $f_{LAr} > 0.8$ and $\langle Q \rangle > 0.8$ and $|\eta| < 2.8$

• $f_{\text{max}} > 0.99$ and $|\eta| < 2$

• $f_{EM} < 0.05$ and $f_{\text{ch}} < 0.05$ and $|\eta| < 2$

• $f_{EM} < 0.05$ and $|\eta| \geq 2$

The BadTight selection introduces an additional criterion $f_{\text{ch}} / f_{\text{max}} < 0.1$ for $|\eta| < 2.4$, thereby increasing the rate of fake jet identification at the expense of cutting additional good jets.

The analysis presented in this dissertation uses LooseGood jets, which are jets that are not identified as BadLoose jets. The efficiency for loose jet identification is greater than 99.5% using anti-$k_T$ jets with $R = 0.4$ [64].
Part III

Dijet Resonance Analysis
Building a Dijet Spectrum

A common final state produced at the LHC will consist of two high-$p_T$ jets, collectively referred to as dijets. For each event, the invariant dijet mass $m_{jj}$ is calculated by summing the four-vectors of the two leading $p_T$ jets. Dijet events at the LHC produced by known SM physics will not generally originate from the decay of a heavy particle of fixed mass. When plotted over a large range of events, this QCD background of known physics results in a smooth, monotonically falling dijet invariant mass spectrum. Searching for such a deviation from a smoothly falling distribution, or a “bump,” is the subject of this search analysis.

Before conducting a search for new physics, an $m_{jj}$ spectrum must be constructed. This begins with selecting reconstructed dijet events in a specified kinematic phase-space with the goal of maximizing a potential signal-to-background ratio. Selection of events proceeds with attention to data quality of the event, jet measurement quality, and analysis cuts implemented to isolate a desired kinematic phase-space. The selected events are accumulated and binned in $m_{jj}$ to give a dijet mass spectrum presented in this chapter. In Chapter 7, the $m_{jj}$ spectrum is analyzed with a data-driven approach to search for unexpected features.
In Chapter 8, a Bayesian-limit setting method is briefly introduced and limits are presented.

The discussion presented in Chapters 6, 7, and 8 represents an overview of the 2015 dijet resonance search analysis. Additional details can be found in [20, 21].

6.1 Data Sample

The results presented in this dissertation are drawn from the ATLAS 2015 Run II dataset of 3.6 fb$^{-1}$ taken at $\sqrt{s} = 13$ TeV. The details of the LHC and detector were described in Chapters 3 and 4. The dataset includes 0.03 fb$^{-1}$ of data where the LHC bunch separation was 50 ns, with the remaining data consisting of 25 ns bunch separation. Additionally, data collected with the IBL inactive is included, which constitutes approximately 0.3 fb$^{-1}$ of the total dataset. Studies described in detail by [21] show that the inclusion of IBL tracking has negligible impact on the high-$p_T$ dijets used in this analysis. This conclusion is consistent with the IBL’s negligible effect on jet energy calibration as discussed in Chapter 5. A full list of runs included in the analysis dataset is listed in [21].

6.2 Monte Carlo Sample

Monte Carlo (MC) is used to simulate QCD background and study theoretical benchmark signal behavior for optimization of analysis techniques. For event reconstruction consistency, simulated MC events were processed using the same reconstruction software that processes ATLAS data.

6.2.1 Standard Model Background

QCD processes were simulated with pythia 8 [65] using the A14 [66] set of tuned parameters for the underlying event and the leading-order (LO) NNPDF2.3 [67] parton distribution functions (PDFs). Using NLOJET++ [68–70], next-to-leading-order (NLO) and electroweak corrections were applied for each $m_{jj}$ bin. NLO corrections are useful for monitoring data and MC agreement to ensure detector operations, event reconstruction, and the analysis procedure are functioning as expected. The background model used for the search phase of this analysis, as described in Chapter 7, relies on a fit to the observed
data and is independent of MC modeling. Detector operation is modeled using GEANT4 [71] within the ATLAS software infrastructure [72].

6.2.2 Benchmark Signals for Cut Optimization Studies

Excited quark signals were simulated using PYTHIA 8 for use in optimizing kinematic analysis cuts. Relevant details of the model are introduced in Chapter 8.

6.3 Trigger Strategy

The data used in this analysis was collected using the lowest unscaled single jet trigger available. An unscaled trigger is one that records every event meeting a trigger threshold; whereas, a pre-scaled trigger is one that records a random sampling of events meeting the trigger threshold and is implemented to maintain a steady, manageable rate of data acquisition. For the 2015 dataset a HLT_j360 trigger seeded from the L1_J100 trigger was used. This trigger chain is expected to remain unscaled for luminosities $\mathcal{L} < 6.5 \times 33 \text{cm}^{-2} \text{s}^{-1}$. Naming conventions for ATLAS single-jet triggers follow “Jnnn” for L1 triggers and “jnnn” for HLT triggers, where “nnn” is the nominal $p_T$ threshold of the trigger jet in GeV. L1 triggers use the EM energy scale, while HLT jets include a calibration sequence similar to off-line calibrations, thereby bringing their energies to the hadronic scale. While all events used in this analysis ultimately pass the HLT_j360 trigger, a pre-selection stage accumulates data inclusively using several single-jet triggers (L1_J75, L1_J100, HLT_j360, HLT_j380, and HLT_j400) and triggers associated with three and four jet topologies (HLT_3j175, HLT_4j85, and HLT_4j100).

6.3.1 Debug Stream

The debug trigger stream, which includes events that could not be sufficiently reconstructed in the allotted trigger time window, was incorporated into the analysis dataset. Debug stream events are an important contribution to ensure the dataset is not biased by a trigger reconstruction effect. In 2015 there was an issue with muon trigger reconstruction, which accounted for many events in the debug stream. These
events are of particular interest in this dijet analysis because high-$p_T$ jets can leak, or punch through, into the muon spectrometer.

6.4 Event Selection

Anti-$k_T$ $R=0.4$ jets were used with JES calibrations applied as described in Chapter 5. These jets were then subjected to event-level requirements. The dijet resonance and angular analyses were accomplished simultaneously and shared many analysis tools; therefore, a baseline event-level selection was first applied to accumulate a general dijet dataset. To pass the baseline cuts, an event must satisfy the following:

- The event is included in the Good Run List (GRL). The GRL lists runs and luminosity blocks for which all relevant detectors were in a good state to reconstruct physics events.*
- There are no event reconstruction errors in the LAr calorimeter.
- There are no event reconstruction errors in the tile calorimeter.
- A complete event build is accomplished.
- The primary vertex, defined as the vertex with the highest $\sum (p_T^{\text{tracks}})^2$, has at least two associated tracks. This cut removes events that did not originate from a beam collision.
- Passes one of the single-jet or multiple-jet triggers listed in Section 6.3.
- Has at least two jets with $p_T > 50$ GeV.
- Has a leading jet $p_T > 200$ GeV.

This less-restrictive dataset was used for many studies, including trigger efficiency studies, before applying a tighter selection for the resonance search analysis. In addition to the above baseline cuts, events used in this dissertation must satisfy the following:

- Passes the HLT_j360 trigger.

*The GRL applied for this analysis was data15_13TeV.periodAllYear_DetStatus-v05-pro19-01_DQDefects-00-01-02_PHYS_StandardGRL_All_Good.xml.
• The three highest $p_T$ jets satisfy the *LooseGood* cleaning criteria.

• Has a leading jet $p_T > 440$ GeV.

• Dijet invariant mass $m_{jj} > 1100$ GeV.

• Center-of-mass decay angle consistent with $|y^*| < 0.6$.

• The leading and sub-leading jets have $|\eta| < 2.8$.

### 6.4.1 Jet Momentum and Dijet Mass Cuts

The HLT$_{j360}$ trigger is 99.5% efficient when the leading jet of the dijet event has $p_T > 409$ GeV. An off-line cut requiring the leading jet to have $p_T > 440$ GeV ensures that kinematic trigger bias is sufficiently avoided. Furthermore, events are required to have $m_{jj} > 1100$ GeV, which is mostly redundant with the leading jet $p_T$ cut, but further serves to offer confidence in triggering efficiency. The trigger plateaus and off-line cuts are summarized in Table 6.1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>jet $p_T$</td>
<td>409</td>
<td>440</td>
</tr>
<tr>
<td>$m_{jj}$</td>
<td>900</td>
<td>1100</td>
</tr>
</tbody>
</table>

*Table 6.1*: Trigger plateau values, as defined by 99.5% trigger efficiency, and off-line cut values. The $m_{jj}$ cut further ensures kinematic trigger bias is removed.

### Pileup Influence on Kinematic Analysis

Pileup jets generally have low $p_T$. Given that all jets used in this analysis are required to have $p_T > 50$ GeV, there is little effect on observables and kinematic distributions. Pileup can affect single jet energy measurements, but this is addressed in jet energy calibration.

### 6.4.2 Jet Cleaning

The *LooseBad* jet quality flag is applied to reject jets consistent with background or calorimeter noise as described in Chapter 5. The three leading-$p_T$ jets are required to pass this cleaning requirement; otherwise,
the event is discarded. This cleaning requirement, rather than simply requiring two leading clean jets, helps protect against inclusion of events where the third jet should be swapped with the sub-leading jet due to a poor reconstruction.

6.4.3 Centrality Requirement

Recall the discussion of Chapter 2, QCD dijet production is dominated by $t$-channel scattering processes, which peaks in production rate at low center-of-mass scattering angles $\theta^*$. New physics resulting in a resonance is expected to be $s$-channel dominant, which results in decays isotropic in $\theta^*$. Therefore, a cut implementing the Lorentz invariant observable $y^*$ was optimized to isolate central decay tracks where signal-to-background ratios are most suitable for a resonant search.

The $y^*$ cut was optimized by comparing $m_{jj}$ distributions for MC-generated QCD background and $q^*$ signals. For a given $y^*$ cut, where events were required to have $|y^*| < y^*_\text{cut}$, and $0 < y^*_\text{cut} < 1.6$, QCD background and signal $m_{jj}$ distributions were generated. The chosen merit of significance to evaluate each proposed cut value was taken to be the quadrature sum of binned $\frac{S}{\sqrt{B}}$. That is,

$$\frac{S}{\sqrt{B}} (|y^*| < y^*_\text{cut}) = \sum_i S_i / \sqrt{B_i}$$

(6.1)

where the sum over $m_{jj}$ bins $i$ is done in quadrature. Figure 6.1 shows the signal significance as a function of proposed $y^*_\text{cut}$ for several $q^*$ signals of different masses. The significance curves for each signal have also been normalized to unit area for direct comparison. The optimal centrality cut is taken to be $|y^*| < 0.6$, which agrees with the most 8 TeV iteration of the dijet resonance analysis search.

6.4.4 Additional cuts considered

A value of particular interest due to its prevalence in dijet kinematic discussions is $y_B$. High mass events of particular interest in this analysis often occur when both incoming partons have a high momentum fraction, resulting in low $y_B$ values. However, cut optimization studies showed no gain by implementing a cut on $y_B$ as shown in Figure 6.2.
Figure 6.1: (a) The $y^*$ cut value is varied and significance curves are calculated and plotted. (b) Each significance curve, pertaining to a specific $q^*$ mass, is normalized to unit area for comparison of all signal optimal values. The optimal centrality cut is approximately $|y^*| < 0.6$.

Figure 6.2: (a) The $y_B^{cut}$ cut value is varied, where events are required to have $|y^B| < y_B^{cut}$ and significance curves are calculated and plotted. (b) Each significance curve, pertaining to a specific $q^*$ mass, is normalized to unit area for comparison of all signal optimal values. No optimal cut value was identified, therefore a $y_B$ was not implemented in the resonance analysis.

Several additional cuts were considered for this analysis including cuts on $p_T^{jj}$, number of jets in the event, number of primary vertices, and a mass-drop value, which was motivated by jet substructure studies [73]. None of these cuts were implemented as they offered no consistently substantial gain in signal sensitivity for $q^*$ resonances.
6.5 Analysis Selection Summary

The analysis detector range is shown in Figure 6.3. The active detector region (shown in blue) for leading and sub-leading jets was restricted to $|\eta| < 2.8$ to ensure the jet areas were fully covered by barrel and end-cap calorimeters. Given this analysis uses high-$p_T$ jets as inputs, the drawing safely assumes jets of zero mass, which allows the assumption $y \approx \eta$. Center-of-mass frame rapidities $y^*$ are drawn as lines of positive unit slope. The analysis cut boundaries of $y^* = \pm 0.6$ are drawn in red and the area between is shaded to reflect the kinematic phase-space of events passing the analysis selection. Boosts are represented as lines of negative unit slope; $y_B = 0$ is drawn in green for reference and represents events with balanced incoming parton momentums.

Figure 6.3: Figure shows the detector and $y^*$ cuts drawn in detector $\eta$. Events with leading and sub-leading jets appearing in red and blue shaded regions of the detector will pass $\eta$ and $y^*$ analysis cuts.

Presented in Table 6.2 is the 2015 analysis cut-flow summary. Table 6.3 provides an abbreviated analysis cut-flow beginning with our HLT_j360 trigger cut. There is approximate agreement between data and MC for event cut percentages in the final stages of the analysis chain.
Table 6.2: Cut-flow for 2015 data events with resonance analysis cuts applied.

<table>
<thead>
<tr>
<th>Selection criteria</th>
<th>$N_{events}$</th>
<th>Relative Change [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>All events</td>
<td>35477718</td>
<td>–</td>
</tr>
<tr>
<td>LAr calorimeter</td>
<td>35398888</td>
<td>−0.22</td>
</tr>
<tr>
<td>Tile calorimeter</td>
<td>35395678</td>
<td>−0.01</td>
</tr>
<tr>
<td>Core rebuild</td>
<td>35393381</td>
<td>−0.01</td>
</tr>
<tr>
<td>Primary vertex</td>
<td>35391453</td>
<td>−0.01</td>
</tr>
<tr>
<td>Inclusive Triggers (OR)</td>
<td>23350594</td>
<td>−34.02</td>
</tr>
<tr>
<td>≥ 2 jets with $p_T &gt; 50$ GeV</td>
<td>23020926</td>
<td>−1.41</td>
</tr>
<tr>
<td>$\text{Jet}^{\text{lead}}_p &gt; 200$ GeV</td>
<td>12740838</td>
<td>−44.66</td>
</tr>
<tr>
<td>Included in GRL</td>
<td>12171027</td>
<td>−4.47</td>
</tr>
<tr>
<td>HLT_j360</td>
<td>11995952</td>
<td>−1.44</td>
</tr>
<tr>
<td>Jet cleaning</td>
<td>11988448</td>
<td>−0.06</td>
</tr>
<tr>
<td>$\text{Jet}^{\text{lead}}_p &gt; 440$ GeV</td>
<td>4979860</td>
<td>−58.46</td>
</tr>
<tr>
<td>$m_{jj} &gt; 1.1$ TeV</td>
<td>2480182</td>
<td>−50.2</td>
</tr>
<tr>
<td>$</td>
<td>y^{*}</td>
<td>&lt; 0.6$</td>
</tr>
</tbody>
</table>

Table 6.3: Cut-flow for MC-generated events with resonance cuts applied.

<table>
<thead>
<tr>
<th>Selection criteria</th>
<th>$N_{events}$</th>
<th>Relative Change [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT_j360</td>
<td>4804917</td>
<td></td>
</tr>
<tr>
<td>$\text{Jet}^{\text{lead}}_p &gt; 440$ GeV</td>
<td>1884597</td>
<td>−60.78</td>
</tr>
<tr>
<td>$m_{jj} &gt; 1.1$ TeV</td>
<td>954174</td>
<td>−49.37</td>
</tr>
<tr>
<td>$</td>
<td>y^{*}</td>
<td>&lt; 0.6$</td>
</tr>
</tbody>
</table>
6.6 Dijet Mass Spectrum

The dijet resonance search is accomplished using a binned data spectrum. The choice of binning is particularly important during the search phase. Three considerations guided binning selection:

- Bins should be at least as wide as detector resolution. Increasing bin width will limit bin-to-bin migration of events due to reconstruction efficiency.

- Bins should be narrower than the expected width of a signal resonance. A signal peak fully contained within one bin would be impossible to characterize. Signal spread over multiple bins also improves the search phase sensitivity, which is designed to consider windows of multiple bins.

- The derived binning should provide a smooth MC generated dijet mass distribution. This is an essential binning feature as the search phase seeks to identify local deviations from an expected smooth, monotonically decreasing QCD background.

Studies used to derive a dijet mass binning must avoid kinematic biases introduced by requiring a leading jet $p_T$ cut. Given an analysis leading jet $p_T$ cut of 440 GeV, a lesser leading jet $p_T$ cut of 350 GeV was used to generate QCD background for binning derivation. The looser $p_T$ requirement compared to the analysis selection removes a kinematic bias on the low mass side of the search phase region of interest by shifting the $m_{jj}$ distribution peak to lower mass. All other cuts remained in place, with the added exception of the $m_{jj}$ cut.
6.6.1 Detector Resolution

Before a dijet binning can be generated, the relevant detector resolution must be determined. The dijet mass resolution is calculated by comparing truth and reconstructed $m_{jj}$ values in MC. Resolution is explicitly given by:

$$R = \frac{\sigma\left(\frac{m_{jj}^{\text{recon}}}{m_{jj}^{\text{true}}}\right)}{\mu\left(\frac{m_{jj}^{\text{recon}}}{m_{jj}^{\text{true}}}\right)}$$  \hspace{1cm} (6.2)

where $\mu$ and $\sigma$ are standard Gaussian parameters. The calculated detector resolution is less than 3% for events with $m_{jj} > 1$ TeV and approximately 2% for events with $m_{jj} > 4$ TeV. Figure 6.4 gives the calculated detector resolution for the range $1.1 > m_{jj} > 8$ TeV. The resolution is assumed to plateau at high-$m_{jj}$ and is extrapolated to higher masses for bins that are generated beyond this range.

![Figure 6.4: The dijet mass resolution is shown.](image)

6.6.2 Dijet Mass Binning Strategy

Having experimented with many binning options to address the aforementioned binning considerations, bin widths were chosen to match detector resolution. For each $m_{jj}$ bin, the bin width divided by bin center
mass value approximately matches detector resolution. This choice allowed for the maximum number of inputs for the background fitting procedure introduced in the search phase. A truncated\(^1\) listing of the dijet mass binning is provided in Table 6.4. Truth and reconstructed dijet spectra using this derived binning are shown in Figure 6.5. As a measure of distribution smoothness, efficiency and purity values were calculated for every bin, as defined by:

\[
\text{Bin Efficiency} = \frac{N_{\text{event}} \cap N_{\text{recon}}}{N_{\text{truth}}} \tag{6.3}
\]

\[
\text{Bin Purity} = \frac{N_{\text{event}} \cap N_{\text{recon}}}{N_{\text{recon}}} \tag{6.4}
\]

A smooth \(m_{jj}\) distribution will generally require these values to smoothly decrease, or remain constant, over the entire dijet mass range. The smoothly decreasing trend of each value, as shown in Figure 6.6, reflects the choice of bin resolutions that match the dijet mass resolution curve. With bins of size \(\sigma\), one would expect these values to begin at approximately 34\%, but recall that resolution was taken to be \(\sigma/\mu\)

\(^1\)Bins below 1.1 \(\text{TeV}\) are not shown because they are eliminated by a trigger efficiency cut requiring \(m_{jj} > 1.1 \text{ TeV}\) and bins greater than 10 \(\text{ TeV}\) are ignored because there were no events in the 2015 dataset requiring such high bins.

Figure 6.5: Binned MC dijet mass spectrum is plotted in bins of reconstructed \(m_{jj}\) versus truth \(m_{jj}\).
<table>
<thead>
<tr>
<th>Bin Ranges [TeV]</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.10 – 1.13</td>
<td>2.60 – 2.66</td>
<td>5.28 – 5.38</td>
</tr>
<tr>
<td>1.13 – 1.17</td>
<td>2.66 – 2.72</td>
<td>5.38 – 5.49</td>
</tr>
<tr>
<td>1.17 – 1.20</td>
<td>2.72 – 2.78</td>
<td>5.49 – 5.60</td>
</tr>
<tr>
<td>1.20 – 1.23</td>
<td>2.78 – 2.84</td>
<td>5.60 – 5.71</td>
</tr>
<tr>
<td>1.23 – 1.27</td>
<td>2.84 – 2.91</td>
<td>5.71 – 5.82</td>
</tr>
<tr>
<td>1.27 – 1.31</td>
<td>2.91 – 2.97</td>
<td>5.82 – 5.93</td>
</tr>
<tr>
<td>1.31 – 1.34</td>
<td>2.97 – 3.03</td>
<td>5.93 – 6.05</td>
</tr>
<tr>
<td>1.34 – 1.38</td>
<td>3.03 – 3.10</td>
<td>6.05 – 6.17</td>
</tr>
<tr>
<td>1.38 – 1.42</td>
<td>3.10 – 3.17</td>
<td>6.17 – 6.29</td>
</tr>
<tr>
<td>1.42 – 1.45</td>
<td>3.17 – 3.24</td>
<td>6.29 – 6.41</td>
</tr>
<tr>
<td>1.45 – 1.49</td>
<td>3.24 – 3.31</td>
<td>6.41 – 6.53</td>
</tr>
<tr>
<td>1.49 – 1.53</td>
<td>3.31 – 3.38</td>
<td>6.53 – 6.66</td>
</tr>
<tr>
<td>1.53 – 1.57</td>
<td>3.38 – 3.45</td>
<td>6.66 – 6.79</td>
</tr>
<tr>
<td>1.57 – 1.61</td>
<td>3.45 – 3.52</td>
<td>6.79 – 6.92</td>
</tr>
<tr>
<td>1.61 – 1.66</td>
<td>3.52 – 3.60</td>
<td>6.92 – 7.05</td>
</tr>
<tr>
<td>1.66 – 1.70</td>
<td>3.60 – 3.67</td>
<td>7.05 – 7.19</td>
</tr>
<tr>
<td>1.70 – 1.74</td>
<td>3.67 – 3.75</td>
<td>7.19 – 7.33</td>
</tr>
<tr>
<td>1.74 – 1.79</td>
<td>3.75 – 3.83</td>
<td>7.33 – 7.47</td>
</tr>
<tr>
<td>1.79 – 1.83</td>
<td>3.83 – 3.91</td>
<td>7.47 – 7.61</td>
</tr>
<tr>
<td>1.83 – 1.88</td>
<td>3.91 – 3.99</td>
<td>7.61 – 7.76</td>
</tr>
<tr>
<td>1.88 – 1.92</td>
<td>3.99 – 4.07</td>
<td>7.76 – 7.90</td>
</tr>
<tr>
<td>1.92 – 1.97</td>
<td>4.07 – 4.15</td>
<td>7.90 – 8.06</td>
</tr>
<tr>
<td>1.97 – 2.02</td>
<td>4.15 – 4.24</td>
<td>8.06 – 8.21</td>
</tr>
<tr>
<td>2.02 – 2.07</td>
<td>4.24 – 4.33</td>
<td>8.21 – 8.36</td>
</tr>
<tr>
<td>2.07 – 2.11</td>
<td>4.33 – 4.41</td>
<td>8.36 – 8.52</td>
</tr>
<tr>
<td>2.11 – 2.16</td>
<td>4.41 – 4.50</td>
<td>8.52 – 8.69</td>
</tr>
<tr>
<td>2.16 – 2.22</td>
<td>4.50 – 4.60</td>
<td>8.69 – 8.85</td>
</tr>
<tr>
<td>2.22 – 2.27</td>
<td>4.60 – 4.69</td>
<td>8.85 – 9.02</td>
</tr>
<tr>
<td>2.27 – 2.32</td>
<td>4.69 – 4.78</td>
<td>9.02 – 9.19</td>
</tr>
<tr>
<td>2.37 – 2.43</td>
<td>4.88 – 4.98</td>
<td>9.37 – 9.54</td>
</tr>
<tr>
<td>2.43 – 2.49</td>
<td>4.98 – 5.07</td>
<td>9.54 – 9.73</td>
</tr>
<tr>
<td>2.49 – 2.54</td>
<td>5.07 – 5.18</td>
<td>9.73 – 9.91</td>
</tr>
<tr>
<td>2.54 – 2.60</td>
<td>5.18 – 5.28</td>
<td>9.91 – 10.10</td>
</tr>
</tbody>
</table>

Table 6.4: Dijet mass binning used to bin 2015 dataset.
and average $m_{jj}^{\text{rec}} / m_{jj}^{\text{true}}$ values differ from unity by a few percent; therefore, the initial purity and efficiency values are lower.

Figure 6.6: Efficiency (a) and purity (b) for each of our derived $m_{jj}$ bins. A smooth, monotonic efficiency curve is expected for a smooth binning distribution.
6.6.3 Expected Signal Widths

Dijet bins should be narrower than expected resonant signal widths of interest. This criteria was shown to be satisfied by the derived binning when compared to $q^*$ signal widths. Several $q^*$ samples were generated of varying mass and compared to the detector resolution. As seen in Figure 6.7, $q^*$ signals exhibit a long radiative tail in the low mass region due to gluon emissions escaping the dijet reconstructions. Therefore, a two-step iterative Gaussian fitting process was applied to evaluate the signal width. The initial fit was accomplished with an assumed mean corresponding to the nominal signal mass. A second Gaussian fit was implemented over the inner $2\sigma$ core of the initial fit with no constraint on the mean. A resulting fit for $M_{q^*} = 4$ TeV is shown in Figure 6.7. Additional $q^*$ mass points are plotted in Appendix A. For each signal, the signal width was taken as the standard deviation of this inner-core Gaussian and divided by the fit mean to give a conservative estimate of signal resolution. Calculated signal resolution values and their comparison to detector resolution are shown in Table 6.5. Notice that the signals tested are indeed wider than the derived binning and detector resolution; furthermore, we should expect to see signal peaks spread over multiple bins.

<table>
<thead>
<tr>
<th>$M_{q^*}$</th>
<th>Signal Width $= \frac{\sigma(m_{jj})}{(m_{jj})}$</th>
<th>$q^*$ width \text{Detector Resolution}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 TeV</td>
<td>0.068</td>
<td>2.8</td>
</tr>
<tr>
<td>2.5 TeV</td>
<td>0.054</td>
<td>2.3</td>
</tr>
<tr>
<td>3 TeV</td>
<td>0.048</td>
<td>2.2</td>
</tr>
<tr>
<td>4 TeV</td>
<td>0.046</td>
<td>2.2</td>
</tr>
<tr>
<td>5 TeV</td>
<td>0.036</td>
<td>1.8</td>
</tr>
<tr>
<td>5.5 TeV</td>
<td>0.038</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Table 6.5: Presented is the signal width for select simulated $q^*$ signal masses. Also, signal width divided by detector resolution is provided. Notice that the $q^*$ signals are multiples of detector resolution and will be distributed over multiple $m_{jj}$ bins. Signal fits are provided in Appendix A.
Figure 6.7: A $M_{q^*} = 4$ TeV resonant signal is simulated and the peak width is conservatively estimated.
6.7 $m_{jj}$ Spectrum for 2015 Data

Upon accumulating all dijet events that pass the aforementioned cuts, the $m_{jj}$ values are binned as discussed in the previous section. The $m_{jj}$ distribution for the 2015 dataset of $3.6 \text{ fb}^{-1}$ is shown in Figure 6.8. Standard Model MC predictions for the dijet invariant mass background are provided in the same figure as an indication that the spectrum behaved as expected. JES one-sigma uncertainty bands are drawn for the MC distribution. Deviations from MC expectation could indicate issues relating to MC generation, detector modeling, or detector operation. Such deviations were seen in early 2015 detector operation and investigations pointed to an incorrectly updated detector timing calibration table that dealt with tower saturation in high-energy jet events.

![Graph showing $m_{jj}$ spectrum for 2015 data, next-to-leading order Standard Model predictions, and jet energy scale uncertainty on Monte Carlo predictions.](image)

**Figure 6.8:** Displayed is the $m_{jj}$ spectrum for 2015 data (black), next-to-leading order Standard Model predictions (blue), and jet energy scale uncertainty on Monte Carlo predictions (light blue shading). The binning reflects the derived dijet binning and the analysis cuts have been applied, including $m_{jj} > 1.1 \text{ TeV}$, leading jet $p_T > 440 \text{ GeV}$ and $|y| < 0.6$.

6.8 Additional Kinematic Distributions

To ensure our detector and triggering system were operating as expected, several kinematic distributions, in addition to energy deposition and jet cleaning variables, were closely monitored during data collection.
Two kinematic distributions of the 2015 analysis dataset are presented in Figure 6.9. Notice that the data and MC agree within the JES uncertainty. Additional kinematic plots are provided in Appendix B.

![Figure 6.9: Leading (a) and sub-leading (b) jet $p_T$ distributions with analysis cuts applied.](image)

6.9 **Highest Mass Dijet Event**

The highest mass dijet event satisfying the resonance analysis event-level requirements is reconstructed in Figures 6.10 and 6.11. The event consisted of two central jets with $p_T \approx 3.2$ TeV resulting in a dijet mass of 6.9 TeV. The total missing transverse momentum for this event is 46 GeV. Also, note that the two jets punched through the calorimeter and deposited energy in the muon spectrometer.
Figure 6.10: Highest 2015 recorded $m_{jj}$ event satisfying the resonant dijet analysis cuts. The $m_{jj} = 6.9$ TeV event consists of two central jets with $p_T \approx 3.2$ TeV.

Figure 6.11: Highest 2015 recorded $m_{jj}$ event satisfying the resonant dijet analysis cuts. The $m_{jj} = 6.9$ TeV event consists of two central jets with $p_T \approx 3.2$ TeV.
The Search Phase

The dijet invariant mass spectrum constructed in the preceding chapter must be closely analyzed for statistically significant resonant features atop Standard Model background production. Such a resonance search requires a confident modeling of Standard Model background processes. The search also requires tools for characterizing the observed deviations from an anticipated background in the 2015 dataset, and a preset threshold for what constitutes a statistically significant discrepancy between our background model and the data.

The goal of the search phase is to answer one question: “What is the probability of observing data at least as extreme as the measured $m_{jj}$ distribution, assuming that the Standard Model is true?” Without foresight of the answer, there must also be a predetermined cutoff for what constitutes a significant result. This chapter describes the tools to measure extremeness of the observed data, methods to limit bias, and presents associated results.
7.1 **Statistical Framework**

This section introduces the tools for quantifying agreement between observed data and an expected background-only hypothesis \( H_0 \). Here, the background hypothesis is assumed to exist, while its determination is the subject of Section 7.2. The statistical tools discussed in this section are implemented in both the determination of an unbiased background estimation, where a bias could exist by chance, and in the case of a resonant signal quantifying the extremeness of the observed data spectrum.

7.1.1 **The BumpHunter Algorithm**

This section describes the BumpHunter algorithm and relies on material from [74].

A shortcoming of \( \chi^2 \) and log-likelihood tests is that they only quantify the significance between observed and expected values for single, independent bins. However, one could imagine, for instance, that three adjacent bins with excesses would be more interesting in a resonance search than a single bin with an excess greater than any other bin, but adjacent to two bins with deficits. BumpHunter addresses this scenario by considering a window, varying in size, of adjacent bins. Simply put, BumpHunter defines a test statistic to quantify the “bumpiness” of the data compared to the background model.

Within a window of adjacent bins \( m \) through \( n \), let \( d_i \) be the observed number of events in bin \( i \) and \( b_i \) be the expected number of events in bin \( i \), as given by the background model. The total number of observed and expected events \( d \) and \( b \), respectively, within a BumpHunter window are given by:

\[
\begin{align*}
    d &= \sum_{i=m}^{n} d_i \\
    b &= \sum_{i=m}^{n} b_i.
\end{align*}
\]  

The observed event count of each window is assumed to be Poisson distributed with characteristic parameter given by \( b \). Recall that a Poisson distribution follows:

\[
p(d, b) = \frac{b^d}{d!} e^{-b}.
\]
For each set of adjacent bins $[m, n]$ the cumulative Poisson probability of obtaining a result at least as significant as the one observed can be calculated as

$$P(d, b) = \begin{cases} 
\sum_{n=d}^{\infty} \frac{n^n}{n!} e^{-b} & \text{for } d \geq b, \\
\sum_{n=0}^{d} \frac{n^n}{n!} e^{-b} & \text{for } d < b.
\end{cases} \quad (7.4)$$

Recalling the properties of Gamma functions, which are conveniently tabulated in ROOT’s TMath class for computational ease, the above expression can be rewritten in the form [74]:

$$P(d, b) = \begin{cases} 
\Gamma(d, b) & \text{for } d \geq b, \\
1 - \Gamma(d + 1, b) & \text{for } d < b.
\end{cases} \quad (7.5)$$

A Poisson probability $P(d, b)$ is calculated for every possible adjacent-bin window size between two bins and half the number of bins in the spectrum. The set of adjacent bins with the smallest probability of arising from a Poisson background fluctuation is the most significant “bump.” A BUMPHUNTER test statistic describing the overall spectrum is then defined as the negative-log of the smallest Poisson probability calculated from consideration of all allowed sets of adjacent bins, given by:

$$t = - \log \min P(d, b). \quad (7.6)$$

This process is visualized for the 2015 dijet data set in Figure 7.1.

BUMPHUNTER is routinely used in Beyond the Standard Model physics analyses throughout ATLAS because of its theory-agnostic capability. In this analysis we have ignored deficits as they are not reflective of the resonance signals of new physics models being considered.

### 7.1.2 Obtaining a $p$-Value

To quantify the agreement between the observed data $x$ and the background-only hypothesis $H_0$, a $p$-value is calculated. A frequentist probability definition inspires the implementation of pseudo-experiments to
etermine how likely one is to observe another spectrum $y$ that is at least as extreme as our observed data $x$. One pseudo-experiment is composed of each bin count in the $m_{jj}$ distribution being randomly drawn from a Poisson distribution with parameter equal to the expected number of events in that bin according to $H_0$. Approximately 1000 pseudo-experiments are generated; for each, a BumpHunter test statistic $t_i$ is calculated. The $p$-value quantifying agreement between data and the background-only hypothesis is then defined as the fraction of events with a test statistic greater than the observed test statistic $t_{obs}$. This procedure is visualized for the 2015 dijet analysis in Figure 7.2.

### 7.2 Background Estimation

There are many possible methodologies for defining a background-only hypothesis, which serves as a control for examining the observed data. Many analyses throughout ATLAS rely on MC simulations to generate distributions of the observable, in this case $m_{jj}$. These simulations take into account all known physics processes relevant to the analysis background and known detector response. The generated distribution
Many pseudo-experiments are generated to quantify the probability of randomly observing a spectrum at least as bumpy as the one observed. The observed BumpHunter test statistic is represented by the red arrow. In this case, approximately 670 of the 1000 pseudo-experiments resulted in a spectrum at least as bumpy as the one observed in data.

The LHC is operating in a new energy regime of 13 TeV collisions. The MC-generated distribution may not be reliable given uncertainty in modeling of QCD jet production in current LHC operating conditions.

ATLAS is building physics jets of higher energy than previously measured. Therefore, time-demanding studies are needed to verify that detector response appropriately models high-energy dijet production.

Recall from Chapter 2, dijet background production is exceptionally common at a proton-proton collider. In the high integrated luminosity era of particle physics at the LHC, it is unrealistic to generate, and use for search purposes, statistically reliable MC samples for the $m_{jj}$ distribution.
To summarize, it is difficult to imagine generating a statistically significant and theoretically confident MC dijet mass distribution. Alternatively, since the CDF dijet searches at the Tevatron it has become common practice to use a data-driven background approach to estimate $m_{jj}$ background using a smooth parameterized function. While the exact form of this background characterized by QCD scattering is unknown, its shape will reflect the convolution of quark and gluon PDFs, kinematic cuts, and detector resolution. The result is a smooth, monotonically falling $m_{jj}$ distribution. Furthermore, a useful background parameterization will not have so many parameters that the function can readily adapt to mask resonant signals, but still have enough parameters to readily capture the shape of the dijet invariant mass distribution.

The choice of background parameterization as a smooth function ultimately restricts this analysis to a resonance-only search. A non-resonant process, such as contact interactions introduced in Chapter 4, will manifest as an increase in event cross-section in the high-mass tail of a dijet distribution and a parameterized function will naturally evolve to fit this increasing tail. The fit function cannot be manually forced to any given value at high mass because this region of SM-only dijet production represents unknown territory in high-energy experimental physics and an interesting search space for some BSM models, such as QBH production. Creating an unbiased background estimation can only be accomplished in the search for resonance-like features.

### 7.2.1 Parameterizing the Background

The standard functional form chosen to parameterize the dijet invariant mass background spectrum in this analysis is given by:

$$f(x) = p_1 (1 - x)^{p_2} + p_3 x + p_4 \ln x + p_5 (\ln x)^2,$$

where $p_i$ are fit parameters and $x \equiv m_{jj}/\sqrt{s}$. Parameters $p_4$ and $p_5$ may be forced to zero, in which case the function is referred to as the “3-parameter” fit function. This background parameterization has been shown [21, 75] by ATLAS and other experiments to effectively model the dijet mass distributions of NLO MC and previous hadron collider data. The 3-parameter fit was implemented in this analysis as it was shown to accurately describe MC simulations of the dijet mass spectrum up to $\approx 10 \text{ fb}^{-1}$.

While the parameterization in Equation 7.7 seems arbitrary, the core terms are identifiable. The $(1 - x)^p$
component mimics common PDF parameterizations while leading-order QCD matrix element calculations motivate the $x^p$ term. Additional terms have proven necessary in some analyses to account for high-mass tails in the dijet spectrum and additional luminosity. For instance, factors of the form $x^{(\ln x)^n}$ were inspired by an inclusive dijet measurement at CDF.

### 7.2.2 Building the Background Estimate

Only the first three parameters of Equation 7.7 were employed in modeling the 2015 dataset. However, the number of parameters necessary is known to change with increasing luminosity and, since the background parameterization is not derived from first principles, the choice of fit function must be made by fitting the data. In a previous version of this analysis [75] a fit function was chosen by fitting one-fourth of the available dataset before unblinding the analysis to the remaining data; ultimately, the full dataset fit was found to be in poor agreement with the dijet mass background and exhibited an oscillating behavior uncharacteristic of theorized resonant physics models. Learning from this lesson, this analysis chose to implement a fit function choice procedure that evolved with accumulated luminosity. Data was fit as it was accumulated and the default fit function was tested against higher-parameterization alternatives using a Wilks likelihood ratio. If the associated Wilks $p$-value dropped below 0.05, the default fit function would be abandoned for one with an additional parameter. This method implemented by the dijet analysis team is introduced for completeness, but the procedure never favored a function over the default 3-parameter fit function. Future analyses, such as the ongoing 2016 dijet resonance search, will likely find that a higher parameterization is preferred as luminosity increases.

### 7.2.3 Removing “bump” Bias in the Background Estimate

The data and background-only hypothesis are assumed to be in agreement if the $p$-value is greater than 0.01, at which point the analysis moves on to the limit setting phase described in the next chapter. If a lesser $p$-value is observed, then an equally “bumpy” spectrum would only occur by statistical fluctuations in fewer than 1% of cases if the data were indeed represented by the background-only hypothesis.

When a $p$-value $\leq 0.01$ is calculated, it becomes necessary to ensure the background-only hypothesis
is not biased by an excess in the data. The set of adjacent bins constituting the greatest deviation from the background hypothesis, plus the neighboring low-mass bin,* are temporarily removed from the spectrum, and a new background estimation is calculated. The BUMPHunter algorithm is deployed and the steps discussed above are repeated to analyze the newly reduced data spectrum. If there is agreement (p-value > 0.01) between the reduced data spectrum and new background estimation, then the newly calculated background parameterization is assumed as an improved background-only hypothesis, \( H_0 \); otherwise, the process of removing the most discrepant bins is repeated. Using the unbiased \( H_0 \), and the full data spectrum, the search phase is repeated. In the case of a true signal, this will likely result in a more significant excess because the original background fit bias was masking some of the signal shape.

7.2.4 Background-only Hypothesis Result

The background-only hypothesis resulting from the aforementioned background estimation strategy is shown in Figure 7.3. Data is shown by filled black points, while \( H_0 \) is illustrated by the solid red line. The most discrepant region identified by the BUMPHunter algorithm is indicated by blue vertical lines. Bin-by-bin significance of the data-fit difference is illustrated in the bottom panel and described in additional detail below.

Displaying Bin-by-bin Significance

Fit result plots throughout this dissertation use the residual to display bin-by-bin statistical significance derived from the difference between data and \( H_0 \). This method of quantifying statistical significance is equivalent to finding the probability that a deviation at least as large as the one observed in the data could occur, assuming each bin is treated as independent and data is Poisson distributed around a mean value given by the background-only hypothesis. This probability, or p-value, will often span a large range of magnitudes making it difficult to visualize. It is often convenient to translate a p-value to a z-value, which is defined as the number of standard deviations to the right of the mean of a Gaussian distribution [76],

---

*The removal of one additional bin on the low-mass side of the region of interest was found to reduce residual bias after a signal was removed in MC tests of the search phase algorithm.
given by:

\[ p\text{-value} = \int_{z\text{-value}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \, dx. \]  

(7.8)

Considering the span of \( p \)-values, it is useful to note that a \( p \)-value of 1.0 would represent complete agreement between data and \( H_0 \), while a \( p \)-value of \( \approx 2.9 \times 10^{-7} \), corresponding to a \( z \)-value of 5 (i.e. “5\( \sigma \) effect”), is conventionally accepted as a discovery-level observation by the high energy physics community \([76, 77]\).

\( z \)-value \( \geq 0 \) corresponds to \( p \)-value \( \leq 0.5 \), whereas a negative \( z \)-value corresponds to a \( p \)-value \( \geq 0.5 \). This analysis searches for bumps in the mass spectrum and it is useful to visualize bins of excess and deficit, thus, only \( z \)-values \( \geq 0 \) are considered and negative \( z \)-values are suppressed to zero. Bin-by-bin significance plots drawn throughout this dissertation only show bins with positive \( z \)-values. The positive or negative direction of the bar corresponds, respectively, to an excess or deficit in the observed data as compared to \( H_0 \).

![Figure 7.3](image)

Figure 7.3: Shown is the reconstructed \( m_{jj} \) spectrum with 2015 data (filled dots) fitted with the dijet fit function described by Equation 7.7 (solid line). The blue vertical lines indicate the region of greatest excess. Bin-by-bin significance of the data-fit difference is shown in the bottom panel, considering statistical uncertainties only.
7.3 Uncertainty on Background Estimation

The data-driven background estimation approach employed in this analysis limits the introduction of uncertainties, which would serve to reduce sensitivity. Alternatively, an MC-generated background estimation would introduce many additional uncertainties, including those associated with QCD modeling of high-energy jets, detector inefficiencies and operation, and limited statistics. The data-driven background estimation introduces two sources of uncertainty: choice of fit function and determination of fit parameters. These uncertainties are visualized in Figure 7.4.

Quantifying Fit Function Choice Uncertainty

The nominal 3-parameter fit is compared to the 4-parameter fit. The uncertainty of fit function choice for each bin is taken as the difference between the two fits to the data, scaled by the RMS of the difference between the two fits and data when pseudo-experiments are generated from the original data for each bin.

Parameter Determination Uncertainty

Many pseudo-experiments are generated from the nominal fit. Each pseudo-data set is fit using the same starting conditions as the observed data. The error on the nominal fit in each bin is defined as the RMS of the difference between the pseudo-experiment function values and the nominal function value.
Figure 7.4: Background only estimation is shown in red. The statistical uncertainty on the fit and the choice of fit parameterization are shown.
7.4 Search Results for 2015 Data

The 2015 dijet invariant mass resonant search result is shown in Figure 7.5. The $p$-value corresponding to the observed extremeness of the 2015 dataset is 0.67, implying that there is no statistically significant deviation from QCD background.

\textbf{Figure 7.5:} The reconstructed dijet mass for data passing event selection criteria is plotted as filled dots. The solid red line is the background estimation. The most discrepant region is identified by vertical blue lines. Two resonant models, $q^*$ and QBH (BlackMax generator), are shown above the fit using a predicted cross-section. The $q^*$ cross-section has been increased by a factor of three for visualization of the signal shape. Bin-by-bin comparisons of the data to the background fit are shown in the middle panel. The bottom panel shows NLO MC background predictions with JES uncertainty bands.
During the search phase of the analysis, as outlined in Chapter 7, no statistically significant deviation from a SM-compatible, smooth $m_{jj}$ spectrum was observed. Given the absence of new phenomena in the observed data spectrum, the natural next step is to set limits on theoretical models describing new physics production of two high-$p_T$ jets at the LHC. These limits seek to answer the following question: “Given the observed data, for what maximum dijet invariant mass can we exclude a given resonant production model?” This chapter will briefly introduce the limit setting approach implemented by the presented analysis and offer mass limits the BSM physics models considered in Chapter 1. Finally, limits are presented for generic Gaussian-shaped signals.
8.1 Overview of Limit Setting Strategy

The presented analysis implemented a Bayesian limit setting method following the methods used in ATLAS Run I dijet searches [13]. Using this Bayesian method, a posterior probability generically given by:

\[
p(\nu, x) \propto \int \mathcal{L}(x|\nu, \tilde{\theta}) \pi(\nu) \prod_i \pi(\theta_i) d\tilde{\theta}
\]

is calculated. The measured data set is denoted as \( x \) and the parameter of interest \( \nu \) is the signal strength. Nuisance parameters are denoted by \( \tilde{\theta} \). Signal strength and nuisance parameter priors are \( \pi(\nu) \) and \( \pi(\theta_i) \), respectively. Note that all of the priors are considered to be independent in this analysis. The integral over all nuisance parameters, referred to as a marginalization, is the most computationally intensive component of limit setting. \( \mathcal{L}(x|\nu, \tilde{\theta}) \) is a likelihood function for signal strength as a function of the nuisance parameters.

The signal strength prior was chosen to be uniform, while nuisance parameter priors, one for each systematic uncertainty, were assumed to have Gaussian distributions. For the background estimation, systematic uncertainties included fit quality and choice of fit function, as introduced in Chapter 7. For signal model templates, the following systematic uncertainties were considered: statistical uncertainties due to a limited number of generated signal events at each mass point; PDF and higher-order correction uncertainties; a luminosity uncertainty of 9%; and jet energy scale uncertainties, which were seen in Chapter 5 to be less than 10%.

The upper limit on the number of signal events was taken as the 95% quantile of the posterior for each generated mass point for a given signal model. An observed limit curve is formed by interpolating logarithmically between the discrete mass points. This number of events is divided by luminosity to obtain an effective signal cross-section. However, this is not an actual cross-section because detector, reconstruction, and analysis efficiencies will lower the true value to a measured effective cross-section. The reported value is cross-section \( \sigma \) times acceptance \( A \), written as \( \sigma \times A \). A signal model is considered to be excluded for all mass values where the theoretical cross-section is greater than the measured cross-section.

Expected limit curves were calculated at each mass point using a mass template generated from the back-
ground fit and signal strength set to zero. All nuisance parameters were included. The expected limit provides information for how much the upper limit of signal could shift due to statistical fluctuations in the background. An ensemble of expected limits was calculated by applying the prescribed Bayesian method to approximately 100 pseudo-datasets generated from the nominal background fit. In this way, frequentist bands of $\pm 1\sigma$ and $\pm 2\sigma$ are generated for Bayesian 95\% credibility-level upper limits with zero signal.

8.2 Benchmark Signal Models

MC-generated signal models at specific mass points were used with the observed data to calculate 95\% credibility-level upper limits on proposed signal cross-section times acceptance ($\sigma \times A$) as a function of mass. Limits for the following models* were calculated and presented in [20]:

- Quantum black holes for ADD scenario with $n = 6$ extra dimensions

- Quantum black holes for RS scenario with $n = 1$ extra dimension

- Excited quarks

These benchmark signal processes were modeled at various mass points to provide templates for limit setting. Benchmark models were generated using the same PDF and tuning parameters as described for QCD process simulations introduced in Section 6.2. Details for each model are summarized below.

8.2.1 Excited Quarks Model

Excited quark mass templates were generated using PYTHIA 8 for various mass points. The excited quark model implemented was described in Section 1.2.1. As in [27], SM quark coupling constants were assumed; $SU(3)$, $SU(2)$, and $U(1)$ coupling multipliers were set to one ($f_3 = f = f' = 1$); and the composite-ness scale was set to the excited quark mass ($\Lambda = M_q^*$). Only excited quarks decaying to a gluon and an up ($q^* \to ug$) or down ($q^* \to dg$) quark are modeled, which accounts for an approximately 85\% branching

*Two additional resonant models were considered in the published analysis results: a $W'$ boson and a leptophobic $Z'$ boson coupling to a dark matter candidate. Details of these models and the newly established limits were provided in [20].
ratio. The signal width, before the inclusion of parton showering effects, is approximately equivalent to
detector resolution, as seen in Equation 1.3 and Figure 6.4.

8.2.2 Quantum Black Holes Model

Two generators, BlackMax [78] and QBH [79], were used to model quantum black hole production at the
LHC. The two generators operate with different physical assumptions concerning black hole production
and decay. Differences concerning physical assumptions between the two generators, outlined in detail
by [21, 75], result in a QBH generator cross-section that is approximately twice the modeled BlackMax
generator cross-section. Model scenarios considered were Arkani-Hamed-Dimopoulos-Dvali (ADD) and
Randall-Sundrum (RS) scenarios. In the ADD scenario, \( n = 6 \) extra dimensions were assumed and the
the fundamental scale of gravity was set to be the mass threshold for black hole production \( M_D = M_{\text{th}}. \)
In the RS scenario, only the QBH generator was implemented and \( n = 1 \) extra dimension was assumed.
These models assume a branching ratio to dijets greater then 96%. The PDFs used were CTEQ6L1 [80].
A general discussion of ADD-scenario black hole production at the LHC was described in Section 1.2.2.

8.3 Limits Using 2015 Dataset

ATLAS Run II observed (assume possibility of signal) and expected (assume zero signal) limits for bench-
mark models are listed in Table 8.1, where signal masses below the limit are excluded at 95% CL. Run I
limits are also provided for context. The significant extension of Run I limits, despite considerably less
luminosity, is due to the increase in center-of-mass collision energy from 8 TeV to 13 TeV.

Limit curves are provided for QBH and \( q^* \) models, respectively, in Figures 8.1 and 8.2. The observed and
expected limit curves show \( \sigma \times A \) given the data. The steeper dashed lines are theoretical values of \( \sigma \times A \) for
the given signal model. Therefore, signal masses with a theoretical cross-section greater than the observed
cross-section (i.e., observed limit curve is below the theory line) are excluded at 95% credibility-level. The
limits presented in Table 8.1 represent the highest mass for which this exclusion holds, as indicated by the
intersection of the limit curve and theory line.

In addition to benchmark signal models, limits are calculated using Gaussian signal shapes. For each
<table>
<thead>
<tr>
<th>Model</th>
<th>Run I Observed</th>
<th>2015 Observed</th>
<th>2015 Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantum Black Holes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADD (BLACKMAX)</td>
<td>5.6 TeV</td>
<td>8.1 TeV</td>
<td>8.1 TeV</td>
</tr>
<tr>
<td>ADD (QBH)</td>
<td>5.7 TeV</td>
<td>8.3 TeV</td>
<td>8.3 TeV</td>
</tr>
<tr>
<td>RS (QBH)</td>
<td>–</td>
<td>5.3 TeV</td>
<td>5.1 TeV</td>
</tr>
<tr>
<td>Excited Quark</td>
<td>4.1 TeV</td>
<td>5.2 TeV</td>
<td>4.9 TeV</td>
</tr>
</tbody>
</table>

Table 8.1: New and previous ATLAS dijet resonance 95\% credibility-level lower limit for QBH and $q^*$ model masses are summarized.

signal, the mass value $m_G$ is the Gaussian mean while the cross-section is the Gaussian standard deviation $\sigma_G$. Results are presented as $\sigma \times A \times BR$, where BR is the branching ratio for signal decay to dijets. These model-independent limits are useful for generalization to theoretical models not explicitly considered in this analysis.
Figure 8.1: The 95% credibility-level upper limits for $\sigma \times A$ are shown for quantum black hole signals as a function of mass. ADD models are denoted by QBH (BM) and QBH (QBH) where $n = 6$ extra dimensions are assumed. A RS model with $n = 1$ extra dimension is denoted by QBH (RS). The observed 95% credibility-level lower limit on QBH mass threshold in a given model is identified as the intersection of the theory and observed curves. Result presented in [20].
Figure 8.2: The 95% credibility-level upper limits for $\sigma \times A$ are shown for $q^*$ signals as a function of mass. The observed 95% credibility-level lower limit on $q^*$ mass is identified as the intersection of the theory and observed curves. Result presented in [20].
Figure 8.3: The 95% credibility-level upper limits for $\sigma \times A \times BR$ are presented for a Gaussian-shaped signal width $\sigma_G$ and mass $m_G$, where the two values refer to the Gaussian standard deviation and mean, respectively. These results are useful for extrapolating analysis results to signal models that were not explicitly considered. Result presented in [20].
Part IV

Looking Forward
There is no statistically significant evidence of new physics phenomena decaying to two high-$p_T$ jets in the 2015 dataset, but there is considerable motivation to keep searching. Increased discovery potential is found in three sources: further refining of analysis techniques, increased collision energy, and increased luminosity.

An example improvement to the current ATLAS dijet analysis technique involves the use of wider jets to capture additional final state radiation. This will allow for a more accurate reconstruction of jet $p_T$. Optimal jet size and $p_T$ threshold for inclusion of soft radiation must be studied in detail. A motivation of this approach is introduced in Appendix C.

The largest gain in discovery potential, and limit setting, in the presented analysis over previous iterations was due to an increase in center-of-mass collision energy. Increasing beam energies from 4 TeV to 6.5 TeV per beam significantly increased parton luminosities, as shown in Figure 9.1; the increase in parton luminosity allows for the higher production rate of massive objects at the LHC. For 2016, LHC collision
energies will remain unchanged, but an increase to $\sqrt{s} = 14$ TeV is under consideration for post-2016 operation. The most immediate gain in sensitivity to exotic dijet resonance production will come from increased luminosity. Cumulative luminosity expectations for the remainder of Run II are $> 30$ fb$^{-1}$ in 2016 and $> 100$ fb$^{-1}$ by 2018. Following Run II, there will be a significant upgrade to allow for doubling of the LHC’s nominal design instantaneous luminosity. Another $> 300$ fb$^{-1}$ is expected to be produced by 2022. The final step in LHC improvements is the High Luminosity LHC (HL-LHC), which intends to offer instantaneous luminosity of $5 \times 10^{34}$ cm$^{-2}$s$^{-1}$ and ultimately accumulate $> 3000$ fb$^{-1}$ of $\sqrt{s} = 14$ TeV collisions.

Rapidly increasing integrated luminosity represents exceptional opportunity for the new discovery. Consider $\sqrt{s} = 14$ TeV production of two dijet resonance benchmark signals [82]: excited quarks and quantum black holes. Proposed 5$\sigma$-discovery luminosities are shown as a function of $q^*$ mass and QBH production threshold mass in Figure 9.2. In the case of no $q^*$ or QBH discovery, expected limits that can be set on the quark compositeness energy scale and the QBH mass threshold are shown for various luminosities in Figure 9.3.

Figure 9.1: The ratios (13 TeV/8 TeV) of LHC parton luminosities are shown as a function of mass. This increase in parton luminosity for massive object production allows further reach into the high-mass region of a dijet mass spectrum, which increases discovery potential. Figure taken from [81].
Figure 9.2: Discovery potential for $q^*$ and QBH benchmark signals with a dijet mass resonance search with the ATLAS detector at $\sqrt{s} = 14$ TeV. Discovery is defined as a 5$\sigma$ excess identified by the search procedure described in Chapter 7. Figures taken from [82].

Figure 9.3: Limit setting potential for $q^*$ and QBH benchmark signals with a dijet mass resonance search with the ATLAS detector at $\sqrt{s} = 14$ TeV. Expected upper limits are set using Bayesian methods described in Chapter 8 and correspond to a 95% credibility level. Figures taken from [82].

While no discovery was made in 2015, the potential for discovery, or greater constriction on theoretical models seeking to address looming questions of the universe, is ever-present on the horizon. This excitement for expedition, for even the potential of discovery, continues to drive an international collaboration of thousands of physicists and encourages the world to keep asking questions and seeking answers.
Example $q^*$ Signal Shapes

Shown in Figure A.1 is a subset of the generated $q^*$ signal templates. Note that this is not the full set used for the limit setting stage discussed in Chapter 8. These mass templates are meant to offer a visual guide for the conservative signal width estimates used in Chapter 6 to demonstrate that a $q^*$ mass peak is expected to be a few multiples wider than the binning resolution. Furthermore, note that the peak mass is shifted downward from the nominal mass peak and there is a long, low-mass radiative tail evident for the signals. This radiative tail differs from an expected quantum black hole signal shape, which is not a true resonance, but rather a sharp turn-on for production at the mass threshold for QBH production, followed by a rapid turn-off due to PDF effects.
Figure A.1: Six $q^*$ signals generated in MC are presented. Peak fits were accomplished with an iterative Gaussian fitting technique to give conservative estimates of the signal width. Analysis cuts were applied and the number of events assumes $\approx 1 \text{ fb}^{-1}$ of data.
Kinematic Distributions

The following kinematics generally show good agreement between the 3.6 fb$^{-1}$ dijet analysis dataset and MC. A scale factor of 3.1 was applied to the MC, which was generated at 1 fb$^{-1}$, to bring it to scale with the data. The difference in the number of expected and observed events, which is encapsulated in the scale factor, is of little interest to this analysis; however, the MC shape is useful to check that the detector is functioning as expected. The following kinematic plots are provided:

- Figure B.1 shows common dijet system plots, including $y^*$, $y_b$, $m_{jj}$, $p_T^{jj}$, etc.
- Figure B.2 shows the angular and $p_T$ balance of the dijet system.
- Figures B.3, B.4, and B.5 give kinematic distributions of the first, second, and third jets, respectively.
- Correlation between dijet invariant mass and the first and second jet kinematic variables is displayed in Figure B.6.
· Figure B.7 gives the average number of interactions per crossing ($\langle \mu \rangle$) and the dependence of various kinematic variables on $\mu$. Note that the assertion made in Chapter 6 that dijet masses formed from jets with $p_T > 50$ GeV are not affected by pile-up is supported.

· Figure B.8 gives the angular correlation between the leading and sub-leading jets with the third jet. Notice that the first and third jets are generally separated in $\eta$ and $\phi$, but the second and third jets are often in close proximity. This angular correlation can motivate “three-jet invariant mass searches” and the use of wide jets.
Figure B.1: Shown are kinematic distributions for the dijet system.
Figure B.2: Shown are kinematic distributions showing angular and transverse momentum balance for the dijet system.
Figure B.3: Shown are kinematic distributions for the leading-$p_T$ jet.
Figure B.4: Shown are kinematic distributions for the subleading-$p_T$ jet.
Figure B.5: Shown are kinematic distributions for the sub-subleading-$p_T$ jet.
Figure B.6: Shown are kinematic distributions displaying the correlation of average dijet invariant mass as a function of leading and sub-leading kinematic variables.
Figure B.7: Shown is the average interactions per crossing, $\langle \mu \rangle$, and the correlation of several quantities with $\langle \mu \rangle$. 
Figure B.8: Shown are the angular correlations between the identified dijets and the third jet.
Preliminary Studies Using Wide Jets

The use of “wide jets” could provide additional improvement to the dijet resonance search analysis. Motivated by methods implemented in the 2015 CMS dijet resonance search [83], an initial study was performed to gauge prospects for sensitivity improvement using wide jets. Below is a general overview of the method used for constructing wide jets for this study, a merit of increased sensitivity, and the proposed source of increased sensitivity. Following these studies, more rigorous approaches have been developed by the dijet analysis team to perform proper re-clustering using the tools described in Chapter 5 in preparation for additional Run II data.

C.1 Simple Model

In this study, wide jets are formed using a simple cone algorithm, where the inputs are anti-$k_T$ $R=0.4$ jets, or narrow jets. Beginning with a list of jets ordered in descending $p_T$, the leading and subleading jets are identified as the typical narrow dijet system. One-by-one, jets within $\Delta R$ of the leading jets are added to
Figure C.1: Cone size is optimized for a simple wide jet model applied to $q^*$ resonances and MC QCD background. Sensitivity increase is plotted with respect to narrow jets.

The closest primary jet. In this way, the leading and subleading jets shift in direction and grow in size as additional final-state radiation is captured. Jet inclusion is accomplished by simple four-vector addition.

Two factors can be optimized in this wide jet scheme: minimum $p_T$ threshold for inputs and $\Delta R$. For this study, the minimum $p_T$ was assumed to be 50 GeV, which allowed for the use of the dijet analysis MC samples. However, $\Delta R$ was optimized using an approach similar to the $y^*$ cut optimization described in Chapter 6. The resulting optimization curves for various $q^*$ signals are shown in Figure C.1. Plotted is relative increase in sensitivity compared to the narrow jets used in the current dijet analysis. As the cone size grows, final state radiation is collected and offers more accurate reconstruction of the hard-scatter event. However, a downward trend manifests as the cone grows to large sizes due to the addition of pile-up. An optimal value for signals at our current dijet $q^*$ limits was chosen to be $\Delta R = 1.2$. 

126
**Figure C.2:** $q^*$ resonances for wide and narrow jets are compared.

### C.2 Improved Sensitivity

Using the simple model described above, one can expect additional sensitivity. As seen in Figure C.1, the sensitivity for a $M_{q^*} = 5$ TeV signal increases approximately 30% for wide jets as compared to the traditional $R = 0.4$ jets used in the analysis presented in this dissertation. This increase is due to a significant sharpening of the dijet resonance peak as shown for two mass points in Figure C.2.
References


[63] Bradley Axen, Garabel Halladjian, Cristina Galea, Zach Marshall, Nils Ruthmann, Caterina Doglioni, Mario Jose Da Cunha Sargedas De Sousa, Patricia Conde Muino, Amelia Maio, Carsten Daniel Burgard, and David Jennens. A measurement of single hadron response using data


