Search for new physics in the dijet invariant mass spectrum at 8 TeV

Katherine Pachal

St Cross College, Oxford





Thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Hilary Term, 2015

Search for new physics in the dijet invariant mass spectrum at 8 TeV

Katherine Pachal
St Cross College, Oxford

Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

January 2015

Abstract

A search for resonance signals from new physics models was performed using the ATLAS dijet invariant mass spectrum at a center-of-mass energy of 8 TeV. The full dataset from the 2012 data-taking period was used, corresponding to an integrated luminosity of 20.3 fb^{-1}. The analysis covered a mass range from 250 GeV to 4.5 TeV and found no evidence of an excess anywhere in the spectrum. Upper mass limits were set at 95% CL on both model-specific benchmark new physics signals and on two different generic resonance forms: a Gaussian shape and a Breit-Wigner shape which accounted for pdf and detector effects. The analysis excluded excited quarks below 4.06, heavy W' bosons below 2.45, colour-octet scalars below 2.70, leptophobic chiral excited bosons W^* below 1.75, and quantum black holes below 5.66 and 5.62 for the QBH and BLACKMAX generators respectively.

Acknowledgements

I would like to thank the Clarendon Fund and the Natural Sciences and Engineering Research Council of Canada for funding my DPhil.

Thank you to my advisor Jeff for his support and advice, for giving me the freedom to explore my interests and seek out new projects, and for many entertaining coffee breaks and literary chats. I am also extremely grateful to Cigdem and Todd for support and guidance above and beyond the call of duty. Your willingness to give of your time and advice has truly shaped my experience of Oxford.

I have been fortunate enough to work with two of the most wonderful postdocs I know, so thank you from the bottom of my heart to Andreas and Caterina. Andreas got me on my feet in the experiment, handled my beginner's ignorance with grace, and gave me as much chocolate as I needed. Caterina has been my guide and inspiration throughout this degree and has taught me almost everything I know. I have no doubt that without her I would not be where I am today. As I start my own postdoctoral work, I will try my hardest to be to my own students what you have been to me.

I am so thankful to my amazing classmates here in Oxford, most especially Nazim and Mireia: living two years with you was a real joy, and I'll miss you so much in this next stage. Thank you to Krish for always being there for me, to Craig, Jim, and Shaun for being the cheeriest next-door office-mates one could hope for, and to everyone for being such fun friends and rewarding colleagues.

I thank Megan for being the most generous and loyal friend anyone could ask for and for brightening up the last year so tremendously. Your encouragement means the world to me.

Thank you to Matt, Max, Christina, Larry, Dani, Tova, Mia, and all my CERNbuddies for the warmth and support of your friendship and for creating a community in which I can feel so at home. The pressure of this field is a small price to pay for the honour of sharing it with you.

Above all, I want to thank Thibaut for his love and encouragement these many years. Your belief in me has made me both a happier person and a better physicist, and I treasure all the adventures we've had together. I can't wait to share with you whatever life and physics have in store for us next.

Finally, I thank my family. Mom, Dad, Owen, Anna: your incredible support and encouragement for all these years have made this achievement possible. This thesis is dedicated to you.

Declaration

Due to the collaborative nature of ATLAS, no analysis is entirely the work of one author. I am greatly indebted to my fantastic analysis group who shared this work with me. My personal contributions to the analysis are outlined below; as the work was so evenly shared, its exact division will not be noted in the body of the thesis. Any results taken from outside the dijet analysis group will be cited appropriately.

Personal contributions

- I was one of two contact people for the 8 TeV analysis between May 2012 and the submission of the paper in summer 2014. As such I was involved in all major analysis decisions, organised meetings, handled communications with the convenors and editorial board, and assisted more junior students.
- All statistical and limit-setting code was written by me. I was responsible for the change to a more rigorous limit-setting system and its implementation in a fast, flexible package.
- I was responsible for determining and validating the background estimation procedure, including testing a range of functions for stability and performance in data and Monte Carlo.
- I helped determine the correct trigger combination and tested various methods of combining the data streams.
- I made a substantial contribution to the definition of the new Breit-Wigner-based generic limits.
- I produced all limit plots and values and the data-to-background fit comparison plot for the paper.

Contents

1 Introduction **1**

I The Background **3**

2 Theoretical framework **4**
2.1 The Standard Model . . . 4
2.1.1 The electroweak sector and symmetry breaking . . . 7
2.1.2 Quantum chromodynamics . . . 9
2.2 Limitations of the Standard Model . . . 13
2.2.1 The hierarchy problem and Higgs mass fine tuning . . . 13
2.2.2 Strong CP problem . . . 14
2.2.3 Dark matter . . . 15
2.2.4 Matter-antimatter asymmetry . . . 17
2.3 Theories of new physics . . . 17
2.3.1 Grand unified theories . . . 17
2.3.2 Excited chiral bosons . . . 19
2.3.3 Fermion compositeness . . . 21
2.3.4 Exotic coloured resonances . . . 22
2.3.5 Extra dimensions, gravity, and black holes . . . 23

3 QCD from hard scattering to particle jets **26**
3.1 Factorisation . . . 26
3.1.1 Anatomy of a hadron collision . . . 27
3.2 The hard scatter . . . 28
3.2.1 Probing the proton structure . . . 28
3.2.2 Evaluating and evolving the parton distribution functions . . . 30
3.2.3 Parton-parton scattering at the LHC . . . 31
3.2.4 Divergences . . . 33
3.3 Final state evolution to the edge of perturbativity . . . 34
3.4 Hadronisation . . . 35
3.5 Jets and jet algorithms . . . 36
3.5.1 Jets as physically meaningful objects . . . 38
3.5.2 Infrared and collinear safety in jet definitions . . . 39
3.5.3 Jet reconstruction algorithms . . . 39

4 The ATLAS Experiment **43**
4.1 The LHC . . . 43
4.1.1 Running conditions in 2012 . . . 46
4.1.2 Pileup . . . 48
4.2 ATLAS Detector Overview . . . 49
4.2.1 Coordinate system . . . 50
4.3 Inner detector . . . 52
4.3.1 Pixel detectors . . . 53
4.3.2 Semiconductor Tracker . . . 54
4.3.3 Transition Radiation Tracker . . . 54
4.4 Electromagnetic calorimeter . . . 55
4.4.1 Electron and photon calibration . . . 58
4.5 Hadronic Calorimeter . . . 59
4.5.1 Tile (barrel) calorimeter . . . 60
4.5.2 Hadronic endcap (HEC) and forward (FCal) calorimeters . . . 62
4.5.3 Calibrating the energy scale . . . 63
4.5.4 Calorimeter performance . . . 64
4.6 Muon system . . . 65
4.7 Triggering . . . 66
4.7.1 The delayed stream in 2012 data . . . 67
4.8 Monte Carlo . . . 68
4.9 Data processing and object reconstruction . . . 69

5 Jets in ATLAS **70**
5.1 Jet definition and inputs . . . 70
5.2 Calibration and performance . . . 71
5.2.1 The EM+JES calibration scheme . . . 72
5.2.2 Performance in 2012 data . . . 75
5.3 Uncertainties . . . 76
5.3.1 Jet energy scale uncertainty . . . 76
5.3.2 Jet energy resolution uncertainty . . . 81

II The Dijet Analysis **84**

6 The Dijet Spectrum **85**
6.1 Trigger strategy . . . 87
6.1.1 Dijet analysis trigger map . . . 89
6.1.2 Trigger combination and event weights . . . 90
6.1.3 Comparison of alternative trigger methods . . . 92
6.1.4 Incorporating the delayed stream . . . 93
6.2 Event selection criteria . . . 99
6.2.1 Good run list . . . 99
6.2.2 Data integrity checks . . . 100
6.3 Jet calibration and selection . . . 101

6.3.1 JES and JER calibration . . . 101
6.3.2 Jet cleaning . . . 102
6.3.3 BCH_CORR_CELL cuts . . . 104
6.3.4 Additional hot tile removal . . . 109
6.3.5 Resistance to pileup . . . 110
6.3.6 Centrality requirement . . . 111
6.4 Summary of analysis selection . . . 112
6.5 The Spectrum . . . 114
6.5.1 Dijet mass resolution . . . 114
6.5.2 Binning strategy . . . 115
6.5.3 Selection of lowest dijet mass for the analysis . . . 117
6.5.4 The spectrum in 8 TeV data . . . 117

7 Searching for Excesses **119**
7.1 Statistical framework . . . 120
7.1.1 Definitions of probability . . . 120
7.1.2 Quantifying frequentist probabilities . . . 121
7.1.3 Comparing a weighted spectrum to a prediction . . . 122
7.1.4 Test statistics . . . 123
7.1.5 Extending test statistics to p-values . . . 128
7.1.6 Representing significance bin-by-bin . . . 130
7.2 The background estimate . . . 132
7.2.1 Parameterisation with a smooth function . . . 133
7.2.2 Creating the estimate . . . 134
7.2.3 The resulting background-only hypothesis . . . 136
7.3 Validating the background-only hypothesis . . . 136
7.3.1 Testing the search phase on blinded data . . . 137
7.3.2 Testing the search phase on Monte Carlo . . . 138
7.4 Validating the function choice . . . 142
7.4.1 Alternate parameterisations . . . 143
7.4.2 k-fold cross-validation in the full dataset . . . 145
7.5 Calculating the error on the background estimate . . . 150
7.5.1 Uncertainty on the fit result . . . 151
7.5.2 Uncertainty on the choice of function . . . 153
7.6 Background stability against signal contamination . . . 154
7.6.1 Signal contamination: large . . . 155
7.6.2 Signal contamination: small . . . 158
7.7 Search results in 8 TeV data . . . 163

8 Systematic Uncertainties **169**
8.1 Luminosity uncertainty . . . 169
8.1.1 Incorporation in dijets . . . 170
8.2 Beam energy uncertainty . . . 171
8.2.1 Effects of beam energy on signal MC . . . 172
8.3 Background uncertainty . . . 173

8.4 Background parameterisation . . . 173
8.5 Jet Energy Scale . . . 176
8.6 Jet Energy Resolution . . . 178
8.7 Angular resolution uncertainties . . . 179
8.7.1 The effects on dijets: Negligible . . . 180
8.8 Theoretical uncertainties . . . 180
8.8.1 Scale uncertainties: Negligible . . . 180
8.8.2 PDF choice uncertainty . . . 181

9 Limit Setting **183**
9.1 Theory of limit setting . . . 183
9.1.1 Bayes' Theorem and its use in signal quantification . . . 184
9.1.2 Selecting a function for the signal prior . . . 189
9.1.3 Upper limits and their distributions . . . 190
9.2 Implementation of Bayesian limits . . . 192
9.2.1 Markov Chain Monte Carlo and the Bayesian Analysis Toolkit . . . 193
9.2.2 Constructing a likelihood for the MCMC . . . 195
9.2.3 The complete dijet calculation . . . 203
9.3 Limits on benchmark models at 8 TeV . . . 205
9.4 Limits on generic resonance shapes . . . 208
9.4.1 Gaussian limits . . . 210
9.4.2 Breit-Wigner limits . . . 211
9.5 Evolution of limits . . . 216

III The Future **219**

10 Dijet-based searches in Run II **220**
10.1 Dijets in Run II . . . 220
10.1.1 Projected performance in early 13 TeV data . . . 220
10.1.2 Data scouting . . . 222
10.2 Related analysis plans . . . 222
10.3 What if? . . . 223

IV Appendices **225**

Bibliography **226**

Chapter 1

Introduction

'None but those who have experienced them can conceive of the enticements of science. In other studies you go as far as others have gone before you, and there is nothing more to know; but in a scientific pursuit there is continual food for discovery and wonder.'

— Mary Shelley, *Frankenstein*

Since taking its modern form in the 1960's, the Standard Model of particle physics has been subjected to rigorous experimental tests and has demonstrated a remarkable ability to describe nature across many orders of magnitude in energy. Despite this success, it has for decades been understood that this theory cannot be the full solution. Unexplained observations, such as the presence of dark matter in the universe and the breaking of certain symmetries in ways not described by the Standard Model, indicate that it may be just one part of a greater system.

The Large Hadron Collider, along with its two most general detectors ATLAS and CMS, was constructed to measure the elusive Higgs Boson, to test the boundaries of the Standard Model, and to discover what might lie beyond. Since 2010, the LHC has been collecting data from proton-proton collisions, first at a center-of-mass energy of 7 tera-electron-volts (TeV), then at 8 TeV. The discovery of the Higgs was announced in July 2012 to tremendous acclaim around the globe. However, the second physics goal has so far come to nothing: the data collected has shown no hints of any particles beyond those predicted by the Standard Model.

The most abundant physics objects produced in proton-proton collisions at the LHC are

jets, collimated sprays of particles resulting from the production of high-energy quarks and gluons. Their high production rate provides ample statistics for searches, while their strong force interactions make them a useful handle by which to search for new strongly interacting physics processes. The analysis presented in this thesis is a search for new physics signatures in a collection of dijet events. A new heavy object decaying to a pair of strongly charged particles would appear as a "bump" at the new particle mass when compared to a background of jet pairs from Standard Model processes. This analysis was performed on the 8 TeV dataset collected during 2012, and provided comprehensive limits on a variety of theories of beyond-the-Standard-Model physics.

The thesis is organised as follows. Chapter 2 introduces the theoretical background of the Standard Model, the objections raised against it, and the proposed solutions examined by the dijet analysis. The strong force and the jets for which it is responsible are explored in Chapter 3. The LHC and the ATLAS experiment are described in Chapter 4, while Chapter 5 describes the measurement and calibration of jets in ATLAS. The creation of the dijet invariant mass spectrum is described in Chapter 6, while Chapter 7 illustrates the statistical process used to determine whether or not it contains signs of new physics. The systematic uncertainties affecting the analysis are explored in Chapter 8 and the mass limits on a range of new physics models set by the results of the search are given in Chapter 9. Finally, Chapter 10 provides an outlook on the possibilities for the dijet analysis in 13 TeV data this year.

Part I

The Background

Chapter 2

Theoretical framework

'Memnon one day conceived of the mad project of becoming perfectly wise. There are few men to whom this folly has not occurred from time to time.'
— Voltaire, *Memnon, ou la sagesse humaine*

The Standard Model of particle physics describes three of the four fundamental forces and incorporates all currently observed particles. It has remained largely unchanged for decades and withstood or adapted to all direct experimental tests[1]. However, a range of concerns both empirical and aesthetic lead the scientific community to believe that this theory may not be the final answer. Many extensions and variations have been hypothesized, a range of which will be discussed here. Most information discussed in this chapter has been drawn from [1–4]. Other sources will be cited in the text.

2.1 The Standard Model

The Standard Model of particle physics is a mathematical framework which describes the strong, weak, and electromagnetic forces: that is, all fundamental forces except gravity, and it incorporates all directly observed elementary particles to date. The Standard Model (SM) is a quantum field theory belonging to the gauge group $U(1)_Y \times SU(2)_L \times SU(3)$. The generators of each group correspond to the spin-1 particles which are carriers of that field. Thus the $U(1)_Y$ group corresponds to a particle $B_\mu(x)$ with quantum number denoted by Y, the *weak*

[1] One exception may be raised to this statement: neutrino masses are known to be non-zero from their observed oscillations. While the simplest Standard Model as defined in this chapter does not contain neutrino masses, they may be added by a simple extension to this theory without the need to introduce new forces or particles. They are therefore not considered a contradiction to the Standard Model in this context.

hypercharge. Three force-carrying particles W_μ^a ($a = 1, 2, 3$) correspond to the $SU(2)_L$ group, and the quantum number of the group is *weak isospin* T. The subscript L indicates the restriction that only left-handed particles carry isospin, a restriction which is imposed by hand to reflect experimental observations. Note that these groups do not exactly correspond to the familiar forces observed in nature: the electromagnetic force carried by the photon and the weak force carried by the W and Z bosons correspond to linear combinations of the B_μ and W_μ fields, and thus the electric charge Q arises from a combination of the hypercharge and the third component of the isospin: $Q = T_3 + Y/2$. The group $SU(2)_L \times U(1)_Y$ as a whole represents the unified electroweak force. Finally, the $SU(3)$ component corresponds to the strong force: its quantum number *colour* takes 3 possible values referred to as *red*, *green*, and *blue* and its 8 force carriers $A_\mu^{\alpha=1,\ldots 8}$ are the gluons. Since the strong force is not unified with the electroweak force in the Standard Model, the fields do not mix and the Lagrangian for the full theory can be separated into $SU(3)$ and $SU(2)_L \times U(1)_Y$ components [1].

Particles fall into two categories: *fermions* with spin 1/2 and *bosons* with spin 0 or 1. The bosons are the force carriers of the previously mentioned fields and are exchanged by fermions to mediate interactions. The photon γ is massless and carries no charge while the electrically neutral Z boson and the charged W^+ and W^- bosons are massive. The 8 gluons mentioned above are generally referred to as carrying two colours each, but this is a simplification: in reality the colour states of gluons are linear combinations of the three colours invariant under all global $U(3)$ transformations. Solving for the decomposition of $U(3)$ to $SU(3) \times U(1))$ allows the definition of 8 distinct colour + anti-colour combinations which are the generators of $SU(3)$ and one colourless singlet, the generator for $U(1)$, for which there is no corresponding particle in nature. The coloured octet of $SU(3)$ corresponds to the 8 physical gluons. The fundamental forces and their corresponding gauge bosons are summarised in Table 2.1.

Fermions are the fundamental constituents of matter and are further subdivided into *quarks* and *leptons*, where only the former interact via the strong force. Leptons come in two types, massive leptons (e, μ, and τ) and neutrinos. The heavy leptons carry electric charge $Q = \pm 1$ while the neutrinos are electrically neutral, although both interact via the weak force. Since neutrinos interact via the weak force alone and do not couple to right-handed particles, only left-handed neutrinos are included in the Standard Model. To each massive lepton

Force	Carriers	Charge carried
Electromagnetic	γ	none
Weak	$W^{\pm}$	weak, electromagnetic
	Z	weak
Strong	g	colour

Table 2.1: Forces of the standard model with the bosons which carry them and their quantum numbers. The weak isospin is given by T and its third component by T_3. Y designates the hypercharge. The electromagnetic charge is Q and colour charge is C.

corresponds one neutrino (ν_e, ν_μ, ν_τ) and weak interactions between the leptons preserve this family number. The quarks (antiquarks) carry electric charge of Q = +2/3 (Q = −1/3), weak isospin, and colour charge. Like the leptons, they come in three generations which differ only by mass and by the decays those masses permit. Each generation is composed of two quarks of different flavours: u and d in the first generation, c and s in the second, and t and b in the third. The fermions are summarised in Table 2.2.

	Families			Quantum numbers				
	1st	2nd	3rd	T	T_3	Y	Q	C
Quarks	u_L	c_L	t_L	1/2	1/2	+1/3	+2/3	(r, g, b)
	d_L	s_L	b_L	1/2	−1/2	1/3	−1/3	(r, g, b)
	u_R	c_R	t_R	0	0	+4/3	+2/3	(r, g, b)
	d_R	s_R	b_R	0	0	−2/3	−1/3	(r, g, b)
Leptons	e_L	μ_L	τ_L	1/2	−1/2	−1	−1	none
	$\nu_{e,L}$	$\nu_{\mu,L}$	$\nu_{\tau,L}$	1/2	1/2	−1	0	none
	e_R	μ_R	τ_R	0	0	−2	−1	none

Table 2.2: Fermions of the standard model with their quantum numbers under the fundamental forces of particle physics [5]. The three families differ in mass but share quantum numbers. Here T is the weak isospin and T_3 its third component, Y is the hypercharge, and Q and C are the physically recognisable electromagnetic and colour charges. The left-handed quarks and leptons transform as doublets while the right-handed ones are singlets. No right-handed neutrinos are included in the Standard Model.

The only scalar boson in the Standard Model, the Higgs boson, is not a force carrier but rather arises as a consequence of symmetry breaking in the Standard Model gauge group. The Higgs field is responsible for the masses of the gauge bosons and fermions via the mechanism discussed in the following section.

2.1.1 The electroweak sector and symmetry breaking

The gauge group of the Standard Model is known only experimentally, with the field content selected to reflect observation. However once the group is defined, the Lagrangian for the electromagnetic and weak force sector can be constructed directly from the permitted interactions of the W_μ and B_μ fields with the fermions. The most general Lagrangian to describe these interactions which is both gauge-invariant and renormalisable is [1, 5]:

$$\begin{aligned}\mathcal{L}_{EW} = &-\frac{1}{4}W^{a\mu\nu}W^a_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\ &+ \sum_{\psi_L}\overline{\psi}_{Lj}i\gamma^\mu D_\mu\psi^j_L + \sum_{\psi_R}\overline{\psi}_{Rj}i\gamma^\mu D_\mu\psi^j_R\,.\end{aligned}$$

Note that the covariant derivative D_μ must reflect the fact that the ψ_L are doublets which couple to W_μ while the ψ_R are singlets and do not:

$$D_\mu\psi_R = (\partial_\mu + ig_1Y_WB_\mu)\psi_R\,, \tag{2.1}$$

whereas

$$D_\mu\psi_R = \left(\mathbf{I}(\partial_\mu + ig_1Y_WB_\mu) + ig_2\frac{\vec{\boldsymbol{\tau}}}{2}\vec{W}_\mu\right)\psi_L\,. \tag{2.2}$$

Here g_1, g_2 are the gauge coupling constants for the $U(1)$ and $SU(2)$ groups respectively, Y_W is the hypercharge of the relevant fermion, and $\mathbf{I}$ and $\vec{\boldsymbol{\tau}}$ are the 2×2 identity and the Pauli matrices. Expanding the above Lagrangian using these covariant derivative definitions reveals terms such as $-g_1Y_W\overline{\psi}_{Lj}\gamma^\mu B_\mu\psi^j_L$, indicating the existence of an interaction vertex between two ψ_L fermions and the B_μ field. The strengths of the interactions thus encoded are governed by the coupling constants in each term.

A comparison of the fields in $\mathcal{L}_{EW}$ to the observed bosons and fermions shows that they cannot be the same: no particle included in $\mathcal{L}_{EW}$ is massive. Mass terms, which take the form $-m^2f^2$ for a field f, cannot be directly added to the Lagrangian because they violate local gauge invariance. The masses must instead arise naturally by forcing the symmetry to break.

Let a scalar complex doublet Φ be introduced with some potential $V(\Phi^\dagger\Phi)$. Allowing the most general gauge-invariant interactions with the fermion sector, the following terms

are added to the full Lagrangian:

$$\mathcal{L}_H = -(D_\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi^\dagger\Phi) - g_a\overline{\psi}_L\Phi^\dagger\psi_{a,R} - g_b\overline{\psi}_L\Phi^\dagger\psi_{b,R}\,, \tag{2.3}$$

where $\psi_{a,R}$ and $\psi_{b,R}$ are the up and down/charged lepton right-handed singlets respectively, and the terms with both ψ and Φ are *Yukawa interactions*. Let the potential be

$$V(\Phi^\dagger\Phi) = -\mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2\,, \tag{2.4}$$

where μ and λ are arbitrary coefficients. Note that this potential is not at its minimum at $\langle\Phi\rangle = 0$ and is in fact minimised when $|\Phi|^2 = -2\frac{\mu^2}{\lambda}$. Any arbitrary point can be selected for the components of Φ which satisfies this requirement: let $(\phi_1, \phi_2) = (+(-4\mu^2/\lambda)^{1/2}, 0) \equiv (v, 0)$ be selected. Then the Lagrangian can be rewritten in terms of the ground state of Φ (let it be ϕ) by using the definition $\phi_A = \phi_1 - v$, $\phi_B = \phi_2$, a change of variable which introduces a range of new terms.

The most critical terms introduced by the expansion are the following. First,

$$V(\Phi^\dagger\Phi) = -\frac{\mu^2v^2}{4} + \mu^2\phi^2 + \lambda v\phi^3 + \frac{\lambda}{4}\phi^4\,. \tag{2.5}$$

This specifies a mass for the ϕ boson of $\mu/\sqrt{2}$ as well as defining 3-ϕ+v (the Higgs field) and 4-ϕ interaction vertices. Second, the gauge boson interactions are affected:

$$(D_\mu\Phi)^\dagger(D_\mu\Phi) = \frac{v^2}{2}\left(g_1Y_WB_\mu + g_2\frac{\vec{\boldsymbol{\tau}}}{2}W_\mu^3\right)^2 + \frac{v^2}{2}(\frac{\vec{\boldsymbol{\tau}}}{2})^2g_2^2(W_\mu^1)^2 + \frac{v^2}{2}(\frac{\vec{\boldsymbol{\tau}}}{2})^2g_2^2(W_\mu^2)^2 + (D_\mu\phi)^\dagger(D_\mu\phi) + \text{interaction terms.} \tag{2.6}$$

The new terms assign masses to three fields: identical masses for the first and second W_μ fields which become the massive $W^\pm$ bosons and a slightly different mass for a linear combination of B_μ and W_μ^3 which corresponds to the Z. Once the bosonic portion of the Lagrangian is expressed with the same change of variables, the orthogonal linear combination can be identified as a 4th and massless gauge boson, the photon. Expansion of the Yukawa terms in the same basis provides mass terms for the fermions of the form $-g_f\frac{v}{\sqrt{2}}\bar{f}_mf_m$, where a separate coupling constant exists for each fermion corresponding to their different masses. Neutrinos do not receive masses in this process due to the lack of any right-handed components.

The above spontaneous symmetry breaking mechanism by which masses are given to certain fields without introducing any massless scalars is known as the Higgs or Brout-Englert-Higgs [2] mechanism [6–11]. The massive ϕ "Higgs" boson predicted by the model was the final undiscovered piece of the Standard Model until July 2012, when ATLAS and CMS announced the observation of a particle consistent with its properties [12, 13]. While it cannot be stated that this is the Higgs boson of the Standard Model, it is now considered certain that it is at least "a" Higgs boson. This remarkable result is a great triumph of the Standard Model.

2.1.2 Quantum chromodynamics

QCD corresponds to the $SU(3)$ group of the Standard Model and is a non-Abelian gauge theory, and as a consequence its force carrier, the gluon, experiences self-interactions. The QCD Lagrangian is given by:

$$\mathcal{L}_{QCD} = \sum_f \bar{\psi}^i_f (i\gamma^\mu \delta_{ij} \partial_\mu - g_s t^C_{ij} A^C_\mu - m_f \delta_{ij}) \psi^j_f - \frac{1}{4} F^A_{\mu\nu} F^{A\,\mu\nu} \,, \tag{2.7}$$

where $F^A_{\mu\nu}$ is the field tensor defined by

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g_s f_{ABC} A^B_\mu A^C_\nu \,. \tag{2.8}$$

The ψ^i are quark field spinors with flavour f and mass m_f. The A^α_μ are gluon fields, with α ranging from 1 to $N_C^2 - 1 = 8$. The t^C_{ij} are the $SU(3)$ field generators, 3×3 matrices with the property that $[t^A, t^B] = i f_{ABC} t^C$. Finally, g_s is the strong coupling constant and is the only free parameter in $\mathcal{L}_{QCD}$ [14, 15]. Its value will be discussed in the following section.

Expanding the above equation, three types of vertices are visible: a $q\bar{q}g$ vertex, a ggg vertex proportional to g_s, and a 4-g vertex proportional to g_s^2. These vertices are illustrated in Figure 2.1.

Any experimental observable must be invariant under transformations in their symmetry group. It can be shown that for $SU(3)$ where quarks and antiquarks rotate using opposite

[2] or Anderson-Higgs or Higgs-Kibble or Anderson-Higgs-Kibble etc. etc. Other frequently connected names of note are Guralnik, t'Hooft, and Hagen. Any interpretation of the resulting naming proliferation is left as an exercise for the reader.

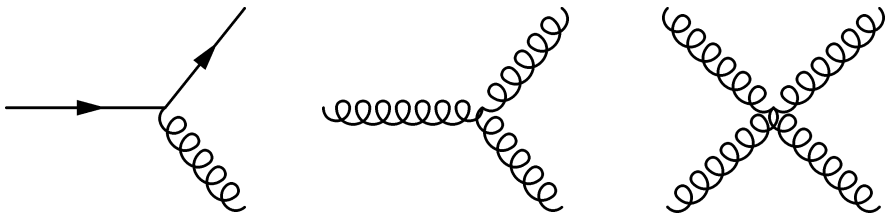


Figure 2.1: Vertices involving the strong force. Curling lines represent gluons and straight lines quarks. The diagrams can be interpreted as (2.1a) a quark radiating a gluon or a gluon splitting to generate a quark-antiquark pair, (2.1b) gluon splitting, and (2.1c) a four-gluon self-interaction vertex.

transformations then individual quarks are not invariant under rotation but the inner product of a q and a $\bar{q}$ is. There are three simple combinations of quarks which are rotationally invariant in $SU(3)$:

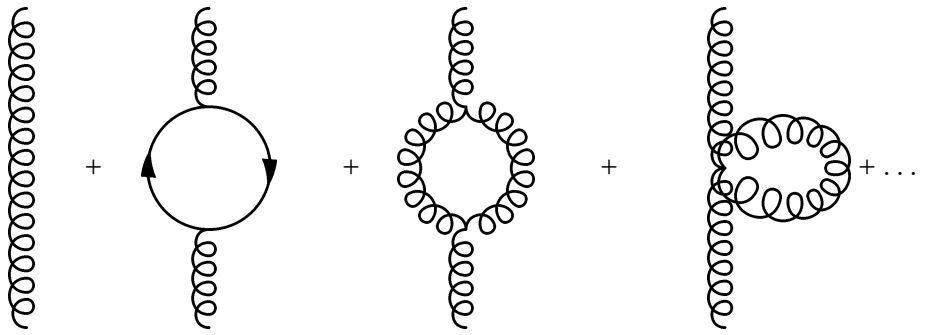
$$q^i \bar{q}_i, \qquad \epsilon^{ijk} q_i q_j q_k, \qquad \epsilon^{ijk} \bar{q}_i \bar{q}_j \bar{q}_k \,. \tag{2.9}$$

Note the colour indices in the above combinations. These are all colour singlets containing either a colour and its anticolour or one of each of three colours or anticolours. Such *colourless* particles are mesons ($q\bar{q}$) and baryons (qqq, $\bar{q}\bar{q}\bar{q}$) and are the only permitted light hadrons and are the only form in which quarks can be observed[3].

Renormalisation, asymptotic freedom and confinement

For QCD as for any theory, the simple 2-point propagator representing the transition of a particle from point x to y sustains higher-order corrections from a variety of loop diagrams. This is illustrated for a sample of one-loop corrections to the strong (gluon) propagator below.

[3]Four-quark colourless states are also possible in theory, e.g. $qq\bar{q}\bar{q}$. Such exotic mesons have been observed at Belle and LHCb [16] since 2007, with similar hints appearing at other experiments. They are, however, highly unstable and rare in comparison to baryons and mesons.



(2.10)

The contribution of all such diagrams across all orders (2-loop, 3-loop, and so forth) provides a diverging infinite sum which can be circumvented by using renormalisation: a (finite) set of parameters of the theory are rescaled to counteract the additions from the loops. The process of renormalisation always introduces a scale (here μ_R) at which the calculations for the rescaling procedure are made. The amplitude for any scattering process is a function of the momentum transfer t only and, since it is a physically measurable quantity, cannot depend on the renormalisation scale. Thus for a generic scattering amplitude $\boldsymbol{A}(t)$,

$$\mu_R \frac{d(t\boldsymbol{A}(t))}{d\mu_R} = 0\,. \tag{2.11}$$

For convenience let the renormalisation scale be chosen such that $t = -\mu_R^2$. The choice of amplitude to be calculated is immaterial, since the dependence of α_s on the scale will be the same in any case. The derivative of such an amplitude will result in an equation like the following for any renormalisable theory, where the *beta function* can be expanded in orders of α_s:

$$\mu_R \frac{d\alpha_s(\mu_R^2)}{d\mu_R} = \beta(\alpha_s) = -\left(b_0\alpha_s^2 + b_1\alpha_s^3 + b_2\alpha_s^4 + \ldots\right), \tag{2.12}$$

The b_i are obtained from the perturbative calculation of the chosen amplitude and correspond to the order of diagrams with i loops. Equations with the above form exist for all renormalisable theories and are known as *renormalisation group equations*, where the individual terms will depend on the renormalisation group of the theory [17]. For QCD, $b_0 = (33-2n_f)/(12\pi)$. Substituting this value and truncating to first order, the resulting differential equation can be solved:

$$\alpha_s(\mu^2) = \frac{\alpha_s(\mu_R^2)}{1 + b_0\alpha_s(\mu_R^2)\ln(\mu^2/\mu_R^2)}\,. \tag{2.13}$$

This defines an α_s which varies with the scale μ, known as a *running coupling*. The measured values of α_s for a variety of scales and a range of experimental observables are plotted in Figure 2.2.

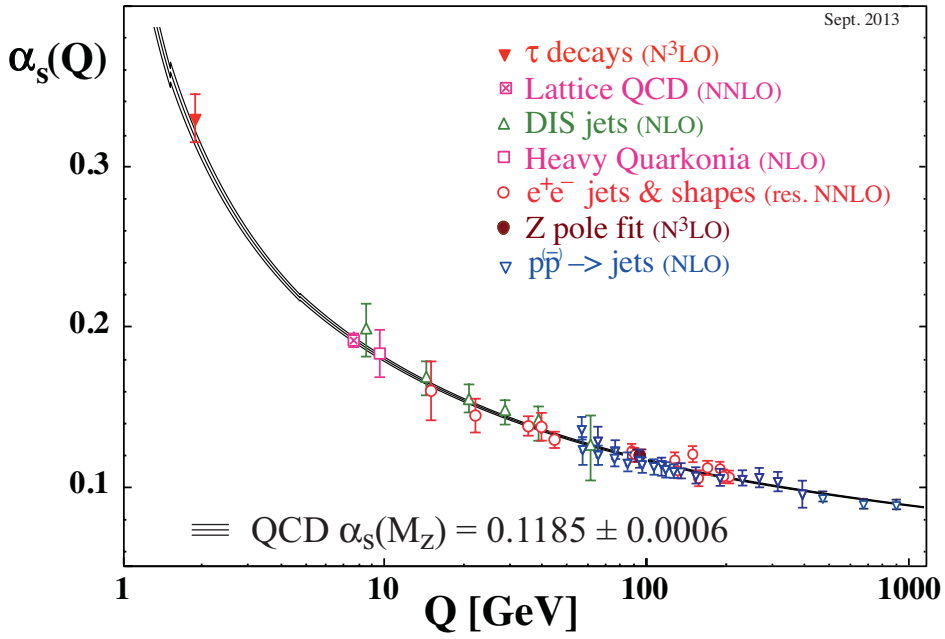


Figure 2.2: Combined measurements of α_s as a function of the scale Q for a variety of experimental sources. NLO, NNLO etc indicate the order of perturbative QCD used in the calculation of the reported value [15].

Note the behaviour of the coupling towards low and high scales. As Equation 2.13 also reflects, the coupling is largest at low scales and lessens dramatically as the scale increases. This behaviour has two critical results, both well established experimentally. *Asymptotic freedom* reflects the fact that if two quarks are very close together the coupling between them is extremely small and they barely interact. However, as two quarks are separated more and more the coupling between them increases until such a point that less energy is required to create two new quarks than to continue separating the original pair. The result will be two separate quark-antiquark pairs. This illustrates *confinement*: due to the strength of the strong force at long distances, no quark can exist alone in the laboratory. They will always be observed in bound states, specifically the colourless hadrons of the previous section.

The value of α_s is usually rephrased in terms of the scale parameter Λ_{QCD}, defined by

$\ln(\Lambda^2_{QCD}) = \ln(\mu^2) - 1/(b_0\alpha_s(\mu^2_R))$:

$$\alpha_s(\mu^2) = \frac{1}{b_0 \ln\left(\frac{\mu^2}{\Lambda^2_{QCD}}\right)}. \tag{2.14}$$

This is of course only a useful scale for a leading-order definition of α_s. The picture gets much more complex once higher-order terms are considered, and the definition of Λ_{QCD} becomes somewhat ambiguous. The critical consideration is that for any calculation expanded in orders of α_s, these terms will only get less significant at higher orders for scales of $\mu > \Lambda_{QCD}$. At lower scales where confinement effects dominate and the higher-order terms become large, QCD is no longer a perturbative theory. The value of Λ_{QCD} cannot be exactly defined, since changes in its scale can be absorbed into the $\mathcal{O}(\alpha_s^2)$ term, but it must be of order 100 MeV to be compatible with the size of hadrons.

The consequences of these behaviours for QCD in proton-proton collisions, and their use in designing experimental observables and search methods, will be explored in the next chapter.

2.2 Limitations of the Standard Model

Despite the great success of the Standard Model in predicting interactions between the known particles, it experiences some tension with both observations which do not fit any prediction of the model and with many scientists' ideas of how a physical theory ought to behave. Certain measurements indicate that either more particles or different interactions must exist than are codified in the SM. The most important outstanding issues are discussed below.

2.2.1 The hierarchy problem and Higgs mass fine tuning

The Standard Model has no way to account for a troubling discrepancy between two underlying scales: the scale of electroweak symmetry breaking at which the SM operates ($\mathcal{O}(100)$ GeV) and the Planck scale at which general relativity operates and on which gravity becomes strong ($\mathcal{O}(10^{16})$ TeV). The title *big hierarchy problem* refers to the difficulty in unifying these scales with any overarching theory without resorting to extreme *fine tuning*.

This difference in scales directly leads to a fine tuning problem in the mass of the Higgs boson. The Higgs mass is directly related to its coupling in the SM Lagrangian via $m_H^2 = 2\mu^2$, but it should also receive corrections due to loop diagrams involving other SM particles. Calculating the contributions from such diagrams indicates that modifications to the mass would follow the expression

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 + \dots \tag{2.15}$$

for fermions, and

$$\Delta m_H^2 = \frac{|\lambda_f|^2}{16\pi^2}\left[\Lambda_{UV}^2 - 2m_S \ln(\Lambda_{UV}/m_S) + \dots\right] \tag{2.16}$$

for bosons. In the above equations the λ_f and λ_S are the Yukawa couplings of the particles and Λ_{UV} is an ultraviolet cutoff scale. Since the Standard Model must be treated mathematically as though it would remain valid up to the scale of gravitational interactions and the introduction of some new overarching theory, this cutoff is of the scale of the Planck mass. These resulting quadratically diverging corrections to the Higgs mass are thus 10^{17} orders of magnitude larger than its observed mass.

It could be that the bare mass of the Higgs is such that all of these loop contributions cancel out and result in the observed mass of 125 GeV. This, however, requires the bare mass and the corrections to match up to the 33^{rd} decimal place – a highly coincidental result. This level of fine tuning is certainly possible, but it is improbable that the final measured mass of the Higgs would be as small as it is.

Introduction of new physics around the TeV scale can solve this problem, either by changing the scale at which gravity applies (this is the case in extra-dimensional models) and thus reducing the cutoff and removing the degree of fine-tuning, or by introducing automatic cancellations between fermionic and bosonic terms, as in supersymmetry. A third solution treats the Higgs boson as composite, and without its behaviour as a fundamental scalar these loop corrections no longer apply.

2.2.2 Strong CP problem

A separate fine-tuning question tackles the lack of CP violation in the strong force, known as the *strong CP problem*. CP violation is naturally permitted in the strong force since no

physical principle prevents the existence of the interaction term:

$$\Delta\mathcal{L}_\theta = -\frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}F^{\alpha\mu\nu}F^{\alpha\lambda\rho}\,. \tag{2.17}$$

This could be added to Equation 2.7 with no violation of invariance. A non-zero value of Θ_3 would result in observable CP violation. So far, however, the best experimental limits from measurements of the neutron electric dipole moment have constrained its value to $\Theta_3 \leq 10^{-9}$ [1]. There is no inherent reason for the value to be so small in the Standard Model, and while simply setting it equal to zero by hand might be a permitted strategy in the case that $SU(3)$ was the only group present, the evident CP violation of the weak force makes this option unacceptable [5].

2.2.3 Dark matter

Although indirect hints at the existence of *dark matter* were discovered as early as the 1930s, the measurement of galactic rotation speeds in 1970 provided the first clear evidence that some additional mass invisible to telescopes must be present in and around galaxies in order to cause the gravitional effects observed. The distribution of rotation speeds across the radii of galaxies adds to anomalous observations of gravitational lensing to indicate that vast amounts of additional unseen mass must be present [18, 19].

Most recent evidence for DM comes from measurements of the cosmic microwave background (CMB) and enables the calculation of the total amount of dark matter in the universe rather than simply the galactic scale [20]. A measurement of the anisotropies observed in the CMB can be used to constrain the parameters of a given cosmological model, including the abundance of matter. The world's leading CMB power spectra using data from the WMAP and Planck satellites lead to a measurement of the dark matter density of $\Omega_c h^2 = 0.1196 \pm 0.0031$ [21, 22]. Compared to the measured baryonic matter density of $\Omega_b h^2 = 0.02207 \pm 0.00033$, this indicates that there is ~ 5.5 times as much dark matter as ordinary matter in the universe.

There is no appropriate DM candidate in the Standard Model. The neutrino was briefly considered but has proven too light and non-abundant to produce the correct relic density. Extremely heavy right-handed neutrinos are a possibility, but difficult to reconcile with the

low masses and mixing requirements of the known neutrinos [23]. Thus a new particle outside the SM is expected to exist. The following properties of such a DM particle can be inferred from cosmological observations and previous collider results:

- It does not interact electromagnetically (hence *dark*). Otherwise it could be observed with standard astronomical procedures.
- It is unlikely to interact via the strong force. It is possible to construct models which are exceptions to this rule, but they are disfavoured due to the fact that such particles have not yet been experimentally observed, and strong interactions would make such particles much easier to detect.
- It must be stable since no SM decay products are observed.
- It must have a *freeze-out temperature* matching the observed relic density. At some early point in the universe thermal equilibrium would have existed between dark matter particles and other (SM) particles. As the universe expanded and matter cooled, this equilibrium was at some point no longer possible and the amount of such particles became essentially fixed. The temperature at which this happens is related to the particle masses and annihilation cross-sections and determines their final density. Thus the dark matter candidate must have a mass and cross section which predict the correct relic density [24].
- It clearly interacts gravitationally, but there are very few constraints on the masses of the dark matter particles themselves.
- Its interactions, or lack thereof, with the weak force are unknown.

This final point provides a possible handle for measuring particulate dark matter. Many particles that fill this role have been postulated, but the current leading candidates are Weakly Interacting Massive Particles (WIMPs), a catchall name for theorised species including sterile neutrinos, axions, a whole barrage of supersymmetric particles, Kaluza-Klein states, and more. The so-called "WIMP miracle" corresponds to the fact that the correct relic density of DM is obtained if the particles are assumed to have a mass of ~ 100 GeV and an annihilation cross-section of the order of a weak force interaction[4]. Dark matter searches are ongoing at the LHC and if particulate dark matter does indeed take a weakly interacting

[4]The assumption that DM particles are weakly interacting accords well with the freeze-out temperature requirement but is also to a non-negligible extent hopeful, if not wishful, thinking. Without weak interactions the chance that DM particles could ever be observed in a laboratory is very small.

form it will hopefully be discoverable in the Run II dataset. The one certainty is that the existence of dark matter poses a serious problem for the Standard Model.

2.2.4 Matter-antimatter asymmetry

One of the most fundamental observable properties of our universe is the fact that it is made entirely of matter. Theories of leptogenesis consistent with the Standard Model indicate that matter and antimatter should have been produced in equal quantities at the Big Bang, yet the vast majority of the universe is matter today. The Standard Model introduces some matter-antimatter asymmetry into the weak interaction through the nonzero mixing angles in the CKM matrix, but this is far too small an effect to account for the asymmetry observable in the universe today [25].

2.3 Theories of new physics

Model-nonspecific searches for physics beyond the Standard Model are often organised by final state rather than by the theory sought. Such *signature-driven* searches tend to be sensitive to a wider range of new physics theories than a search tuned specifically to the predicted particles of a preferred theory. The analysis discussed in this thesis searches for new physics which produces two hard objects charged under the strong force: two quarks, two gluons, or a quark and a gluon. Many models predict this final state, and five were chosen as *benchmarks* for the analysis. That is, they are well-understood models which provide theorists with clear interpretations of the experimental data. The selected models derive from a range of theoretical motivation and cover all three possible final states. The BSM theories used as benchmarks in the analysis are introduced here.

2.3.1 Grand unified theories

A grand unified theory (GUT) is one which places the Standard Model within a larger gauge group, thus unifying QCD with the electroweak sector and transforming the three separate coupling parameters into facets of a single parameter. Such theories provide several aesthetic benefits: they reduce the number of free parameters in the model by defining the relationship between the couplings, explain the quantisation of electric charge, and eliminate the chiral

(triangle) anomaly by placing leptons and quarks in the same family [26]. The simplest option and the first proposed GUT was $SU(5)$, but a basic model based on this group predicts proton decay and specific relationships between quark and lepton masses, both of which contradict the observed data [27]. Many more complex models now exist, favourites of which include $SO(10)$ or E_6.Any decomposition of these models which introduces a new factor of $SU(2)$ simultaneously introduces new charged and neutral gauge bosons $W_R^{\pm}$ and Z_R^0. Observational limits constrain the masses of these new particles to be several times heavier than their known SM counterparts, but they could fall well within the reach of the LHC.

The collection of new heavy charged and neutral bosons generically referred to as W' and Z' provide a handle to search for such gauge-group extending theories which are conveniently unconstrained by the exact model generating them. A search using new massive bosons as a benchmark is thus sensitive to more than one model and the exact form of new boson used to constrain these models becomes moot. The benchmark selected in this search to constrain the variety of theories is the "sequential standard model" (SSM) W', that is, a left-handed W' sharing the same couplings as the Standard Model W and varying only in mass [28, 29].

A naïve transcription of the SSM bosons leaves SM-like couplings to quarks and leptons but necessitates the addition of a $W'^{\pm}W^{\mp}Z$ vertex:

$$\begin{aligned} V_{W'^{\pm}W^{\mp}Z} &= V_{W^+W^-Z'} \\ &= ie\cot\theta_W \cdot [g_{\mu\nu}(q-p)_\lambda + g_{\mu\lambda}(p-r)_\nu + g_{\nu\lambda}(r-q)_\mu] \end{aligned} \tag{2.18}$$

The branching ratio associated with this process increases dramatically with mass, meaning that $W' \to WZ$ would become the dominant process for $M_{W'} \gtrsim 250$ GeV. However, in the majority of extended gauge models this factor is absent up to the point of symmetry breaking and beyond that point is suppressed by a factor on the order of $M_W^2/M_{W'}^2$. This leads to a greatly reduced branching ratio to WZ, as visible in Figure 2.3. A higher branching ratio can be introduced through interference between the W' and the SM W, but is neglected in the model used for this analysis for simplicity. Other ATLAS searches cover the $\ell\nu$ and WZ decay channels of the W', with the leading limit currently resulting from the leptonic decay, where a decrease in branching ratio is counterbalanced by a decrease in background

production rates [30]. The results of the collection of searches are complementary, providing a thorough assessment of a SSM W' scenario.

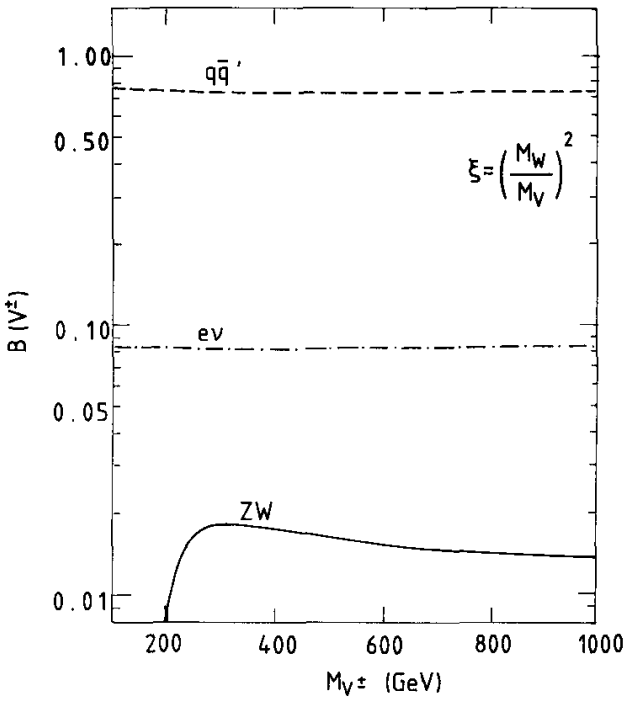


Figure 2.3: Branching ratios of the W' model used as a benchmark for the search. The coupling of W' to W and Z is suppressed by a factor $M_W^2/M_{W'}^2$, making quark pair production the dominant process [28].

2.3.2 Excited chiral bosons

One specific GUT-like modification of the SM involves an $SU(3)_W \times U(1)_W$ extension to electroweak group. Spontaneous breaking of this symmetry group can take a variety of forms as explored in [31] but always results in a light Higgs doublet and the introduction of new heavy chiral spin-1 bosons. These are equivalent in many ways to the heavy gauge bosons represented by W' but transform differently. Where a W' transforms under the real (triple adjoint) representation (1/2, 1/2) of a spin-1 particle like the Standard Model W, the W^* transforms via the inequivalent chiral representations (1, 0) and (0, 1). The full theory is developed in [31, 32] but the primary consequence of the chiral representation is that the W^* to first order only shares magnetic-type interactions with the Standard Model particles.

A simplified reference model is used wherein the magnetic interactions of the heavy chiral

particle W^* are specified by the following Lagrangian [33]:

$$\mathcal{L}_{ref,W^*} = \frac{g}{M_{W^*}}(\sin\theta_X \bar{u_L}\sigma^{\mu\nu}d_R + \frac{2}{\sqrt{3}}\cos\theta_X \bar{\mu_R}\sigma^{\mu\nu}d_L + \sin\theta_X \bar{\nu_L}\theta^{\mu\nu}e_R)\partial_\mu W^{*+}_{\mu\nu} + \text{h.c.} \tag{2.19}$$

Here the first and third terms correspond to the coupling of the W^* to left-handed up-type and right-handed down-type fermions while the second term describes its coupling to right-handed up-type and left-handed down-type fermions. There is no equivalent of the third term for this second type of interaction because while the model introduces a right-handed neutrino, it is assumed to be extremely heavy and thus decoupled at the weak scale on which this interaction occurs [34]. In order to facilitate comparisons with the W' heavy boson family, the coupling constants are fixed in such a way as to ensure the total width of the W^* is the same as that of a W' of the same mass. The remaining free parameter is the mixing angle θ_X: this determines the relative coupling of the W^* to quarks and to leptons. The sample W^* model used to search for chiral theories of this type in the following analysis fixes $\sin\theta_X = 0$ to simulate a leptophobic W^* with the quark-based final state of interest, but also includes a simpler interpretation corresponding to leptophilic ($\sin\theta_X = 1$) models for comparison.

The cross section for single production of W^* particles and their decay to a pair of fermions is explored in [35] and shown to have the following dependence on azimuthal angle of the product fermions:

$$\frac{d\sigma(q\bar{q} \to W^* \to f\bar{f})}{d\cos\theta} \propto \cos^2\theta\,. \tag{2.20}$$

That is, unlike the gauge bosons (and heavy gauge bosons such as W') which have the highest production cross-sections for final state fermions perpendicular to the beam axis, the W^* cross-section decreases near $\theta = \frac{\pi}{2}$ and has its maximum near $\theta \sim \frac{\pi}{4}$ [33]. For lighter masses a strong longitudinal boost in the final particles is possible and the cross-section more closely resembles that of a W', but the differences become more pronounced as the W^* mass increases [36]. The experimental consequences of this behaviour on the search for such models will be explored later.

2.3.3 Fermion compositeness

The appearance of generations of fermions, identical but for mass and with a large variation in that mass, poses a puzzle to which one solution would be a deeper layer of structure. If quarks were not fundamental particles but rather bound states of sub-particles (the most common term is *preons*), then the observed quark masses and mixing angles could be explained in much the same way that the properties of hadrons are understood through the combinations of quarks [24, 37].

A directly observable consequence of compositeness would be the presence of high mass excited quarks, q^*. Assuming a simple model of excited quarks with spin and isospin 1/2, then the form of the coupling between these quarks, ordinary fermions, and gauge bosons is fixed by gauge invariance and the branching ratios of the q^* can be calculated. Under the condition that the q^* mass $M^* > m_{W,Z}$ (known from current experimental results), the dominant branching ratio is via the strong force to a normal quark and a gluon: $q^* \to qg$. Fixing the three free parameters of the model equal to one another (f), the total width of the q^* can be calculated:

$$\Gamma(q^*) \simeq 0.04 f^2 M^* \,. \tag{2.21}$$

The ratio 0.04 is an analysis choice made to ensure that for $f \sim 1$ the q^* is a narrow resonance. The exact model used and the parameter values assumed in simulation are described in [37, 38].

Pair production of excited quarks is suppressed in comparison to single production. Production cross sections for the processes $pp/p\bar{p} \to q^* \to qg$ can be calculated and compared to the expected background in the two-parton final state channel. The results, for a range of q^* masses with the parameter values used by the simplified model discussed in [37], are shown in Figure 2.4. Note that a shift of the model free parameters away from 1 would change the width of the observable q^* resonance but not its peak value. Figure 2.4 corresponds to a $pp/p\bar{p}$ collider with a centre-of-mass energy of 40 TeV and therefore does not make exact predictions for the LHC, but the qualitative behaviour illustrates the value of the excited quark as a benchmark for a search for fermion compositeness: with its high production cross section and branching ratio, combined with its narrow width, resonances due to excited quarks are expected to be significantly measurable compared to the QCD background.

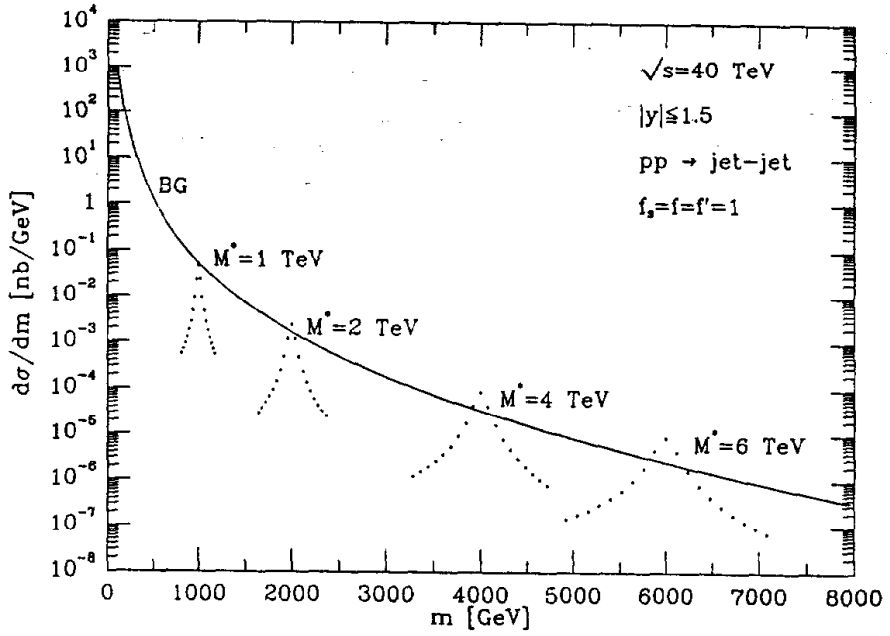


Figure 2.4: Resonant peaks from excited quark production at a range of masses in the final state with two partons. The peaks are compared to the expected background from QCD production of two-parton final states at the same energy and integrated luminosity.

2.3.4 Exotic coloured resonances

Generic new resonances with colour charge can be produced by many theories beyond the Standard Model. Excited quarks are one such resonance, but others are predicted by certain SUSY models, technicolour, extra dimensional models, and many others. These resonances can be treated most generically when categorised by their group theory decomposition and relation to light partons, as shown in [39]. In order to complement the benchmark models previously selected ($W' \to q\bar{q}'$, $W^* \to q\bar{q}'$, $q^* \to qg$) a coloured resonance was sought which would produce a gluon pair in the final state.

Gluon-gluon fusion can produce coloured resonances with the following possible quantum numbers: $(SU(3), SU(2))^J_Q = (\mathbf{8} \oplus \mathbf{8},\ 1)^{0,1,2}_0$. The possible interactions between gluons, a spin-0 (scalar) octet S_8 and a spin-2 (tensor) octet T_8 are specified by

$$\mathcal{L}_{gg8} = g_{QCD} d^{ABC} \left(\frac{\kappa_S}{\Lambda_S} S_8^A F^B_{\mu\nu} F^{C,\mu\nu} + \frac{\kappa_T}{\Lambda_T} (T_8^{A,\mu\sigma} F^B_{\mu\nu} F^{C,\nu}_{\sigma} + f T_8^{A,\sigma}{}_{\sigma} F^{B,\mu\nu} F^C_{\mu\nu}) \right), \qquad (2.22)$$

where the couplings $\kappa_{S,T}$ are of order 1, d^{ABC} is the $SU(3)$ isoscalar factor, and $\Lambda_{S,T}$ are

the scales of the new physics. A similar Lagrangian can be constructed for CP-odd colour octets as described in [39].

Of these possible new resonances, the CP-even colour-octet scalar ($s8$) described by the first term of Equation 2.22 was selected to act as the benchmark for a new physics search in the two-gluon final state. Such resonances are predicted by technicolor models [40, 41], although such theories are largely disfavoured by LHC Run I data [42]. They are also, however, predicted by models of universal extra dimensions. Extra dimensional theories will be discussed in the following section, while specific examples of models predicting colour-octet scalars can be found in [43, 44].

2.3.5 Extra dimensions, gravity, and black holes

One solution to the hierarchy problem postulates that gravity is indeed at the same scale as the remaining forces but is diluted into a number of additional spatial dimensions [45, 46]. A variety of models exist to describe this, varying in the number and topology of the proposed additional dimensions. Arkani-Hamed, Dimopoulos, and Dvali (ADD) models hypothesise multiple flat extra dimensions while Randall-Sundrum (RS) models hypothesise a single, warped extra dimension. The selected benchmark is based on the ADD model, but similar arguments can hold for special RS scenarios as well.

In the case that the universe does contain some number of extra dimensions, it should be possible to produce black holes near the TeV scale. The quantum-gravity energy scale $M_D = [M_{Pl}^2/(8\pi r_c^n)]^{1/(2+n)}$ [47] corresponding to the production of micro black holes decreases as the number n (and size r_c) of these large extra dimensions increases, reaching a mass scale of 1 TeV or below for $n \geq 5$.

The Schwarzschild radius of a black hole of mass M_{BH} existing in a $(4+n)$-dimensional spacetime is given by:

$$r_S = \frac{1}{\sqrt{\pi} M_D} \left(\frac{M_{BH}}{M_D} \left(\frac{8\Gamma(\frac{n+3}{2})}{n+2} \right) \right)^{\frac{1}{n+1}} , \tag{2.23}$$

where Γ is the complete Gamma function and the extra dimensions have size $>> r_S$ [48, 49]. For two colliding protons to form a black hole, their impact parameter must be less than this radius. It follows that the production cross-section of a black hole from two colliding

partons with centre-of-mass energy E scales like so:

$$\sigma_{ij\to BH}(E) \approx \pi r_s^2(E) = \frac{1}{M_D^2}\left(\frac{E}{M_D}\left(\frac{8\Gamma(\frac{n+3}{2})}{n+2}\right)\right)^{\frac{2}{n+1}} . \tag{2.24}$$

Note that this cross-section grows rapidly with energy and no small coupling constants are present: for $M_D \sim 1$ TeV black holes should be produced in great numbers.

However, due to the fact that not all energy carried in a quark or gluon is available for black hole creation and since the probability of a parton carrying a given energy decreases rapidly as that energy grows, the production of semiclassical black holes (with $M_{BH} >> M_D$) at the LHC is disfavoured given preexisting experimental limits [50]. Black hole-like objects could be produced in this TeV mass range using the argument above, but given that their masses will be on the same order as M_D, these objects must exhibit quantum rather than classical behaviour. No quantum gravity theory exists as yet, but quantum black holes (QBH) can be approximately described using a variation on the classical predictions [47].

Among the primary phenomenological differences between semiclassical micro black holes and quantum black holes is that the latter are permitted to have colour and electromagnetic charge. They are thus defined by their mass, spin, and charges. To calculate an exact scattering cross-section values for the QBH quantum numbers must be selected, but the general behaviour and favoured decays can be predicted using an approximation that does not rely on any specific choice.

In the semiclassical situation, one would expect isotropic decay of these micro black holes via Hawking radiation to a many-particle final state. This would require black holes to be produced substantially above their mass threshold in order to have sufficient entropy for such a decay [51]. However it is expected that simpler, transverse final states would become accessible below the actual thermal black hole production threshold. In fact, the average number of particles in the final state can be approximated using a variety of methods, all of which are consistent within order 1 (see for example [47, 48, 52]). Using the method of [47] and $n = 6$ dimensions,

$$\langle N \rangle \sim \frac{4\pi\rho k(6)}{8}\left(\frac{M_{BH}}{M_D}\right)^{8/7} , \tag{2.25}$$

where $\rho \sim 0.34$ results from the decay calculation method and

$$k(n) = \left(2^n \pi^{(n-3)/2} \frac{\Gamma(\frac{n+3}{2})}{2+n} \right)^{\frac{1}{1+n}} . \tag{2.26}$$

For an ADD black hole with $n = 6$, the above calculation shows that $\langle N \rangle > 2$ only for $M_{BH} > 1.5 M_D$ [51]. Therefore for all QBHs produced near the mass threshold in the TeV range, the 2-body final state is dominant. Production of these states is expected to begin once the M_D energy threshold is reached, and to drop off again above that point as a result of PDF effects. The result is a resonance-like production of two-body final states near M_D. Since the QBH can be charged and, even if it is not, arises from the interaction of coloured particles, decays to two partons are strongly favoured. For the ADD $n = 6$ quantum black hole used as a benchmark for the study of extra dimensions in this analysis, the branching ratio to a pair of partons is estimated to be in the range from 84% to 99%.

Chapter 3

QCD from hard scattering to particle jets

'An unalterable and unquestioned law of the musical world required that the German text of French operas sung by Swedish artists should be translated into Italian for the clearer understanding of English-speaking audiences.'
— Edith Wharton, *The Age of Innocence*

In the previous chapter the QCD Lagrangian was defined and its consequences for confinement and asymptotic freedom were briefly explored. Here the use of perturbative QCD (pQCD) for calculating processes at high scale and the complexities arising from the non-perturbativity of the theory at long distances will be considered. A method will be defined for interpreting the collimated clusters of hadrons resulting from the high-energy production of a parton.

3.1 Factorisation

The computation of a cross-section for a hadronic process naturally depends on the scale μ of the interaction, the masses involved, and the renormalisation scale μ_R. The renormalisation scale will appear in the calculation in both the forms μ/μ_R and μ_R/m. Even allowing a perturbative α_s, for a large μ_R one of these two ratios must become large and the perturbative expansion for the cross-section calculation will fail. The computation then contains both perturbative and non-perturbative behaviours simultaneously. The *factorisation theorem* states that the cross-section for a scattering process between hadrons A and B leading to

the final state X can be calculated by separating the high-scale, perturbative part of the interaction from the low-scale, non-perturbative part and treating them independently [53].

This fact can be understood intuitively by considering the interaction between two relativistically moving hadrons with high momentum transfer. Within the hadron the partons experience asymptotic freedom, and the time dilation for the relativistically moving hadrons is sufficient that any virtual particles will be nearly constant during the collision and the parton momenta will not change substantially. Thus the hadrons interact not as single particles but rather as collections of free partons of fixed momentum. The cross-section for the collision can then be computed by using the cross-section for the single parton-parton interaction of interest, weighting it with the probability for each parton to carry a given proportion of the hadron's momentum, and then integrating over the possible combinations [54]. For the collision of hadrons A and B,

$$\sigma_{AB\to X} = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) \times d\hat{\sigma}_{ab\to X}(\alpha_s(\mu_R^2), \mu^2)\,. \tag{3.1}$$

Here the cross-section is perturbatively calculable and depends only on the normalisation scale and the scale μ of the interaction. The $f_{i/I}(x_i)$ are the *parton distribution functions* (pdfs) representing the probability density for a parton of type i to carry a fraction x_i of the total momentum of parton I. The summation is over all types of partons in the two colliding hadrons.

3.1.1 Anatomy of a hadron collision

The full process of a hadron collision begins with two relativistic hadrons and ends with showers of mesons and baryons at much lower energies. The complete interaction can be described by decomposing it into the following stages:

Hard scattering This is the interaction of two quarks represented by the perturbatively calculable cross-section $\hat{\sigma}_{ab\to X}$ in Equation 3.1. It produces high-energy final state partons which may or may not radiate additional particles (*parton showering*): these various final states with and without additional radiation are all calculable.

Underlying event The soft interactions of the remaining partons in the colliding protons are non-perturbative and are included in the calculation of the total cross section by

the summation over pdfs.

Hadronisation The behaviour of outgoing partons from the hard process becomes dominated at longer distances by non-perturbative QCD effects. Confinement ensures that they must form hadrons; the additional energy from the strong force at large ranges is used to do so.

A simulated pp collision illustrating these various interactions is shown in Figure 3.1.

3.2 The hard scatter

In order to calculate the cross section for production of any final state partons, the two components of Equation 3.1 must be separately understood. This section will begin with a treatment of the parton distribution functions, then examine the hard scattering cross section for the partons involved in a sample process.

3.2.1 Probing the proton structure

Splitting functions describe the probability distribution for a parton to radiate another parton. There are four unique splitting functions which correspond to the following processes, where z is the fraction of the initial quark momentum carried by one of the products: $P_{qq}(z)$ corresponds to the splitting $q \to q(z)g(1-z)$, $P_{gq}(z)$ to $q \to g(z)q(1-z)$, $P_{qg}(z)$ to $g \to q(z)\bar{q}(1-z)$, and $P_{gg}(z)$ to $g \to g(z)g(1-z)$. The splitting functions can be directly calculated from the matrix elements for the relevant vertices and can be defined as perturbation series in α_s as follows:

$$P_{gg}(z, \alpha_s) = P_{gg}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{gg}^{(1)}(z) + \dots \tag{3.2}$$

A full solution for all four splitting functions to leading and next-to-leading order in α_s can be found in Reference [56]. The evolution of the parton content of a hadron depends upon these splitting functions as the scale and momentum transfer change.

Structure functions $F(x, \mu^2)$ are process-dependent quantities which can be directly measured in the laboratory and correspond to the internal structure of the proton as observed by probes of various types (electrons, neutrinos, ...). There are two structure functions in cases without parity-violating interactions, and they are related the parton distribution functions through coefficients $C_{2,i}^{(n)}$ which can be computed from the perturbative predictions

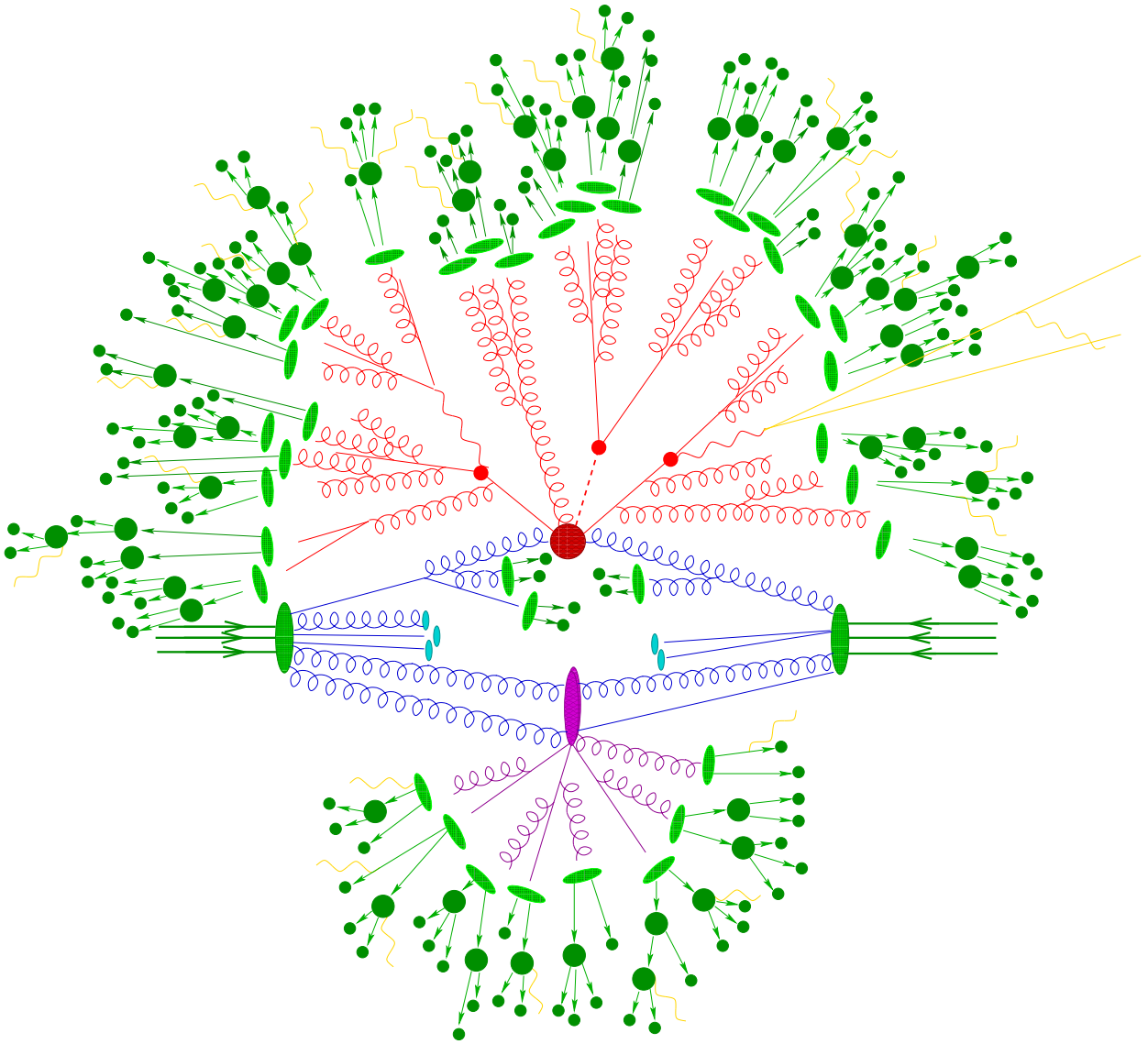


Figure 3.1: Simulation of a sample $pp \to ttH$ event. The incoming partons are in blue. The hard scattering process indicated by the large circle and the ensuing parton showering are drawn in red and are perturbative, high-energy QCD processes. The underlying event interactions are shown in purple. Hadronis are indicated in green: these are both the initial protons and those which begin to form once particles have lost sufficient energy that their QCD interactions are no longer perturbative [55].

for scattering at each order:

$$F_2(x,\mu^2) = x\sum_{n=0}^{\infty}\frac{\alpha_s^n(\mu_R^2)}{(2\pi)^n}\sum_{i=q,g}\int_x^1\frac{dz}{z}C_{2,i}^{(n)}(z,\mu^2,\mu_R^2,\mu_F^2)f_{i/p}(\frac{x}{z},\mu_F^2) + O(\lambda/\mu^2)\,; \tag{3.3}$$

$$F_L(x,\mu^2) = 0\,. \tag{3.4}$$

The pdfs and coefficients both here gain a dependence on the *factorisation scale* μ_F. This dependence arises through handling collinear radiation from the partons: in general, emissions with transverse momenta above μ_F can be accounted for by the process-defining coefficients $C_{2,i}^{(n)}$ while softer emissions must be absorbed into the definitions of the pdfs. Since the structure functions are measurable observables they must not be scale dependent, and indeed with a true infinite sum the μ_F dependence of the coefficients and pdfs cancel out exactly. At a limited order N, however, a dependence on the factorisation scale of order α_S^{N+1} must be introduced.

3.2.2 Evaluating and evolving the parton distribution functions

From the factorisation theorem, one can derive the following equations, known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [57–60]. These are the renormalisation group equations for the parton distributions functions and define their dependence on scale:

$$\frac{\partial q_i(x,\mu^2)}{\partial \log \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dz}{z}\left[P_{q_iq_j}(z,\alpha_s)q_j(\frac{x}{z},\mu^2) + P_{q_ig}(z,\alpha_s)g(\frac{x}{z},\mu^2)\right], \tag{3.5}$$

$$\frac{\partial q_i(x,\mu^2)}{\partial \log \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dz}{z}\left[P_{gq_i}(z,\alpha_s)q_j(\frac{x}{z},\mu^2) + P_{gg}(z,\alpha_s)g(\frac{x}{z},\mu^2)\right]. \tag{3.6}$$

Solving these implicit equations provides expressions for the parton distribution functions but involves a renormalisation-like process which results in a dependence on the factorisation scale interpretable in the same manner as for the structure functions.

The absolute amplitude of parton distribution functions cannot be calculated. However, the relationship between the measurable proton structure functions and the pdfs, given in Equation 3.3, can be used to calculate the value of a pdf at a fixed scale from a laboratory result. Once obtained for a particular scale, the pdf amplitude can be evolved to other scales to provide a full description of hadron contents using the scale dependence defined by the DGLAP equations. Results are typically combined across a range of experimental results to produce a set of parton distribution functions for use in further experiments. Figure 3.2 provides an example of a next-to-next-to-leading order set of pdf calculations produced by the NNPDF collaboration [61, 62].

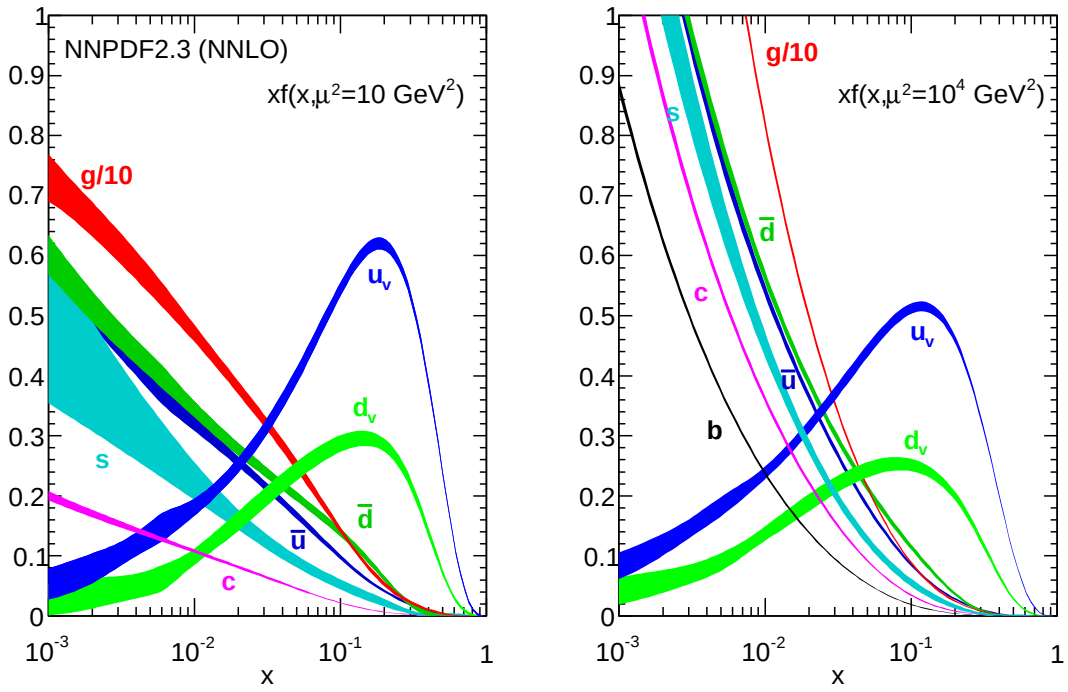


Figure 3.2: Measurements of pdfs incorporating LHC data, from the NNPDF collaboration. Results are shown for scales $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2 (right). See full results in Ref. [62]

3.2.3 Parton-parton scattering at the LHC

The cross-section for pp collision producing two partons (*dijet* production) for an arbitrary hard scatter process is given by the following expression:

$$\frac{d^3\sigma}{dy_3 dy_4 dp_T^2} = \frac{1}{16\pi s^2} \sum_{a,b,c,d=q,\bar{q},g} \frac{f_{a/A}(x_a,\mu^2)}{x_a} \frac{f_{b/B}(x_b,\mu^2)}{x_b} \times \sum |\mathcal{M}(ab \to cd)|^2 \frac{1}{1+\delta_{cd}} . \tag{3.7}$$

In order to obtain the full and final cross-section for dijets the contributions must be taken into account for all possible $2 \to 2$ processes. At leading order, these correspond to the diagrams shown in Figure 3.3 and the matrix elements given in Table 3.1.

The more detailed description obtained by considering additional effects from higher-order terms reveals some intriguing consequences of the process. A small sample of NLO contributions to the above process are shown in Figure 3.4. Note that these include some loop diagram modifications of the dijet production mechanism as well as some processes

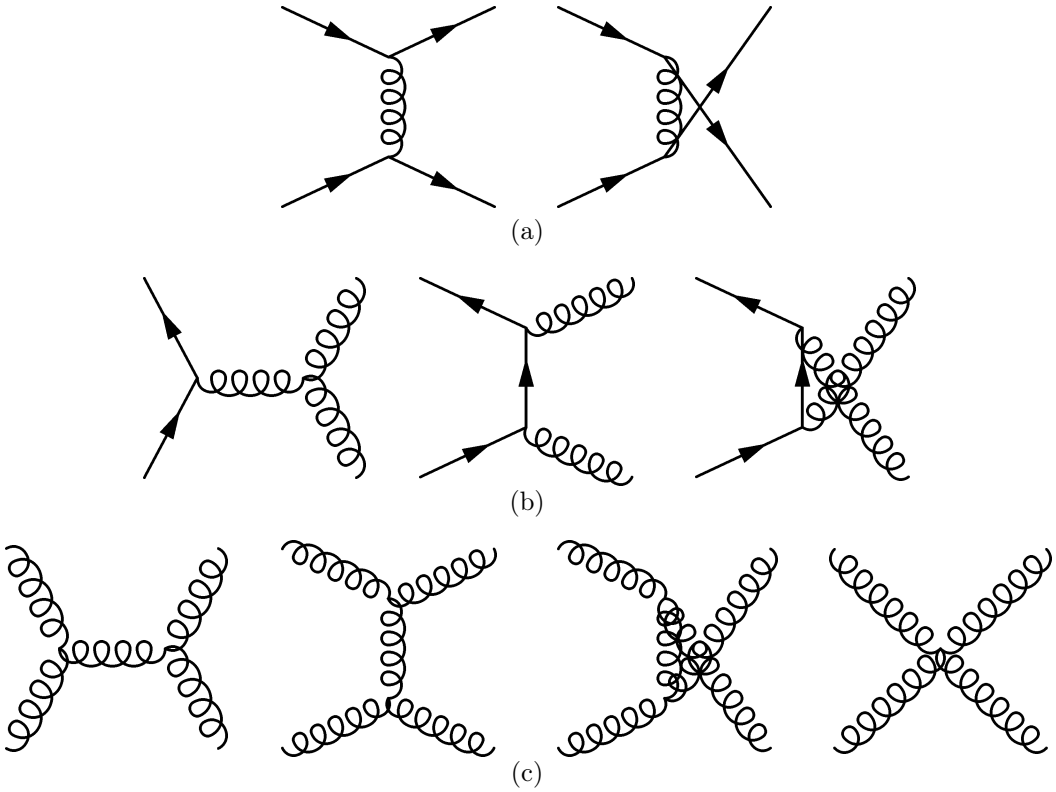


Figure 3.3: Leading-order processes for production of two partons. Diagrams are sorted into (3.3a) $qq \to qq$, (3.3b) $qq \to gg$ and (3.3c) $gg \to gg$ processes. All diagrams can be rotated, allowing (for instance) $qg \to qg$ processes. Adapted from [56].

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

Table 3.1: Matrix elements for all leading-order Feynman diagrams corresponding to dijet production in a pp collider. Evaluation of the matrix elements at $\theta^* = \pi/2$ is also given. At central angles, then, the dijet production is dominated by $gg \to gg$ production, followed at much lower rates by $gq \to gq$ and $qq \to qq$. Taken from Reference [56].

involving the radiation of an additional new gluon. An attempt to calculate the matrix element for one of these diagrams shows that the result is divergent for loops like those in Figure 3.4 and for the emission of an additional gluon when it is soft or parallel to the parent

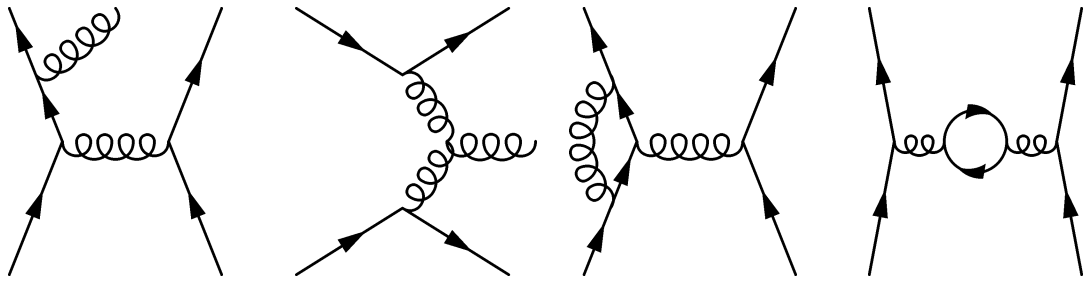


Figure 3.4: Next-to-leading-order processes beginning with two partons which contribute to the higher-order dijet production calculations. Adapted from [63].

particle. The effects of these divergences are discussed in the next section.

3.2.4 Divergences

Divergences actually arise in QCD cross section calculations in a variety of ways. The two most important for experimentalists are those involving the emission of real particles. *Infrared divergences* result from the radiation of a very soft gluon, since cross-section calculations are ill-defined as the energy of one parton approaches zero. When an emitted gluon is parallel to its parent particle, the result is not meaningfully distinguishable and causes a *collinear divergence* in the calculation.

Fortunately, these divergences can be handled such that a meaningful cross-section can still be computed. The Kinoshita-Lee-Nauenberg theorem uses unitarity arguments to demonstrate that the infrared divergences from loops cancel out those from legs as long as all diagrams are included [64–66]. Collinear divergences in the initial state are absorbed into the description of the proton structure through the definition of the parton distribution functions.

Difficulties could still arise, however, depending on the quantity chosen as the observable of interest for a cross-section calculation. If the observable is sensitive to collinear or infrared radiation then it would distinguish between the various diagrams considered in the above calculation and the issue of divergences would reemerge. The solution is based on a Bloch-Nordsieck-like summation: as long as the final state is sufficiently inclusive that it covers cases both with and without radiation, the divergences cancel in the sum of such states and the process can be quantitatively predicted [67]. An approximate way of understanding this

result is to consider that the cross-section for a process with the emission of an additional gluon is about the same as the reduction in the cross-section for a process without that gluon, leading to a constant total cross-section [1].

Thus cross-sections are calculable in the case that the final state is as inclusive as possible and all diagrams for ISR, FSR, and loop corrections at a given order are included. An observable is acceptable as a final product of a QCD calculation if it meets the following criteria [63]:

- it is *infrared-safe*: if an emitted gluon is very soft, the observable should not change it if is removed.
- it is *collinear-safe*: if an emitted gluon is parallel to its parent particle, the observable will be the same regardless of whether they are treated as two separate particles or a single one.

3.3 Final state evolution to the edge of perturbativity

Between the production of a parton in the hard scatter and the point at which non-perturbative QCD effects take over, it can be expected to radiate (or "fragment") numerous times. This purely probabilistic process nonetheless needs to be well-described in order to simulate a QCD interaction and arrive at a meaningful observable.

The description of *parton showering* is built from the same parton splitting functions used to describe hadron contents in Subsection 3.2.1. Let a parton a give rise to b and c with energy fraction $z = E_b/E_a$. The probabilistic splitting function $P_{a\to b,c}(z)$ describes the relative likelihood of a splitting to a fixed b, c, and z, and thus the differential probability of splitting to any allowed b and c at a scale given by $t = \ln(\mu^2/\Lambda^2)$ is:

$$d\mathcal{P}_a = \sum_{b,c} \frac{\alpha_s}{2\pi} P_{a\to bc}(z) dt dz \,. \tag{3.8}$$

Define the probability density for a specific branching occurring at a scale δt at any z:

$$I_{a\to bc}(t) = \int_{-z(t)}^{+z(t)} dz \frac{\alpha_s}{2\pi} P_{a\to bc}(z) \,. \tag{3.9}$$

Then the outcome that no branching occurs during a scale evolution interval δt can be

described by combining the above two definitions: $P_{no}(t, \delta t) = 1 - \sum_{b,c} I_{a\to bc}(t)\delta t$. If the evolution of parton a is considered from an initial scale t_0 to a later scale t, the probability that it would not have branched during that time is the product of probabilities that it would not have branched at any δt during that time. Letting $\delta t \to 0$, the Sudakov form factor is defined:

$$\mathcal{P}_{no}(t_0, t) = \Delta_{a\to b,c}(t_0, t) = \exp\left(-\int_{t_0}^{t} dt' \sum_{b,c} I_{a\to bc}(t')\right) \tag{3.10}$$

The total probability of parton a branching to any b, c at a time t is therefore the product of the probability of that splitting occurring and the probability that no other splitting occurred up to that point:

$$\begin{aligned}\frac{d\mathcal{P}_a}{dt} &= \left(\sum_{b,c} I_{a\to bc}(t)\right)\Delta_{a\to b,c}(t_0, t) \\ &= \left(\sum_{b,c} I_{a\to bc}(t)\right)\exp\left(-\int_{t_0}^{t} dt' \sum_{b,c} I_{a\to bc}(t')\right).\end{aligned} \tag{3.11}$$

This equation can be used to describe the probabilistic nature of parton splitting between the hard scattering process and the limit of perturbative QCD, and is used in simulations of QCD scattering processes [68].

3.4 Hadronisation

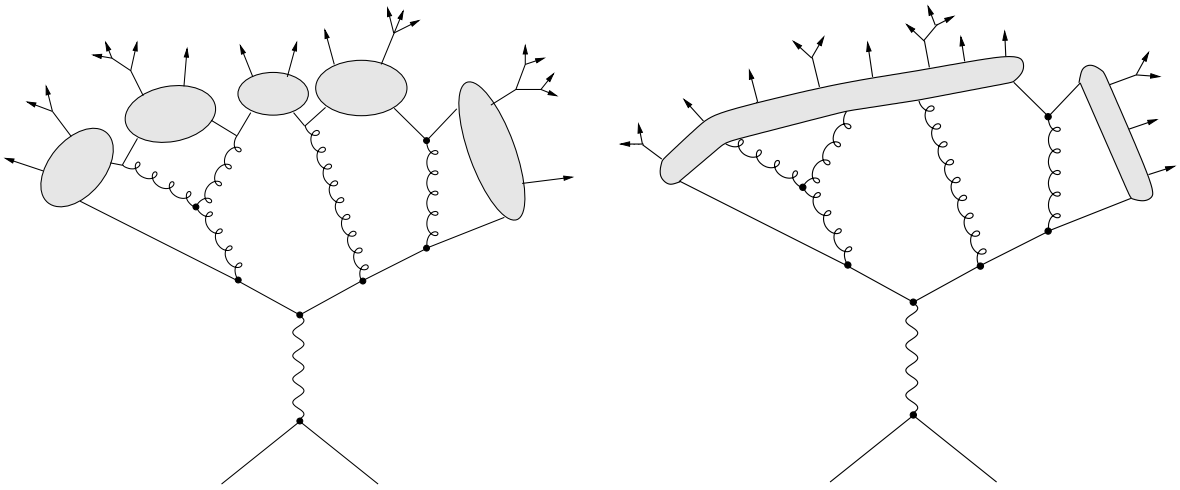
Once the partons produced in the showering process reach sufficiently large distances and low energies that QCD no longer behaves perturbatively, confinement takes effect. By a process of *hadronisation*, baryons and mesons are formed from the previously independent partons. It is these final state colourless particles which are observable in the laboratory.

Because hadronisation is a non-perturbative process, no exact description of it has been developed. These two assumptions can nonetheless be made: first, that the flow of energy and flavour at the particle level should follow the behaviour at the parton level, since hadronisation is a low-scale process with small energy exchange; and second, that some non-perturbative α_s can be defined for very small scales [69]. These assumptions have been used to build two specific models of hadronisation which are routinely used in simulations of jet formation:

Cluster model Gluons are split into $q\bar{q}$ pairs and colourless clusters of nearby quarks are formed. These clusters, behaving as proto-hadrons, then decay to the final-state hadrons [70].

String model This model treats quarks as connected by a gluonic "string" with a linearly increasing potential as the partons stretch farther and farther apart. Beyond some energy the string will break and a new parton forms at each raw end. When a colour-neutral set of partons are near each other a hadron is formed [71].

The resulting particle shower will vary slightly between the two methods, as in the example in Figure 3.5. The string model is used for the simulations in this analysis.



(a) Cluster model (b) String model

Figure 3.5: Hadronisation models. Figures from [69].

3.5 Jets and jet algorithms

The final state stable hadrons produced in the previous step are the only experimentally observable objects to result from the production of the initial parton. Some link must be drawn between the two. Since momentum is conserved throughout the showering and hadronisation process, if a parton has a sufficiently large momentum then its product particles will also tend to maintain a significant fraction of their momenta in that direction. The final-state hadrons produced will be grouped together in a spray of particles in the same region of the detector: the higher-momentum the original parton, the more tightly collimated the final products.

This provides the basis for a link between observable particles and the initial state: a *jet* can be formed by gathering together the final-state particles according to some algorithm and combining their energies and momenta into a single newly-defined object. The simplest example is to simply use a cone: let a circle be drawn around the hadrons collected in some detector and traced back to a point at the hard scattering. Then the momenta and energies of the hadrons falling inside the circle can be summed to provide a measurement of the E and p of the original hadron. The identification of a jet object with the original parton is somewhat ambiguous and the measured jet properties will never exactly correspond to the hard scattering, but jets do provide the necessary link between the theoretical predictions of QCD at the parton level and the experimental observables of final-state hadrons. The idea of a jet was first theoretically motivated in Reference [72]; for an in-depth discussion see Reference [73]. An illustration of jet formation from hard scatter through parton showering and hadronisation, resulting in a collimated cluster of particles reaching a detector, is shown in Figure 3.6.

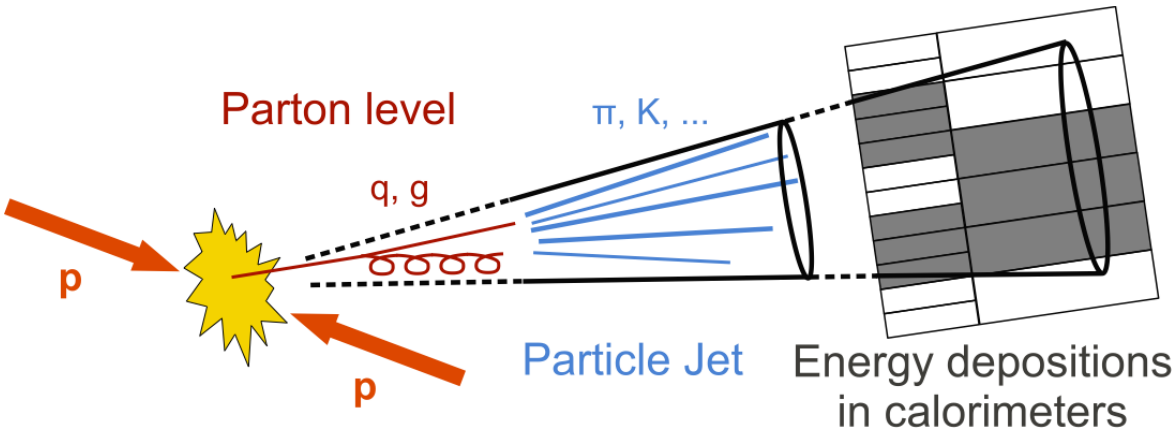


Figure 3.6: The formation of a particle jet. A parton ejected by a hard scattering showers and eventually hadronises, forming a collimated and approximately cone-shaped jet of stable, light particles which can be experimentally detected. Image from [74].

A *jet algorithm* defines the procedure by which final-state observables are clustered into a jet object: the definition used above of a cone containing the hadrons to include is one example. There are in fact a number of requirements for jet algorithms which determine their stability and safety as an observable and their use as a description of theoretical objects.

3.5.1 Jets as physically meaningful objects

There are a wide variety of possible jet reconstruction algorithms, of which a few commonly used ones will be discussed in the following section. No particular choice is "correct", so long as it satisfies a set of agreed requirements: the idea of a jet is not fixed and the object does not exist outside of its algorithmic definition. In order to fully define a jet object, only the following items are required, as formalised in the Les Houches Accord of 2007 [75]:

Jet algorithm A set of rules for selecting the objects to be included in the jet and determining when the jet is complete

Jet algorithm parameters Fixed values for any free parameters in the jet algorithm, such as the radius of the completed jet

Recombination scheme An algorithm for combining the energies and momenta of the constituent particles to determine those of the jet

These three components combine to create a *jet definition*. However not all jet definitions are equally useful. The primary requirement is that jet definitions must be stable against infrared and collinear radiation, a topic which will be discussed in the following section. In addition, a jet definition should be simple to implement for both theorists and experimentalists and should be well-defined and predict a finite cross-section at any order in perturbation theory (see the Snowmass accord [76]).

It is difficult to define the extent to which any particular jet definition matches the prediction of the original parton by QCD. Part of the challenge arises from the fact that the initial parton is an ill-defined, even physically meaningless quantity: parton showering or boosts that bring several partons very near to each other could change the number of jets found in the event without necessarily becoming a more or less useful description of the initial state. Additional questions arise because of the restrictive nature of jet definitions, such as a maximum reach beyond which two constituents will not be considered part of the same jet. This is necessary for a well-defined object, yet for any definition there will be cases where radiation from the correct initial parton falls outside the jet and is not included, or where hadrons resulting from two initial partons fall within a single jet: no single definition is ideal for all situations. Hadronisation itself presents a further challenge due to its non-perturbative nature. One way to estimate the relationship between partons and jets is to

simulate jet formation from an initial parton, reconstruct the jet from simulated final-state particles, and compare the result to the initial parton, but this method is highly dependent on the order of the perturbative calculation and the treatment of the hadronisation effects. Although this prevents a determination of the exact correspondence between the jet object and the parton, it is at least a useful tool for examining the differences between various jet algorithms [77].

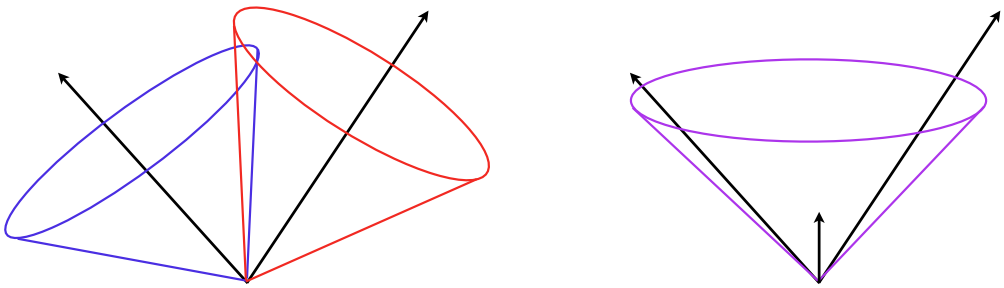
3.5.2 Infrared and collinear safety in jet definitions

To confidently link a theorized process to an observable jet, the jet definition must not be sensitive to the addition of collinear or infrared radiation. These radiative processes are statistically random and do not change the underlying physics of a scenario, so they should not change its interpretation through jets. Furthermore, although detectors do not have infinite resolution and may not be sensitive to these radiative processes, this cannot be relied upon and means that a jet algorithm which is sensitive to them will not be able to form a reliable link between theoretical predictions and experimental observables [77].

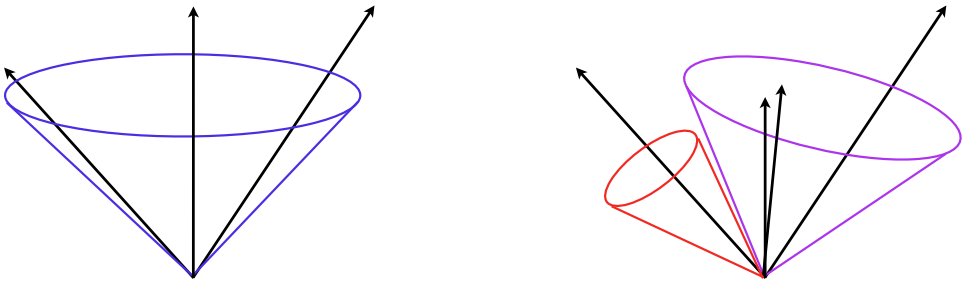
Infrared and collinear (IRC) safety in the context of jet algorithms can be defined thus: if one particle in the jet radiates a gluon which is either very soft (infrared) or parallel to the direction of the parent particle, then the recreated jet should not change. An example of unsafe behaviour of each type is illustrated in Figure 3.7. As an example, simple cone algorithms where a single seed such as the most energetic measured particle is used to define the centre of a circular jet are not IRC safe. This can be clearly illustrated by picturing the most energetic particle radiating a collinear gluon with half its initial energy. The physical properties of the scenario would be identical but if some other constituent particle now had more than half the original leading energy, the new jet would be centred around this particle instead and would include a different set of particles.

3.5.3 Jet reconstruction algorithms

A jet algorithm defines a process for determining which of a collection of objects should be included in a jet and clustering these *constituents* together. The algorithm can apply to any type of object with a well-defined momentum and energy, such as partons, hadrons, particle tracks, or energy deposits in a calorimeter. Generally, the constituents are grouped using a



(a)



(b)

Figure 3.7: Illustration of jet algorithms which are unsafe to infrared radiation (3.7a) and collinear radiation (3.7b). In the infrared-unsafe case, two particles which are treated as separate jets in the first image are combined into a single jet because of the radiation of a soft gluon geometrically between them. In the collinear-unsafe case, one jet becomes two after a single parton splits into two collinear partons with decreased energy. Based on an illustration in [78].

distance parameter defining the closeness of two objects and a "distance-to-beam" parameter providing a measure of the extent to which an object belongs in the jet versus acting as a standalone object.

Most high-energy physics applications have taken to using *sequential recombination algorithms* instead of cone-based algorithms. Here out of a total of n constituents each possible pairing is considered and the distance between them calculated. The two objects separated by the smallest distance are combined to form a single constituent with properties given by the selected recombination scheme. The process is repeated with the now $n-1$ constituents, and the iteration continues until the smallest remaining distance between two constituents is larger than the distance-to-beam, when the process stops and the combined objects are

taken as the final jet collection.

Define the following measures d_{ij} for the distance between constituents i and j and d_{iB} for the distance-to-beam of constituent i:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2},$$
$$d_{iB} = k_{ti}^{2p}. \tag{3.12}$$

Here k_{ti} and k_{tj} are the transverse momenta of the two constituents. $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is a geometrical distance term indicating the separation of the constituents in rapidity and azimuthal angle. Note that Equation 3.12 relates Δ_{ij} to a constant R: this jet algorithm parameter is related to the overall cutoff size of the jet, although the exact nature of that relationship depends on the value of p. The quantity p controls the dependence of the jet algorithm on dynamics (the transverse momenta) without which it would be purely geometrical. Instead of considering this a jet algorithm parameter it actually distinguishes between different algorithms, as its effects on the physical behaviour are so formative.

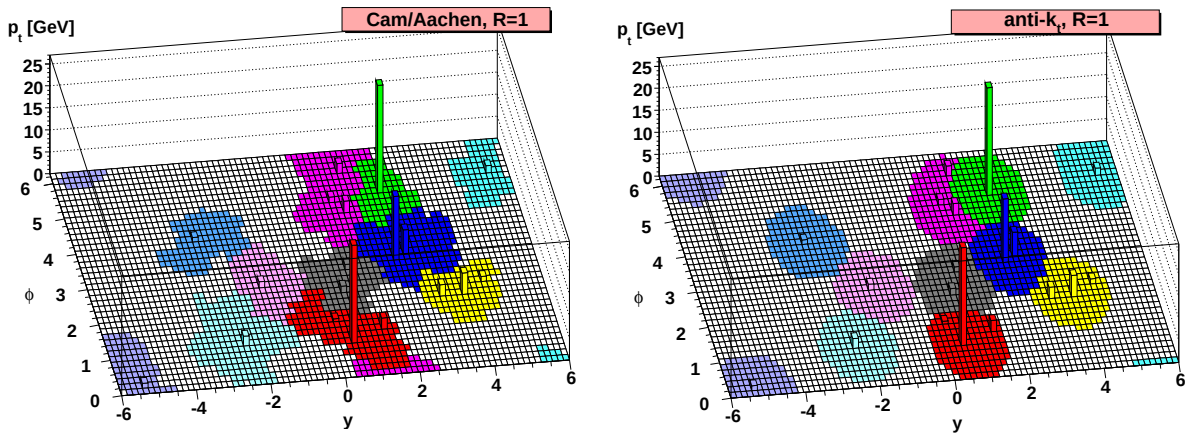
A refined version of the earliest sequential adaptive algorithm, known as the k_t algorithm, uses $p = 1$ [79]. With $d_{ij} \propto \min(k_{ti}^2, k_{tj}^2)$, the jet clustering begins from the softest components. This clustering sequence mimics to some extent the parton showering which took place, making k_t a popular choice for theorists. The k_t algorithm was used consistently at HERA and occasionally at CDF and D0, but is disfavoured by experimentalists today due to the sprawling, irregular shapes of the jets it produces.

Another common variant is the Cambridge-Aachen (CA) algorithm, which sets $p = 0$ in Equation 3.12 and uses a distance-to-beam of R [80]. This is a strictly geometrical clustering since all dependence on momentum is removed. The constituents nearest in Δ_{ij}^2 are clustered until the smallest remaining distance is R, at which point the jet is complete. Cambridge-Aachen jets are also somewhat irregular in shape. This is one of the two main jet algorithms used by the LHC experiments.

The second main jet algorithm is anti-k_t, which fixes $p = -1$ [81, 82]. Here, the distance parameter decreases as constituent momentum increases, causing soft components to be clustered with hard ones long before they are clustered with each other. The resulting

jets are relatively circular, since a single hard component would gather to itself all soft components within a distance R and then be prevented from expanding beyond that radius. When two hard particles are not near enough to be clustered but form two separate jets, the soft constituents in between are geometrically divided along a curve depending on the relative momenta of the two hard constituents. Unlike the k_t and Cambridge-Aachen algorithms, this clustering sequence is not analogous to QCD splitting and so does not provide any insight into the jet formation or substructure. However, its reliance on the hardest instead of softest constituents and its regular shape give it many of the same advantages of cone-jet algorithms without the need to sacrifice IRC safety.

The differences between the two most common algorithms can be seen in in Figure 3.8. The same constituents are used in both images and reclustered according to the different algorithms. This illustrates how algorithm selection affects jet shapes and the constituents included in each jet. In some cases the total number of jets can also change.



(a) Cambridge-Aachen algorithm (b) Anti-k_t algorithm

Figure 3.8: Comparison of the Cambridge-Aachen (3.8a) and anti-k_t (3.8b) jet clustering algorithms. Note the irregularity of the jet shapes in the CA cases compared to the anti-k_t case [77].

The most common recombination scheme is the *four-momentum recombination scheme*, used for all jets addressed in this analysis. Here, the the final four-momentum of a jet is calculated by summing together the four-momenta of all its constituents, resulting in jets with a defined mass. With a jet definition now complete, QCD $2 \to 2$ processes can be linked to a final observable state.

Chapter 4

The ATLAS Experiment

"Data! Data! Data!" he cried impatiently. "I can't make bricks without clay."'
— Sir A. Conan Doyle, *The Adventures of Sherlock Holmes*

Each step taken by the physics community to examine its theories of physics on a smaller scale necessitates a machine capable of producing proportionally higher energies. The discovery of the top quark in 1995 and the τ neutrino in 2000, both at the Tevatron, left the Higgs boson as the only remaining missing piece of the Standard Model. By the late 1980s and early 1990s, there were theoretical constraints on its mass suggesting that it ought to be on the order of a few hundreds of GeV. There was also a strong drive to explore the myriad hypotheses for physics beyond the standard model, many of which were also expected to be less than 1 TeV in mass.

4.1 The LHC

In the early 1950s, following closely on the end of the second world war, a collection of US and European physicists proposed an international research laboratory that would bring European science back to the cutting edge and would encourage collaboration amongst the countries so recently at war. The proposal would eventually come before UNESCO in late 1951, where a resolution was taken to create a European Council for Nuclear Research – in French, Conseil européen pour la recherche nucléaire, or CERN. The first eleven nations joined early in the following year, and in 1954 the new laboratory was founded.

Over the next five decades CERN would slowly grow to become a leader in nuclear

research. It hosted an enormous range of experiments over the years at an increasingly complex set of accelerators. The proton synchrotron (PS), CERN's oldest major accelerator, was used in conjunction with the giant bubble chambers of the early 1970's to discover the first evidence of neutral currents in 1974. At the super proton synchrotron (SPS) a few years later, the UA1 and UA2 collaborations discovered the W and Z bosons and firmly established Europe as a leading power in particle physics research. In 1989 a new electron collider was constructed, the most powerful lepton collider to date, in a tunnel 27 km in circumference spanning the French/Swiss border near Geneva. The Large Electron-Positron Collider (LEP) provided valuable measurements of known particle masses and couplings and ran throughout the 1990's in an attempt to beat its rival, Fermilab's Tevatron, to the discovery of the Standard Model Higgs boson. Neither experiment made the great discovery, and their days were numbered as designs were finalised for a new accelerator. LEP was shut down in 2000 to make way for the new giant of particle physics, which would use its original tunnel: the Large Hadron Collider, or LHC.

The LHC, like LEP, is thus 27 km in circumference. It accelerates beams made of clusters of protons (*bunches*) in opposite directions and allows them to intersect at four out of eight equally spaced crossing points around the ring. Each of these four points hosts a major experiment (see Figure 4.1).There are three main figures of merit which define an accelerator's abilities: centre-of-mass energy, luminosity, and the type of particle to be accelerated. The centre-of-mass energy determines the maximum mass of the product particles and defines the cross sections of the various interactions which will occur. Instantaneous luminosity defines the rate of interactions and thus the number of events of any given process which can be collected: $N_{events} = \mathcal{L}\sigma_{events}$, where N_{events} is the number of events collected per second [83]. For a beam with a Gaussian profile and a revolution frequency f_r containing n_b bunches of N_b particles each, the luminosity can be calculated thus:

$$\mathcal{L} = f_r N_b^2 n_b \frac{\gamma_r}{4\pi\epsilon_n\beta^*} F\,. \tag{4.1}$$

where $\gamma_r = E/m$ is the relativistic factor for the protons. The normalised transverse beam emittance ϵ_n is a measure of the distance between the protons and how different their momenta are: a small emittance means tightly packed protons with similar momenta. β^* is defined by the magnet layout and corresponds to the narrowness of the beam; like emittance,

it has units of length. Finally, F is a geometric factor which defines the loss in luminosity due to the crossing angle of the beams. The LHC has been designed to achieve a maximum centre-of-mass energy of 14 TeV and a luminosity of 10^{34} cm^{-2}s^{-1} [83]. Since its first data taking in 2009, the LHC has been slowly increasing both its energy and luminosity towards this goal.

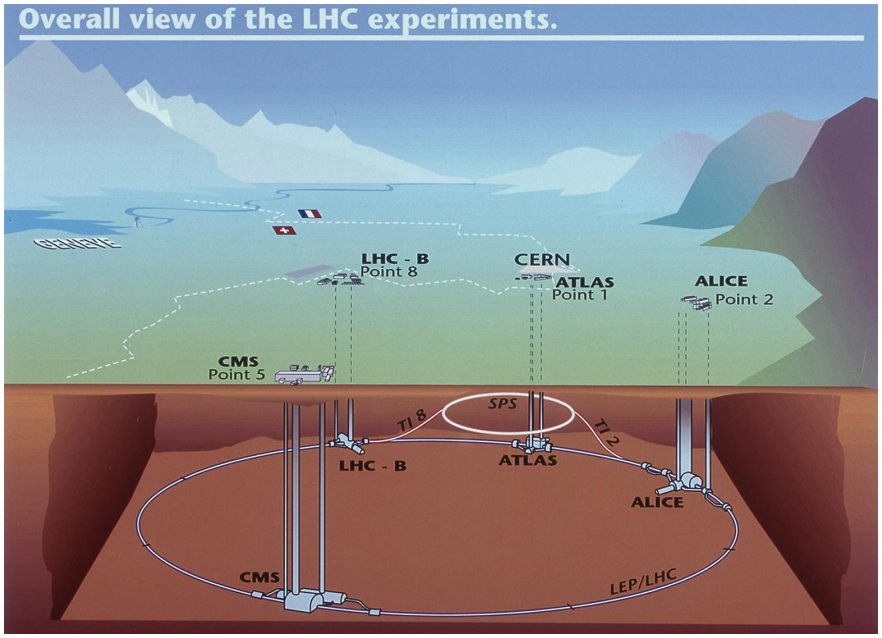


Figure 4.1: A cutaway perspective on the LHC site showing the location of the accelerator and main experimental caverns relative to surface landmarks [84]

Bunches of protons are injected into the LHC from a series of feeder rings and accelerated until they reach the operating energy. The LHC radio frequency cavities not only accelerate the protons in the beam but also ensure that they are grouped into tight *bunches*. Since radio frequency acceleration is a pulsed system which requires the particles to enter each cavity at a particular phase, a continuous beam is not possible and tightly packed bunches are used instead [85].

The majority of the LHC ring is composed of 1232 dipole magnets. Since the radial acceleration of the protons is enforced by the dipoles surrounding the beamlines, the magnetic field of the magnets varies with the centre-of-mass energy. At design energy, the dipoles will generate a magnetic field of 8.33 Tesla [86]. Each such magnet carries two beam pipes embedded first in superconducting coils and then in a mass of steel which is cooled during operation to 1.9 degrees Kelvin using superfluid helium. The cold mass is surrounded by insulation and a shielding case [87]. A dipole magnet in cross-section is shown in Figure 4.2.

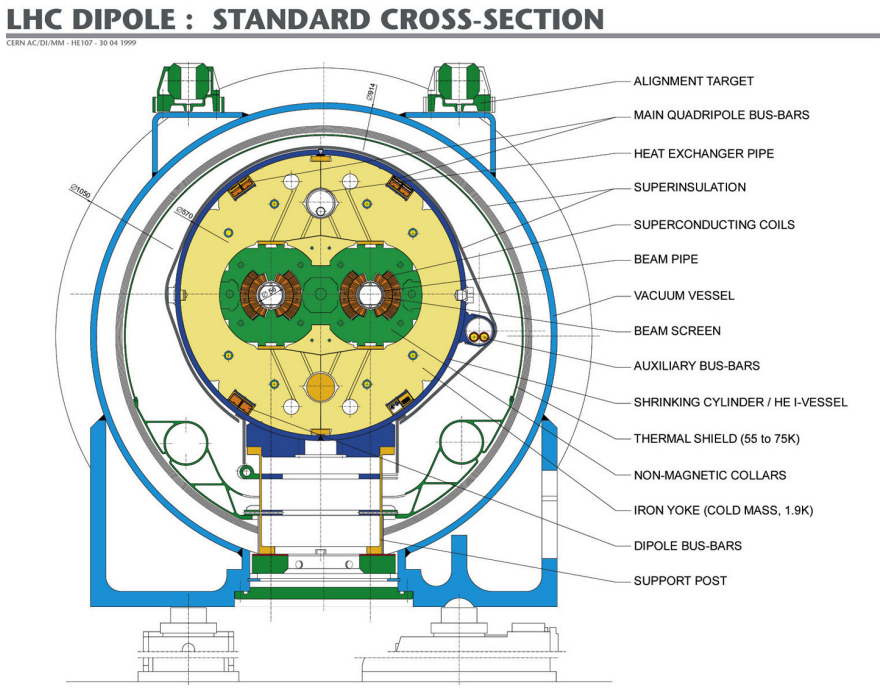


Figure 4.2: Cross-section of an LHC dipole magnet. The two beam pipes are visible in the centre, surrounded by the thin coils of superconducting wire which provide the magnetic dipole field. These are in turn surrounded by support structures and cooling and insulation systems, including a cold mass of iron and a final layer of shielding preserving a vacuum within the machine. The dipole field is created to force the radial acceleration of the protons travelling in opposite directions through the two beam pipes [88].

Interspersed at regular intervals between the dipole magnets are 386 quadrupole magnets and 382 pairs of sextupole magnets. The former provide beam focusing effects in opposite planes to counteract the tendency of the proton beam to disperse in the transverse plane. The latter assist in focusing to counteract the effect of the slightly different energies of the protons within a bunch and their resulting different focal lengths in a quadrupole field. A variety of other magnets in much smaller numbers are also distributed around the ring and serve as additional correctors.

4.1.1 Running conditions in 2012

Throughout 2012 the LHC operated at a center-of-mass energy of 8 TeV, just over half of the accelerator's design energy of 14 TeV. The instantaneous luminosity was similarly above half of its design, but the run conditions varied dramatically from the nominal in order to achieve it. Luminosity can be controlled in two ways (see Equation 4.1): either the number of bunches or the density of protons in each bunch can be altered. At design luminosity,

the LHC would carry 2808 bunches, each with 1.15×10^{11} protons and separated by 25 ns. During the 2012 data taking period it operated instead with 1374 bunches, each with an average intensity of $1.6 - 1.7 \times 10^{11}$ protons and with a spacing of 50 ns. The result was a peak luminosity of 7.7×10^{33} cm^{-2}s^{-1} as compared to a design peak of 1×10^{34} cm^{-2}s^{-1}, but with much higher proton density [89]. The consequences will be explored in the following section.

The recorded data is divided into *runs*, corresponding to data generated within the same proton fill in the LHC and lasting on average about 12 hours. Runs are further subdivided by ATLAS into *luminosity blocks*, which correspond to the smallest amount of data over which the integrated luminosity can be reliably determined [90]. Run conditions such as the state of LHC magnets or the accelerator temperature are treated as constant over a single luminosity block.

The total integrated luminosity delivered by the LHC during 2012 was 22.8 fb^{-1}, of which ATLAS successfully recorded 21.3 fb^{-1}. Figure 4.3 shows the rate of increase of delivered and recorded luminosity over the year. The uncertainty on the luminosity recorded by ATLAS in 2012 is 2.8% [91].

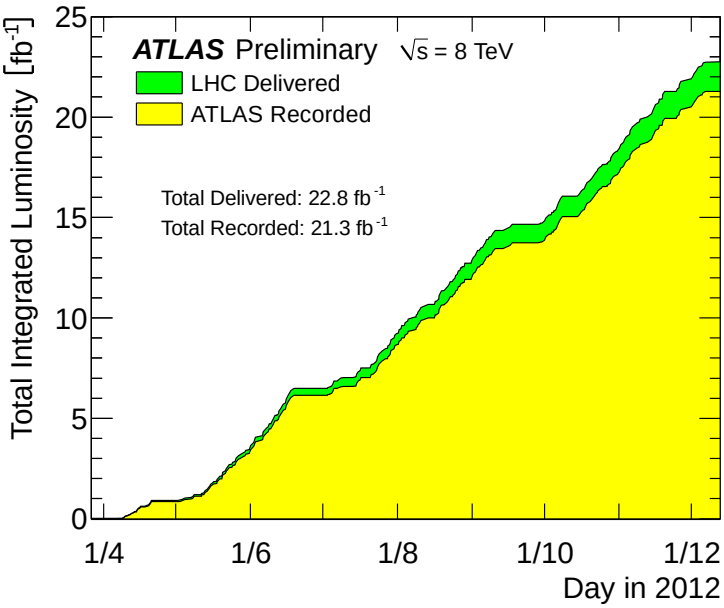


Figure 4.3: Integrated luminosity, delivered and recorded, as a function of time during 2012 [92].

4.1.2 Pileup

A result of high luminosity is that several proton collisions will be recorded within a single event. This is *pileup* and results from multiple pairs of protons colliding within a small time. There are two types of pileup with different origins:

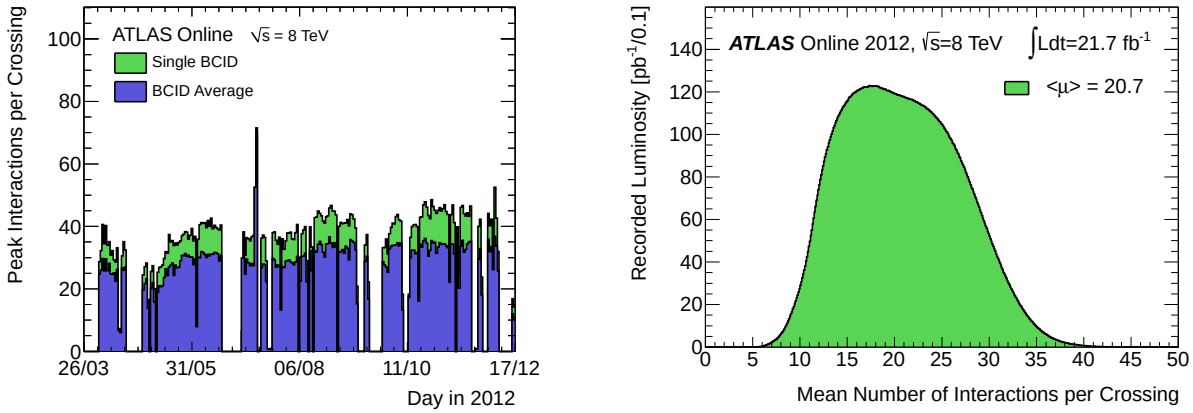
- *In-time pileup*: Multiple particle collisions occur within a single bunch crossing and the products are all saved as one event. This is a result of dense proton bunches.
- *Out-of-time pileup*: Particle collisions occurring in different bunch crossings are nonetheless close enough in time that the products of different collisions still in the detector get recorded within a single event. This is a result of closely spaced bunches.

Pileup can be predicted from the run conditions using the following formula. If μ represents the average number of collisions per bunch crossing, then

$$\mu = \frac{\mathcal{L}\sigma_{inelastic}}{n_b f_r}\,, \tag{4.2}$$

where n_b and f_r are the number and frequency of bunches as before and $\sigma_{inelastic}$ is the inelastic cross-section for pp collisions [93]. The value of μ can be predicted for each bunch crossing from this formula. Often a more useful measure is the average of μ over a given luminosity block (denoted $\langle\mu\rangle$).

In 2012 run conditions pileup is a highly significant effect. The LHC's initial design specifications intended it to achieve a maximum luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ with one bunch collision every 25ns [83]. This would imply pileup of order 19 interactions per bunch crossing if all of the bunches were evenly spaced, although since bunches are grouped together into closely-spaced *bunch trains*, the true expected value was closer to 23 [86, 94]. The ATLAS detector was designed to cope with this level of pileup. However, in 2012 data taking, the 50 ns bunch separation and increased bunch density caused in-time pileup far above the design specifications, with $\langle\mu\rangle$ reaching 40 by the end of the year. Figure 4.4 summarises the pileup conditions in 2012.



(a) Maximum number of events in a beam crossing and maximum average interactions per luminosity block $\langle\mu\rangle$ per day in 2012.

(b) Recorded luminosity as a function of $\langle\mu\rangle$ over the course of 2012. This indicates the relative frequency of various pileup levels in the full dataset.

Figure 4.4: Pileup conditions in the 2012 dataset. The plots indicate the distribution of average pileup by time and by luminosity collected [92].

4.2 ATLAS Detector Overview

ATLAS, **A** **T**oroidal **L**HC **A**pparatu**S**[1], is a 7000-tonne detector situated at LHC Point 1. It is multi-purpose in its conception, designed for both standard model measurements and searches for new physics. It is symmetric with respect to an interaction point at the center of the detector and is designed to be most sensitive to particles produced moving transverse to the beam line. A diagram of the full detector is given in Figure 4.5.

The inner detector is situated closest to the collision point and is designed for tracking purposes. It measures momenta and vertex locations of particles leaving the interaction point and is a critical component of particle identification. There are three components to the inner detector: from the inside out, these are the pixel layers, the semiconductor tracker, and the transition radiation tracker. The inner detector will be discussed in detail in Section 4.3. Surrounding the inner detector are two calorimeters, one for electromagnetic showers and the other for hadronic showers. These are the most critical components for measuring jet properties and so will be thoroughly examined in Sections 4.4 and 4.5. Finally, the muon spectrometer surrounds the entire detector as its outermost layer, and is discussed

[1] Attempting to uncover the history of this name is surprisingly challenging. We are given to understand that the "Toroidal" in this sense refers to the toroidal magnet system rather than to the overall shape of the detector, but have been unable to substantiate this claim. As for the capitalisation of the final "S", the only way to redeem this outrageous backronym is to assume it is a reference to the G.R.O.S.S. club of *Calvin and Hobbes* fame.

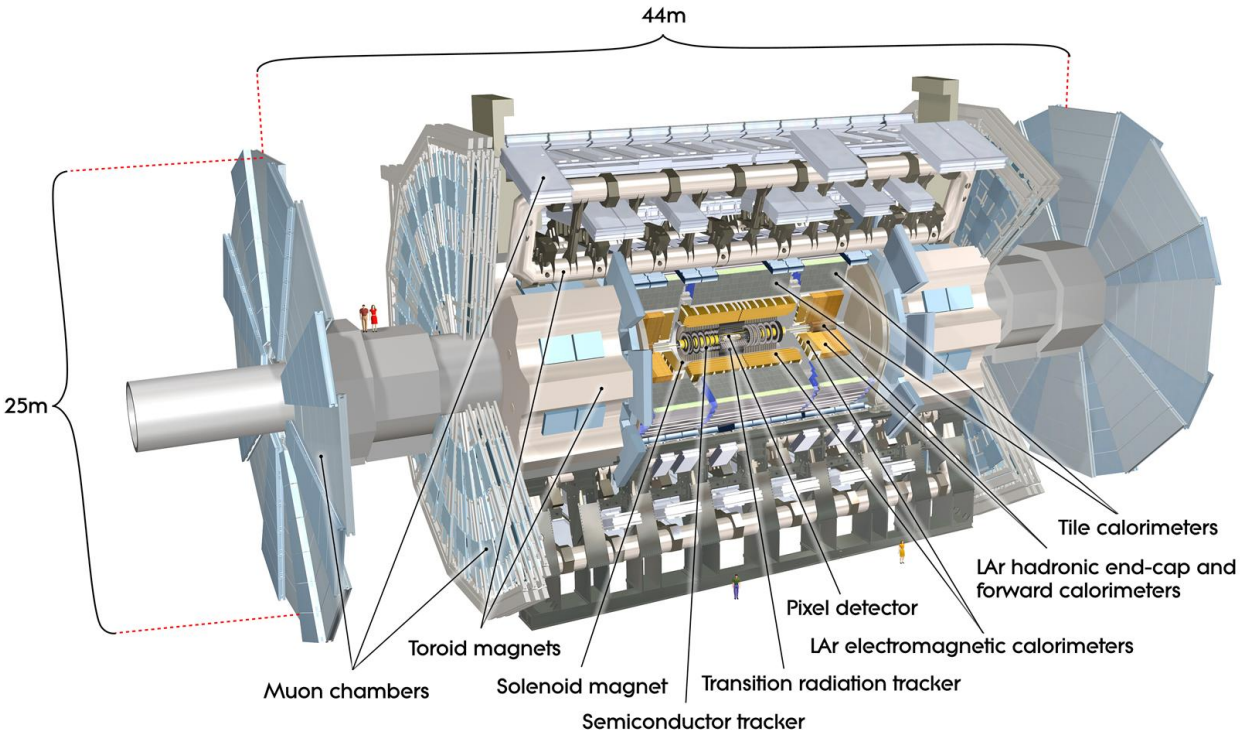


Figure 4.5: CGI schematic of the full ATLAS detector, with labelled components. The muon system is most clearly visible as the outermost layer of the detector[3]. Immediately inside those are the toroidal magnets which accelerate muons along the beam axis direction as they pass through. The calorimeters follow, then the solenoid magnets which accelerate particles in the tracking system, and finally the inner detector with its tracking equipment sits at the centre, nearest to the beamline [95].

in Section 4.6.

The ATLAS detector contains several critical magnet systems which generate fields designed to curve the tracks of particles in the detector [90]. First is a solenoidal 2 T magnet situated outside of the inner detector and providing a longitudinal field which causes particles trajectories to curve around the direction of the beamline. Outside the calorimeter, the three toroidal magnets which give ATLAS its name provide axial acceleration to the particles entering the muon detector system [96].

4.2.1 Coordinate system

ATLAS makes use of a right-handed coordinate system to define directions within the detector. The x-axis lies perpendicular to the beamline with $\hat{x}$ directed towards the centre of

the LHC ring. The z-axis lies along the beamline, with the $\hat{z}$ directed from Point 1 towards Point 8; that is, to a viewer in the ATLAS control room, towards Geneva. The detector sides are labeled such that side-A lies in the positive z direction and side-C lies in the negative z direction. Lastly the y-axis, perpendicular to both, is nearly vertical (though off by $\sim 0.7\,\text{deg}$ due to the slant of the LHC ring between Points 1 and 5) [97].

A cylindrical coordinate system is also routinely used in ATLAS, which defines $r = \sqrt{x^2 + y^2}$ and uses ϕ for the polar angle. The azimuthal angle is represented by θ but is rarely used in practice. Instead, ATLAS uses the pseudorapidity η:

$$\eta = -\ln(\tan\frac{\theta}{2}) \tag{4.3}$$

This variable runs from 0 in the $z = 0$ plane to $\pm\infty$ along the beamline. The rectangular and cylindrical coordinate systems are illustrated in Figure 4.6.

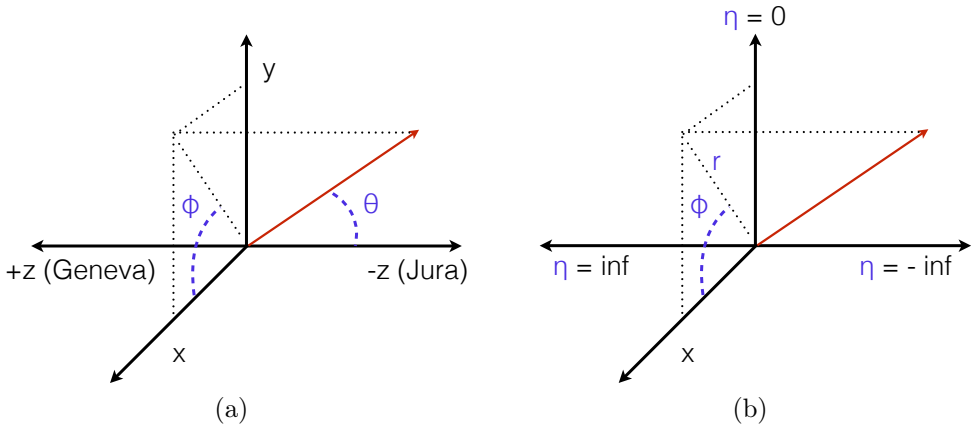


Figure 4.6: Coordinate system for the ATLAS detector. The standard right-handed rectangular coordinate system and its angles are shown in 4.6a while the cylindrical coordinate system and the pseudorapidity are shown in 4.6.

A further variable, rapidity, is used for massive object such as jets. For such an object with energy E and longitudinal momentum p_z, the rapidity y (not to be confused with the coordinate y) is defined thus [90]:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right). \tag{4.4}$$

Differences in rapidity between two particles are invariant under Lorentz boosts along the z-axis, allowing this variable to provide a measurement of the degree to which two jets are

back-to-back in their rest frame. Pseudorapidity has the advantage of ease of measurement, making it the standard choice for ATLAS applications. However, it neglects object masses while rapidity accounts for them, making y a more suitable choice for certain jet quantities.

4.3 Inner detector

The inner detector (ID) covers an angular range $|\eta| < 2.5$ and was designed to provide a transverse momentum (p_{T}) resolution of $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\% \times p_{\mathrm{T}}$ [GeV] $\oplus$ 1% and a transverse impact parameter resolution of down to 11 μm for central, high-momentum particles [98]. Tracks of transverse momentum 0.4 GeV and above are used in offline event reconstruction, but in certain circumstances the ID is sensitive to tracks with p_{T} as low as 0.1 GeV [90]. In order to provide the most accurate measurement of track locations near the interaction point (IP), the inner detector primarily uses silicon semi-conductor tracking detectors. These are in the form of pixels nearest the beam (the pixel detectors) and of microstrips at a slighter greater radius (the semiconductor tracker). A straw tube tracker provides additional location information at a lower cost (the transition radiation tracker). All these components are designed and laid out to ensure appropriate coverage of all tracks: any particle passing through the inner detector should cross at least three pixel layers, four microstrip layers, and leave on average 36 hits in the transition radiation tracker [98]. The relative arrangement of these components is shown in Figure 4.7 and their coverage and design resolutions are detailed in Table 4.1.

Detector system	η coverage	Design resolution $\sigma(\mu\mathrm{m})$
Pixels, barrel	$\lvert\eta\rvert < 1.7$	R-ϕ = 12, z = 66
Pixels, b-layer	$\lvert\eta\rvert < 2.5$	R-ϕ = 12, z = 66
Pixels, endcap	$1.7 < \lvert\eta\rvert < 2.5$	R-ϕ = 12, z = 77
SCT, barrel	$\lvert\eta < 1.4\rvert$	R-ϕ = 16, z = 580
SCT, endcap	$1.4 < \lvert\eta\rvert < 2.5$	R-ϕ = 16, z = 580
TRT, barrel	$\lvert\eta\rvert < 0.70$	170 per straw
TRT, endcap	$0.70 < \lvert\eta\rvert < 2.5$	170 per straw

Table 4.1: Components of the inner detector with their coverage in η and spatial resolution.

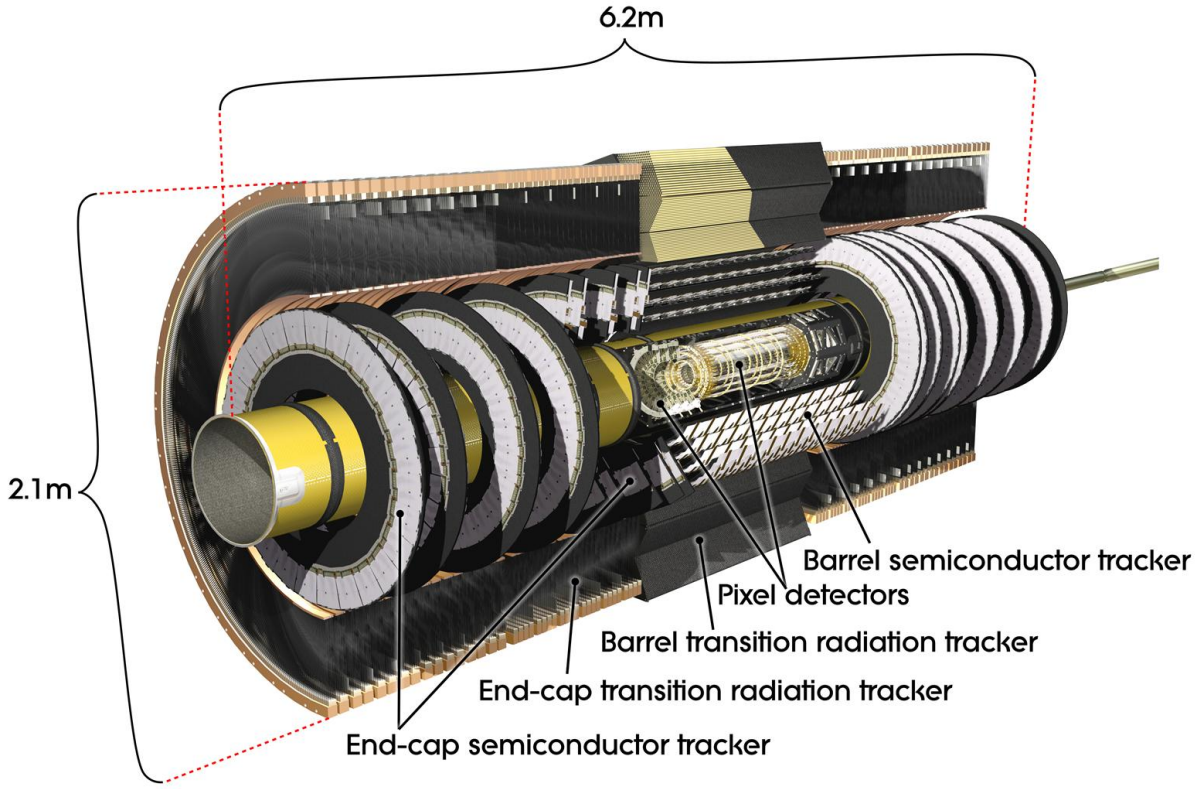


Figure 4.7: CGI schematic of the inner detector, with labelled components. The SCT, TRT, and pixel detectors are labeled [99].

4.3.1 Pixel detectors

The pixel system consists of three cylindrical layers surrounding the beam pipe and six disk layers serving as endcaps, three at each end of the barrel section. The innermost layer, or b-layer, is affixed directly to the beam pipe and has an average radius of 50.5 mm, while the second and third layers of the barrel are located at radii of 88.5 and 122.5 mm respectively. The disk layers are situated at z positions of 495, 580, and 650 mm. The sensing unit of the detector is a pixel 50 μm in ϕ by 400 μm in r/z. These are assembled in groups of $\sim 40,000$, together with various readout electronics, into the 1744 "modules" which comprise the detector's active area [100]. The pixel detector, given its location as the innermost layer of the detector, has sustained the most radiation damage during Run I operation. A new "insertable b-layer", more radiation-hard than the current one and located closer to a new, thinner beryllium beam pipe, is being added to the inner detector during the first long shutdown [101, 102]. Further details of the pixel detector design can be found

in References [90, 103].

During 2012 operation, the pixel detector had a data taking efficiency of 99.9% in the active channels, but by the end of Run I data taking only 95% of modules remained active [104].

4.3.2 Semiconductor Tracker

The semiconductor tracker (SCT) follows the same physical principles as the pixel detector, but is composed of larger silicon strips (80 μm × 12cm) rather than pixels. It also comprises a barrel region and two endcaps, in this case consisting of four concentric central layers located at radii of 299, 371, 443, and 514 mm and nine disks on each end at z positions of 853.8, 934, 1091.5, 1299.9, 1399.7, 1771.4, 2115.2, 2505, and 2720.2 mm. The arrangement of the components is such that any particle leaving a primary vertex within 152 mm of the centre of the detector should pass through all four layers of the SCT [105, 106]. These hits provide sufficient information to reconstruct the track momentum with better than 30% precision up to p_{T} = 500 GeV. Each SCT module consists of two layers of silicon strips off parallel from one another by a small angle (40 mrad), allowing the calculation the z coordinate of the vertex within 1 mm and the resolution of two separate tracks within 200 μm [98].

More than 99% of the channels in the detector remained active at the end of Run I, contributing to an efficiency of 99.9% in the barrel (99.6% in the endcaps) [107]. The combination of the SCT and pixel detectors measured the transverse impact parameter of high p_{T} tracks with a resolution of 10 μm over the full $|\eta| \leq 2.5$ range during 2012 data taking [108].

4.3.3 Transition Radiation Tracker

The transition radiation tracker (TRT) is a drift chamber with inner radius 56 cm and outer radius 108 cm designed to assist in track and vertex reconstruction and electron identification. A barrel section 114 cm long and composed of more than 50,000 axially arranged drift tubes (straws) is sandwiched by two endcap regions each consisting of over 122,000 radially arranged straws divided amongst 20 independent wheels. Each polyimide straw is 4 mm in diameter with a thin external aluminum layer acting as a cathode and an anode formed

from a central gold-plated tungsten wire [109]. The drift tubes are filled with a gas mixture comprised mainly of xenon and CO_2. During operation, the wire is grounded and the outer case held at 1530 V. The anode is split in z into two electrically independent wires by a glass separator, allowing electrical readout from both ends of the straw and thus increasing measurement precision [110,111]. In between the straws is polypropylene as either fibres or a film. Transition radiation is generated when a relativistically moving particle passes through the boundary between materials of different permittivities, in this case between the straws and the polypropylene filler, and the radiation takes the form of a collection of photons. Since electrons are much lighter than pions they can produce transition radiation from far lower energies, allowing the TRT to distinguish between the two particle types.

During Run I, the TRT delivered a track position measurement accuracy of between 100 μm and 130 μm, depending on track p_{T} and location and the amount of pileup [112]. This is sufficient to increase the accuracy with which a track location is known from the two innermost layers, although the TRT can only provide information about a track in the plane perpendicular to the straws. The TRT also reduces the frequency of fake tracks by confirming those found in the inner detector using a greater number of layers and hits.

4.4 Electromagnetic calorimeter

Directly outside the solenoidal magnet system is the liquid argon-based electromagnetic calorimeter, used to measure energy deposits from jets, photons, and electrons, and to contribute to the measurement of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). Electromagnetic showering is dominated at energies $\gtrsim 1$ GeV by two processes: bremsstrahlung, in which an electron radiates a photon, and pair production, in which a photon produces an e^+e^- pair. Since electrons and photons undergo these processes as a result of interactions with atomic nuclei, calorimeters employ layers of high-density *absorbers* to instigate showering. As the resulting particles from the showering fall below a critical energy E_c, ionisation dominates instead and the shower formation ceases. The depth of absorber required to fully contain a shower depends on the material used and the energy of the incoming particles. Absorbing materials can be described by their *radiation length* X_0, which is the distance an electron must travel in the material for its energy to be reduced by a factor e. For example, 6.91 radiation lengths of material would be required to reduce the energy of a particle by 99.9%.

Described in great detail in [113], the electromagnetic sampling calorimeter alternates lead absorbers with active gaps consisting of liquid argon between copper electrodes held at a high voltage difference. Charged particles shower as they traverse the absorber, creating a spray of particles. As each charged particle then passes through the argon, it ionises the liquid and releases electrons in quantities proportional to its energy. These electrons are accelerated across the narrow active gap by the copper electrodes, which collect the resulting pulse and transmit it to the readout electronics.

The electromagnetic calorimeter consists of a barrel section covering the range $|\eta| < 1.475$ and two endcap sections covering the range $1.375 < |\eta| < 3.2$. The full granularity of all EM calorimeter components is shown in Table 4.2. The barrel section is at least 22 radiation lengths in thickness at all points, while the endcaps have a thickness of at least 24 X_0 for $|\eta| > 1.475$ and running as high as 36 X_0 at the smallest angles. An accordion design was chosen for the electrodes and absorbers, since it allows full coverage in ϕ with no cracks and permits a high granularity. Figure 4.8 illustrates this geometry in detail. The absorbers are composed of lead sheets 1.13 to 1.53 mm in thickness, coated on each side by a stainless steel layer of width 0.2 mm which provides mechanical support. In between the absorbers are electrodes consisting of three layers of copper separated by insulating sheets: the outer two layers are held at high voltage while the inner is used to read off the signal. Honeycomb-shaped spacers maintain a separation between the lead absorbers and the electrodes. In both the barrel and the endcaps, the modules are arranged such that the absorbers extend radially from the beamline, with the crests of the accordion folds running parallel to the beamline in the barrel and extending radially in the endcaps [114,115]. Liquid argon flows continuously through the remaining space in the calorimeter, ionised by passing particles and thus providing the signal. Since this active material is constantly replaced, the electromagnetic calorimeter suffers no radiation damage over the lifetime of the experiment.

A particle will have already passed through between 1.8 X_0 ($\eta = 0$) and 4.4 X_0 ($|\eta| = 1.5$) of upstream material in the inner detector and its services before reaching the electromagnetic calorimeter. Where these interactions would otherwise reduce the resolution of the calorimeter, a presampler is added to correct for the loss. This detector sits closer to the beamline and is essentially a single layer of LAr active material, with no absorber, used to estimate the amount of energy already dissipated before particles reach the calorimeter

Subdetector	Region	Granularity
EM Barrel	Presampler, $\|\eta\| < 1.52$	$\Delta\eta \times \Delta\phi = 0.025 \times 0.1$
	Layer 1, $\|\eta\| < 1.4$	$\Delta\eta \times \Delta\phi = 0.025/8 \times 2\pi/64$
	Layer 1, $1.4 < \|\eta\| < 1.475$	$\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/64$
	Layer 2, $\|\eta\| < 1.4$	$\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/256$
	Layer 2, $1.4 < \|\eta\| < 1.475$	$\Delta\eta \times \Delta\phi = 0.075 \times 2\pi/256$
	Layer 3	$\Delta\eta \times \Delta\phi = 0.05 \times 2\pi/256$
EM Endcap	Presampler, $\|\eta\| < 1.52$	$\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/64$
	Layer 1, $1.375 < \|\eta\| < 1.425$	$\Delta\eta \times \Delta\phi = 0.050 \times 2\pi/64$
	Layer 1, $1.425 < \|\eta\| < 1.5$	$\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/64$
	Layer 1, $1.5 < \|\eta\| < 1.8$	$\Delta\eta \times \Delta\phi = 0.025/8 \times 2\pi/64$
	Layer 1, $1.8 < \|\eta\| < 2.0$	$\Delta\eta \times \Delta\phi = 0.025/6 \times 2\pi/64$
	Layer 1, $2.0 < \|\eta\| < 2.4$	$\Delta\eta \times \Delta\phi = 0.025/4 \times 2\pi/64$
	Layer 1, $2.4 < \|\eta\| < 2.5$	$\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/64$
	Layer 1, $2.5 < \|\eta\|$	$\Delta\eta \times \Delta\phi = 0.1 \times 2\pi/64$
	Layer 2, $1.375 < \|\eta\| < 1.425$	$\Delta\eta \times \Delta\phi = 0.050 \times 2\pi/256$
	Layer 2, $1.425 < \|\eta\| < 2.5$	$\Delta\eta \times \Delta\phi = 0.020 \times 2\pi/256$
	Layer 2, $2.5 < \|\eta\|$	$\Delta\eta \times \Delta\phi = 0.1 \times 2\pi/64$
	Layer 3, $1.5 < \|\eta\| < 2.5$	$\Delta\eta \times \Delta\phi = 0.050 \times 2\pi/256$

Table 4.2: Geometry and granulatrity of the various electromagnetic calorimeter components. Electromagnetic barrel calorimeter properties from [116]; presampler info from [117]. EM endcap details from [115].

proper. The presampler's active layer is 11 mm in depth in the barrel region and 2 mm in the endcaps and extends up to $|\eta| < 1.8$ [117].

The pulse generated in the copper electrodes by the passage of a particle is proportional to the energy of the particle and decreases linearly with time. Each pulse is read out by a front-end board (FEB), one of which corresponds to 128 individual channels in the calorimeter. The pulse is first amplified in three gain ratios (1, 10, 100), then undergoes a bi-polar shaping resulting in the final shape illustrated in Figure 4.9. After shaping, the pulse is sampled every 25 ns (40 MHz) by a Switched Capacitor Array (SCA) and the measurements stored until the level-1 trigger decision on the event has been determined. If the event passes the trigger, the most appropriate gain for the pulse is selected and the stored samples are passed to an analog-to-digital converter (ADC) to produce the final digital signal [119]. Once corrected for detector issues such as fluctuating voltages in different calorimeter regions, the

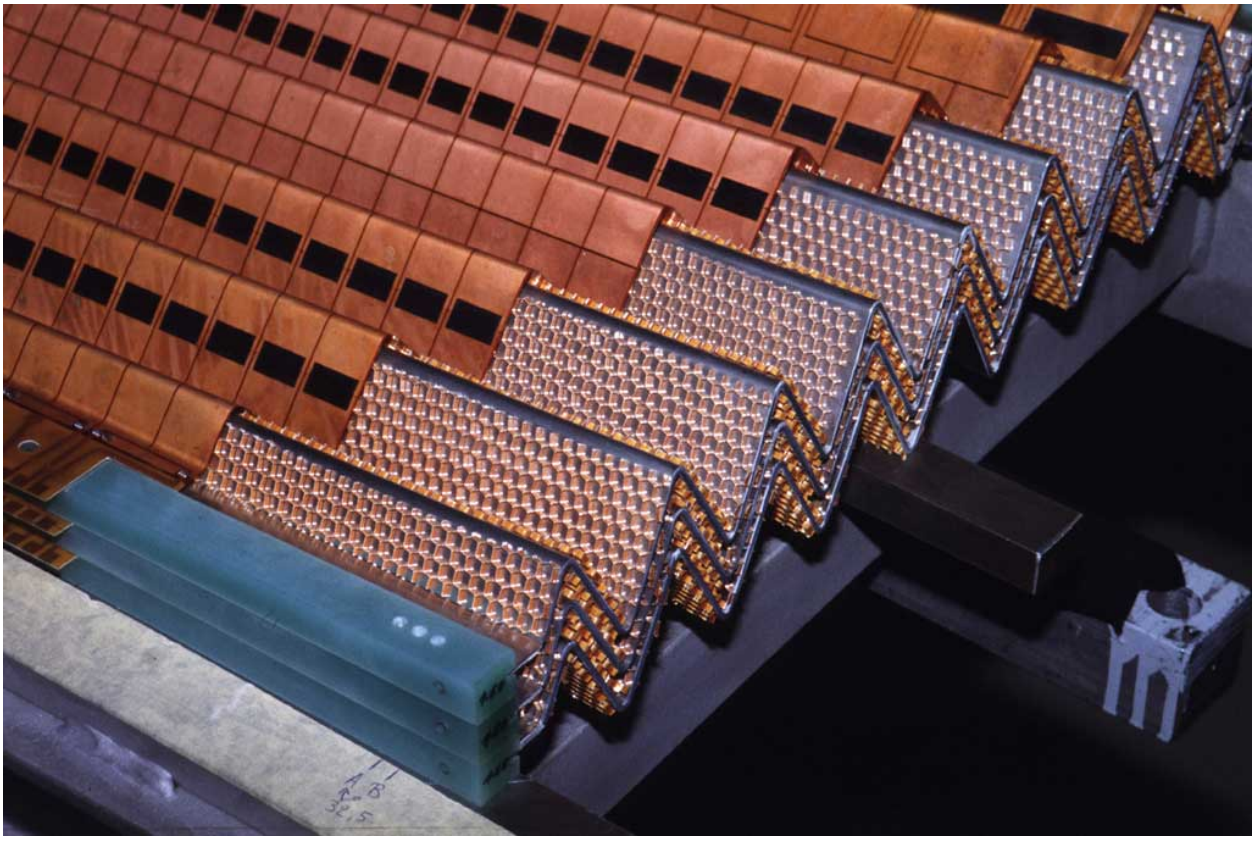


Figure 4.8: Detail of a section of the electromagnetic calorimeter barrel region, showing the absorbers and spacers arranged in "accordion" fashion. The edge visible in the lower left sits nearest the beamline. The rippling of the layers enables the calorimeter to provide full coverage in ϕ [118].

measured energy deposit in each cell is saved and passed into the offline processes of object reconstruction and calibration.

4.4.1 Electron and photon calibration

The visible energy in a single cell of the EM calorimeter can be calculated from the current passed out of the readout system, the conversion factor between current and energy, the gain chosen for readout, and some coefficients derived from a description of the pulse shape and noise. The visible energy does not, however, account for energy deposited in the passive material of the calorimeter. It must be scaled to the full deposited energy by the sampling fraction, defined as the fraction of energy deposited in active regions. The sampling fraction varies with η, reflecting the changing composition of the detector. Calorimeter cells having passed this two-stage energy calibration are referred to as *EM-scale* and are the building

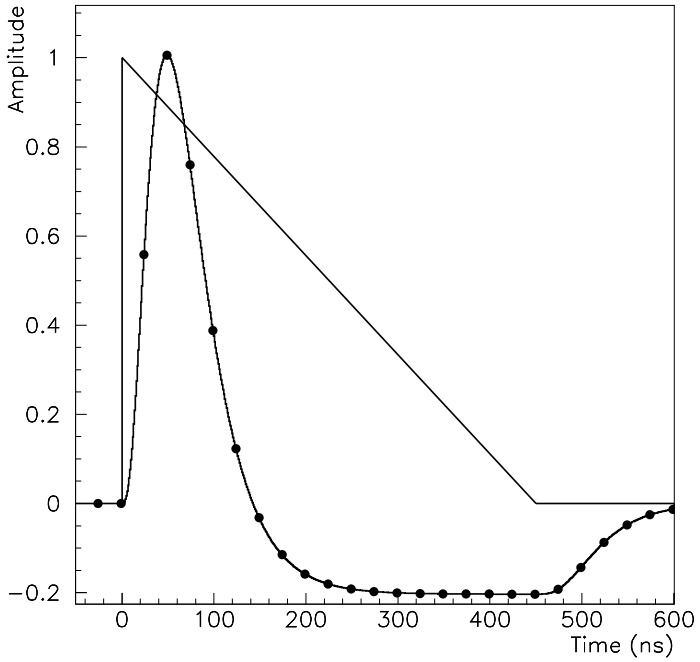


Figure 4.9: Signal shape as read out from the barrel modules (triangular line) and after shaping by the front-end board electronics (curved line). Sampling by SCA chips is indicated by the black points [120].

blocks for jets as discussed in Chapter 5 [121].

4.5 Hadronic Calorimeter

Only approximately 2/3 of the hadronic shower from a strongly charged particle is contained within the electromagnetic calorimeter. The remainder is stopped by the hadronic calorimeter (HCAL), a system composed of both scintillating tile and liquid argon components ~ 8-12 interaction lengths deep and situated immediately around the ECal. Hadronic particles shower differently in matter than electromagnetic particles due to their strong force interactions: some *invisible energy* always results from incoming hadron energy contributing to nuclear breakup rather than lost in collision. The fraction f_{em} of hadron energy deposited through electromagnetic showers of secondary particles increases with energy according to $f_{\mathrm{em}} = 1 - E(\mathrm{GeV})^{-0.15}$. Of the energy deposited through strong interactions, the greatest fraction is via ionisation (40-60%) followed by nuclear breakup and recoil of target nuclei and nuclear fragments, accounting for 30-45% of the shower energy with the exact value depending on the absorber material. Neutron generation accounts for 10-15% and photon

generation through fission about 3% [122]. Hadronic calorimeters require a significant absorber thickness to contain showers, e.g. 77 cm of iron to contain 95% of a shower, requiring them to be sampling rather than homogeneous calorimeters in most cases. In the liquid argon components of the hadronic calorimeter, the electromagnetic shower components and nuclear ionisation are measured as in the electromagnetic calorimeter; invisible energy loss is accounted for in later calibration. The plastic scintillators of the tile calorimeter measure energy via molecular excitation. Incoming particles excite molecules of a primary active material, which emit UV light on de-excitation. This is in turn absorbed and re-emitted at a lower wavelength by a secondary active material, a necessary conversion to allow the light to propagate more than a few mm and exit the tile.

The HCAL sits within the return path of magnetic flux from the solenoid, but most flux is carried by the steel girders which support the hardware. Consequently, the maximum field within the active calorimeter regions is only 6 ± 2 mT, leading to an increased light output from the scintillators of $\sim 1\%$. Radiation effects are minimal at this distance from the interaction point, but another issue arises: the cryostat housing the EM calorimeter is 2-3 interaction lengths thick, causing substantial energy loss before particles reach the HCal. A correction for this loss is applied and is able to restore linearity of the energy response. Table 4.3 shows the geometry and granularity of all components of the hadronic calorimeter.

Subdetector	Region	Granularity
Tile barrel	Barrel, $\|\eta\| < 1.0$	$\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
	EB, $0.8 < \|\eta\| < 1.7$	$\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
Hadronic endcap	$\|\eta\| < 2.5$	$\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
	$\|\eta\| > 2.5$	$\Delta\eta \times \Delta\phi = 0.2 \times 0.2$
Forward calorimeter	FCAL1, $3.1 < \|\eta\| < 4.9$	$\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
	FCAL2 & FCAL3, $3.1 < \|\eta\| < 4.9$	$\Delta\eta \times \Delta\phi = 0.2 \times 0.2$

Table 4.3: Geometry and granulatrity of the various hadronic calorimeter components. Hadronic end-cap properties from [123]. Forward calorimeter properties from [113, 124].

4.5.1 Tile (barrel) calorimeter

The central calorimeter region consists of a tile-scintillator detector composed of a barrel region ($0 \leq |\eta| < 1.0$) and two extended barrel (EB, $0.8 < |\eta| < 1.7$) sections [125]. It is in three

layers, together spanning radii between 2288 and 4230 mm [126]. Each of the 64 modules which compose the barrel HCal contains scintillating tiles as the active medium embedded in a steel absorber. The tiles are 3mm thick, and are arranged such that this narrow edge runs in the z direction. Polystyrene is the base matrix of the tiles, doped with paraterphenyl as a scintillator and various wave-shifting dyes. The barrel tile calorimeter has a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ($\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ in the third layer), substantially coarser than the ECal. The less fine segmentation is a design choice which allows for the weighting techniques used to restore energy response linearity and which also facilitates cuts on energy leakage into the HCAL for improved electron and photon identification. The arrangement of the absorbers and spacers in a module, together with the readout components, are shown in Figure 4.10.

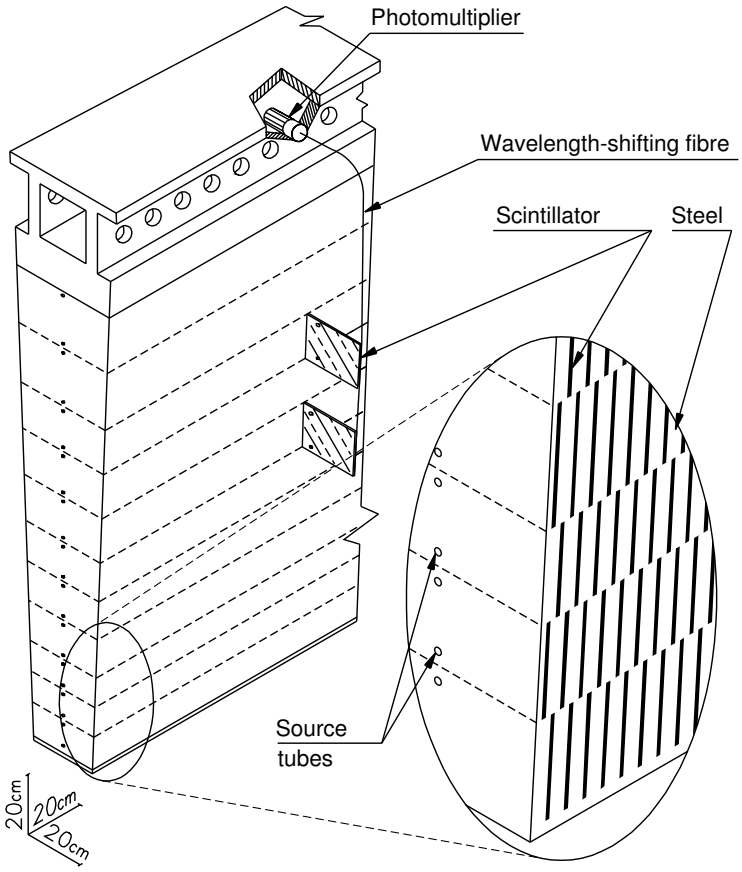


Figure 4.10: Schematic of a module of the ATLAS Tile calorimeter, showing the arrangement of the scintillating tiles and the placement of the readout components [125]. The source tubes are access channels for performing cesium calibration tests (see Section 4.5.3).

Particles entering the hadronic calorimeter create scintillation light proportionally to their energy. One optical fibre per one or two scintillating tiles then collects this light and carries

it to a PMT in the steel girders supporting the calorimeter. The signal which leaves the PMT is an analog current pulse with a width of order 18 ns full width at half maximum (FWHM) and an amplitude proportional to the energy deposition [127]. This signal is passed through a pulse shaper and a compressor, and is split into two samples with an amplitude ratio of 64: the larger sample is generally used unless sufficient energy is deposited to saturate the measurement, in which case the smaller is available [128]. The result is two parallel unipolar pulses with 50 ns FWHM, which maintain the original charge to amplitude relationship. Digitisation of the pulses is performed by analog-to-digital converters; the digital signal is sampled 7 times and the resultant measurements sent to the read-out electronics for trigger use.

4.5.2 Hadronic endcap (HEC) and forward (FCal) calorimeters

The more forward regions of the detector experience much higher particle flux than the low-η regions and must therefore be instrumented with hadronic calorimeters more radiation-hard than the tile calorimeter. Instead of scintillators, this region is instrumented with liquid argon sampling calorimeters. Two subsystems fill this purpose: the hadronic endcaps (HEC) cover the range $1.5 < |\eta| < 3.2$ while the forward calorimeters (FCal) cover the range $3.1 < |\eta| < 4.9$ [113, 123]. These detectors operate on the same principles as the electromagnetic calorimeter and share its readout electronic design.

The HEC consists of two wheels at each end of ATLAS, positioned at $z = 4.28$m (HEC1) and $z = 5.13$m (HEC2) and contained within the same cryostat which holds the ECal endcaps. Copper absorbers, 25 25-mm thick plates in HEC1 and 17 50-mm thick plates in HEC2, are layered in the z direction and separated by 8.5 mm active gaps. To ensure stability, each active gap is subdivided into four regions by a central readout sandwiched between two high-voltage sources.

The forward calorimeter also shares the endcap cryostat and begins at $z = 4.7$ m. It consists of three circular modules, one electromagnetic (FCAL1) with a copper absorber and two hadronic (FCAL2 & FCAL3) with tungsten absorbers, each 45 cm thick. The forward calorimeter must be extremely dense as a result of its location between the electromagnetic endcaps/HEC and the beamline: as much radiation as possible must be absorbed by these modules to prevent it from spilling out and manifesting as pileup in the nearby detector

components. The requisite density is achieved by placing the liquid argon active materials in longitudinal channels cut into a matrix of the absorber rather than in spaces between plates. A rod maintained at high voltage is placed in the centre of each channel while the surrounding matrix is grounded. The narrow and carefully controlled space between the rod and matrix then serves as the active gap. These spaces are as small as 250 μm in FCAL1 and 375 μm in FCAL2/FCAL3. A schematic of the rod and channel design is shown in Figure 4.11.

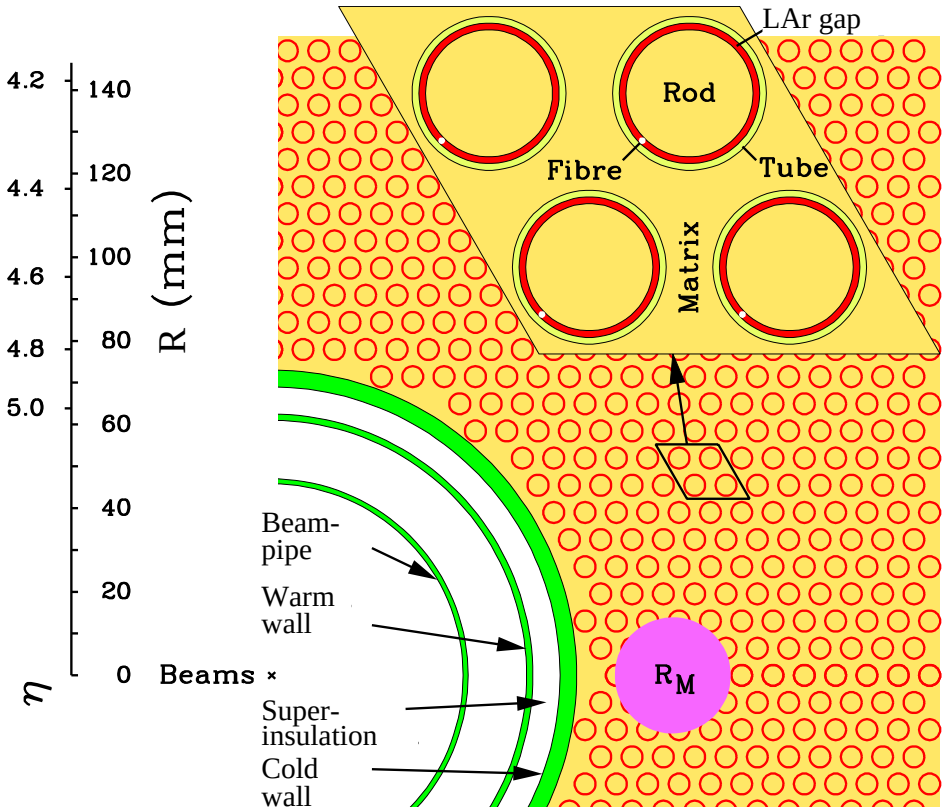


Figure 4.11: Layout of the longitudinal channels and high-voltage rods in the hadronic forward calorimeter [113].

4.5.3 Calibrating the energy scale

Cell energy from the tile calorimeter is calculated according to [129]:

$$E[\mathrm{GeV}] = \mathrm{A[ADC]} \cdot \mathrm{C_{ADC\to pC}} \cdot \mathrm{C_{Cs}} \cdot \mathrm{C_{laser}} \cdot \mathrm{C_{pC\to GeV}} \,. \tag{4.5}$$

The factors in the above equation are as follows:

- $A[ADC]$ is the readout pulse amplitude in ADC counts

- $C_{ADC \to pC}$ is a conversion factor, translating ADC counts into equivalent charge. It is calculated from regular tests of the readout electronics. A capacitor is discharged into the ADC, providing a relationship between measured amplitude and a known input charge.
- C_{Cs} results from approximately monthly tests of the TileCal using a cesium source to measure response differences between various channels and equalize the response of all the cells.
- C_{laser} originates from a measurement of the gain of each PMT using a laser scan performed more frequently than the cesium test.
- $C_{pC \to \mathrm{GeV}}$ is a conversion factor between charge and energy determined during the test beam measurements of 2001-2003.

The above calibrations were checked during the beam tests used to measure $C_{pC \to \mathrm{GeV}}$, while early cesium scans allowed the calibration of individual cells and PMTs to a constant response. The consistency of the response was tested by measurements with both test beam and cosmic muons, returning a non-uniformity of $2-3\%$ within a single layer of the TileCal and a maximum difference between layers of 4%. The calorimeter response to single hadrons is tested by selecting well-isolated tracks in the ID which the ECal identifies as minimum ionising particles, and recording the ratio E/p of the TileCal-measured energy to the momentum from the tracking detector. The E/p agreement between data and MC is found to be within 5% for the 2011 dataset [130].

The resulting measurement of the cell energy and energy uncertainty correspond to the EM scale. The ATLAS calorimeters are *non-compensating*, meaning that they respond differently to electromagnetic and hadronic showers. This is a common condition, since hadronic showers contain a mixture of electromagnetic and strong force-dominated subshowers which vary in the proportion of energy carried. The calorimeter response is lower for hadronic showers, and conversion from the energy scale of electromagnetic showers to that of hadronic ones is a complex process. It is the topic of the following chapter.

4.5.4 Calorimeter performance

The liquid argon and tile calorimeters operated at a good data quality efficiency of 99.1% and 99.6% respectively during 2012 data taking. For the liquid argon detector, this was the best

result yet, with incidents of high voltage trips in the detector reduced by 50% between 2011 and 2012, as well as a substantial improvement in the ability to tag coherent noise bursts. Several front-end boards became disconnected in the endcap due to a cooling leak, but the loss of FEBs was minimal compared to previous years, and older losses were accounted for in the detector simulations [131].

The tile calorimeter also observed an increase in efficiency between 2011 and 2012, in this case resulting from the replacement of a large number of low-voltage power supplies (LVPS) between the two years. Throughout Run I, the tile calorimeter has suffered the loss of various cells due to the failure of the LVPS supplying each. The cells cannot be reactivated during the run and the corresponding regions of the detector are *masked* for analysis purposes. Half as many cells were lost in 2012 as 2011 due to the replacement LVPS units [132].

4.6 Muon system

The outermost layer of ATLAS is the muon spectrometer (MS). Built directly around the toroidal magnets, the MS sits in a toroidal field of between 2 (central) and 8 (end-caps) T-m. A muon exiting the calorimeter system is accelerated in the $\pm z$ direction by this magnetic field, causing a curved trajectory that enables a precise measurement of its momentum. The muon momentum is measured at three separate points along its trajectory, permitting a tracking measurement with a design resolution of 50 μm.

The majority of the muon spectrometer is instrumented with monitored drift tubes (MDTs) arranged in double layers of 3 or 4 tubes of 30 mm diameter. The tubes, each with a central anode wire and filled with an argon, CO_2, and water vapour mixture, provide a high-precision tracking measurement but with a slow drift time of 700 ns. In the endcap regions, the MDTs are replaced with cathode strip chambers better able to cope with the high backgrounds arising from radiation near the beampipe. The muon triggering system requires a faster detector than either the MDTs or the cathode strip chambers, and so resistive plate chambers (RPCs) and thin gap chambers (TGCs) are added in the barrel and endcap regions respectively. These deliver a signal with much lower spatial resolution but with a timing spread of 15-25 ns, permitting the tagging of individual beam crossings [90, 133].

Measurements in 2011 *pp* collision data show that the transverse momentum resolution

as a function of p_{T} for muons in the barrel of the muon spectrometer is [134]

$$\frac{\sigma(p)}{p} = \frac{0.25 \text{ TeV}}{p_{\mathrm{T}}} \oplus 0.0327 \oplus 0.168 \text{ TeV}^{-1} \cdot \mathrm{p_T} \,. \tag{4.6}$$

The uncertainty on the momentum scale is within ±0.05%, and the total muon reconstruction efficiency in 2011 and 2012 data was close to 99% for $|\eta| < 2.5$ [135].

4.7 Triggering

With a bunch crossing every 25 ns, ATLAS generates data at a rate of 40 MHz in nominal operating conditions, with 1 to 2 GB of data per event. This is vastly more data than can be written to disk, and the majority of the events will contain little of interest. In order to reduce the data rate to a manageable quantity while ensuring that the events recorded are those which contain interesting physics, a three-tiered *trigger* system is implemented. The first level (L1) is hardware-based. With 2.5 μs to decide whether an event is worth storing, the L1 trigger reduces the data rate to 75,000 events per second. The high-level triggers (HLT) have longer to process the data and use more complete event information to further reduce the rate to the 200 events per second finally written to disk [90].

The L1 trigger uses reduced-granularity information from the calorimeters and the quick-response portions of the muon spectrometer to make rough identifications of certain physics objects: jets, electrons/photons, muons, and τ decays. It also responds to a large transverse energy or missing transverse energy signal. Decisions made by the L1 trigger involve only the presence or multiplicity of certain signals, but the locational information of the objects which fired the trigger remains in the processors and is passed on to the higher-level triggers if the event is kept.

Two trigger levels process the information passing L1: the L2 trigger level balances a high rejection level with low-precision, high speed algorithms and few computing requirements, while the event filter (EF) trigger level uses slower and more computing-intensive higher-precision algorithms with a lower rejection potential to make final decisions. The L2 trigger uses only a small fraction of the event information for decision making: it only considers geometrical areas flagged as regions of interest (ROIs) by the L1 trigger and typically making up about 2% of the full detector. Decisions at this level reduce the data rate enough (to

about 5 Gbytes/s) that the event can be fully reconstructed before passing it to the EF trigger. Analysis at this final level has ~ 4 s available to consider each event and thus can use complex algorithms. All events kept by the EF triggers are written to disk [136].

A *trigger chain* is a set of L1, L2, and EF triggers which can define the sequence of tests by which an event was accepted. Data which passes a trigger chain is read into a *stream*, which sorts events by type: *jet*, *muon*, etc. The *debug* stream is for events which could not be sufficiently processed during the allocated trigger time, but which have the possibility to be interesting: this is processed more slowly later and is often included in analyses because it tends to contain a few events of every type.

For some physics processes, including QCD jet production, the reduced rate is not sufficient to accommodate all the interesting events. Jets are produced in such abundance, especially at low energies, that only a subset of jet events can be kept. To reduce the rate further for triggers sensitive to such processes, a *prescale* is introduced. Only a fraction 1/(prescale) of events are recorded, decreasing the rate to a manageable level. Prescales can be adjusted on a run-by-run basis if desired, allowing triggers to adapt to run conditions.

Specific jet triggers are designed to identify events with at least one high-p_{T} jet originating from the hard interaction, and occupy about 10% of the full trigger bandwidth. The L1 calorimeter trigger considers sliding windows of $\Delta\eta \times \Delta\phi = 0.8 \times 0.8$ to locate regions with large energy deposits. If the transverse energy passes some threshold, a region of interest is defined around the calorimeter deposit and the event is passed to the HLT. L2 runs a simple iterative cone algorithm at full detector granularity to construct basic jet objects which are then compared to E_{T}, η, and quality requirements. Events which pass are available to the EF with full calorimeter data. The EF trigger uses offline event reconstruction and calibration algorithms and is able to reconstruct jets using the anti-k_t algorithm. Differences between trigger and offline jets, then, arise mostly from the fact that the final calibration is not available at the time of data-taking [137].

4.7.1 The delayed stream in 2012 data

In 2012, some data was recorded which passed a trigger chain but which was not reconstructed immediately due to insufficient CPU. The data in this *delayed stream* was reconstructed after the 2013 shutdown. The data selected to fill the additional trigger space

was lower-priority than the immediately reconstructed data: in the jet case, additional events were recorded in a p_{T} region covered by standard triggers but with a lower prescale. The delayed stream trigger `EF_j220_a10tcem`, which was unprescaled and had a lower p_{T} threshold of 220 GeV, became fully efficient at 333 GeV while the lowest unprescaled jet trigger in the normal (promptly-reconstructed) data stream became fully efficient at 460 GeV. `EF_j220_a10tcem` recorded 17.33 fb^{-1} of data during 2012, so its inclusion in the analysis dataset provided substantial statistical sensitivity.

4.8 Monte Carlo

Many processes, ranging from the determination of experimental resolutions to the estimation of analysis background rates, require a prediction based on the Standard Model and directly comparable to data. Simulated physics events (Monte Carlo, or MC) provide this tool. An event generator encodes the predictions of QED and QCD for pp collisions and simulates the range of possible outcomes over many collisions, resulting in a collection of events at the level of final-state particles. For jet physics, most available predictions are just at leading order. The initial matrix element can be calculated directly via a user-specified pdf selection. Parton showering is simulated probabilistically via the splitting functions and some hadronisation model is then used to generate the stable particles of the final state. Contributions from the underlying event and pileup interactions are also included in the simulations.

Pythia is the default MC generator for this analysis [138]. Internal generator settings governing the modelling of showering, hadronisation, fragmentation, multiple parton interactions, and so on are determined by *tunes* of the generator to a wide range of experimental data. The MC used in this analysis employs the AU2 tune [139] with Pythia version 8.160 and the CT10 pdf set [140]. Since it is a leading-order event generator in all processes, additional scaling factors (*k-factors*) were added in the analysis to correct the predictions to match NLO results. These were calculated by comparing the Pythia predictions in each bin of the dijet spectrum to the predictions from a next-to-leading-order generator. For the comparison, the NLOJET++ v4.1.3 generator [141–143] was used with the (NLO) CTEQ6.6 pdf set [144].

In order to simulate the interactions between the final-state hadrons and the ATLAS

detector, the MC events are passed to the GEANT4 software toolkit [145]. GEANT4 simulates the interactions of particles with the detector, taking into account the scattering of the particles in different materials and using the lifetimes and branching ratios of each particle to probabilistically select its decay path. Modelling here accounts for the different densities and interaction lengths of the various subdetector components and even simulates particle interaction with the readout electronics. The result is a set of digitised signals from each simulated detector component taking exactly the same form as data readout from the real ATLAS detector. These signals are then processed using the same reconstruction software used for data.

4.9 Data processing and object reconstruction

Data processing is handled on a large scale by the LHC Computing Grid, which distributes data to sites worldwide for processing and storage. ATLAS datasets are always maintained in multiple copies at different Grid sites to protect against any data loss. [146].

Measured data from the ATLAS detector and simulated data from GEANT4 are both passed into the same reconstruction software. ATLAS uses a C++ framework called ATHENA to reconstruct the various detector signals and identify them as physics objects. The first step is to reconstruct data from the *express* stream, a subset consisting of about 5% of the incoming data, using the best available calibration estimates. A more detailed preliminary calibration obtained from the express stream is then applied to the remainder of the data, which is processed beginning about 24 hours after its collection [147]. The details of jet object reconstruction and calibration are discussed in the following chapter.

A critical ingredient in data storage and processing is *metadata*, information about the LHC and ATLAS conditions at the time of data-taking. For instance, the state of the LHC magnets, the detector component temperatures, and the presence of any dead regions in the detector subsystems should all be accounted for in data reconstruction. Trigger conditions are also critical. Using the relevant metadata, each luminosity block is assigned a *data quality flag* to indicate the reliability of its data. *Good run lists* (GRLs) of runs and luminosity blocks safe for physics analyses are then assembled by filtering on specific data quality flags [148]. Although these can be user-produced, it was recommended in 2012 to use centrally-produced GRLs for a reliable outcome and consistency across analysis groups.

Chapter 5

Jets in ATLAS

'Nothing has ever really happened until it has been recorded.'
— Virginia Woolf

5.1 Jet definition and inputs

Jets in ATLAS can be defined using a variety of objects either at the tracking or calorimeter level or at a higher level of reconstruction, and jets in Monte Carlo can additionally be defined using simulations at the parton or particle level. Jets made using reconstructed tracks and momenta in the inner detector as inputs are called *track jets* and are often used for calibration purposes. *Truth jets* can be formed in Monte Carlo using stable particle-level constituents as input four-vectors to the same jet reconstruction algorithm used on experimental quantities. Because these provide a measurable relationship between initial parton properties and those of the final jet object, they are an important tool in jet calibration and performance studies. Finally, jets used for analysis are most commonly constructed from calorimeter measurements, and although calorimeter cells can be directly input into jet clustering algorithms, ATLAS uses a higher-level object, the *topocluster*, instead.

A topocluster is a grouping of neighbouring calorimeter cells distinguished by significant energy deposits, and thus reflects the energy flow from a shower developing in the calorimeter. All cells above a high threshold energy-to-noise ratio r_{seed} are selected to seed topocluster development. Neighbouring cells with ratios above a threshold r_{cell} are added to the cluster, and if their ratio is significant enough to pass an intermediate threshold $r_{\text{neighbour}}$ then their

neighbours are also considered. Here, neighbours can be defined within the same calorimeter layer as a seed cell or can span layers and even subsystems. Final clusters are kept if their total transverse energy is above some cutoff. The topocluster creation process is designed to suppress noise and pileup by requiring a minimum energy to expand clusters, but includes softer radiation in the perimeters to prevent loss of jet energy. Given the high energy density in the calorimeter, energy deposits from different particles can be expected to merge during this step, and so a topocluster splitting step is also included. Cells constituting local maxima are selected and used to seed a re-clustering using only the initially selected cells and without any cluster merging, giving the final set of separate topoclusters [149]. Each topocluster is treated as a massless four-vector with energy $\sum E_{\text{cell}}$ and a direction defined from the detector centre to the energy-weighted centre of the topocluster.

Jet reconstruction algorithms are implemented in FastJet, which clusters any collection of four-vectors according to specified parameters [150]. In standard ATLAS data processing, two jet radius parameters are commonly used: $R = 0.4$ and $R = 0.6$. Other jet sizes are occasionally available, and the collaboration's increasing use of "fat" jets has popularised the choice of $\Delta R = 1.0$-jets. This analysis makes use of anti-k_t $R = 0.6$ jets.

5.2 Calibration and performance

Jets immediately reconstructed from calorimeter topoclusters do not necessarily reflect the true energy of the original parton. Perhaps the most important concern is that the cells and topoclusters themselves accurately reflect the energy that would be deposited by electromagnetic showers but do not properly account for hadronic showers. Other factors which disrupt the energy prediction result from detector geometry and the jet reconstruction algorithms. The following factors contribute to jet energy mismeasurement and are addressed by *jet calibration* [151]:

- **Non-compensating hadronic calorimeter**: This refers to the discrepancy between the energy measurement of electromagnetic showers and hadronic showers, where the hadronic calorimeter does not properly account for the latter.
- **Dead material**: Some of the energy of a jet may land in non-responsive parts of the detector.

- **Leakage**: Some energy may fall outside the calorimeter altogether, such as any showers which punch through the calorimeter and into the muon system.
- **Out of jet cone**: Particles which, in a Monte Carlo comparison, belong in the truth jet but fall outside the reconstructed calorimeter jet are a source of energy loss due to the jet reconstruction process.
- **Thresholds and reconstruction efficiency**: Some energy belonging in the jet may be missed during the clustering and reconstruction due to falling below the noise threshold requirements for cluster formation.

Two jet collections are produced before the bulk of the calibration is undertaken. One collection, the electromagnetic (EM) jets, is formed directly from EM-scale topoclusters. A second, local cell signal weighting (LCW) jets, makes a preliminary calibration at the topocluster level before the jet reconstruction is performed. The LCW calibration uses shower shape, depth, and energy density to assign topoclusters to electromagnetic or hadronic showers and then adjusts those identified as hadronic for dead material, leakage, and non-compensation [152]. Both EM and LCW jets still need to undergo a detailed, dedicated calibration process at the level of the entire jet in order to achieve the final desired energy resolution. The following 4-step process is the current recommended ATLAS calibration procedure and the resulting jet collections are EM+JES and LCW+JES jets. This analysis uses EM+JES jets and the relevant terminology will be followed below.

5.2.1 The EM+JES calibration scheme

The EM+JES calibration scheme was introduced in 2010 and can be found in detail in Ref. [151]. It returns observed jets as near as possible to the energy scale of corresponding truth jets simulated at particle level. This four-stage process is schematically illustrated in Figure 5.1 while individual stages are described in the following sections.

Pileup correction

In this stage, the average contribution to a jet's energy from pileup is calculated and the jet is modified accordingly. Monte Carlo events are used to determine a correlation between the number of primary vertices in the event, the bunch spacing, and the energy deposited on average in the calorimeter as a function of η. The expected energy density is calculated

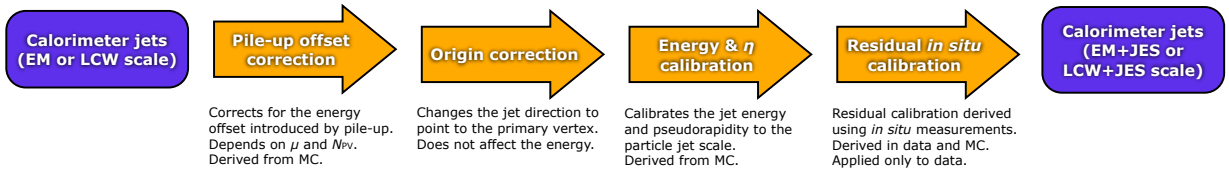


Figure 5.1: Flow of the EM+JES calibration scheme. EM or LCW-scale topoclusters are used to form jets, which then undergo a four-step process to return them to the hadronic energy scale and to correct the direction of the jet four-vector [153]

for each possible N_{PV}. Combined with a measure of jet area, an offset correction to the jet's energy is calculated and applied to return it to a reference scale corresponding to $N_{PV} = 1$ [154, 155].

Jet origin correction

Initially, the directional component assigned to calorimeter cells and topoclusters originates at the geometrical centre of ATLAS detector, and as a result so do EM-scale jets. In actuality (non-pileup) jets originate from the primary vertex, which may be some distance away. The four-momenta of each topocluster are adjusted to point to the true origin and their kinematic properties adjusted accordingly, and then the four-momentum of the jet is recalculated by summing the corrected values. While the jet direction is altered, its energy is not. This correction improves the jet angular resolution.

Jet energy and η corrections

This, the most complex stage of the calibration, returns the jet energy and η to the scale of truth MC jets. The calibration to be applied is derived as follows [151].

A Monte Carlo sample is selected for the process which does not simulate pileup, since those effects were already accounted for in the first calibration step. From the sample, an MC calorimeter jet is used for the calculation if it is isolated and matched to a truth jet ($\Delta R < 0.3$ between the two jets). The EM-scale jet energy response is defined for these matched MC jets as:

$$\mathcal{R}^{\text{jet}}_{\text{EM}} = E^{\text{jet}}_{\text{EM}} / E^{\text{jet}}_{\text{truth}} \,. \tag{5.1}$$

The response is calculated for all such jets and binned in truth jet energy and η_{det}, where

the detector-based rather than origin-corrected η value is used in order to produce a result which corresponds most directly to a region of the calorimeter. Within each $(E^{\text{jet}}_{\text{truth}}, \eta_{\text{det}})$ bin, the $\mathcal{R}^{\text{jet}}_{\text{EM}}$ distribution is fitted with a Gaussian whose mean is used to define $\langle \mathcal{R}^{\text{jet}}_{\text{EM}} \rangle$. This quantity depends on the mean reconstructed jet energy $\langle E^{\text{jet}}_{\text{EM}} \rangle$ in the same E_{truth} bin and so varies with jet energy. The shape of its distribution for an η bin k across the range of E_{truth} values is parameterised by fitting it with the following function:

$$\mathcal{F}_{\text{calib},k}(E^{\text{jet}}_{\text{EM}}) = \sum_{i=0}^{N_{\text{max}}} a_{i,k} (\ln E^{\text{jet}}_{\text{EM}})^i_k \,. \tag{5.2}$$

Here, the a_i are free parameters of the fit and the number of terms in the function N_{max} can be up to 6 depending on the goodness of fit. Now for a measured EM jet of a given η_{det} and E^{EM}_{jet}, its energy can be corrected to the truth jet scale by dividing it by the value of $\mathcal{F}_{\text{calib},k}$ in the corresponding η_{det} bin at the appropriate energy:

$$E^{\text{jet}}_{\text{EM+JES}} = \frac{E^{\text{jet}}_{\text{EM}}}{\mathcal{F}_{\text{calib}}(E^{\text{jet}}_{\text{EM}})\big|_{k=\eta_{\text{det}}}} \,. \tag{5.3}$$

Residual *in situ* calibration

The final calibration stage accounts for differences between data and MC. Using a well-measured set of objects as a reference, the correct jet energy can be derived from conservation of transverse momentum. In the central region of the detector ($\eta < 1.2$), the direct p_{T} balance between a jet and a Z boson can estimate the remaining jet correction up to $p_{\text{T}} \sim 200$ GeV. Up to 800 GeV, a photon can be balanced against the total recoil of the jet and the proton remnant. Finally, for jets up to the TeV scale, a very high-p_{T} jet can be balanced against a group of lower-p_{T} jets already calibrated with the preceding two methods. For jets in the forward calorimeter regions, the transverse momentum balance method draws a comparison to well-calibrated central jets, and thus the calibration is extended across the full η_{det} range. The same balance techniques are applied on data and MC and the difference between the two is taken as the residual *in situ* correction and applied to the data [156]:

$$\text{Correction} = 1/\mathcal{R}(p^{\text{jet}}_{\text{T}}, \eta) = \frac{\langle p^{\text{jet}}_{\text{T}} / p^{\text{ref}}_{\text{T}} \rangle_{\text{MC}}}{\langle p^{\text{jet}}_{\text{T}} / p^{\text{ref}}_{\text{T}} \rangle_{\text{data}}} \,. \tag{5.4}$$

5.2.2 Performance in 2012 data

A full jet energy scale calibration is not yet complete for 8 TeV data and so the final 7 TeV calibration results are the most current estimate available of the calibration performance [153]. Figure 5.2 illustrates the average jet response for EM jets across the range of η_{det} for five energy bins using a 7 TeV MC dataset. The inverse of the values shown are the average jet energy scale correction for jets in that bin. As the trend between differing jet energies shows, the JES correction shrinks with increasing jet energy.

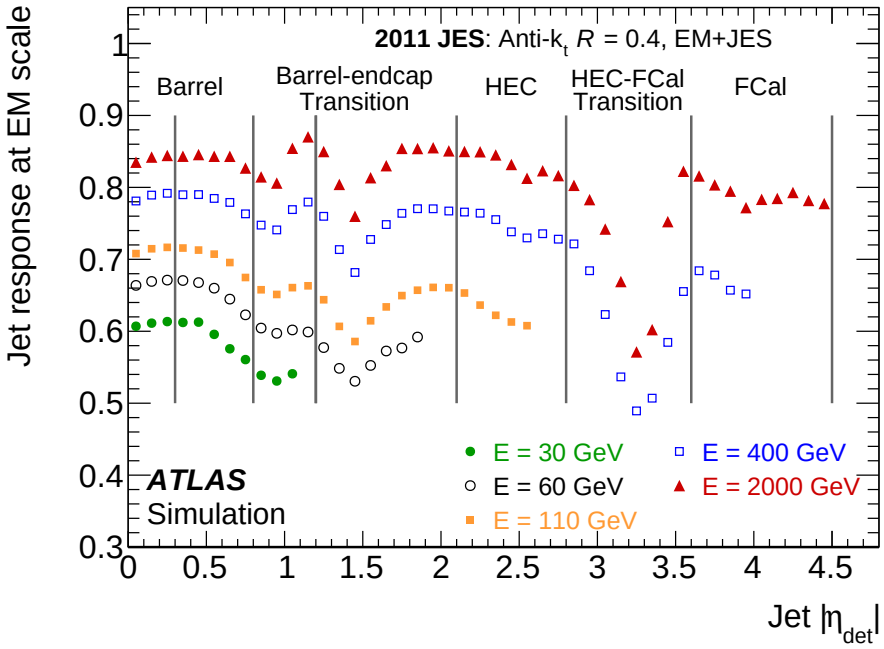


Figure 5.2: Average jet response $\mathcal{R}^{\text{jet}}_{\text{EM}}$ is shown for a range of jet energies across the full η range of the detector. EM-scale jets from a 7 TeV sample were used to compute the calibration shown [153].

Not only the energy correction is critical for a well-measured jet based analysis: the uncertainties on the calibration must be well-understood. Jet energy scale, and to a lesser degree jet energy resolution, uncertainties are among the largest sources of systematic uncertainty in most jet-based analyses. The following section discusses current methods for estimating the uncertainty and gives the values used in analysing the 2012 dataset.

5.3 Uncertainties

5.3.1 Jet energy scale uncertainty

The largest uncertainty for most jet analyses is associated with the jet energy scale calibration process. The uncertainty on the jet energy scale after the *in situ* calibration can be estimated by varying the parameters of the various constituent methods and observing the impact on the final correction. The result is a set of many uncertainties which vary in size across different η and p_{T} regions rather than a single unified systematic.

Z-jet p_{T} balance

Measurement of the jet-Z boson p_{T} balance is binned in Zp_{T} = $(p_{\mathrm{T}}^{\mathrm{ref}})$ and the azimuthal distance between the jet and the Z boson, $\Delta\Phi$. The correction is computed in one of two ways: for $p_{\mathrm{T}}^{\mathrm{ref}} < 35$ GeV, the distribution of $\mathcal{R}_{Z,\mathrm{jet}} = p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}}$ is sensitive to the jet selection threshold $p_{\mathrm{T}}^{\mathrm{jet}} > 12$ GeV. Thus for $(p_{\mathrm{T}}^{\mathrm{ref}},\ \Delta\phi)$ bins with p_{T} below 35 GeV, the mean value of $\mathcal{R}_{Z,\mathrm{jet}}$ is obtained by fitting its distribution with a smooth Poissonian fit function multiplied by a turn-on curve and taking the mean of the fitted Poisson. Following the fits, the distribution of Poisson widths across all $p_{\mathrm{T}}^{\mathrm{ref}}$ bins for a given $\Delta\phi$ is parameterised by its own function. Finally the $R_{Z,\mathrm{jet}}$ distribution in each bin is re-fit while fixing the Poisson width to the value obtained from the parameterisation. For bins with $p_{\mathrm{T}}^{\mathrm{ref}} > 35$ GeV, the jet balance is found by computing an arithmetic mean with no fit required. Due to sensitivity to additional radiation when the jet and Z are exactly back-to-back, the estimate of $R_{Z,\mathrm{jet}}$ for $\Delta\phi = \pi$ is computed by linearly extrapolating from nearby bins. [157].

The Z-jet p_{T} balance method is sensitive to seven sources of uncertainty:

1. Fitting procedure: uncertainties arising from the fit procedure are estimated by propagating the uncertainty on the width parameterisation through the final fit and measuring the deviation in the resulting mean.
2. Extrapolation procedure: the limit of the linear fit is changed and the alteration in the predicted $\Delta\phi = \pi$ bin used to set an uncertainty on the process.
3. Radiation suppression: an uncertainty results from the veto of a second jet in the dataset. Its magnitude is found by varying the jet veto cut and comparing the p_{T}

balance in data and MC.

4. Out-of-cone radiation and underlying event: the Z truly balances all event momentum, not just that of the jet. The uncertainty due to missing or additional energy in the jet is estimated from the change in p_{T} balance when considering all event energy in data and MC.
5. Additional pileup: the p_{T} balance is compared across samples with varying N_{PV} and $\langle\mu\rangle$ and the difference used to estimate the uncertainty.
6. Electron energy scale: the Z boson p_{T} measurement depends on electron energy scale. Deviation in the ratio at different E_e scales provides an uncertainty.
7. MC generator: uncertainties are calculated by comparing the p_{T} balance between Pythia and Alpgen MC.

Statistical uncertainties are also included. The total uncertainty is typically ~ 1 to 2% and no more than 10% for this method.

γ-jet p_{T} balance

Instead of a direct balance, the missing transverse momentum projection fraction (MPF) is used to compute the γ-jet correction. The missing energy projected onto the photon direction is used as a stand-in for the proton remnant:

$$\mathcal{R}_{\mathrm{MPF}} = 1 + \frac{\vec{p_{\mathrm{T}}}^{\gamma} \cdot \vec{E}_{\mathrm{T}}^{\mathrm{miss}}}{|p_{\mathrm{T}}^{\gamma}|^2} \,. \tag{5.5}$$

$\mathcal{R}_{\mathrm{MPF}}$ is binned in p_{T}^{γ} and the mean computed by fitting the distribution: in the first bin a Poisson distribution is used while all further bins use a Gaussian function. The difference in MPF response between data and MC is of order 1-2% and depends on differences in multi-jet event backgrounds, out-of-cone energy distributions, and response to the low p_{T}^{γ} cutoff [158]. The significant systematic uncertainties in this calibration are as follows (no description is given when the systematic is analogous to the Z-jet case):

1. Radiation suppression.
2. Jet energy resolution (JER). The uncertainty on JER in data is used to determine an additional uncertainty in MC.
3. Photon purity. An uncertainty on the proportion of jets mis-identified as photons is

estimated by comparing the responses in signal and background events for a samples of known purity and taking the difference in the comparison between data and MC as the error band.

4. Out-of-cone radiation and underlying event.
5. Additional pileup.
6. Photon energy scale. The γ p_{T} measurement depends on the photon energy scale. The uncertainty from the photon calibration is directly propagated to the response.
7. MC generator.

Statistical uncertainties are also included. The total uncertainty is below 1% over most energies, rising to $\sim 2.5\%$ at very low p_{T}^{γ}.

Multijet balance

The multijet balance technique determines the energy scale of one high-p_{T} jet by balancing it against the total $\vec{p_{\mathrm{T}}}$ of a recoil system of well-calibrated lower-p_{T} jets. The response variable MJB = $|p_{\mathrm{T}}^{\mathrm{leading}}|/|p_{\mathrm{T}}^{\mathrm{recoil}}|$ is calculated for data and MC to determine the jet energy scale for events with subleading jets of known JES, and the uncertainties on the scale for the recoil system are propagated through to the new higher-p_{T} JES. By then relaxing the cut on subleading jet p_{T} to include the jets with newly calibrated JES, recoil systems of higher p_{T} can be considered and higher-p_{T} leading jets calibrated. The process of computing a new JES, propagating uncertainties from the recoil system, and relaxing the subleading jet requirements is iterated until a jet energy scale is available across the full range of jet p_{T} [159].

The most significant sources of uncertainty are as follows:

1. Absolute JES uncertainty. The JES calibration uncertainty for the jets in the recoil system is propagated through the MJB calculation and affects the uncertainty on the leading jet. All uncertainties in the Z+jet and γ+jet calibrations affect this.
2. Relative JES uncertainty. Relative jet response uncertainties from a dijet η-intercalibration are similarly propagated to the leading jet.
3. Close-by effects. Jet response depends on the proximity to another jet, and the low-p_{T} jets in the recoil system will be close enough together that the effects can be dramatic.

The calorimeter response is defined as the p_{T} ratio between matched calorimeter jets and track jets. The ratio of reponse between isolated and non-isolated jets is then compared between data and MC, and the difference used to determine the uncertainty on the jet energy scale.

4. Analysis cuts. Any effects on MJB resulting from analysis cuts which affect data and MC differently introduce a systematic. Cuts are varied and the result on the MJB data-to-MC ratio used to define the uncertainty.
5. Underlying event (UE), fragmentation, initial and final state radiation (ISR/FSR). Modelling of these processes by the MC may not accurately reflect data and can affect analysis cuts and kinematic variables. The size of the effect is estimated by comparing the nominal MC samples to an alternate generator.

A total of 9 separate uncertainty components result from the effects in the preceding list. The η-intercalibration step wherein central jets are used to balance those in the forward region of the detector to derive a jet energy scale across the entire η range follows the same principles but is often listed separately due to its different region of influence.

Other sources of JES uncertainty

A few remaining uncertainty components separate from the *in situ* techniques also affect the JES. The early stage of correcting MC for pileup introduces two systematic uncertainties for N_{PV} and $\langle\mu\rangle$. The effects of close-by jets on all stages of the calibration are also considered. Here, the relative response for track jets matched to calorimeter jets, binned in ΔR to the nearest other jet, provides a measurement of calorimeter response to close-by scenarios. Additionally, uncertainty components arise from the structural differences between jets originating from different partons. The flavour composition and flavour response JES uncertainty components account for the differences in composition and calorimeter response between quark and gluon jets as described by Pythia and Herwig.

The final uncertainty source is the most relevant for very high-p_{T} jets. The single hadron response calibration is used when jets are of such high p_{T} that no other method can reliably estimate the energy scale. In this case, jets are treated as collections of energy depositions from individual hadrons and the known calorimeter response to single particles is used to generate an estimate for the particle cluster energies. Single isolated hadrons in minimum

bias data samples provide a measure of the jet energy scale per particle. Then the type of particle for each energy deposit in a jet is identified and the single-particle response used to calibrate the energy. Comparing data to MC provides an uncertainty on the calibration which increases with jet p_{T} [160]. The single hadron estimate is only used for jets with $p_{\mathrm{T}} > 1$ TeV, as the jet balance methods provide a better estimate below that point.

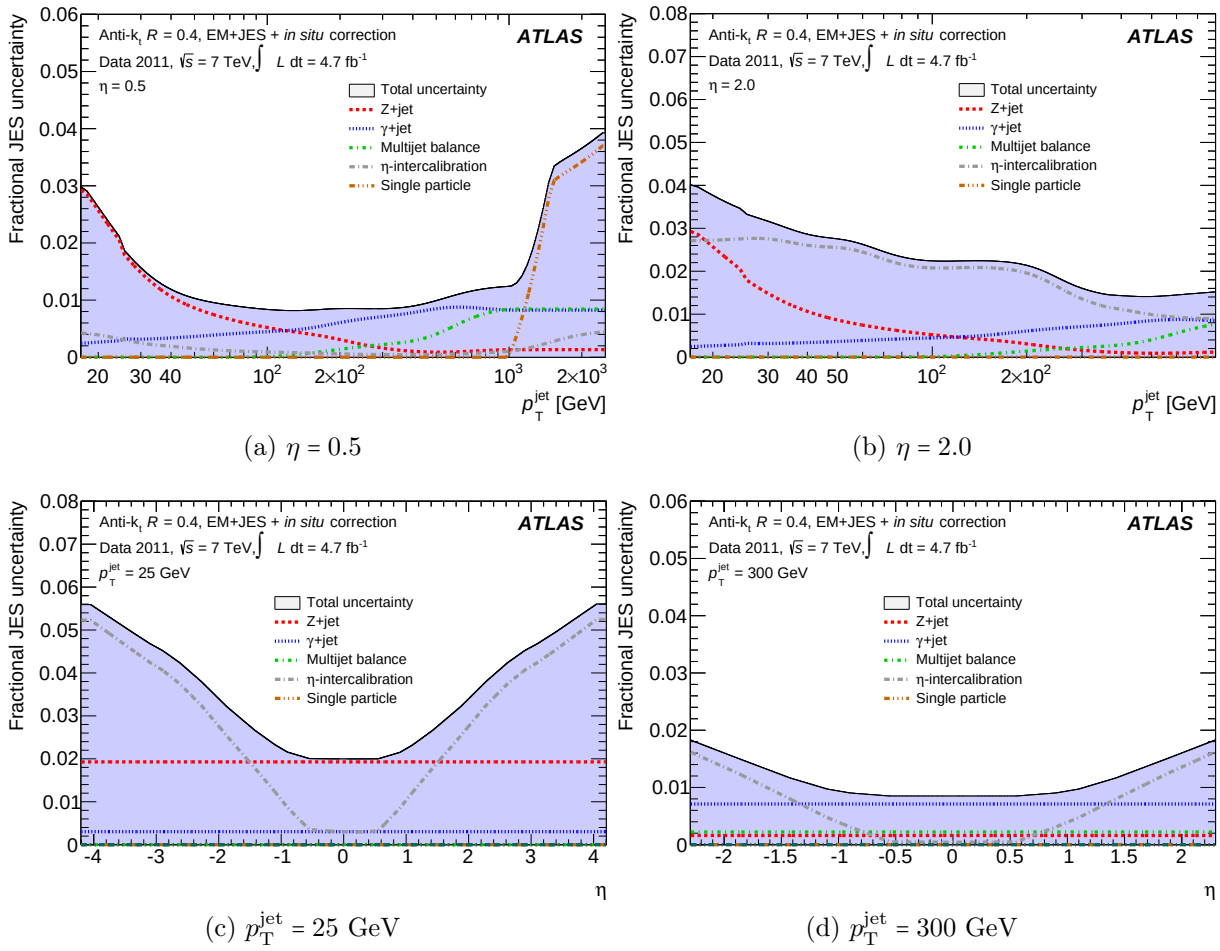
Total

In total, there are 54 individual systematic uncertainty components used to describe the *in situ* jet energy scale calibration. In some cases several components arise from a single physical cause. The uncertainty components can be classified into four types: detector, model, statistics/methodology, or mixed. In general, detector or statistics-based uncertainties will not be correlated, but model-dependent ones may be. Due to the increased calculation time and presence of spurious degrees of freedom when components have little effect, not all 54 components will usually be handled separately in an analysis. For such cases, several sets of reduced systematics and their correlations are made available. A decorrelated set of uncertainty components is generated from the original covariance matrix and the weaker components combined in quadrature to make a new reduced set of components. Three such sets are generated, with stronger and weaker assumed correlation between the components: these can be used to test analysis dependence on correlations.

The total jet energy scale uncertainty for EM+JES-calibrated, $R = 0.6$ jet p_{T} values across the η range of the inner detector is shown in Table 5.1. Figure 5.3 illustrates the relative contributions to the jet energy scale systematic uncertainty from each of the calibration steps. Contributions from the single-particle uncertainty turn on above $p_{\mathrm{T}}^{\mathrm{jet}} \sim 1$ TeV, while the η-intercalibration uncertainty takes effect for forward jets.

5.3.2 Jet energy resolution uncertainty

An uncertainty on the jet energy resolution $\sigma(E)/E$ arises when the detector resolution is not correctly simulated in MC. The scale of the uncertainty can be measured *in situ* from the jet balance in events with high-p_{T} jets using two techniques: the *dijet balance method* and the *bisector method* [161]. A combination of the results from these methods is then compared to truth information in MC as a closure test.



(a) $\eta = 0.5$ (b) $\eta = 2.0$

(c) $p_{\mathrm{T}}^{\mathrm{jet}} = 25$ GeV (d) $p_{\mathrm{T}}^{\mathrm{jet}} = 300$ GeV

Figure 5.3: Fractional jet energy scale uncertainty for $R = 0.4$ jets calibrated using the EM+JES method. The top two plots show uncertainty versus jet p_{T} for central jets (5.3a) and forward jets (5.3b). The lower two plots show uncertainty versus η for low-p_{T} jets (5.3c) and high-p_{T} jets (5.3d) [153].

η region	$p_{\mathrm{T}}^{\mathrm{jet}} = 20$ GeV	$p_{\mathrm{T}}^{\mathrm{jet}} = 40$ GeV	$p_{\mathrm{T}}^{\mathrm{jet}} = 200$ GeV	$p_{\mathrm{T}}^{\mathrm{jet}} = 800$ GeV	$p_{\mathrm{T}}^{\mathrm{jet}} = 1.5$ TeV
$\lvert\eta\rvert = 0.1$	2.7%	1.4%	0.8%	1.8%	3.3%
$\lvert\eta\rvert = 0.5$	2.7%	1.5%	0.8%	1.8%	3.3%
$\lvert\eta\rvert = 1.0$	2.8%	1.6%	0.9%	1.8%	3.3%
$\lvert\eta\rvert = 1.5$	3.0%	1.9%	1.3%	1.9%	3.3%
$\lvert\eta\rvert = 2.0$	3.6%	2.6%	1.9%	2.0%	–

Table 5.1: Total size of the jet energy scale uncertainty for $R = 0.6$ jets after the EM+JES calibration in 7 TeV data. Only the η regions corresponding to the inner detector are shown [153].

The dijet balance method exploits the fact that jet p_{T} resolution and energy resolution are the same at fixed η by measuring the p_{T} resolution uncertainty. It determines the dependence of the p_{T} balance between the two leading jets of a dijet system on the presence of additional

jets. The p_{T} asymmetry is defined as

$$A(p_{\mathrm{T},1}, p_{\mathrm{T},2}) \equiv \frac{p_{\mathrm{T},1} - p_{\mathrm{T},2}}{p_{\mathrm{T},1} + p_{\mathrm{T},2}} \,. \tag{5.6}$$

In a sample of jets where $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ are randomly selected between the two leading jets, A will be normally distributed with width

$$\sigma(A) = \frac{\sqrt{\sigma^2(p_{\mathrm{T},1}) + \sigma^2(p_{\mathrm{T},2})}}{\langle p_{\mathrm{T},1} + p_{\mathrm{T},2} \rangle} \simeq \frac{\sigma(p_{\mathrm{T}})}{\sqrt{2} p_{\mathrm{T}}} \,. \tag{5.7}$$

The p_{T} resolution and thus energy resolution are measured directly from a fit to A [162].

Two corrections are applied to the dijet balance method. The first accounts for soft radiation in the calorimeter by testing the change in $\sigma(A)$ with increasing p_{T} of the third jet. Extrapolating the trend back to $p_{\mathrm{T},3} = 0$, a correction factor for events with soft additional radiation is defined. As $p_{\mathrm{T},1}$ increases the correction becomes less significant, decreasing from 25% for 50 GeV jets to 5% for $p_{\mathrm{T},1} \sim 400$ GeV. A second correction accounts for the particle-level p_{T} difference between the two leading jets: not only the calorimeter jets but particle level jets should be balanced, and these two jet types differ due to soft QCD effects and out-of-cone radiation. Corrections for these discrepancies are calculated as in the calorimeter-level jet soft radiation treatment. This correction ranges from 10% for 40 GeV jets to 2% for jets above 400 GeV.

The bisector method complements this, examining the projection of $\vec{P}_{\mathrm{T}} = \vec{p_{\mathrm{T},1}} + \vec{p_{\mathrm{T},2}}$ onto the plane perpendicular to the bisector of the angle between a dijet pair. $\vec{P}_{\mathrm{T}}$ is 0 for a perfectly balanced event; deviations can be decomposed into the bisector direction $\vec{\eta}$ and the perpendicular component ($\vec{\Psi}$). The biggest contribution to $\vec{P}_{\mathrm{T}}$, initial-state radiation, is expected to be isotropic in Ψ and η. Factors affecting jet energy resolution such as FSR and contamination from 3-jet events, however, will affect $P_{\mathrm{T},\Psi}$ more than $P_{\mathrm{T},\eta}$ since the former is the difference of two large numbers while the latter is the sum of two small ones. The difference in variance of the two components becomes a measurement of the energy resolution:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{\sqrt{(\sigma_{\Psi}^{\mathrm{calo}})^2 - \sigma_{\eta}^{\mathrm{calo}})^2}}{\sqrt{2} p_{\mathrm{T}} \sqrt{\langle \cos \delta\phi_{12} \rangle}} \,. \tag{5.8}$$

Fits to σ_{Ψ} and σ_{η} in dijet data show that the resolution in η is approximately constant with

$\bar{p_T}$ while the resolution in Ψ deteriorates with increasing $\bar{p_T}$. The chief uncertainty is the precision with which the assumption $\sigma_{\Psi}^{\text{particle}} = \sigma_{\eta}^{\text{particle}}$ can be checked in data, while JES and pileup make smaller contributions. The overall uncertainty on the *in situ* methods decreases with increasing p_T to around 1.1% for central rapidity `AntiKt6` jets with $p_T > 760$ GeV.

The dominant systematic uncertainties on the measurement are found by combining the *in situ* measurements with truth information. First, the closure uncertainty is the precision with which the truth jet response distribution matches the *in situ* resolution measurements for truth-matched MC calorimeter jets. Second, an additional uncertainty measures agreement between the result in data and the reconstruction-level resolution measurement in MC. Both uncertainties are within 10%, while all other sources are substantially smaller. The total uncertainties on the jet energy resolution for `AntiKt6` jets are given in Table 5.2 [161].

Rapidity range	Total systematic uncertainty		
	Low p_T	Med. p_T	High p_T
$0 \le \lvert y \rvert < 0.8$	12%	10%	11%
$0.8 \le \lvert y \rvert < 1.2$	12%	10%	13%
$1.2 \le \lvert y \rvert < 2.1$	14%	12%	14%
$2.1 \le \lvert y \rvert < 2.8$	15%	13%	18%

Table 5.2: Total jet energy resolution (*in situ*, closure, data/MC agreement, smaller contributions) for `AntiKt6` jets across p_T and y for 2012 data.

Part II

The Dijet Analysis

Chapter 6

The Dijet Spectrum

"Explain all that," said the Mock Turtle.
"No, no! The adventures first," said the Gryphon in an impatient tone: "explanations take such a dreadful time."'
— Lewis Carroll, *Alice's Adventures in Wonderland*

New physics models with the capacity to produce a final state consisting of only two high-energy jets can be explored by investigating the dijet invariant mass spectrum. In events produced by the Standard Model only, the two highest-energy (highest-p_{T}) jets in the event will in general not come from the decay of a single particle at a fixed mass, but will be the combined result of random, usually QCD, processes. In the case of new physics, whether from a resonance, a new interaction between particles, or a newly accessible force such as TeV-scale gravity, the two leading jets in the event could proceed directly from this phenomenon. In that case, their invariant mass would provide a measure of the scale of the new process.

The invariant mass of a dijet system, m_{jj}, is calculated by summing the four-vectors of the leading and subleading jets. A spectrum formed by plotting this quantity over a large range of selected SM events will create a smoothly falling spectrum dominated by QCD jet production. Resonance-like new physics processes, including those discussed in Chapter 2, where a pair of jets are produced would appear as a "bump" above the m_{jj} spectrum at the resonance mass. The dijet resonance analysis uses a data-driven method to describe this smoothly falling background and searches for bumps appearing in the data when compared to the background estimation. If no excesses are found, limits are set on the highest mass

or cross section of new narrow resonances which could remain hidden in the data.

Searches for new physics using the dijet mass spectrum have a long history at collider experiments, dating back to the 1980's at early CERN colliders and the Tevatron. The UA1, UA2, CDF, and D0 collaborations all published at least one version of such a measurement (see for example [163–166]). At the LHC, it has been a staple of the Exotic search menu in both major experiments since the very earliest data, and the analysis in fact provided the first search paper published on collision data in ATLAS. None of these searches has yet shown an excess large enough to constitute a challenge to the SM. The full set of CMS and ATLAS published results for dijet searches can be found in References [167–177].

The analysis described in this thesis has recently been submitted to Phys. Rev. D. and the paper can be found in Reference [178]. This is the final dijet analysis result to be released on Run I data, and as such has been used as an opportunity to extend the scope of the analysis beyond any previous version. This iteration extends the range of masses considered from a previous lower boundary in 8 TeV data of 1 TeV down to 250 GeV, and includes an additional dataset reconstructed after the run completed (the *delayed stream*, to be discussed in detail below). The set of new physics models considered in the limit setting was extended to include an example of each possible final states (qg, gg, qq) and to include quantum black holes. A new, more extensive treatment of systematic uncertainties was introduced, allowing a highly detailed treatment of the sources of uncertainty. Finally, in addition to the traditional generic resonance signal represented by a gaussian functional form, a set of more physically motivated Breit-Wigner based generic signals were added to the limit setting, providing a new tool for reinterpretation.

Construction of the m_{jj} spectrum begins with rejecting events that do not meet the appropriate analysis cuts. The definition of these selection criteria, and the calibrations of the data which in some stages accompany the selections, are detailed below. Chapter 7 then explores the background estimation method and process of searching the spectrum for bump-like excesses. In Chapter 8, the sources of systematic uncertainties are defined and their effects calculated. Finally, in Chapter 9, the Bayesian limit-setting process is described and the resulting limits on the various models are shown.

6.1 Trigger strategy

The various terms which will be used in the detailed triggering discussion below are now defined.

fulfill trigger requirement

The event satisfies the criteria of the trigger (trigger chain).

prescale

The prescale factor p_i, commonly "prescale", defines the rate at which a trigger records events. Out of all events which would otherwise satisfy the trigger requirements, only 1 in every p_i, selected at random, will actually be recorded. The prescale factor is fixed for a given trigger and luminosity block but will often change over the course of a run to maintain a steady rate of data.

unprescaled trigger

A trigger which records every event (i.e. has $p_i = 1$).

passes trigger (fires trigger)

The event both fulfils the trigger requirement and was not discarded because of the prescale. These are the events which are actually recorded.

fully efficient trigger

A trigger is fully efficient at p_{T} values where, discounting the prescale, at least 99.5% of events which should pass the trigger indeed do.

effective luminosity

The luminosity equivalent to all the events collected by a trigger, which depends on the prescale at the time of recording. For example: if 1 pb^{-1} of data is collected with no prescale, and 10 pb^{-1} of data are collected with $p_i = 10$, the effective luminosity is 2 pb^{-1}.

average prescale

The ratio of total luminosity of data to the effective luminosity of the trigger is that trigger's average prescale. In the above example, this is $(11\mathrm{pb}^{-1})/(2\mathrm{pb}^{-1}) = 5.5$.

trigger weight

The weight given to an event in the analysis to account for the trigger prescales. The method for calculating this value will be described below.

normal stream

Data recorded by ATLAS to be promptly reconstructed. This is standard ATLAS data and would need no title at all; it is here defined only in opposition to the:

delayed stream

Data recorded but not immediately reconstructed due to computing constraints.

The first stage in the selection is to check that the event has passed an appropriate trigger. The analysis makes use of single-jet triggers: these select events that have at least one transverse energy deposit in the calorimeter with a magnitude larger than the threshold of the trigger in question. In order to maintain an acceptably low rate of data in all triggers, those with lower p_{T} thresholds have increasingly high *prescale factors* (p_i): only randomly selected events which would otherwise pass the trigger are recorded, at an overall rate of $1/p_i$.

A one-to-one map between the offline p_{T} of the leading jet and a single corresponding trigger is established. In order to maximise the statistics available to the analysis, this mapping employs the trigger with the lowest prescale which is fully efficient at a given p_{T}. An event is selected if the trigger corresponding to the leading jet p_{T}, or any other fully efficient trigger (selected from those with lower p_{T} thresholds in the trigger map), fired on the event. The event is then given a weight determined by the total recorded luminosity of that combination of triggers. This method of selecting and weighting events will be further detailed in the next section.

The statistical power of the m_{jj} spectrum depends not on the number of events which occurred (the smooth spectrum visible after reweighting events according to the recorded luminosity) but on the number of events recorded (the simple count of events which pass the triggers). As a result, the sensitivity of the search in a mass region can be greatly improved by access to more events recorded by a trigger with a lower prescale than would otherwise be available. The delayed stream, discussed in Section 4.7.1, contained data recorded by a trigger with $p_i = 1$ and a p_{T} threshold lower by nearly 130 GeV than the first trigger with $p_i = 1$ in the promptly reconstructed (*normal*) data. The addition of events recorded by the delayed stream drastically increased the sensitivity of the analysis in the corresponding m_{jj} range.

6.1.1 Dijet analysis trigger map

The analysis makes selections based on 11 event-filter level triggers in the promptly reconstructed dataset and 1 in the delayed stream. These are shown in Table 6.1, along with the effective luminosity collected by each trigger during 2012 data taking, and the leading jet p_{T} above which the trigger is considered fully efficient.

Single-jet trigger	Integrated luminosity (fb)	p_{T} at 99.5% efficiency (GeV)
`EF_j145_a4tchad_ht700_L2FS`	20.34	460
`EF_j280_a4tchad`	1.17	411
`EF_j220_a4tchad`	0.26	341
`EF_j180_a4tchad`	7.90E-02	273
`EF_j145_a4tchad`	3.64E-02	241
`EF_j110_a4tchad`	9.84E-03	185
`EF_j80_a4tchad`	2.32E-03	135
`EF_j55_a4tchad`	4.43E-04	99
`EF_j35_a4tchad`	4.54E-04	59
`EF_j25_a4tchad`	7.87E-05	47
`EF_j15_a4tchad`	1.48E-05	27
`EF_j220_a10tcem_delayed`	17.33	333

Table 6.1: Normal stream and delayed stream triggers used in the analysis. The second column reports the recorded integrated luminosity of each trigger (more precisely, trigger chain) while the third column indicates the p_{T} above which the trigger is fully efficient for a calibrated anti-k_t6 jet.

Effective luminosity of each trigger is calculated by the `iLumiCalc` tool, an ATLAS interface which computes the integrated luminosity for a given trigger and set of luminosity blocks from values stored in the COOL metadata database [179, 180]. The tool calculates an estimated integrated luminosity for each LB from the duration of the luminosity block and its estimated average instantaneous luminosity, then sums the results for all luminosity blocks during which the requested trigger was active. The returned value has an associated uncertainty due to the precision with which the instantaneous luminosity can be known and whose effects on the analysis are considered in Section 8.1. Any reported luminosity for a dataset in this analysis is the value reported by `iLumiCalc` for the appropriate GRL (luminosity blocks) and trigger.

The second field in each trigger name indicates the nominal p_{T} threshold in GeV: the trigger threshold used by the analysis is in each case much higher. This results from the fact that the jet triggers are cutting on EM-scale objects in the trigger while the objects considered at the analysis level have been calibrated to the hadronic scale, dramatically increasing their energies. For the reported thresholds, a standard tag-and-probe method used calibrated jets to determine the first bin in which each trigger is $> 99.5\%$ efficient:

1. An unbiased spectrum is produced for comparison by selecting events for which the leading jet fires a trigger with a lower threshold than the trigger to be examined.
2. If the leading jet also passed the requirements for the trigger to be examined, the event was added to a second spectrum
3. The ratio of the the two spectra shows the percentage of jets passing the unbiased trigger which also passed the trigger under examination. The first bin in which this ratio is $> 99.5\%$ is the first bin for which the trigger is considered fully efficient.

The result for each trigger is listed in the final column of Table 6.1.

6.1.2 Trigger combination and event weights

The weight of a single event is calculated using the "inclusion method" described in detail in [181]. For a jet with a given p_{T}, there will be in all cases except the very lowest-energy jets a range of triggers which are fully efficient for that jet. Although the mapping only assigns one trigger to the p_{T} of the jet, the random nature of the prescaling ensures that any trigger with a lower threshold may also have fired on the event. Similarly, the trigger nominally responsible for this p_{T} may not have fired. The probability of the event passing a trigger with lower threshold goes proportionally to the inverse of the trigger prescale, but many events will indeed fire a lower trigger than the intended. For a visual representation of this situation, see Figure 6.1. Only triggers which are fully efficient for the jet in question are considered, preventing the possibility of using an overly high weight if a different jet in the event actually fired the trigger.

When only a subset of events are recorded, each event must be given a weight determined by the rate at which identical events are rejected in order to create a smooth spectrum recreating the shape of the data which would be obtained with no prescales. The probability

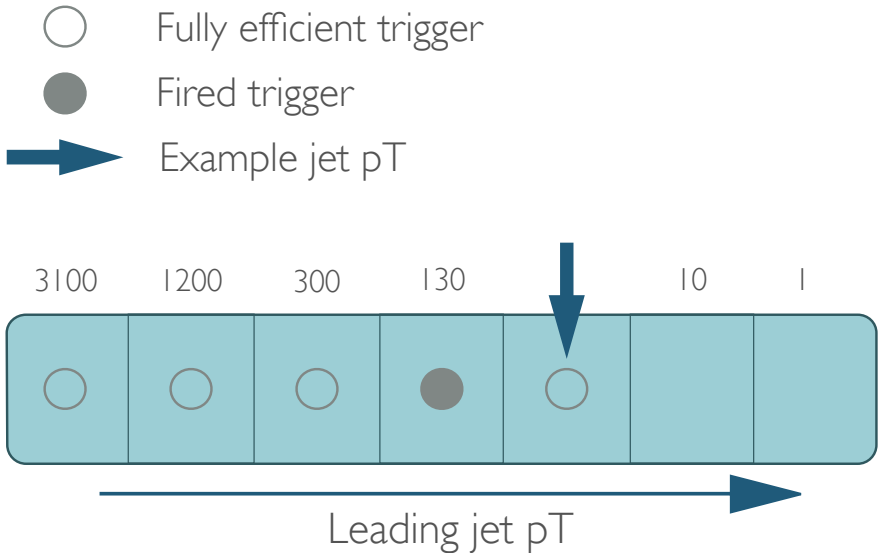


Figure 6.1: Schematic illustrating the relationship between event weights and the trigger map for the inclusion trigger method. Each box is the jet p_{T} region corresponding to one trigger in the map, where it is the lowest prescaled fully efficient trigger. The numbers above the box represent a rough prescale corresponding to the highest p_{T} triggers used in the analysis. If an event has a leading jet p_{T} indicated by the arrow, all triggers up to and including the range in which it falls are fully efficient, and any of them may have been fired by the event, with probability inversely proportional to the prescale. In this example, the first efficient trigger did not fire, but the second did. The weight assigned to the event will depend on all fully efficient triggers, since the probability of observing this event is the same regardless of which trigger truly recorded it.

that no trigger will fire on a jet is the combined probability of each fully efficient trigger missing it:

$$p_{\mathrm{trig}} = \Pi_i \left(1 - \frac{1}{\langle p_i \rangle}\right), \tag{6.1}$$

where the $\langle p_i \rangle$ are the average prescales of all the fully efficient triggers. The probability that at least one trigger accepts the event is then of course $1 - \langle p_{\mathrm{trig}} \rangle$.

Since two jets with the same p_{T} are indistinguishable to the analysis regardless of which trigger they fired, and since the probability of rejection also depends only on which triggers are fully efficient rather than on which ones fired, it should be expected that the assigned trigger weight will also reflect only this information. And indeed, the appropriate weight for each event passing a trigger in the map is the inverse of the probability of its doing so, and depends only on the efficient triggers for that jet:

$$w = \frac{1}{1 - \Pi_i \left(1 - \frac{1}{\langle p_i \rangle}\right)}. \tag{6.2}$$

Each event passing the analysis selection is given the above weight when added to the invariant mass spectrum. When an event with weight w_i is added to the bin, the bin error is adjusted accordingly to account for the fact that a single recorded event was responsible for the contribution: if δb is the error on a bin with content b and an event with weight w is added,

$$b_{new} = b + w \tag{6.3}$$

$$\delta b_{\text{new}} = \sqrt{(\delta b)^2 + w^2} \tag{6.4}$$

The statistical power of a bin in the spectrum is not equal to the sum of the weights but is rather a reflection of the number of events which actually contributed to the bin, a quantity now incorporated in the bin error. It has been shown that the sum of a set of weighted Poisson distributions can be described by a scaled Poisson distribution ([182] provides a summary). This distribution defines an equivalent raw number of events as the number of events which would contribute to a bin with the same weighted bin content $\sum w_i$ if every event had the same weight. Thus an effective number of events N_{eff}, to be interpreted as the statistical power in the bin, can be calculated as follows:

$$N_{eff} = \frac{N^2_{\text{weighted}}}{\sigma^2_{\text{weighted}}}, \tag{6.5}$$

where N_{weighted} is the content of a bin in the weighted distribution and σ_{weighted} is the uncertainty on this weighted bin content. Each of the N_{eff} effective events has an effective weight given by:

$$w_{eff} = \frac{\sigma^2_{\text{weighted}}}{N_{\text{weighted}}}. \tag{6.6}$$

This differs from w in Equations 6.3: rather than taking a different value for each event added it is the single average value that would be taken if all events had the same weight.

6.1.3 Comparison of alternative trigger methods

Several simpler triggering methods were initially considered, but were rejected in favour of the inclusion method because of the additional sensitivity it offers. The following two methods were also tested:

Division method

Only those events passing the single specific trigger corresponding to the appropriate p_{T} range in the trigger map will be kept. Each event is given a trigger weight $w = 1/\langle p \rangle$ in accordance with the fact that only a single trigger could have recorded it.

Two-trigger method

A simplified version of the inclusive method, but where only the two lowest-prescaled fully efficient triggers are considered. The event weight is identical to the inclusive method event weight, but the product runs only over these two triggers.

The effective number of events retained by each trigger method is shown in Figure 6.2. As is to be expected, the inclusive method retains the most raw events and produces the spectrum with the greatest statistical sensitivity. The difference is evident at low dijet invariant mass, where the trigger prescales have the greatest effect, while the three methods coincide for $m_{jj} > 1$ TeV.

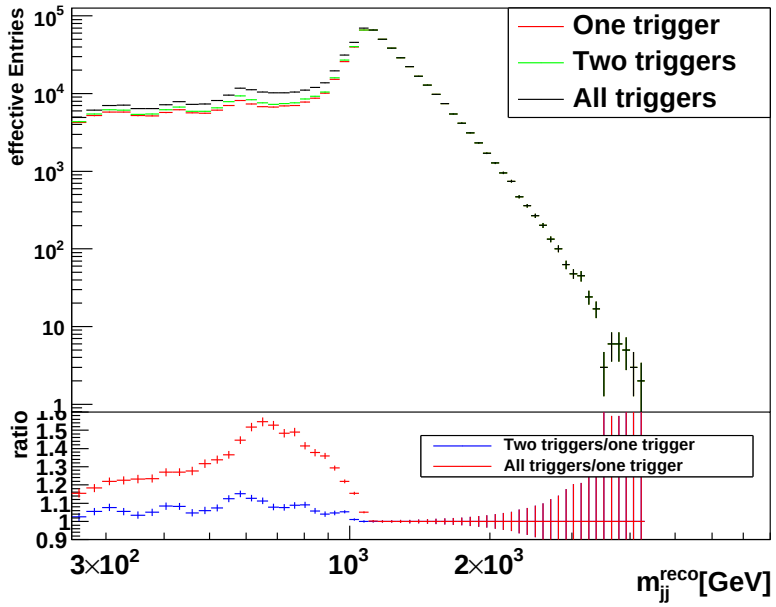


Figure 6.2: Comparison of effective statistics in the m_{jj}spectrum when it is generated using the inclusion method, the division method (labeled "One trigger") and the two trigger method.

6.1.4 Incorporating the delayed stream

In order to combine events recorded by the delayed stream trigger `EF_j220_a10tcem_delayed` with the normal dataset, any possible overlaps must be accounted for. The delayed

stream was only active for part of the time during which normal data was being recorded, but when both were active, any events accepted by the normal stream were also collected by the delayed trigger. In order to ensure that no events were double-counted, the full dataset was split into three independent categories such that each recorded event would fall into exactly one. They will be distinguished by the following terms:

normal-only stream

Data recorded in the normal stream at a time when the delayed stream was inactive. All events in this category must be unique to the normal stream, as the delayed triggers were not recording.

delayed-only stream

All data recorded by the delayed stream triggers.

overlap stream

Events which were measured while the delayed stream was active, but for which the leading jet $p_{\mathrm{T}} < 333$ GeV.

A custom good run list was made for each of the three previous categories of data, and the effective prescales of the triggers in each category were calculated from the luminosity blocks specified in the GRLs. With the prescales known, event weights for each trigger active in a category were then computed using Equation 6.2. The next stage, the combination of the three streams into a single spectrum, presents more possibilities.

Several options are available for forming a combined spectrum, depending on the interpretation of the different streams. First, the three streams could be kept separate and analysed independently. However, this would have required a substantial change in the statistical analysis procedure, since it relies on a single background description from a smooth fit over the full mass range and the three spectra are highly different shapes, with the overlap spectrum only existing below 333 GeV, such that a single background description would not be possible. No such method was attempted.

Second, one could assume that all events recorded form a single measurement of the dijet spectrum. The calculation of event weights from trigger prescales makes the assumption that the sum of all weighted events provides a measurement of the shape of the dijet spectrum with 20 fb^{-1} of data. Every event is required to achieve this spectrum, and their weights are

simply summed in each bin. The delayed stream cannot be left out of this "sum method" calculation since its presence is assumed in calculating the effective trigger prescales. This interpretation has the drawback that N_{eff} in a bin increases very slowly with additional low-prescale events. In fact, a higher-precision measurement in a single bin would be obtained by ignoring contributions from very highly prescaled triggers altogether. Let an example be the combination of two toy datasets with the following average weights and effective number of events in a bin: $N^1_{eff} = 100, w^1_{eff} = 10$ and $N^2_{eff} = 10000, w^2_{eff} = 1$. Following the procedure for combining weighted events,

$$\begin{aligned} \text{Bin content} &= 10000 * 1 + 100 * 10 = 11000 \\ \text{Error} &= \sqrt{10000 * 1^2 + 100 * 10^2} = 141.42 \\ \text{and thus} \qquad N_{eff} &= 6050 \\ w_{eff} &= 1.8\,. \end{aligned}$$

The effective number of events would be improved by leaving out the second dataset altogether. This option was present in the analysis, where the normal data stream could be abandoned for masses where the delayed stream provided better sensitivity, but the choice to discard data is clearly sub-optimal. It is also an indication that the simple addition method is not appropriate, if it favours abandoning data.

Another possibility is to treat the normal stream and the overlap plus delayed-only streams as two separate, independent measurements of the spectrum shape at 20 fb^{-1}, where one measurement has higher statistical precision than the other. When two measurements arise from different probability distributions with known variances σ_i, the weighted average and its variance are given by:

$$\bar{x} = \frac{\sum_i x_i \sigma_i^{-2}}{\sum_i \sigma_i^{-2}}, \qquad \sigma(\bar{x})^2 = \frac{1}{\sum_i \sigma_i^{-2}}\,. \tag{6.7}$$

Returning to the toy model, experiment 1 accounts for 10000/11000 of the total measurement. 10000 events with weight 1.1 produce a measurement with bin content 11000 and error 110. Experiment 2 accounts for 1000/11000 of the measurement, giving bin content 11000 and bin error 1100. The final estimate in this bin is the weighted average:

$$\bar{x} = \frac{\sum_i x_i \sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \qquad \sigma(\bar{x})^2 = \frac{1}{\sum_i \sigma_i^{-2}}$$
$$= \frac{11000/(110)^2 + 11000/(1100)^2}{110^{-2} + 1100^{-2}} \qquad \sigma(\bar{x}) = \frac{1}{\sqrt{110^{-2} + 1100^{-2}}}$$
$$= 11000 \qquad = 109.45$$

This treatment provides a gain of 1.67 compared to the simple sum above. This will be the "weighted average" method, and is the one chosen by the analysis.

Extending this concept to the dijet spectrum, the physical quantity of interest is the *event rate*, the number of dijet events observed per unit of luminosity collected. The separate datasets, each now treated as corresponding to some partial luminosity $\mathcal{L}_i$, measure to varying precision the number of dijet events per fb^{-1}. Let N_{eff} and w_{eff} be defined as before and $\mathcal{L}_i$ be obtained from iLumiCalc for the triggers and GRL of each dataset. $N_{eff,i} = \sigma_{i,eff}^2$ by definition, since the effective number of events represents the true statistical power of the sample. Then the procedure for obtaining the weighted average is:

- Calculate rates in each sample: $rate = N_{i,eff}/\mathcal{L}_i$, $\sigma_{rate} = \sigma_{i,eff}/\mathcal{L}_i$
- Calculate weights: $w_i = 1/\sigma_i^2$
- Calculate weighted average from rates using 6.7
- The weighted average is the effective rate of the combined dataset ($\mathcal{L} = \sum_i \mathcal{L}_i$)
- Scale rate and error by total luminosity to obtain the final estimate.

The power of the method is demonstrated on a bin from a sample dijet spectrum in Table 6.2, where the statistical sensitivity improves by a factor of 5.7.

Two issues arises with the naïve implementation of the weighted average as laid out above. First, the mean rate is not defined for bins with zero content. Second, Poisson-distributed data is not well-described by Gaussian (χ^2) averaging. The solution is to use a slight modification of the strategy. A result equivalent in principle but maintaining the Poissonian nature of the event counts is found by introducing an effective luminosity. For a spectrum i with equivalent total luminosity $\mathcal{L}^i$ in a bin with content b, $\mathcal{L}^i_{eff} = \mathcal{L}^i * N^i_{eff}/b^i$. That is, the effective luminosity is the luminosity to which the dataset would be equivalent if all events had weight 1. With this conversion in mind, events from all constituent datasets are equivalent at $w_i = 1$ and can be directly added to represent a dataset equivalent to a

Quantity	Value	
	Normal stream	Overlap+Delay stream
Raw events	27369	207287
Bin content	131152	207287
Bin error	1637.54	455.288
	Sum method	
Bin content	338439	
Bin error	1699.65	
N_{eff}	39649.6	
	Weighted average method	
Luminosity (fb^{-1})	7.78	12.49
Rate	16.86 ± 0.21 pb	16.58 ± 0.036 pb
New weight	22.68	771.6
Mean rate	16.59 events/pb^{-1}	
Error on mean	0.035 events/pb^{-1}	
Bin content	337500	
N_{eff}	224676	

Table 6.2: Comparison of statistical power between sum method and weighted average method of stream combination for a sample bin. The gain in statistical sensitivity between the two methods is $N_{eff}^{mean}/N_{eff}^{sum} = 5.7$.

total effective luminosity:

$$N_{eff}^{tot} = \sum_i N_{eff}^i \,, \tag{6.8}$$

$$\mathcal{L}_{eff}^{tot} = \sum_i \mathcal{L}_{eff}^i \,. \tag{6.9}$$

$$\tag{6.10}$$

These values can then be scaled up by the ratio of total to effective luminosity, giving a final bin content $B^i = \mathcal{L}_{eff}^{tot} \cdot (\sum \mathcal{L}^i)/(\sum \mathcal{L}_{eff}^i)$: that is, the effective luminosity of the sample scaled by the ratio of the total luminosity of the ATLAS dataset to the sum of the effective luminosities. As in the definition of N_{eff}, this acts as a mean, defining the case where all events are equivalent to some average value. However, by avoiding a Gaussian calculation and employing a sum instead, the Poisson behaviour is preserved. Divergences can be avoided by giving bins with zero content a non-zero error, but the result is the same regardless, with the

zero bin content maintained in the final spectrum. Unlike the naïve weighted mean method, this result is identical to the addition method in bins with no prescale and retains the integer bin contents critical to a Poissonian statistical analysis of the spectrum.

Figure 6.3 compares the statistical power of the final spectrum in these cases:

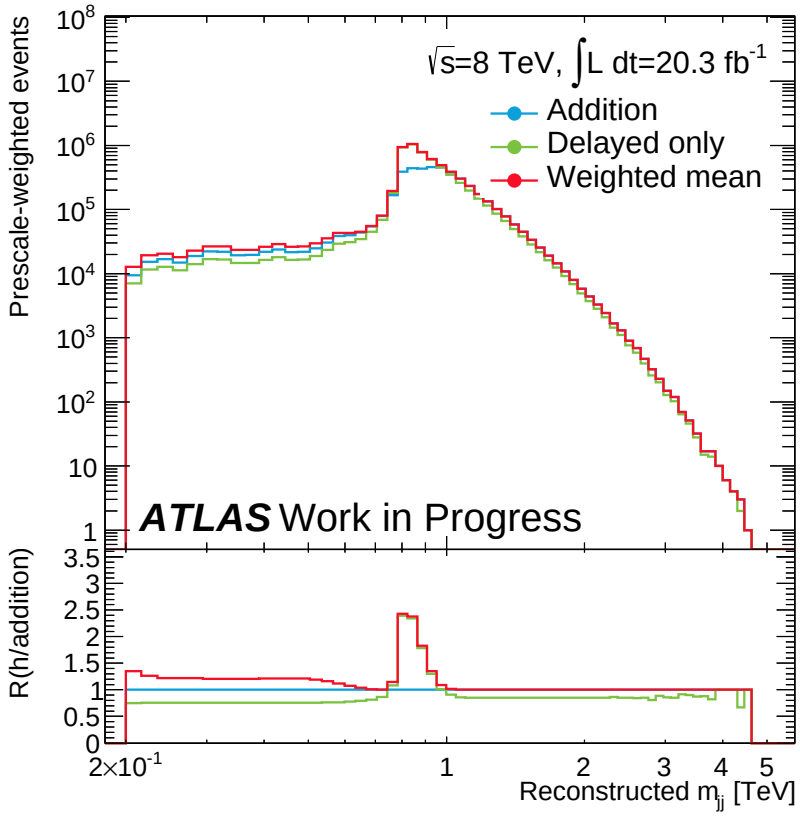


Figure 6.3: Comparison of effective statistics in the final m_{jj} spectrum when the normal and delayed components are treated as integral parts of the same measurement (red line) versus when they are considered independent measurements and combined (purple line). The third proposed option, considering only the delayed stream, is illustrated in green.

Time-dependent inefficiencies in the dijet spectrum complicate the combination. In this analysis, an inefficiency due to removal of tripped tiles, discussed in the next chapter, affects the data streams differently since the delayed stream was not active during all data-taking periods. This means the two spectra do not correspond to the same total luminosity. To address this, a correction for the inefficiency was applied to the luminosity on a bin-by-bin level, returning both spectra to the same scale. The procedure for combining spectra was modified as follows.

1. Derive a bin-by-bin correction for any time-dependent losses in efficiency. For details

on the relevant inefficiency measurement and correction, see Section 6.3.3.

2. Apply the correction to the datasets and cross-check them for consistency. The ratio between the normal and delayed spectra should be flat if the time-dependent differences have been removed. This condition held for the correction factors used in this analysis.
3. After a combination using the efficiency-corrected luminosities, scale back to the full data luminosity.

6.2 Event selection criteria

The first cut on events entering the analysis required that the leading jet passed the trigger map described above. Several further requirements on the basic criteria of the event were made before the jets were examined.

6.2.1 Good run list

The standard method of ensuring good data quality for all events included in an analysis makes use of a *good run list* (GRL). As briefly introduced in Section 4.9, these lists are created and released by the ATLAS Data Quality (DQ) group and are based on the data quality flags released by the various detector subsystems. DQ flags operate on a stoplight system (good, flawed, bad) and an individual luminosity block may have a good DQ flag for one process or subsystem but be labeled bad by another. Because different physics events rely on different systems, DQ flags can be used to assemble separate collections of good data depending on the main physics objects in question: electrons, jets, etc. The most general good run lists will ensure that only universally good data is included and will be centrally applicable to all physics analyses. Each GRL contains a list of run and luminosity block combinations which match the DQ requirements specified at their generation. An event is said to pass the GRL when it is included in one of the selected luminosity blocks. Official GRLs are approved by Physics Coordination and recommended for analysis use [148]. The dijet analysis uses the following offical, object-nonspecific GRL: `data12_8TeV.periodAllYear_DetStatus-v61-pro14-02_DQDefects-00-01-00_PHYS_StandardGRL_All_Good.xml`

ATHENA also includes tools which enable users to generate new lists by taking the overlap or difference of existing GRLS. Analysers can also generate their own GRLs from existing data. These methods were used to produce custom GRLs as subsets of the selected offi-

cial GRL and corresponding to the normal, overlap, and delayed streams. First the event selection code with the above trigger map was run over the normal and delayed streams separately, using the above-listed centrally produced GRL. A custom GRL was then generated from the events kept in each stream. Those luminosity blocks which were only recorded in the normal stream were selected using the difference tool and form the GRL which defines the normal-only stream. Those which were recorded in both were used for the overlap and delayed-only streams, as each individual event would only qualify for one of those categories depending on the p_{T} of the leading jet. Since the initial run over the data included a selection on the centrally-produced GRL, only run numbers and luminosity blocks which passed the official GRL were recorded in the secondary ones.

6.2.2 Data integrity checks

Removal of bad and corrupted events followed the central recommendations as laid out in [183]. Each event stores various flags indicating the reliability of the detector information and data reconstruction. Of concern for the dijet analysis are any flags related to the calorimeter, which identify possible issues with obtaining reliable jet information. The `larError` flag indicates the presence of noise or loss of data integrity in the liquid argon calorimeter: `larError` = 0 is a good event, `larError` = 1 contains noise in the calorimeter and `larError` = 2 indicates a data integrity problem. Noise is considered acceptable but events with data integrity problems are vetoed in accordance with the recommendations found in [184]. A smaller number of events were corrupted in the Tile calorimeter instead, and can be vetoed by eliminating any events with `tileError` = 2.

A TTC (Timing, Trigger, and Control system [185]) restart is a procedure used as a last resort to recover data-taking capacity without stopping a run. It consists of putting the trigger system on hold while completely restarting and reconfiguring some malfunctioning portion of the detector. This allows recovery of the data-taking during a run, but often leads to incomplete information in the luminosity block immediately following the TTC restart. Any such incomplete events are flagged and vetoed.

Finally, Tile calorimeter modules occasionally suffer power *trips*, where the low-voltage power supply to the module is briefly cut off. The effective loss of data quality is small, but in 2012 data it is standard practice to remove them and is the more conservative course of

action. For the dijet analysis, then, events overlapping with trips are identified and vetoed using the `TileTripReader` tool [186].

6.3 Jet calibration and selection

The dijet analysis uses anti-k_t jets with radius $R = 0.6$. Although $R = 0.4$ jets are standard in ATLAS, the broader jets have been historically used by the dijet analysis to ensure that none of the jet radiation is lost as a result of small jet radius. A comparison of the two jet sizes is performed by calculating the sensitivity of the data spectrum to reconstructed peaks in simulated q^* events and comparing the results between the two jet sizes. Sensitivity is defined as a signal-to-noise ratio:

$$S = \frac{s}{\sqrt{\delta s + \delta b}}, \tag{6.11}$$

where s and b are signal and background events in each bin and δs and δb are their respective uncertainties. It was found that the performance of the two jet radii was almost identical, and if anything the sensitivity was slightly higher for the $R = 0.6$ jets. Thus with performance dictating no preference and a choice between adhering to ATLAS common practice versus remaining consistent with the previous Run I dijet analyses, it was decided that the larger jets would be used to facilitate the comparison of this result with earlier publications.

For any event to be considered in the analysis, it must contain at least two jets with $y < 2.8$. This $|\eta|$ requirement is an ATLAS standard, ensuring that both jets fall within the fully sensitive region of the detector. Events with two $R = 0.6$ jets satisfying this angular requirement are then further subjected to the calibration, cleaning, and jet-based cuts described below.

6.3.1 JES and JER calibration

All methods used in the calibration of the jets have been described in detail in Chapter 5. Initially, the jets are calibrated using the EM+JES scheme. The calibration process is then completed using the described *in situ* methods: the $|\eta|$ response is equalised using dijet momentum balancing while the energy response is calibrated using momentum balancing between γ/Z and jets and between jets of differing p_{T}. The $\Delta R = 0.6$ jets, after calibration to the hadronic energy scale, were measured to have an energy scale uncertainty of less than

3.3% across all p_{T} and η values accessible to the dijet analysis.

6.3.2 Jet cleaning

All jets retained in the final spectrum are subjected to quality tests in order to reject those jets which come from non-collision backgrounds instead of from the hard scattering process the analysis wishes to probe. These include jets arising from various secondary interactions of the beam, cosmic ray muons, and random calorimeter noise. A variety of sets of requirements have been developed by ATLAS using calibrated $R = 0.4$ anti-k_t jets, allowing analyses to select the requirements appropriate for their event topologies and efficiency requirements. These are defined in [151] and [187]; only the details relevant to this analysis will be elaborated here.

Most excess calorimeter noise is removed during data quality checks, but some will remain and must be removed by further requirements. The quantities which are most helpful in distinguishing calorimeter noise from real calorimeter hits are as follows:

$\boldsymbol{f_{\mathrm{HEC}}}$ Fraction of the jet's energy which was deposited in the hadronic calorimeter.

$\boldsymbol{f_{\mathrm{EM}}}$ Fraction of the jet's energy which was deposited in the electromagnetic calorimeter.

$\boldsymbol{Q_{cell}^{\mathrm{LAr}}}$ Quadratic difference between actual and expected pulse shape in a liquid argon cell.

$\boldsymbol{f_{\mathrm{Q}}^{\mathrm{LAr}}}$ Fraction of energy deposited in the liquid argon calorimeter for which signal shape quality was poor ($Q_{cell}^{\mathrm{LAr}} > 4000$).

$\boldsymbol{f_{\mathrm{Q}}^{\mathrm{HEC}}}$ Fraction of energy deposited in the hadronic calorimeter for which signal shape quality was poor ($Q_{cell}^{\mathrm{HEC}} > 4000$).

$\boldsymbol{\langle Q \rangle}$ Average jet quality. This is the normalised energy-weighted average of Q_{cell}^{LAr} for calorimeter cells in the jet.

The spurious calorimeter energy will be in the form of either *sporadic noise bursts* or *coherent noise*. Sporadic noise bursts occur in the HEC and are caused by only a few cells which make up most of the jet energy. These cases can be identified by large values of f_{HEC}, $f_{\mathrm{Q}}^{\mathrm{HEC}}$, and $\langle Q \rangle$. Nearby cells will display an apparent negative energy E_{neg} due to capacitive coupling between the channels. Coherent noise, in contrast, occurs in the electromagnetic calorimeter causing cells with high reconstructed energy but low reconstructed quality and is characterised by large values of f_{EM}, $f_{\mathrm{Q}}^{\mathrm{LAr}}$, and $\langle Q \rangle$.

Cosmic rays and beam background are real particles but arise from the wrong source, and so can be expected to be out of time with the actual collision event. The following variables are helpful in characterising such jets:

t_{jet} Jet time, defined as the average of the times corresponding to the energy deposits in the jet weighted by the square of the cell energies.

f_{CH} Jet charged fraction, $\sum |p_{\mathrm{T}}^{\mathrm{track}}|/p_{\mathrm{T}}^{\mathrm{jet}}$

The jet time, defined with respect to the event time defined by the trigger, is the most sensitive to these backgrounds which arise independently from the event. Since a jet from a collision source would leave energy in the innermost layers of the calorimeter, f_{EM} is also a meaningful distinguishing variable. This is combined with a selection on f_{CH}, as real jets would normally leave tracks [188].

All of these criteria combine to identify *bad* jets according to a set of criteria varying in tightness. The two loosest selections used in 2011 and 2012 data are defined in Table 6.3 [189, 190].

Radiation targeted	`Looser`	`Loose`
Sporadic noise bursts (HEC)	($f_{\mathrm{HEC}} > 0.5$ and $\lvert f_{\mathrm{Q}}^{\mathrm{HEC}}\rvert > 0.5$ and $\langle Q\rangle > 0.8$) or $\lvert E_{\mathrm{neg}\rvert} > 60$ GeV	`Looser`, or ($f_{\mathrm{HEC}} > 0.5$ and $\lvert f_{\mathrm{Q}}^{\mathrm{HEC}}\rvert > 0.5$
Coherent noise (EM)	$f_{\mathrm{EM}} > 0.95$ and $\lvert f_{\mathrm{Q}}^{\mathrm{LAr}}\rvert > 0.8$ and $\langle Q\rangle > 0.8$ and $\lvert\eta\rvert < 2.8$	`Looser`, or $f_{\mathrm{EM}} > 0.95$ and $\lvert f_{\mathrm{Q}}^{\mathrm{LAr}}\rvert > 0.8$ and $\lvert\eta\rvert < 2.8$
Cosmics & beam radiation	($f_{\mathrm{max}} > 0.99$ and $\lvert\eta\rvert < 2$) or ($f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{CH}} < 0.05$ and $\lvert\eta\rvert < 2$) or ($f_{\mathrm{EM}} < 0.05$ and $\lvert\eta\rvert \geq 2$)	`Looser`, or $\lvert t_{\mathrm{jet}}\rvert > 25$ ns

Table 6.3: `Looser` and `loose` jet quality criteria used for the rejection of bad jets in 2012 data.

The dijet analysis uses the ATLAS recommendation and defines bad jets as those which pass the `looser` criteria. In 2011 data, the `looser` selection criteria were 99.8% efficient for all jets with $p_{\mathrm{T}} > 20$ GeV and had a rejection rate of 37.8% for jets with $p_{\mathrm{T}} > 150$ GeV, rising to 68.6% for jets with $p_{\mathrm{T}} > 500$ GeV. Both the leading and subleading jet are required to pass the `looser` bad jet criteria.

Ugly jets are those which arise from a real source but which fall into a geographical location such that they will be badly reconstructed. These correspond to jets landing in the region between the barrel and end-cap of the calorimeter and jets which fall in a *masked* calorimeter region and are as a result badly reconstructed. Ugly jets are defined by two variables: the size of the `BCH_CORR_CELL` correction (discussed in detail in the next section), for which a large value indicates a poorly-reconstructed jet near a masked tile; and the location of the jet with respect to the tile calorimeter gap. Because ugly jets are those for which the given energy was poorly reconstructed, it was not sufficient to require that neither of the two leading jets be ugly. In addition, no other jet in the event with $p_{\mathrm{T}} > 0.30 \cdot p_{\mathrm{T}}^{subl}$ must be ugly, to protect against cases where the true leading or subleading jet swapped with a third jet due to having an underestimated energy.

When analysing Monte Carlo events, a single cleaning cut was used: $0.0 < p_{\mathrm{T}}^{\mathrm{avg}}/p_{\mathrm{T}}^{\mathrm{lead,\ truth}} < 1.4$, where $p_{\mathrm{T}}^{\mathrm{avg}}$ is the mean p_{T} of the leading and subleading jets. This cut suppresses events which migrate to higher bins through pileup effects and, due to high MC event weights, dominate the spectrum and distort its shape [191]. Ugly jet removal should also have been applied in MC as it depends on jet location: this oversight is being remedied in the 13 TeV search.

6.3.3 `BCH_CORR_CELL` cuts

Certain modules in the barrel of the tile calorimeter were masked due to the failure of their power supplies during the 2012 data-taking period, and many individual cells were also masked temporarily or permanently. Only one module was masked during the entire year: LBA05 in the barrel of the tile calorimeter, covering the $\eta - \phi$ region $0.0 \leq \eta \leq 0.8$ and $\frac{4\pi}{32} \leq \phi \leq \frac{5\pi}{32}$. An additional substantial effect comes from module LBC16, which was deactivated from June 2012 until the end of Run I. A list of masked modules and the percentage of 2012 data lost in each is shown in Table 6.4.

The `BCH_CORR_CELL` correction, automatically applied at the calorimeter cluster level for all data, is intended to compensate for the masked modules. This correction approximates the energy of a masked cell by assigning it the average of the energies of all its neighbouring cells. The correction is automatically performed during cell reconstruction before jet formation or hadronic energy recalibration [192].

Module	Coordinates (with edges)	Masked from	$\mathcal{L}$ [fb^{-1}]	Fraction
LBA05	-0.1 $\leq \eta <$ 1.0, 2.95 $\leq \phi <$ 5.89	200805	20.28	100%
LBA16	-1.0 $\leq \eta <$ 0.1, 1.37 $\leq \phi <$ 1.67	204265	16.72	82%
LBA18	-0.1 $\leq \eta <$ 1.0, 1.57 $\leq \phi <$ 1.87	211522	7.25	36%
LBA35	-0.1 $\leq \eta <$ 1.0, - 3.04 $\leq \phi <$ - 2.74	212619	5.95	29%
EBC01	0.8 $\leq \eta <$ 1.7, - 0.98 $\leq \phi <$ 0.196	Intermittent	0.82	4%
Others	–	Intermittent	< 0.1	<0.05%

Table 6.4: Locations and luminosities of significant masked modules during 2012 data taking. The second column is the dimensions of the module and the cells bordering it. Each module not labeled "intermittent" is masked beginning from the run in the third column until the end of 2012. The final two columns show the integrated luminosity collected by each cell and the equivalent percentage loss in 2012 data.

It was observed in 2012 that the applied `BCH_CORR_CELL` correction overestimates the total energy of those jets which fall on the edges of the masked regions while underestimating those in the middle. When the highest energy deposit from a jet falls in the masked cell, the residual energy around the edges employed for calculating the correction is unrepresentatively small, while on the other hand if the highest-energy cell is at the edge of the masked cell and only residual energy falls into it, the averaged energy is too large. The more collimated the jet, the more pronounced the effect: hence, it is most troublesome for high-p_{T} jets such as those employed in the dijet analysis. The over- and under-corrections are clearly visible in Figure 6.4, which illustrates the average magnitude of the applied `BCH_CORR_CELL` correction for jets over the full $\eta - \phi$ range.

The effect of the over/under-correction is different for the leading and subleading jets. This is in part because a jet with higher p_{T} will suffer a greater overcorrection, and to some extent also because a sufficiently overcorrected jet can become the leading jet by virtue of that alone, when it would otherwise be subleading. Near the edges of cells, larger values of the `BCH_CORR_CELL` correction correspond to observably larger asymmetries between the p_{T} of the leading and subleading jets. At the centre of masked cells the situation is more complex: with a sufficiently large `BCH_CORR_CELL` correction the asymmetry will again increase, but with the smaller corrections typical of jets in the masked region, the undercorrection can be sufficiently severe that the asymmetry will be negative.

This asymmetry can be used to measure the size of the error arising from the `BCH_CORR_`

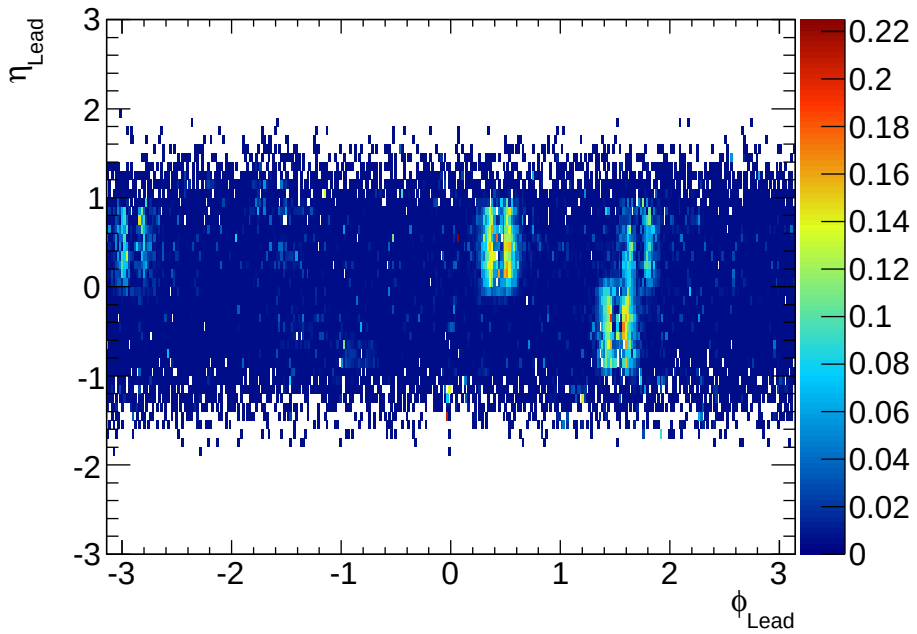


Figure 6.4: Average size of the BCH_CORR_CELL correction applied to jets in the Tile calorimater, mapped over η and ϕ. The undercorrection at the centers of the dead modules and overcorrections near their edges are clearly visible.

CELL correction by employing the two jets in a tag-and-probe fashion. When one jet, either leading or subleading, falls in a masked region, it is considered the probe, while the jet which balances it is the tag. An asymmetry measure A_{p_T} is then defined as:

$$A_{p_T} = 2 \cdot \frac{p_T^{probe} - p_T^{tag}}{p_T^{probe} + p_T^{tag}} . \tag{6.12}$$

The values of the asymmetry would in an ideal case be normally distributed around zero; a measure of the overcorrection effect can therefore be obtained by comparing the mean values of the asymmetry distribution for all events where the probe jet falls on the edge of a given masked module to those events where it falls in the centre. The comparison of these distributions reveals that the difference in means can be as high as 13%. The results of the asymmetry measurement for the LBC16 cell are shown in Figure 6.5.

The most reliable method found to eliminate this effect is to remove the affected jets altogether. Several alternate methods were also suggested and tested, including adding an uncertainty to the jet p_T, applying a further correction to the affected jets to counteract the response, removing the BCH_CORR_CELL correction altogether, or vetoing events with either a large value of the BCH_CORR_CELL correction or with very large p_T asymmetries. These were

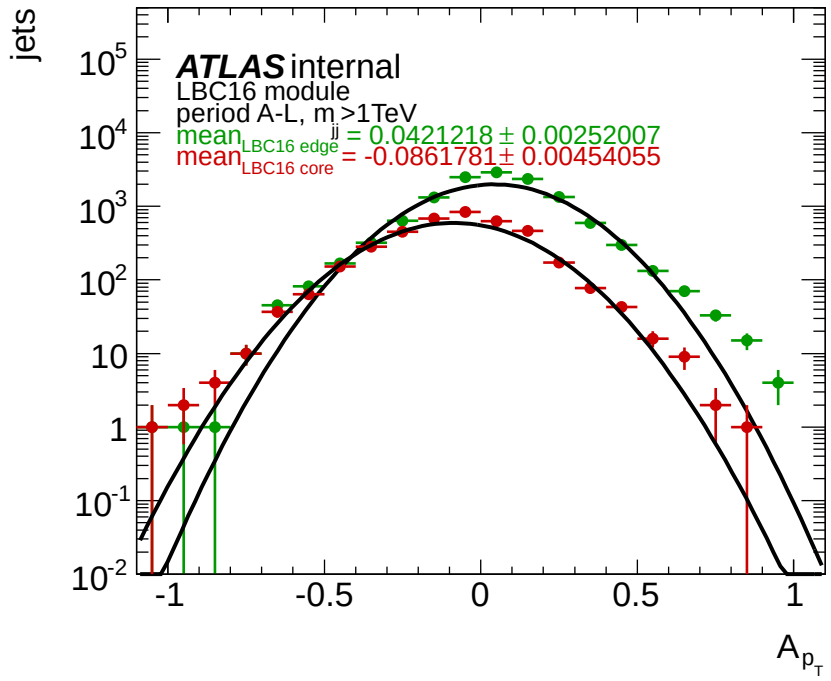


Figure 6.5: Asymmetry between tag and probe jet p_{T} for cases with one jet in the vicinity of the LBC16 Tile calorimeter cell. The difference in means of the asymmetry between cases where the tag jet falls on the edge of the masked region versus in its centre gives an indication of the size of the error introduced by the over- and under- corrections. In this case, the means differ by 13%.

all found to be either technically unnecessarily challenging or else could not be relied upon not to introduce any unexpected features into the spectrum. As a result, it was decided to simply remove the miscalibrated events.

A tool has been developed to perform this cleaning of the affected jets [193]. It defines a border region surrounding a masked module as the modules immediately bordering it in ϕ and as a border of size $\eta = 0.1$ (0.2) for modules in the Tile barrel (extended barrel) region. Two methods have been developed to define whether a jet is considered in need of removal.

`IsBadMediumBCH`

: This depends on both the location of a jet and the size of the `BCH_CORR_CELL` correction applied to it. Any jet falling in the core of a masked region will be flagged. Jets which fall in the border of a masked region are only flagged if their values of `emfrac` and `BCH_CORR_CELL` are such that the jet p_{T} is likely to have been overestimated: jets are vetoed where $(2 - \texttt{emfrac})/10 < \texttt{BCH_CORR_CELL} < 1 - 2.5 \cdot \texttt{emfrac}$.

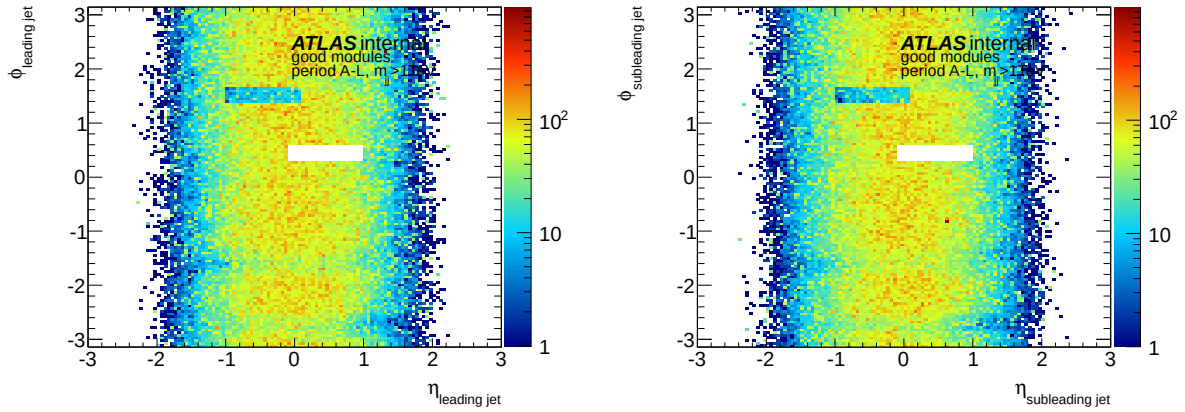
`IsBadTightBCH`

Strictly geometrical, this value rejects any jet which falls in either the core of a masked region or the border surrounding it.

The dijet analysis opted to use the more stringent criteria and reject jets according to the `IsBadTightBCH` criteria. A combination of the `BCHCleaningTool` and `TileTripReader` packages is used to determine whether one of the jets is pointing to a masked region [193,194]. A criterion for vetoing events was sought such that not only would the leading and subleading jets fall outside of masked regions, but also so that one could guarantee that there was not a "true" subleading jet which had fallen into a masked region, been undercorrected, and as a result dropped below the third-leading jet in p_{T}. Truth studies exploring the jet response in MC show that the worst underestimate of p_{T} gives a jet response of 0.5. Therefore, the analysis cuts require that neither the leading nor subleading jet, nor any other jet with $p_{\mathrm{T}} \geq 0.30 \cdot p_{\mathrm{T}}^{subl}$, should fail the `IsBadTightBCH` criteria.

This cut introduces a complication for signal MC, which was produced without knowledge of any masked tiles except for module LBA05. To ensure compatibility between signal MC and the background distribution, events were removed during signal processing to mimic the effects of the `BCH_CORR_CELL` cut. Since the correction varies with jet p_{T} the effects could not be accounted for by simply scaling by the inefficiency. Instead a geometrical cut was applied on jets falling within the same cells in the simulation. The time-dependence of the masked tiles in data was imitated by randomly selecting run and luminosity block numbers from the GRL and passing them to the `TileTripReader` along with the geometrical information on the jets. Thus the proportion of events for which a given tile was masked in data would be reproduced in MC, although on a random selection of events.

Maps of the $\eta - \phi$ distributions of the leading and subleading jets after this cut are given in Figure 6.6. The entirely empty space corresponds to LBA05, masked at all times in 2012 data taking, and the lighter patch corresponds to LBC16, which was masked for the second-longest duration. The average `BCH_CORR_CELL` correction values mapped to $\eta - \phi$ locations are shown after the cut in Figure 6.7. No hotspots are now visible: only the empty regions where jets have been removed geometrically can be distinguished.



(a) Leading jet (b) Subleading jet

Figure 6.6: Map of jet locations in η and ϕ after geometric removal of all jets falling in the centres or edges of masked tile regions. Leading jet locations in 6.6a; subleading jet locations in 6.6b

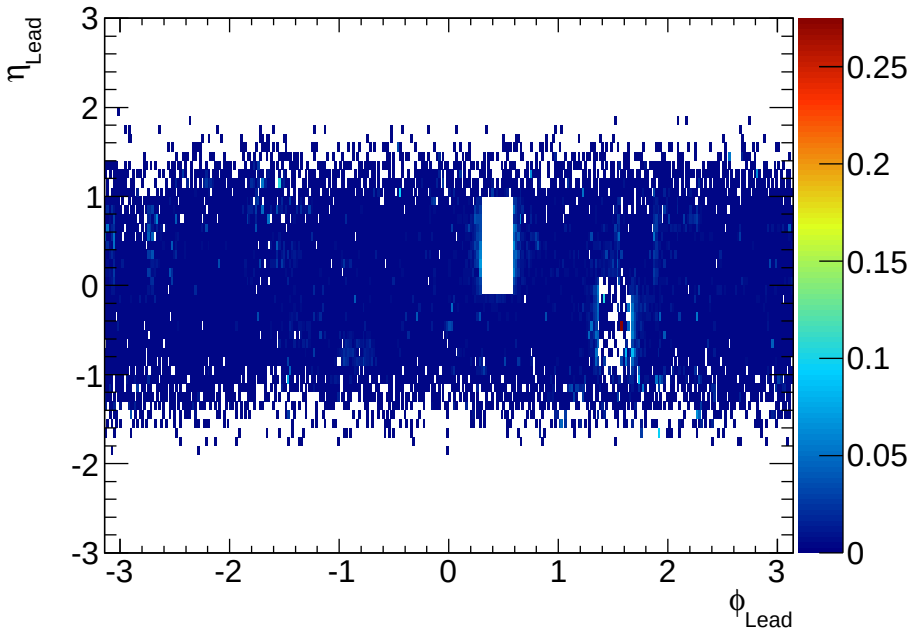


Figure 6.7: Average value of `BCH_CORR_CELL` correction mapped over η, ϕ for all jets remaining after geometric removal of those in or near masked tiles. The missing regions corresponding to LBA05 and LBC16 are visible, but no unusually large or small values of the correction remain.

6.3.4 Additional hot tile removal

One calorimeter tile was dead during several runs of data taking but was not masked in reconstruction, and as a result affected jets must be removed by hand. For the affected runs (202660, 202668, 202712, 202740, 202965, 202987, 202991, 203027, 203169), any jet in the

region $-0.2 < \eta < -0.1$, $2.65 < \phi < 2.75$ is considered affected if a large fraction of its energy was deposited in the second layer of the calorimeter ($E_2/\sum E_{\text{layers}} > 0.6$).

6.3.5 Resistance to pileup

Pileup corresponds directly to the number of primary vertices in an event, so the relationship between N_{PV} and the various kinematic variables relevant to the spectrum (p_{T}^{lead}, p_{T}^{subl}, m_{jj}, ...) will indicate the degree to which pileup is affecting the measurement. There are three ways in which pileup in an event can affect the measured value of m_{jj}:

1. The energy of the jets is increased, so m_{jj} is larger. The value of m_{jj} and the individual jet transverse momenta would depend on N_{PV}.
2. Extra jets can be added to the event, replacing the leading or subleading jet. The jet multiplicity in the event would depend on m_{jj}.
3. A third jet can have its energy sufficiently modified by pileup that it spuriously becomes the leading or subleading jet. This can be checked in Monte Carlo by calculating the fraction of leading and subleading jets which can be truth-matched to the true primary vertex. Some mismatching will always occur as a result of detector reconstruction inefficiency, but as a result the percentage of failed matches can be used to set an upper boundary on the effects of pileup, even if not a direct value for that effect.

Modifications in individual jet energy, as in the first case, are addressed during the jet calibration process and so have already been accounted for. The second and third cases become most relevant when the leading or subleading jet is of very low energy. Since pileup jets arising from the additional softer collisions in a bunch crossing, they are very low in p_{T} and will only replace the leading and subleading jet or substantially modify a third jet at a similarly low mass scale. The analysis therefore chose to address pileup through a cut on leading and subleading jet p_{T}. Events are only accepted when both jets have a $p_{\text{T}} > 50$ GeV.

The following plots illustrate the correlations between N_{PV} and the event's kinematic variables for jets with $p_{\text{T}} > 50$ GeV. Each variable is flat with respect to N_{PV}, indicating that the 50 GeV cut is sufficient to remove the effects of pileup on the m_{jj} spectrum.

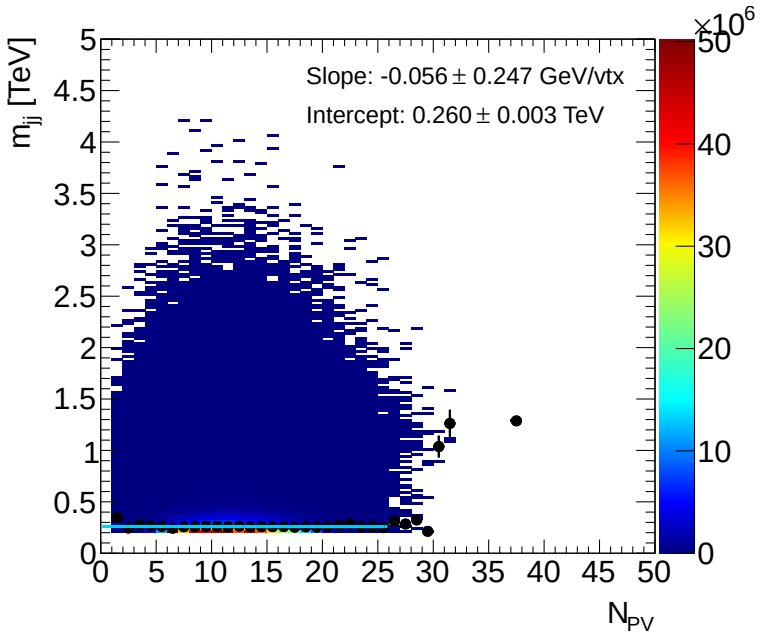


Figure 6.8: Independence of m_{jj} and pileup after jet p_{T} cut.

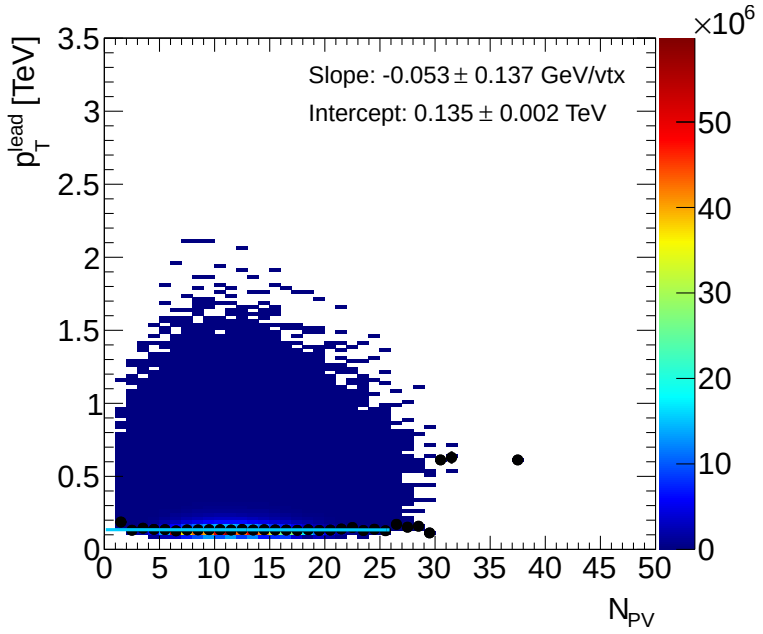


Figure 6.9: Independence of p_{T}^{lead} and pileup after jet p_{T} cut.

6.3.6 Centrality requirement

An additional cut is imposed on the dijet system to suppress QCD in favour of new physics. Most BSM models are expected to produce jets isotropically or centrally, while QCD has a dijet cross-section peaking for small values of θ. Thus the signal-to-background ratio can be increased through a cut on the rapidity difference between the jets $y^* = |y_1 - y_2|$. Note

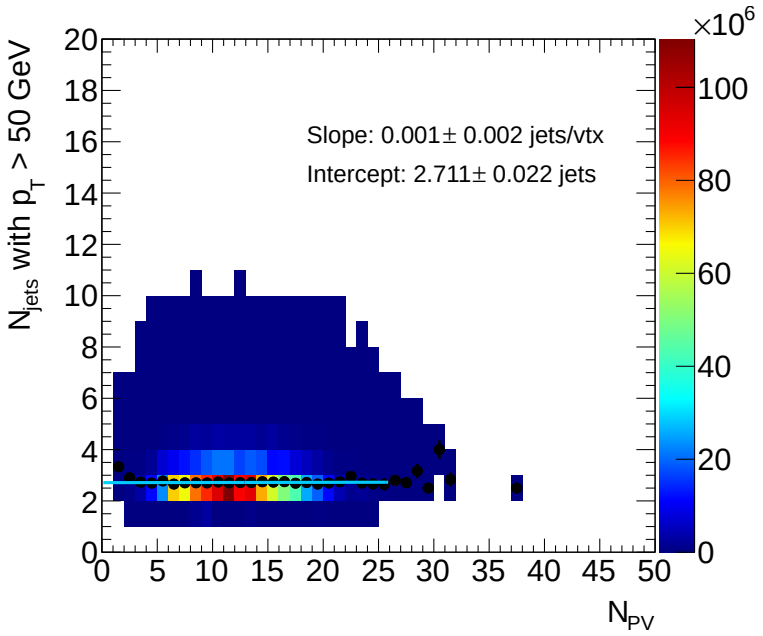


Figure 6.10: Number of jets above 50 GeV versus pileup.

rapidity is used instead of pseudorapidity for this cut due to its invariance for boosts of massive particles along the beam direction. For this analysis, every model except the W^* peaks at $y = 0$. A cut of $y^* < 0.6$ is selected, optimised for the centrally-produced signal models but reducing the sensitivity of the W^*.

6.4 Summary of analysis selection

The full set of cuts applied in the analysis involve all requirements discussed above in addition to a few basic requirements on the event validity. The BCH cleaning cut is included in the ugly jet rejection count, while the cut labeled "The following list summarises the cuts in order of application.

- If the spectrum is to be blinded, 3 out of every 4 events are held back at the first stage.
- Jets are recalibrated here. The leading jet must pass the trigger map requirements as detailed in Section 6.1.
- Events are rejected if they are not contained in a run or luminosity block in the GRL.
- The primary vertex must have more than one associated track: $N_{\mathrm{PV}}^{\mathrm{trck}} > 1$. This basic data quality cut restricts the data to events with real physical processes occurring.
- Events with calorimeter integrity issues or tripped modules are rejected. These apply only to a few runs and the effects were already accounted for in the GRL and

reprocessing used in the final paper.

- The `BCH_CORR_CELL` cut is applied to remove events where the leading or subleading jet, or any other jet with $p_{\mathrm{T}} > 0.3p_{\mathrm{T}}^2$ falls in a masked module. Events where the leading or subleading jet is ugly are also rejected here.
- The event is rejected if either the leading or subleading jet is a bad jet (matching the `BadLooser` definition)
- Events where either the leading or subleading jet has $p_{\mathrm{T}} < 50$ GeV are rejected. This is the pileup suppression cut.
- Events are rejected if the leading-subleading dijet system has $y^* > 0.6$.
- Events are rejected if the leading-subleading dijet system has $m_{jj} < 250$ GeV. This ensures that the spectrum is not biased by the pileup rejection cut.
- The remaining events make up the dijet spectrum to be analysed for signs of new physics.

Table 6.5 records the number of events in the normal and delayed stream for each stage in the cut application. The final row records the number of events entering the final dijet mass spectrum.

Cut requirements	Number of events (normal)	Number of events (delayed+overlap)
Before cuts	871647227	417493026
Passes trigger map	12886319	33967850
Passes good run list	9918952	32461435
Primary vertex check	9918894	32461223
No calorimeter errors	9863909	32349156
y of leading jets	9422143	32289087
Tile hotspot check	9422143	32289087
Ugly jet check, BCH	8339494	28580208
Bad jet check	8334537	28563192
Jet p_{T} cut	7610460	28516483
y^* cut	4396317	16014278
m_{jj} cut	4259455	16013126

Table 6.5: Event selection criteria and the remaining number of events in the normal and delayed data streams.

6.5 The Spectrum

6.5.1 Dijet mass resolution

The experimental resolution was calculated using QCD dijet samples generated in PYTHIA 8.175 using the AU2 tune obtained from ATLAS data [138, 139]. The dijet mass resolution is defined as the width-to-mean ratio of a Gaussian fit to the $m_{jj}^{reco}/m_{jj}^{truth}$ distribution for truth-matched leading and subleading jet pairs. Truth matching is performed by examining all truth jets within $\Delta R < 0.3$ of a reconstructed jet and selecting the one with the most closely matching p_{T} value. A set of truth mass bins was defined to ensure high statistics in each. All dijet mass pairs in the same bin are used to form a distribution of $m_{jj}^{reco}/m_{jj}^{truth}$, and a Gaussian is fit to the peak region to determine the resolution $\sigma_{m_{jj}}/\mu_{m_{jj}}$, where μ and σ are the mean and width returned by the fit. The measurement shows that the dijet mass resolution is 8% at $m_{jj} \simeq 250$ GeV, below 5% from ~ 1 TeV upwards, and falls to slightly under 4% above 3 TeV. The calculated dijet mass resolution is shown in Figure 6.11. One point is shown for each truth mass bin.

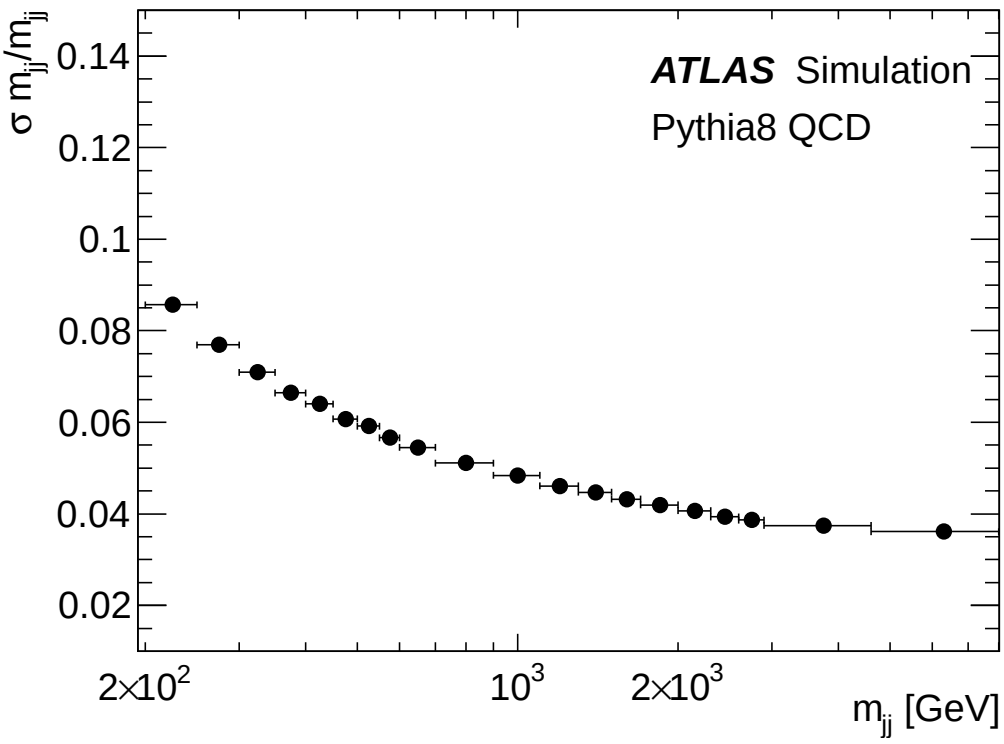


Figure 6.11: Detector resolution in m_{jj} calculated from PYTHIA8 MC samples. Each point represents a tested truth mass bin.

6.5.2 Binning strategy

The binning selected for the dijet spectrum was carefully selected by balancing the two most important criteria:

- Bins must be narrower than the expected width of a signal peak. An entire peak in a single bin would be impossible to characterise, and narrower bins improve the analysis sensitivity to excesses.
- It is preferred that bins be wider than the experimental resolution. The wider they are, the less migration of events between bins in m_{jj} as a result of the detector resolution will occur.

Values of signal peak width were studied in MC and used in conjunction with the dijet resolution described above to determine the optimal bin widths. Two PYTHIA 8 signals were used for determining the possible NP width: the q^* and W' samples were selected, as they are among the narrowest models used and thus set the tightest restrictions on the binning. A range of masses up to 3.2 TeV was used for the W' measurement, while the q^* was considered up to a mass of 5 TeV.

The signal widths are derived from fits to the peak region of the reconstructed MC, since the quantity of interest is the observed width of signal expected in the data. All the signal samples display a low-mass radiative tail to some degree, meaning only a restricted range in m_{jj} can be used for the peak width determination. The fits are also highly dependent on the starting parameters, so the mean is given a start value equal to the truth mass of the generated particles. The Gaussian width parameter takes as a start value the difference in m_{jj} between the peak location and the highest-mass events in the sample. From these initial parameters a fit is run and the width of the signal peak obtained. As is to be expected from the nonzero intrinsic width of the models, this is always larger than the detector resolution at the same value of m_{jj}.

The measured widths of the q^* and W' mass points are shown in Figure 6.12 alongside the dijet resolution. The ideal bin widths fall between the two curves.

The binning range was required to extend well beyond the highest mass points appearing in the data spectrum in order to accommodate those signal samples (particularly black holes)

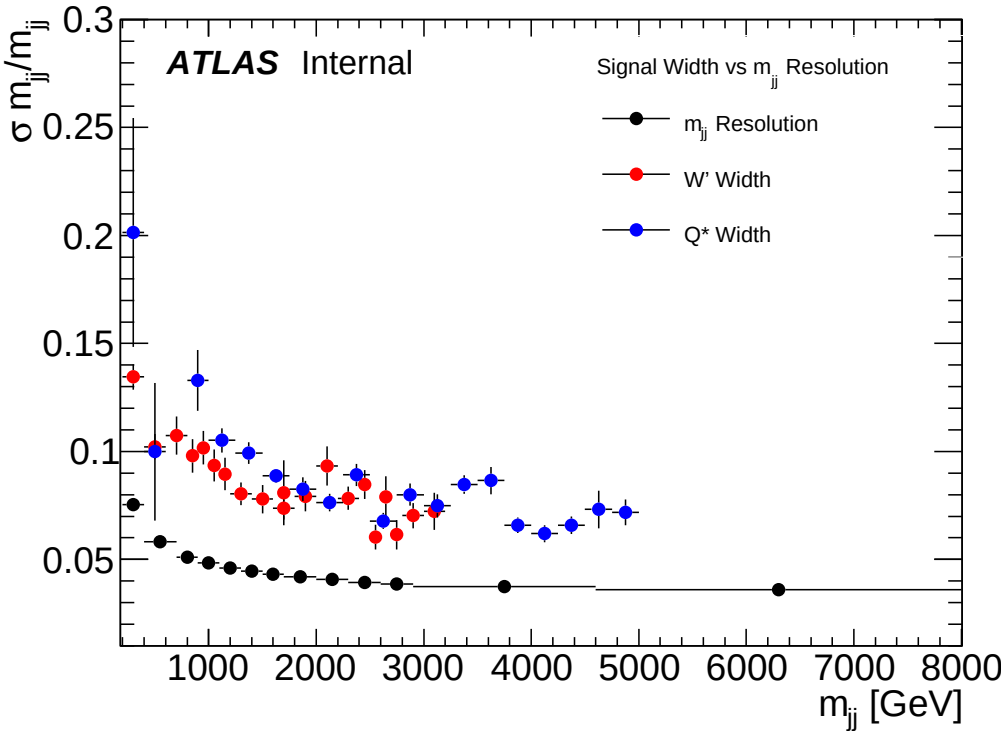


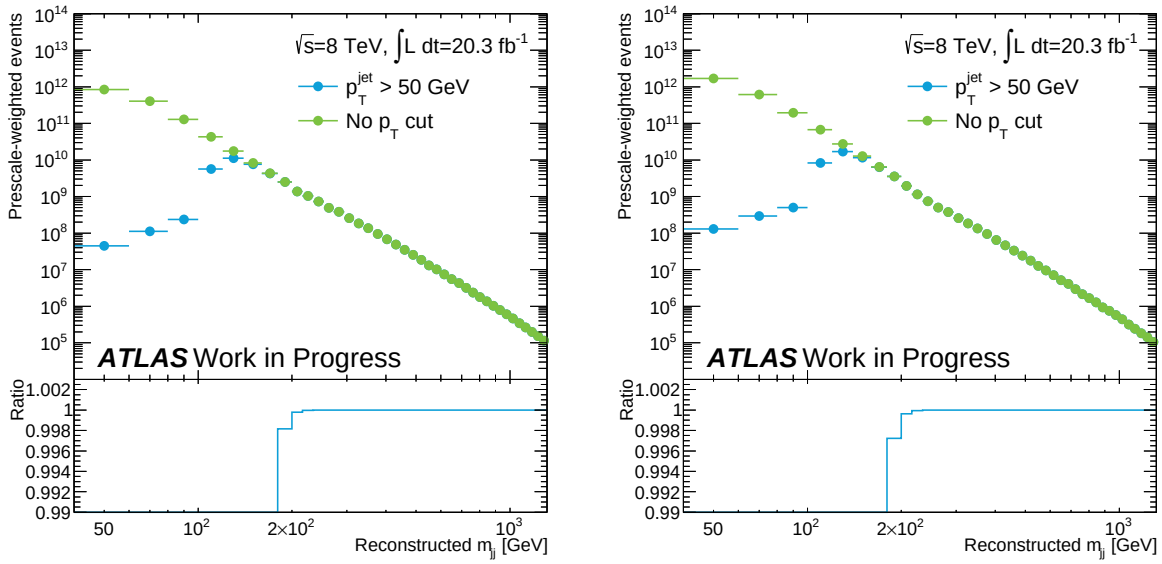
Figure 6.12: A comparison between signal peak widths and the detector resolution in m_{jj} all calculated from PYTHIA 8 MC samples. The black points illustrate the detector resolution, where each represents a tested truth mass bin. The red and blue points correspond to the widths of reconstructed signal distributions at the specified truth masses for the W' and q^* samples. Due to its decay via a quark pair, the W' is the narrowest signal considered in the analysis. Bin widths between this and the detector resolution are preferred for the m_{jj} spectrum.

for which extremely high-mass events would appear. As a result, the binning calculations have been extrapolated as far as the centre-of-mass energy to ensure that no possible event would be missed by falling outside the permitted bin ranges. The final bin edges for the analysis selected from these studies, in GeV, are:

[20, 234, 253, 272, 294, 316, 339, 364, 390, 417, 445, 474, 504, 535, 566, 599, 633, 668, 705, 743, 782, 822, 864, 907, 952, 999, 1048, 1098, 1150, 1203, 1259, 1316, 1376, 1437, 1501, 1567, 1635, 1706, 1779, 1854, 1932, 2012, 2095, 2181, 2269, 2360, 2454, 2551, 2650, 2753, 2860, 2970, 3084, 3202, 3324, 3450, 3581, 3716, 3855, 3999, 4149, 4303, 4463, 4629, 4800, 4977, 5160, 5350, 5546, 5748, 5958, 6174, 6397, 6628, 6866, 7112, 7366, 7628, 7898, 8177]

6.5.3 Selection of lowest dijet mass for the analysis

The pileup reduction cut on the individual jet masses does not result in a sharp cut on the m_{jj} distribution. Instead, due to the difference in p_{T} between the leading and subleading jets, it creates a smooth turn-on curve for the distribution which only reproduces the unbiased shape of the spectrum above some much higher mass. The ratio of the m_{jj} spectrum with and without the cut $p_{\mathrm{T}}^{lead}, p_{\mathrm{T}}^{subl} > 50$ GeV was calculated for both data and QCD MC. Figure 6.13 shows the results: in both real and simulated data, the effect of the cuts becomes negligible for m_{jj} masses above 250 GeV. As a result, the region of the spectrum used for comparisons begins at 253 GeV, which corresponds to the first bin edge falling above 250 GeV.



(a) Data (b) Monte Carlo

Figure 6.13: Bias introduced into the m_{jj} spectrum by requiring $p_{\mathrm{T}}^{lead}, p_{\mathrm{T}}^{subl} > 50$ GeV. In both data (6.13a) and Monte Carlo (6.13b), the impact of these cuts is negligible for $m_{jj} >$ 250 GeV, so the smoothly falling spectrum shape can be analysed above this point.

6.5.4 The spectrum in 8 TeV data

The final mass spectrum, after the application of all the above cuts and using the specified binning and starting m_{jj} values, is shown in Figure 6.14. This spectrum includes the normal stream, delayed stream, and debug stream for the full 2012 dataset.

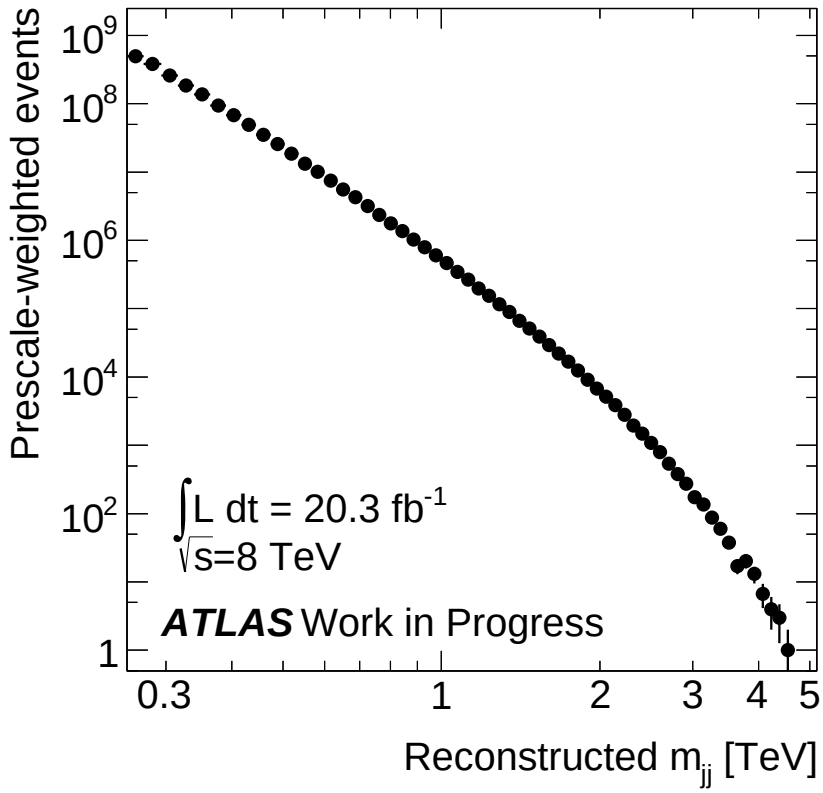


Figure 6.14: The m_{jj} spectrum in the full 2012 dataset, including the normal, delayed, and debug data streams. The lowest bin displayed begins at 253 GeV.

Chapter 7

Searching for Excesses

'Surprises are foolish things. The pleasure is not enhanced, and the inconvenience is often considerable.'

— Jane Austen, *Emma*

The dijet invariant mass spectrum derived in the preceding chapter must be carefully analysed to determine whether there are any statistically significant indications that a new resonant particle is contributing to the data. This requires an expected form for the distribution representing the shape it could be expected to take if only Standard Model processes were present. It also depends upon a set of pre-defined statistical tools for measuring the degree of discrepancy between the data and this expected distribution, tools which can be used to quantify the consistency of the data observed with the Standard Model. Thus it can be established whether or not any new physics has been observed.

The portion of the analysis which performs this comparison is the *search phase*, which seeks to answer the question "What is the probability of observing data at least as extreme as the measured spectrum, given that the Standard Model is true?" Results with lower probability than an agreed cutoff value are considered inconsistent with the Standard Model and may point to new physics. The following section will define the measures of probability and "extremeness" required to compare the data to a prediction. The following sections discuss the creation of the Standard Model-only prediction.

7.1 Statistical framework

Let f_b be some smoothly falling spectrum which describes the dijet invariant mass spectrum after analysis selection in the case that only Standard Model processes are present. This can then be used as a hypothesis describing the underlying shape of the dijet distribution with no new physics: the *background-only hypothesis* or *null hypothesis*, H_0. If this hypothesis correctly describes the natural world, then the observed data would be nothing but a set of random statistical fluctuations away from f_b. The validity of H_0 can then be tested by determining the probability of obtaining the data spectrum as a fluctuation in the background-only hypothesis. To determine this a rigorous definition of probability must be established.

7.1.1 Definitions of probability

There are two common ways of defining probability which differ in their philosophy. Both frequentist and Bayesian methods are used in ATLAS and both provide valid, though different, interpretations of probability. The method to be used can be selected to match the form of the answer sought [195].

The classical or *frequentist* perspective on probability is strictly a statement on the frequency of a certain outcome given a large number N of repeated experiments. Thus the probability p of a certain result is the fraction of times that result occurs as $N \to \infty$. This definition only holds in the case that many repeated trials are available, so no probability can be assigned to any single event. It also makes it impossible to define a frequentist probability for a statement which is either true or false and is not an experimental outcome in and of itself. A frequentist answer exists for the question, "What is the probability of observing a measurement at least as extreme as the Higgs search result, given that the Higgs does or does not exist?": a large number of possible experimental outcomes for those two cases can be calculated and the probability is the fraction of those which are more extreme than the measured result. However, no frequentist answer can be given to the question, "Does the Higgs boson exist?" An answer to this is either true or false and would not change in either case; thus in a frequentist framework it is meaningless to interpret this in terms of probabilities. The same can be said of parameter values: a parameter has a fixed unknown

value about which no probabilistic statement can be made.

The *Bayesian* definition of probability is based on the plausibility of a hypothesis and depends on prior information. In a scientific context a Bayesian analysis can compute the probability of some statement H being true (i.e. "The Higgs boson exists" or "ϕ has a value of 7.2") given a measurement x by using the conditional probability $p(x|H)$ and the prior degree of belief in H. In many cases the outcome of such a calculation will assign a similar or identical probability to an event as would a frequentist method, but the mentality behind the statement remains very different. A Bayesian methodology makes it possible to assign probabilities to individual events, parameter values, and hypotheses with true or false outcomes.

Frequentist methods are the most common in high energy physics by a substantial margin for a variety of reasons. Because of the perceived subjectivity of Bayesian thought and because of the technique's dependence on quantifying a prior degree of belief, frequentist searches are considered more reliably objective [196]. The computation times involved can be less. However, both methods are mathematically valid and both can be handled in clear and objective ways. As long as the approach used to obtain a number is clear to the reader, either choice is valid for an analysis. The real difference lies in the issue being explored, and the dijet analysis employs both methods. A Bayesian approach is used to set limits on benchmark signal models and is described in detail in Chapter 9. For the search phase the result sought is a statement about the probability of obtaining a certain outcome (the measured dijet invariant mass spectrum) given a certain hypothesis (the Standard Model). This question is simply handled by using a frequentist approach.

7.1.2 Quantifying frequentist probabilities

A frequentist hypothesis test determines the consistency between H and the observed experimental outcome x by fixing in advance a value of probability α below which the hypothesis will be rejected as too discrepant. Specifically, if the observation falls in a space of possible outcomes w such that

$$P(x \in w|H) \leq \alpha \tag{7.1}$$

then H is rejected; otherwise it is considered compatible with the observation [197]. In the case that an alternate hypothesis of interest is available, H_1, the above definition can be used to define a specific critical region w for which there is the greatest possible distinction between the hypotheses.

However, in the case of a search-based analysis no second hypothesis is commonly available. Rather than a direct comparison it is desirable to determine the agreement between the observation and H_0 alone. Although it will eventually become necessary to use the spectrum to set limits on specific model hypotheses, the most general, theory-agnostic approach is to define a test which compares the spectrum to the Standard Model only without recourse to any comparison.

Define a *test statistic* T to be any numerical quantity which describes the compatibility between x and H_0 and which increases monotonically with decreasing compatibility. The *p-value* of this test statistic is the probability of obtaining a value at least as extreme as the observed $T = t_0$ given H_0: that is, equal to or higher than the observed value.

$$p = P(T \geq t_0 | H_0) \,. \tag{7.2}$$

Thus a small p-value corresponds to poor agreement between the observation and the hypothesis H_0 [197, 198].

A variety of test statistics are available to the analysers and will be discussed in detail in Section 7.1.4. However, any numerical comparison between x and H_0 must account for the fact that the observation, whether it is data or Monte Carlo, is composed of weighted events. The methods used for such comparisons will be discussed before the test statistics are defined.

7.1.3 Comparing a weighted spectrum to a prediction

It is often necessary to examine a spectrum containing weighted events. Any spectrum created from Monte Carlo samples will exhibit this behaviour. In previous iterations of the analysis the data distribution consisted only of events with weight $w = 1$, but due to the use of multiple triggers with different prescales, that is no longer the case and the data spectrum too consists of weighted events. Either way, the contents of a bin may be substantially

different from the number of raw events that were used to fill it.

This creates a problem for all statistical tests which assume that the contents of a given bin are Poisson distributed observations and which use this to determine the probability of observing a given value. If, for example, a bin has content 100 and is to be compared to a predicted value of 115, the observed value has a much lower probability if the bin was filled using 100 events of weight 1 than if it was filled using 2 events of weight 50. A Poisson-distributed variable is one for which the probability of an observation x with expected value λ is:

$$p(x|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{7.3}$$

In the first case the true probability would be $p(100|115) = 0.014$ while in the second case it would be $p(2|2.3) = 0.265$.

The chosen solution is to use the effective number of events defined in Section 6.1.2. N_{eff} is a Poisson-distributed quantity which has the average statistical power of the events in the bin, and can therefore be used in place of the bin content for statistical tests to obtain the correct measurement. The theoretical value to which a measurement will be compared must be scaled to the same precision for the comparison, as the prediction of 115 counts is scaled down by 50 for the weighted example above. Thus any prediction to be compared to a weighted spectrum is scaled by $1/w_{eff}$ before the calculation. Throughout the remainder of this document, it can be assumed that any comparison between data and a prediction makes use of the effective statistics in data and an appropriately scaled version of the hypothesis unless explicitly stated otherwise.

7.1.4 Test statistics

Three different values are used as test statistics in the search phase to quantify the level of agreement between an observation and a spectrum serving as a hypothesis. These test statistics return a single value which characterises the entire spectrum and which increases monotonically with increasing discrepancy between observation and prediction, making these statistics useful for defining p-values.

The χ^2 measurement

Pearson's χ^2 test characterises the probability that a set of observed values arose as random sampling of a null hypothesis [199]. It is defined as the sum in quadrature of the differences between observations and expectations, normalised to the expected value. For a comparison where the observation and prediction are both binned histograms with contents d_i and b_i respectively in bin i, then:

$$\chi^2 = \sum_i \frac{(d_i - b_i)^2}{b_i} \,. \tag{7.4}$$

This statistic is derived assuming that measured values are normally distributed around the theoretical value, which for Poisson distributed event counts is only a reasonable assumption for bins with more than ~ 5 events. The χ^2 comparing the dijet distribution to a model is therefore unreliable in the tail region but correct elsewhere.

The χ^2 statistic is often used to define goodness of fit. For a single measurement d of a normally distributed variable with standard deviation σ and mean b the probability is:

$$P(d) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(d-b)^2}{2\sigma^2}} \,. \tag{7.5}$$

Thus for a collection of measurements with various predicted means and variances, the combined probability of any observation is the product of the individual probabilities:

$$\begin{aligned} P(\text{obs}) &= \prod_i \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(d_i - b_i)^2}{2\sigma_i^2}\right) \\ &= \frac{1}{\sqrt{2\pi}\prod_i \sigma_i} \exp\left(-\frac{1}{2}\sum_i \frac{(d_i - b_i)^2}{\sigma_i^2}\right) \end{aligned} \tag{7.6}$$

The maximum probability is obtained by minimising the quantity $\sum_i \frac{(d_i-b_i)^2}{\sigma_i^2}$; that is, the χ^2.

The "reduced χ^2" value $\frac{\chi^2}{\text{NDF}}$ is often used to test goodness-of-fit. Here NDF is the number of degrees of freedom in the fit and is given by:

$$\text{NDF} = n - N - 1 \tag{7.7}$$

where the histogram under consideration has n bins and the fitting model has N parameters. Using the reduced χ^2 rather than its complete value can make deviations from an ideal fit

appear deceptively small, but has the advantage of making it possible to directly compare fit quality between cases with different numbers of bins and fit parameters.

The Log Likelihood

Let a data distribution be described by some set of parameters $\boldsymbol{\theta}$. A likelihood $\mathcal{L}$ denotes the conditional probability of a set of parameters given a fixed observation. This is equal to the probability of the specific observation given that fixed set of parameters:

$$\mathcal{L}(\boldsymbol{\theta}|x) = P(x|\boldsymbol{\theta})\,. \tag{7.8}$$

A likelihood is not, however, a true probability distribution since the integral of the likelihood over all parameter values is not fixed to 1. The likelihood should be interpreted as a function of the parameter values for fixed data [200].

Let the set of parameters being tested lead to a predicted value of b_i in bin i. In the case of normally distributed measurements, $\mathcal{L}$ reduces to the above expression for a χ^2 probability and $-2\ln\mathcal{L}$ is the χ^2 test statistic. The likelihood can, however, be calculated for data following any distribution. In the case of Poisson-distributed data such as the histogram bin contents in the dijet analysis, this is the product of the Poisson probability in each bin (Equation 7.3) over all bins:

$$\mathcal{L}(\boldsymbol{\theta}|x) = \prod_i \frac{b_i^{d_i} e^{-b_i}}{d_i!} \tag{7.9}$$

Hereafter referred to as the *log likelihood*, the value of $-2\ln\mathcal{L}$ equivalent to the χ^2 test statistic for Poisson-distributed data is:

$$\begin{aligned} -2\ln\mathcal{L}(b) &= -2\ln\prod_i \frac{b_i^{d_i} e^{-b_i}}{d_i!} \\ &= -2\sum_i \ln\frac{b_i^{d_i} e^{-b_i}}{d_i!} \\ &= -2\sum_i \left(d_i\ln(b_i) - b_i - \ln(d_i!)\right) \\ &= -2\sum_i \left(d_i\ln(b_i) - b_i\right)\,, \end{aligned} \tag{7.10}$$

ignoring terms which do not depend on the parameterisation. This quantity is a single number which quantifies the agreement between data and the prediction and which increases

with greater discrepancy. It is used by the analysis as another possible test statistic.

The log likelihood can also be used to fit parameters. Proposed by Fisher in 1922 as an improvement on the method of least squares, maximum likelihood estimation (MLE) selects model parameters to be those which maximise the likelihood function or, equivalently, minimise $-\ln \mathcal{L}$ [201, 202]. Unlike the χ^2 approach, however, the value of the likelihood itself does not have a direct interpretation in terms of goodness of fit, since it varies with the data and model being used.

The BumpHunter

Both the χ^2 and log likelihood tests quantify the difference between observation and model in individual bins, but neither describes the behaviour between neighbouring bins at all. Instead, each bin is completely independent such that the same value of the test statistic would be obtained if bins with excesses and deficits alternated as if the bins were arranged from greatest excess to greatest deficit. Clearly, however, the interpretations attached to these two cases would be quite different. In a resonance analysis like the dijet search, three adjacent bins with large excesses in each might indicate new physics where three bins with a large excess, large deficit, and large excess would produce the same χ^2 but would be of much less physical interest. A third test statistic has therefore been selected to quantify the "bumpiness" of the data compared to the prediction.

The BumpHunter defines a test statistic sensitive to this behaviour [203]. For a set of adjacent bins (say bins numbered in $[m, n]$) a value t is calculated as the Poisson probability of obtaining a result at least as significant as the one observed (the cumulative Poisson probability) if the set of bins were combined into one. Let d_i and b_i be the observed and expected number of events in bin i respectively. Define:

$$d = \sum_{i=m}^{n} d_i \quad \text{and} \quad b = \sum_{i=m}^{n} b_i \,. \tag{7.11}$$

Then when $d < b$ the test statistic is the probability of obtaining a greater deficit than the observed d, and when $d > b$ it is the probability of obtaining a greater excess. Assuming the

observed counts are Poisson distributed,

$$t = \begin{cases} \sum_{n=0}^{d} \frac{b^n}{n!} e^{-b} & \text{for } d < b \\ \sum_{n=d}^{\infty} \frac{b^n}{n!} e^{-b} & \text{for } b \geq d\,. \end{cases} \tag{7.12}$$

The above expression can be represented in terms of gamma functions. Recall $\Gamma(n) = (n-1)!$, $\gamma(n,x) + \Gamma(n,x) = \Gamma(n)$, and $\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$ [204]. The normalised integer forms of the incomplete lower and upper gamma functions are:

$$\begin{aligned} \Gamma(n,x) &= \frac{1}{\Gamma(n)}(n-1)!e^{-x}\sum_{m=0}^{n-1}\frac{x^m}{m!} \\ &= e^{-x}\sum_{m=0}^{n-1}\frac{x^m}{m!} \end{aligned} \tag{7.13}$$

$$\begin{aligned} \gamma(n,x) &= \frac{1}{\Gamma(n)}(n-1)!\left(1 - e^{-x}\sum_{m=0}^{n-1}\frac{x^m}{m!}\right) \\ &= 1 - e^{-x}\sum_{m=0}^{n-1}\frac{x^m}{m!} \end{aligned} \tag{7.14}$$

Then for $d \geq b$,

$$\begin{aligned} t &= e^{-b}\sum_{n=d}^{\infty}\frac{b^n}{n!} \\ &= e^{-b}\left(\sum_{n=0}^{\infty}\frac{b^n}{n!} - \sum_{n=0}^{d-1}\frac{b^n}{n!}\right) \\ &= e^{-b}e^{b} - e^{-b}\sum_{n=0}^{d-1}\frac{b^n}{n!} \\ &= \gamma(d,b) \end{aligned} \tag{7.15}$$

and for $d < b$,

$$\begin{aligned} t &= e^{-b}\sum_{n=0}^{d}\frac{b^n}{n!} \\ &= \Gamma(d+1,b) \\ &= 1 - \gamma(d+1,b) \end{aligned} \tag{7.16}$$

Thus the final definition of the probability in each window is:

$$t = \begin{cases} 1 - \gamma(d+1, b) & \text{for } d < b \\ \gamma(d, b) & \text{for } b \geq d\,. \end{cases} \tag{7.17}$$

Where the χ^2 or log likelihood values for this bin range keep the probability in each bin separate, this value accounts for the direction of neighbouring fluctuations by looking at the overall excess or deficit in the region.

For every possible range of bins (*window*) within a specified minimum and maximum width, t is recalculated. The possible windows are found by looping over all widths between a minimum and maximum number of bins, then over every possible starting bin for a window of that width. For this analysis, the minimum window width was fixed to 2 bins and the maximum to half the number of bins in the spectrum, values considered reasonable physical limits for the resonances of interest. The narrowest signals explored in the analysis are significantly smaller in width than one-half the spectrum, so this selection was certain to cover all cases of interest. A test statistic describing the overall spectrum is defined as the negative log of the smallest probability obtained for any window. That is,

$$t_0 = -\log t_{\text{min}}\,. \tag{7.18}$$

The BumpHunter approach is popular in BSM analyses because it is entirely theory-agnostic, requiring no MC template to search for a signal. For dedicated searches such as the Higgs analysis, where a well-defined behaviour is expected, this is an unnecessary feature, but for signature-driven searches the generality is a positive feature. In this analysis, only excesses were considered when calculating the most discrepant region. Deficits were ignored as they did not reflect the expected behaviour of any new physics models under consideration.

7.1.5 Extending test statistics to p-values

Once a test statistic has been selected, a p-value quantifying the agreement between the observation and prediction must be calculated. This is done using *pseudoexperiments*: since the frequentist approach defines probability as the fraction of time a certain result would occur in a repeated number of experiments if the hypothesis were true, some simulation of

those experiments must be produced. The goal is to test the hypothesis H_0 by determining how likely the observed spectrum would be to arise by chance alone. Toy data is thrown using H_0 as a guide, in this case by taking a random draw in each bin from a Poisson distribution with parameter equal to the expected bin content from H_0. One pseudoexperiment is a full set of one random draw per bin. The selected test statistic is then calculated to compare each pseudoexperiment to the hypothesis. The fraction of these cases for which the test statistic is more extreme than in data can then be easily computed.

To summarise, given an observation x and a hypothesis H_0 to be tested, the p-value is calculated using the following steps:

1. Calculate the value of the test statistic t_0 which compares H_0 to x. This is our *observed value*.
2. Generate many pseudoexperiments y_i from H_0 to represent a range of possible experimental outcomes in the case that H_0 accurately describes the underlying physical properties of the system.
3. Calculate test statistics t_i comparing H_0 to each y_i in turn.
4. Calculate the fraction of t_i for which $T \geq t_0$, as shown in 7.2. This is the p-value.

The dijet analysis uses the BumpHunter test statistic to create the reported p-value and determine whether new physics has been observed. When calculating a p-value with this statistic, the selected region of most significant excess will be different for different pseudoexperiments, both in mass range and in window width.

Trials factors and the look-elsewhere effect

Different statistical tests are sensitive to different features of the spectrum. For instance, the BumpHunter is sensitive to whether bin by bin discrepancies are consecutive while the log L and χ^2 tests are not. An analysis benefits by having a range of statistical tests to quantify various types of behaviours and provide context for one another. However, since a spectrum may register an extreme p-value with one test and not with another, this creates a dangerous possibility: any observation could potentially be used to exclude any hypothesis on the grounds of some test, as long as enough different tests were considered [203].

This *trials factor* has an equivalent interpretation in the location examined by a statistical

test. A statistical test which considers a single bin could reject the null hypothesis based on an extreme value in that bin, even if the remainder of the spectrum was entirely well-behaved. In this sense each bin constitutes a trial, and the probability of observing a discrepancy in any one of them must be considered rather than only the probability of a given observed excess. This geometrical interpretation is common to all searches even if only a single statistical test is used, and is termed the *look-elsewhere effect*.

The BumpHunter is selected as the statistical test which determines whether or not a signal is present in the dijet mass spectrum not only because it measures the behaviour of greatest interest for a resonance search ("bumpiness") but because it reports a p-value which has already accounted for the look-elsewhere effect. First, every possible location for an excess is considered (all permitted widths of bins in all possible locations) and a p-value for a given test statistic (the Poisson probability) defined for each. Selecting simply the value with the smallest p-value would not properly account for the trials factor of the many locations, so instead this smallest value is itself used as a test statistic for a *hypertest*. The selection of pseudoexperiments performed will all have minimum Poisson p-values corresponding to different bump locations, and thus the hypertest defined from the union of the smallest p -values from all of the individual tests properly accounts for the look-elsewhere effect. By making the BumpHunter the only test on which the background-only hypothesis can be rejected, even though other tests are also performed, the trials effect of other statistical interpretations is also circumvented.

7.1.6 Representing significance bin-by-bin

In addition to single-valued statistical tests representing the overall degree of agreement between two spectra, it is beneficial to have a representation of this comparison on a bin-by-bin level. Such tools facilitate a visual interpretation of the comparison and can pinpoint specific regions of unusual discrepancy. Three such bin-by-bin significance measures have been implemented in the dijet analysis and are often plotted simultaneously with the data and fit distributions to provide a visual indication of the discrepancies. These are all local measures of significance, but will be referred to by distinct terms throughout this chapter to prevent confusion. For a bin with observed content d (error δd) and predicted content b (error δb), these three measurements are defined as follows.

Relative difference

This is a simple ratio comparing the difference between estimate and observation to the scale of the prediction:

$$r = \frac{(d-b)}{b} \tag{7.19}$$

Significance

This compares the difference between the estimate and observation to the size of the total uncertainty in that bin, giving an indication of the significance of the discrepancy in terms of multiples of the error bars.

$$s = \frac{(d-b)}{\sqrt{\delta d^2 + \delta b^2}} \tag{7.20}$$

Residual

This expresses the difference between estimate and observation in terms of the number of standard deviations of difference between them, assuming Poisson distributed data. The mathematical definition is given in the following section.

The residual

The residual is the most commonly used bin-by-bin visualisation in the dijet analysis and is described in detail in [205]. In each bin, a measure p is defined as the probability of measuring a discrepancy between data and background at least as large as the one observed. This is recognisable as the quantity given in 7.17, only the values compared are in a single bin rather than a window of varying length. The measured probability p can then be transformed into a z-value, defined as the number of standard deviations on the right of a Gaussian which would correspond to the same probability. Mathematically,

$$p = \int_{z\text{-value}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \tag{7.21}$$

Any bins with a z-value less than zero show no difference of any interest, while those with a z-value of more than two or three indicate a substantial discrepancy. For clarity of interpretation, one would like the sign of the z-value drawn in residual plots to depend on whether the data falls above or below the hypothesis, not on how significant the difference is. Therefore, when plotted, any bins with a negative z-value are drawn at exactly zero,

while those with a positive z-value are drawn positive or negative depending on whether an excess or a deficit is observed.

7.2 The background estimate

There are several strategies available for defining a background-only hypothesis to which the observed data can be compared. Many analyses employ MC distributions generated across all physics processes which are expected to contribute to their background and use the resulting shapes to model the expected data distribution. This method effectively captures current knowledge about the behaviour of background processes, but requires substantially more statistics in MC than data in order to make any strong statements. This method was employed in several very early dijet searches at D0 and CDF, but is not practical in the LHC era, where integrated luminosities are so high that it is simply not possible to produce enough MC to describe the background with this method. Furthermore, the MC may not have an entirely reliable shape since the behaviour of the underlying distribution may not be fully understood, as is the case with QCD jet production. Therefore relying on MC to adequately describe the behaviour of the background in the dijet analysis is not a practical choice.

Instead, it is common practice to parameterise the background using a smooth functional form. The dijet pairs from non-resonant sources which make up the invariant mass spectrum in the case that no new physics is present are a *combinatorial background*: they arise from the pairing of uncorrelated jets arising not from a single source but from other (usually QCD) SM processes in the event [206]. The selection of the two highest-p_{T} jets in each event where both have a p_{T} drawn at random from a smoothly falling exponential distribution will create a combinatorial background with the same traits. Thus although the exact form of the physics distribution underlying the SM dijet invariant mass spectrum is not known, it will have the following traits:

1. It will be a monotonically decreasing function with an approximately exponential form, and
2. There is some intrinsic number of degrees of freedom in the dijet production process which are reflected in the PDF.

A phenomenological description of the distribution is therefore commonly sought in the form of a polynomial or exponential function which can describe this background and which must exhibit matching properties. The function must be monotonically decreasing and should have the appropriate number of parameters to match the approximate number of degrees of freedom present in the data. Note that the approximate number of degrees of freedom may not match those of the intrinsic dijet production process: many finer features of the data may only appear at high statistics, so a large number of parameters can over-fit a smaller amount of data dominated by statistical fluctuations where it would nonetheless appropriately describe the underlying distribution. On the other hand, a function with relatively few parameters fitting a large amount of data will lack the flexibility to describe the shape of the distribution appropriately.

These considerations alone are not enough to ensure a good choice of background parameterisation. The function must also be guaranteed not to disguise new physics if it were present. That is, it must not adapt around resonance-like structures in the spectrum. The case of *thresholds*, non-resonant new physics manifesting as an increasing number of events in the highest-mass region of the spectrum, cannot be addressed when the background is parameterised in this way. The function will adapt to smooth changes in shape of the tail and cannot be forced to do otherwise since the true shape of the distribution for SM-only processes in the high-mass tail is not known. Therefore, while the background description can be made unbiased in the presence of resonant signals this is not possible with thresholds.

7.2.1 Parameterisation with a smooth function

The function chosen to parameterise the background, referred to as the "standard" or "dijet" function, has the following form:

$$f(x) = p_0(1-x)^{p_1}x^{p_2+p_3\ln x}\,, \tag{7.22}$$

where $x \equiv m_{jj}/\sqrt{s}$ is the ratio of the dijet invariant mass to the center of mass energy and the p_i are free parameters. This functional form was first used by CDF in a 2009 dijet resonance search [207]. Since the UA2 collaboration first used a simple parameterisation in 1990 to describe a falling invariant mass spectrum, successive runs and experiments have repeatedly reached the limit of the descriptive power of one function and been forced to extend it with

additional terms or to select a new form. Two main factors affect this evolving description: increased luminosity and increased reach in invariant mass. With increased luminosity, statistical errors shrink and finer features of the QCD spectrum become perceptible, leading to deviations between the measured spectrum and a simple functional form. The extended mass range with each higher-energy search forces the function to describe regions of QCD with previously unconstrained behaviour, and more recently developed functional forms have accumulated parameters allowing the function to adapt to the high-mass tails [208].

7.2.2 Creating the estimate

The dijet function is fit to the invariant mass distribution using MINUIT [209]. The Poisson-based maximum likelihood estimator was chosen over the least-squares (χ^2) estimator as the quantity to be minimised, since it accounts correctly for the data distribution in the high-mass region with low bin content.

MINUIT provides a selection of minimisation algorithms, of which the MIGRAD + SIMPLEX algorithm was used to find the background estimate. This preliminary fit attempt is made with MIGRAD, which is the most reliable of MINUIT's algorithms for most applications. It uses a steepest-descent algorithm, making it sensitive to the precision with which the first derivatives are known. In the case that the MIGRAD fit fails, a SIMPLEX fit is performed instead. This is not reliant on first derivative calculations and is robust in most situations, but is substantially slower than MIGRAD and does not provide any estimate of the parameter uncertainties.

The minimisation procedure finds the values of the function parameters which maximise the likelihood, denoted $\hat{\boldsymbol{\theta}}$. A confidence interval in each parameter space will correspond to a range of possible values within which the true value of θ will fall in a fraction $1 - \alpha$ of cases. One way to denote this is by the variance of the parameter in cases where the results are believed to follow a Gaussian distribution. Since in the large sample limit maximum likelihood estimators converge to a Gaussian pdf, the confidence intervals on the parameters in the current analysis correspond to their variances [210].

The full search phase algorithm

The BumpHunter p-value is used to quantify the level of agreement between the data and the fit result. Some cutoff p-value must be selected as the parameter α below which we consider the background-only hypothesis to be rejected, as in Equation 7.1. In the dijet analysis the chosen value is $\alpha = 0.01$. If the measured p-value falls below 0.01, it means that the probability of having observed a result at least as bumpy as the real data as a result of statistical fluctuations around H_0 alone is 1 in 100. This is very far from the discovery significance of 5σ but is an acceptable condition for rejection of the proposed background-only hypothesis: results below this value would indicate either that there was a truly new result in the data, or that the background hypothesis from the fit function had failed, or that there was an undiscovered reconstruction issue with the data, or some other such analysis problem. At any rate, should $p < 0.01$ be observed, the analysis would not proceed to the limit-setting phase since the result would not be sufficiently compatible with the background-only hypothesis.

To determine whether a signal is present, the spectrum must be compared to a background-only hypothesis. This means that the background estimate for the spectrum must not change if a signal is present: if the background estimate deviates to accommodate a resonant excess, the p-value of the Standard Model hypothesis will be incorrectly calculated. Such biases can be removed by excluding the range of bins around a significant observed excess from the fit and recalculating the background, iteratively adding bins to the exclusion until the remaining spectrum is well-described by the background. The BumpHunter p-value can be used as a measure of the agreement between data and fit, and to indicate the region of most discrepancy so that it can be removed and the fit recalculated until a stable result is reached. The calculation of uncertainties on the fit result and the tests of its stability against signal contamination will be discussed in detail in Sections 7.5.1 and 7.6.

The search phase algorithm in total is as follows:

1. Fit the observed dijet mass spectrum using the function in Equation 7.22.
2. Calculate the BumpHunter p-value relating the data to the fit obtained in step 1.
3. If $p > 0.01$, stop. No new physics has been observed and the background estimate is compatible with the data.

4. Otherwise, exclude the bins which the BumpHunter selected as the region of greatest excess from the fit and recalculate a background estimate. Check its BumpHunter p -value.
5. As long as this $p < 0.01$, add more bins to the exclusion according to the following rules, re-fit, and re-calculate the p-value

   - If the new most discrepant region selected by the BumpHunter is adjacent to the excluded region, it indicates that enough of the signal peak still remains outside the window to bias the background estimate. Add one more bin to the exclusion on the side of the window indicated by the BumpHunter.
   - Otherwise, there is nothing to indicate where the worst problem lies, so one bin on each side of the exclusion window is added to it.

   Since only one or occasionally two bins are added to the exclusion window at each iteration, the danger of eliminating valid statistical fluctuations is minimised.

7.2.3 The resulting background-only hypothesis

The background-only hypothesis which resulted from fitting the full dijet invariant mass spectrum according to the above algorithm is shown in Figure 7.1. The black points indicate the data and the red line illustrates H_0. The residual is drawn below the primary plot, indicating the significance in standard deviations of the excess or deficit in each bin. Two blue vertical lines indicate the region of greatest excess.

7.3 Validating the background-only hypothesis

It is not possible to guarantee in advance that the algorithm for defining a background-only hypothesis in a dataset will accurately describe that spectrum. Only a fit to the dataset itself will truly answer the question. The dijet analysis, however, was performed *blind*. The background estimation method had to be fixed in advance using the best tests available. Two options were available:

- The background estimation algorithm could be tested in the blinded dataset, 1/4 of the full dataset or 5.06 fb^{-1}. Because statistical uncertainties would be halved between this and the full dataset, some underlying features could become measurable in the

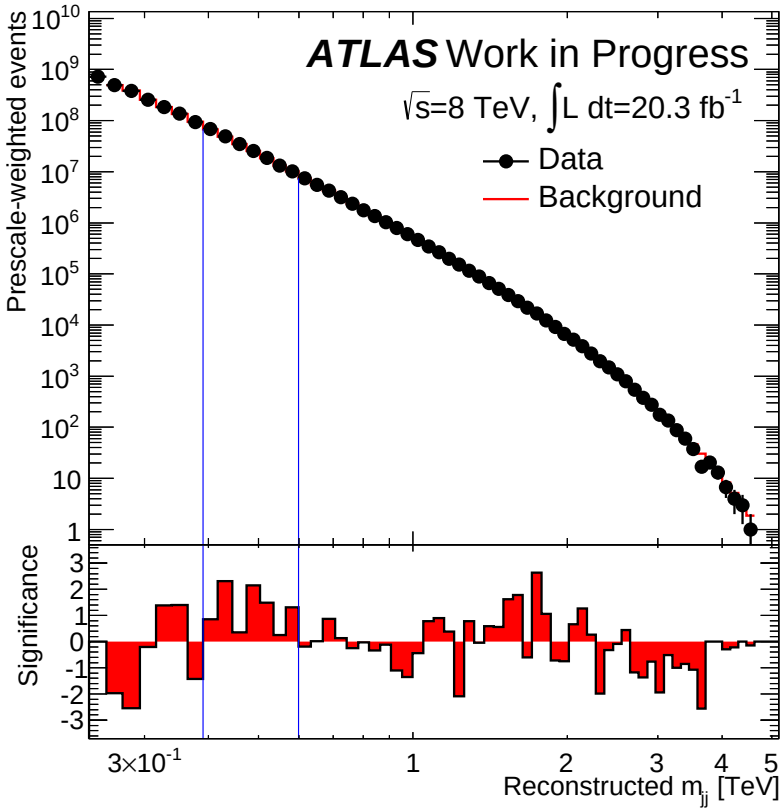


Figure 7.1: The reconstructed dijet mass distribution (filled points) fitted with the dijet function (solid line). The bin-by-bin significance of the data–background difference is shown in the bottom panel. The blue lines indicate the region of greatest excess.

full dataset which were invisible in the quarter, but the test would at least guarantee that the fit function could converge over the entire mass range tested and described the small spectrum well.

- The background estimation could be tested in simulated data.

7.3.1 Testing the search phase on blinded data

A true blinded analysis fixes all cuts and statistical approaches without observing the region in which a signal is expected to lie. The dijet analysis, which examines a single complete spectrum with the expectation that a signal could arise anywhere, imitates this idea by fixing all analysis choices while observing only a *partial dataset.* Since no new physics was observed in 13 fb^{-1} of 8 TeV data (see ATLAS CONF note [211]), then a smaller dataset than that is known to be not sensitive to any new physics. The blinded dataset was to be selected randomly and could contain substantial data not included in the 13 fb^{-1} analysis. For safety,

a partial dataset of 1/4 the full integrated luminosity was selected, which therefore had 1/2 the statistical sensitivity of the full dataset. Any unexpected features of this spectrum can therefore be examined as performance issues without raising any concerns that they might instead result from new physics. Only this spectrum with an integrated luminosity of 5.06 fb^{-1} was used in testing and finalising the search phase algorithm. The result of the fitting algorithm on the blinded dataset is shown in Figure 7.2. The reduced χ^2 of this fit was 0.955 and the BumpHunter p-value returned was 0.15. The log likelihood and χ^2 test statistics returned p-values of 0.44 and 0.54 respectively. The search phase algorithm was thus found to describe the partial dataset well.

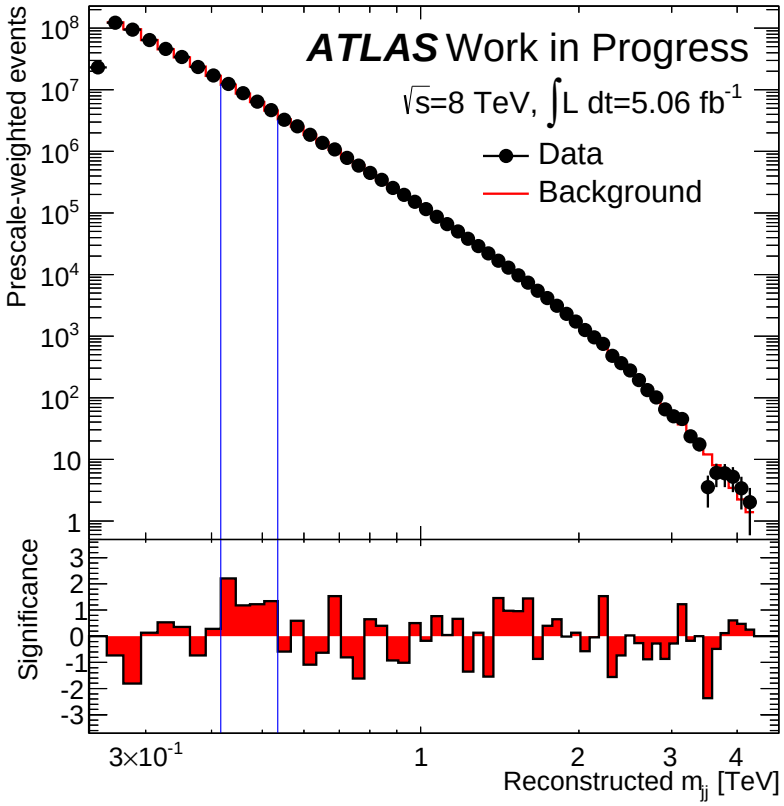


Figure 7.2: Fit to the blinded (partial) dataset. Every fourth event was kept to generate the spectrum shown by the filled points, while the fit result is shown in red. The bin-by-bin significance of the data–background difference is shown in the bottom panel. The blue lines indicate the region of greatest excess. The first bin drawn is still biased by the jet p_{T} cuts and thus not included in the fit.

7.3.2 Testing the search phase on Monte Carlo

The event selection was performed on a centrally produced Pythia8 CT10 dijet QCD sample and a simulated m_{jj} spectrum created. Despite having the same overall shape as data, such

a spectrum does not have a similar distribution of statistics and so provides no information on whether the search algorithm would be expected to accurately describe the full dataset. ATLAS MC samples are generated *flat*: that is, in order to maintain sufficient statistical sensitivity to perform analyses over the entire energy range of a physical process, MC events are produced at the same rate for all p_{T} values in a collection of datasets sliced by p_{T}. The results are reweighted by slice and by event within a slice to produce a final spectrum with the same shape as the true physical process but a very different distribution of statistical significances. The PYTHIA dijet sample has far higher statistics in the high-mass tail and lower statistics at low m_{jj} than does the data.

It should be clearly stated that even with the correct statistics, the success of the background estimate in MC would still be insufficient to predict its success in data. The following reasons have been discussed [212]:

- The MC does not adequately account for the experimental and detector effects seen in data.
- The MC is leading-order. K-factors were applied to make it compatible with next-to-leading-order predictions, but it ought to be at least at NNLO accuracy to make a meaningful prediction of the data shape.
- There are insufficient MC statistics in the central samples to match the statistical sensitivity of data at low m_{jj}.

Nonetheless, a failed attempt to parameterise data-like Monte Carlo statistics would at least provide a clear answer, where success provides tacit support for the background estimation process. Several methods were therefore explored for generating a simulated sample with data-like statistics.

The standard method for producing a data-like spectrum is to fit the full MC distribution and use the fit result as a probability distribution function from which to generate the desired amount of pseudodata. This clearly cannot provide information on stability of the fit function if the function itself was used in the generation of the spectrum. The alternatives are to either generate toys from a precise pdf obtained without the fit function or to extract a data-like distribution from the central MC production.

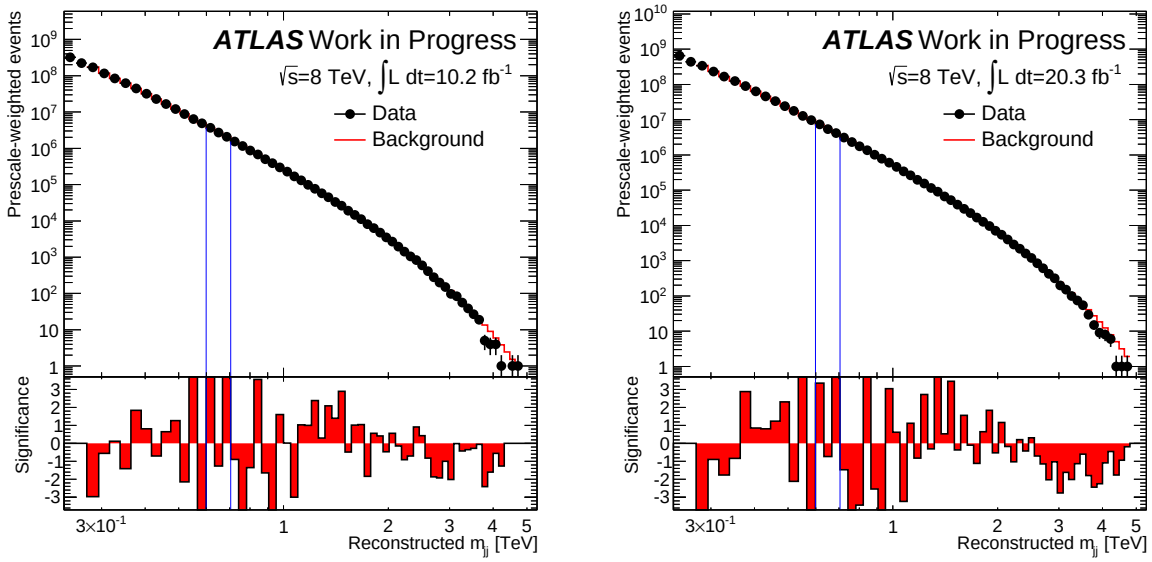
Generating pseudodata from parton-level distributions

A high-statistics sample of Pythia NLO MC was generated and used as a shape prediction for the distribution of pseudodata. However, when a selection of pseudodata equivalent to 20.3 fb^{-1} was generated using this NLO spectrum as a pdf, a fit to this partial spectrum revealed significant discrepancies in the low-mass region. This spectrum, its fit, and the corresponding residual are shown in Figure 7.3. While the overall shape is accurately described by the fit function, the significance of the discrepancies in bins at low mass are larger than physically expected. Repeatedly generation of different pseudodata spectra showed that for any random spectrum, the same bins were excessively high or low. This behaviour indicated that the large discrepancies were caused by the initial spectrum being insufficiently smooth to serve as an accurate pdf for the generation of such large numbers of events. Small statistical fluctuations in the NLO distribution were being propagated through the pseudodata generation and affecting the final distributions. Therefore, while this test demonstrated that the underlying shape of the NLO distribution could be well described by the background estimation procedure, the original curve was insufficiently precise to produce reliable data-like results.

Generating data-like MC spectra from ATLAS samples

An alternative is to select events from the central ATLAS MC samples such that they are have the same distribution of statistical power as the data. The Pythia dijet samples were used as inputs. Events from the MC samples used, which are generated such that the number of events is roughly flat across all p_{T} values, have an associated event-wise weight w_i to return the samples to a physical shape. Moreover, to normalise each slice to the desired luminosity $\mathcal{L}$, each full p_{T} sliced sample is given a weight $w_s = \mathcal{L} \cdot \sigma \cdot \varepsilon_{\text{filter}} / N_{\text{evt}}$. Here, σ and $\varepsilon_{\text{filter}}$ are the cross-section and filter efficiency of the MC generation and N_{evt} is the number of events in the sample. When the average expected prescale in a bin is p, then the fraction of events in a bin which should be selected from the full MC samples is:

$$\begin{aligned} f_i &= \frac{w_i \cdot w_s}{p_i} \\ &= \frac{w_i \cdot \mathcal{L} \cdot \sigma \cdot \varepsilon_{\text{filter}}}{N_{\text{evt}} \cdot p} \,. \end{aligned} \tag{7.23}$$



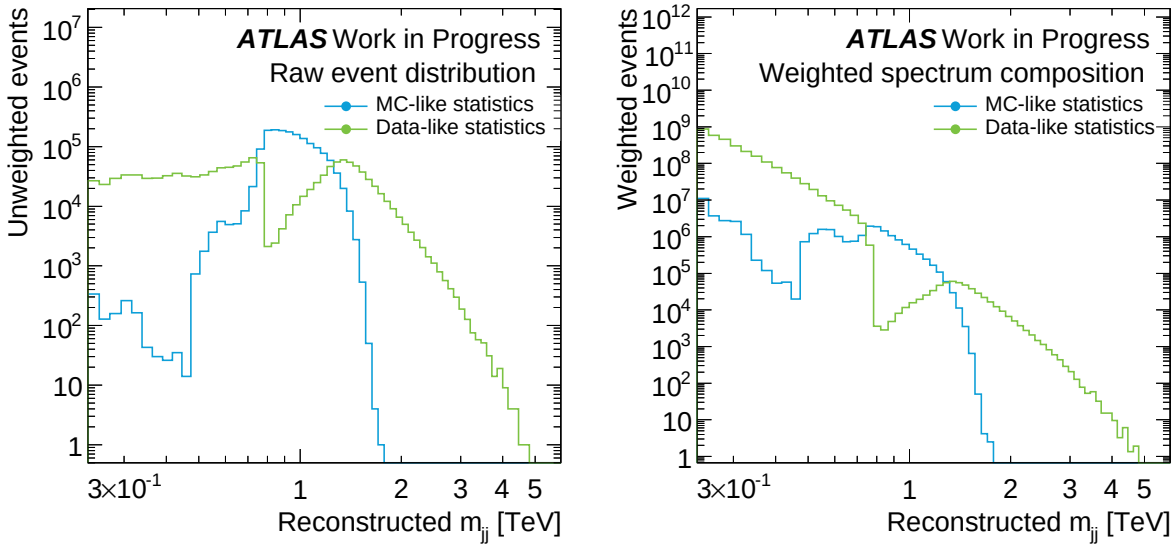
(a) Pseudodata spectrum, 10 fb^{-1}

(b) Pseudodata spectrum, 20 fb^{-1}

Figure 7.3: Fits are shown to two independent random pseudodata distributions drawn from a sample of NLO truth MC. While the two spectra differ in the high and low mass regions, the same bins are excessively high and low in both spectra in the region between 600 GeV and 1 TeV. The high-statistics truth sample is insufficiently smooth, so fluctuations in the truth distribution are being propagated to the pseudodata. These fluctuations are constraining the fit in the centre of the spectrum and causing a poor description of the tails.

If $f_i < 1$, then there would be sufficient MC statistics to reproduce data if all events were similarly weighted. A random number generator selects a value $r \in [0, 1]$ from a uniform distribution, and the event is kept if $r < f_i$. The event is added to the final spectrum with weight $1/p_i$, identical to a data event. If $f_i > 1$, the best alternative is to keep all available MC events and give them each an MC-appropriate weight to create the desired spectrum shape. At 20 fb^{-1}, slices JZ4W and below have regions with insufficient MC statistics to reproduce a data-like distribution. Spectra resulting from this process are thus composed of regions with events of weight 1 distributed like data ("data-like") and regions with event weights greater than 1 distributed nearly uniformly ("MC-like"). The combination does not perfectly reproduce the expected full data spectrum but is the best approximation available with limited MC samples. Figure 7.4 illustrates the regions of different statistical power within a spectrum selected to imitate a 20 fb^{-1} data sample.

Despite the variable statistics, a fit to the data-like selected MC spectrum at 20 fb^{-1} provides some critical information beyond that obtainable from the blinded spectrum. First, a successful fit demonstrates that the function can be expected to converge over the entire



(a) Weighted distribution by mc-like and data-like contributions

(b) Number of raw events treated as data-like and MC-like, by bin

Figure 7.4: Composition of the spectrum created by selecting MC events to mimic data statistics. Events where the MC statistics were insufficient to imitate data are illustrated in blue while events treated exactly like data are shown in green. Figure 7.4a shows the weighted spectrum divided into contributions of each type, while Figure 7.4b gives the raw number of events of each type per bin.

m_{jj} range from 250 GeV to above 4 TeV. Second, since data-like behaviour is accurately reproduced in the high-mass tail, a successful fit shows that the dijet function can adapt to the expected shape of a high-mass observation, the region least constrained by previous studies. This does not prove the data shape, as it would hide high-mass "tails" of new physics, but ensures a robust fit for the resonance search. The fit to the 20 fb^{-1} data-like MC spectrum is shown in Figure 7.5. This spectrum is indeed well described by the dijet function over its entire range and the function adapts successfully to the new kinematic range in the high-mass tail.

7.4 Validating the function choice

The choice of smooth background parameterisation needed to be decided upon while the data was still blinded. With only 1/4 of the dataset available, the standard function and a range of other functions, both historical and newly proposed, were tested and their quality of fits examined. No function with fewer than 4 parameters produced an adequate fit in the blinded dataset.

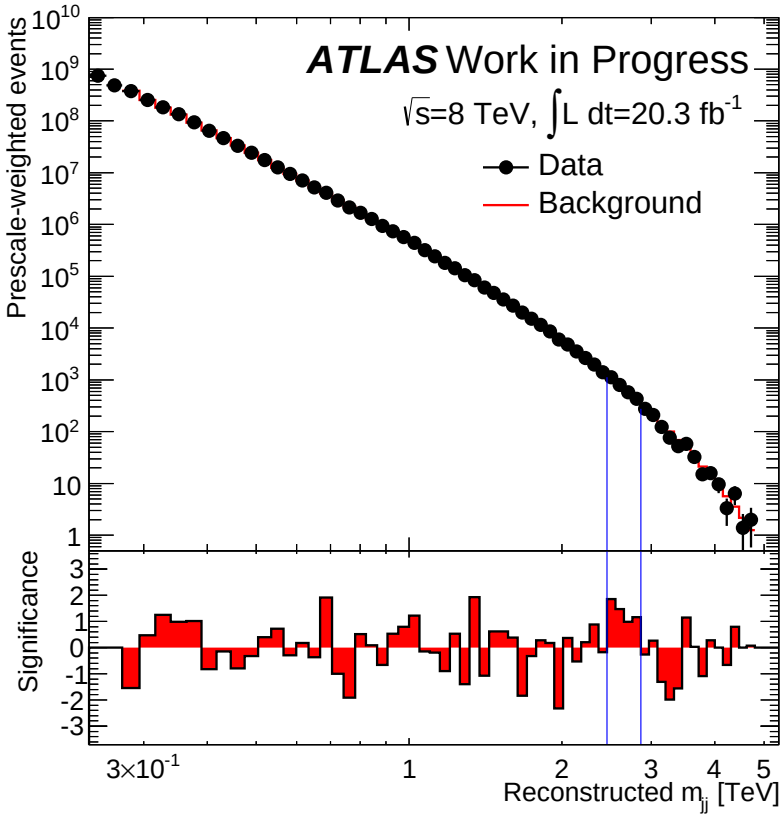


Figure 7.5: Fit to a spectrum with mixed data-like and MC-like statistics created by selecting events from the PYTHIA8 MC samples.

7.4.1 Alternate parameterisations

Table 7.1 summarises the tested functions. Any function with an unlisted experiment, year and citation has no published precedent and was developed by the analysis group to test the effect of additional parameters on the dijet function. The resulting fits to 1/4 of the full dataset are shown in Figure 7.6 while the reduced χ^2 and the p-value using each of the three test statistics are listed in Table 7.2.

Functions I, II, III, and V exhibited poor χ^2/NDF values and BUMPHUNTER $p = 0$ on the blinded dataset and thus cannot be relied upon to adequately describe the full 8 TeV dataset. The low p-values and weak goodness of fit tests indicate that the functions are not able to adapt to the shape of the data, and from the figures this can be seen particularly in the tails and in the mass region near 700 GeV where the data statistics are highest. The only remaining functions with a previous publication history was the standard dijet function. It, however, demonstrated worse agreement with the spectrum than did the func-

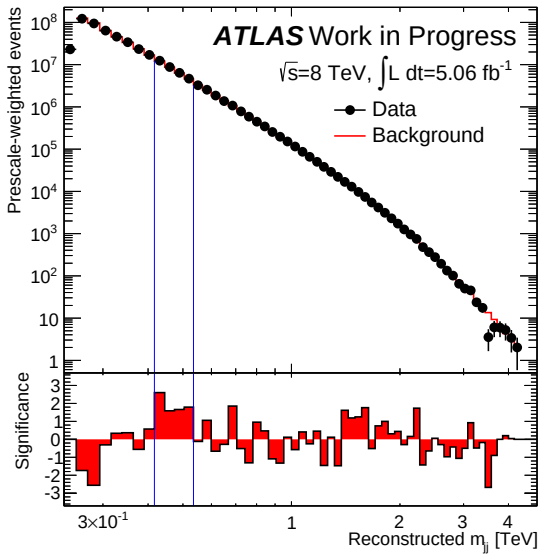
ID number	Function form	DOF	Experiment	Year	Citation
I	$\frac{p_0}{m^{p_1}} e^{-(p_2 m + p_3 m^2)}$	4	UA2	1990	[213]
II	$\frac{p_0}{m^{p_1}}(1 - \frac{m}{\sqrt{s}})^{p_2}$	3	CDF	1995	[214]
III	$\frac{p_0}{m^{p_1}}(1 - \frac{m}{\sqrt{s}} + \frac{p_3 m^2}{s})^{p_2}$	4	CDF	1997	[215]
IV	$p_0(1-x)^{p_1} x^{p_2 + p_3 \ln(x)}$ for $x = \frac{m}{\sqrt{s}}$	4	CDF	2009	[207]
V	$(1-x)^{p_0} x^{p_1 + p_2 \ln(x)}$ for $x = \frac{m}{\sqrt{s}}$	3	ATLAS	2014	[216]
VI	$p_0(1-x)^{p_1} x^{p_2 + p_3 \ln(x)}$ for $x = \frac{m}{p_4}$	5	–	–	–
VII	$p_0(1-x)^{p_1} x^{p_2 + p_3 \ln(x) + p_4 \ln(x)^2}$ for $x = \frac{m}{\sqrt{s}}$	5	–	–	–
VIII	$p_0(1-x)^{p_1} x^{p_2 + p_3 \ln(x) + p_4 \ln(x)^2}$ for $x = \frac{m}{p_5}$	6	–	–	–

Table 7.1: Selection of smooth functional forms used or proposed for describing the dijet combinatorial background. The first column shows the functional form, the second lists the number of parameters (not necessarily independent) and thus the number of degrees of freedom of the function, the third and fourth columns give the experiment and year when this function was first used in a dijet search, and the fifth column provides a reference for the first publication using the function. Where the function has no publication history the final columns are left blank.

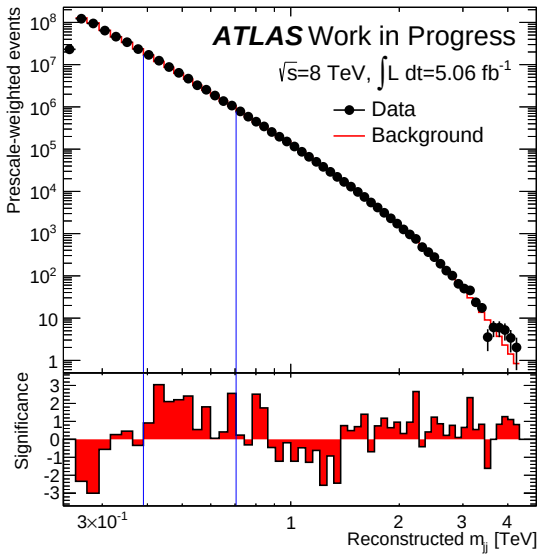
Fit function	χ^2/NDF	χ^2 p-value	log $\mathcal{L}$ p-value	BumpHunter p-value
I	1.24	0.063	0.097	0.0
II	2.00	0.0	0.0	0.0
III	1.32	0.033	0.050	0.0
IV	0.959	0.48	0.53	0.18
V	4.83	0.0	0.0	0.0
VI	0.947	0.51	0.62	0.37
VI	0.879	0.64	0.71	0.83
VIII	0.215	0.73	0.43	0.83

Table 7.2: Measurements of goodness of fit for eight functions to the m_{jj} spectrum in partial 2012 data. p-values are calculated using 1000 pseudoexperiments.

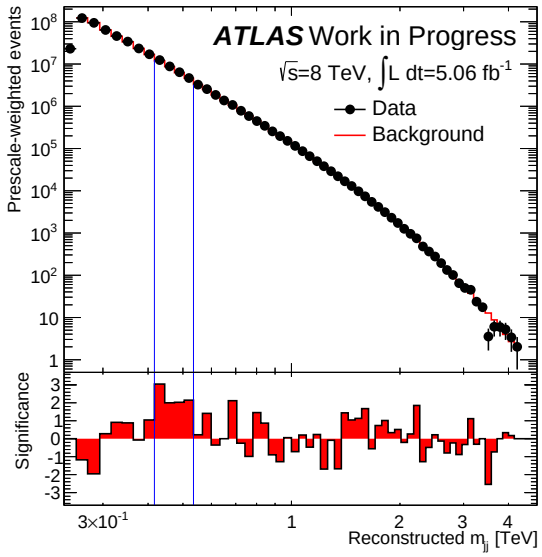
tions with higher numbers of parameters. While testing the various functions in blinded data adequately demonstrated which ones were not suitable for use in the full analysis, it did not provide enough information to determine which of the remaining functions was the most appropriate choice. For that, a cross-validation as described in the next section was performed on the best-performing 3-parameter function (Function II), the four-parameter standard dijet function, and the three new functions.



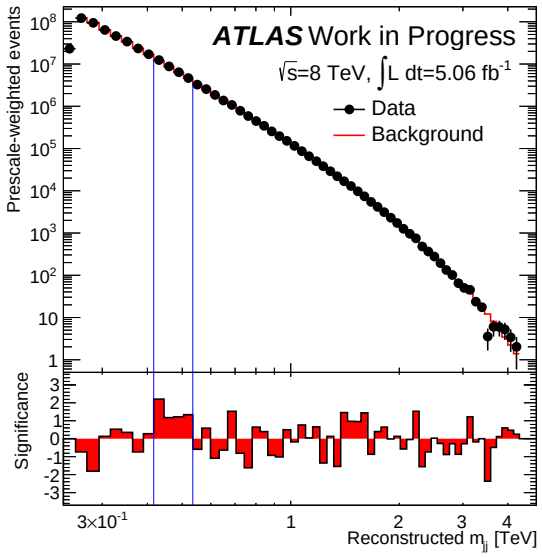
(a) Function I



(b) Function II



(c) Function II

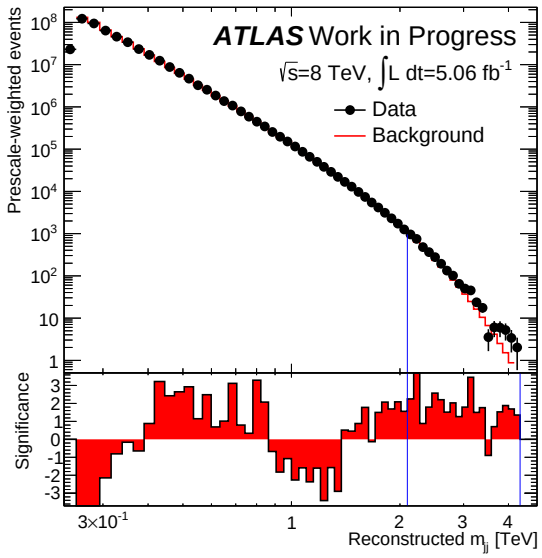


(d) Function IV (Dijet)

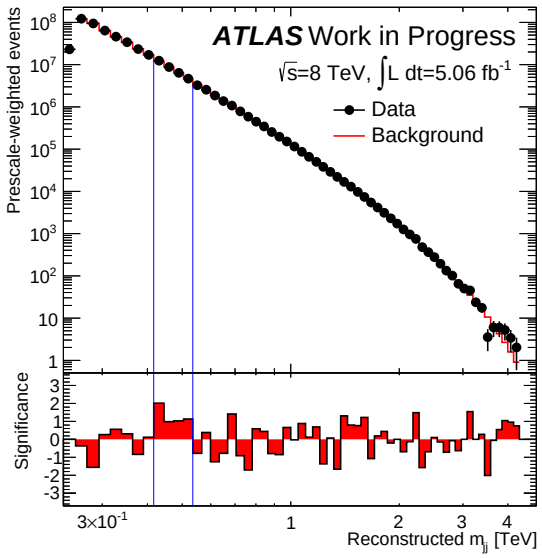
Figure 7.6: Fit, residual, and BumpHunter selected interval using the functions whose forms are given in Table 7.1. The blue lines indicate the region of most extreme excess. All results are shown in partial data (1/4 of the 2012 dataset).

7.4.2 k-fold cross-validation in the full dataset

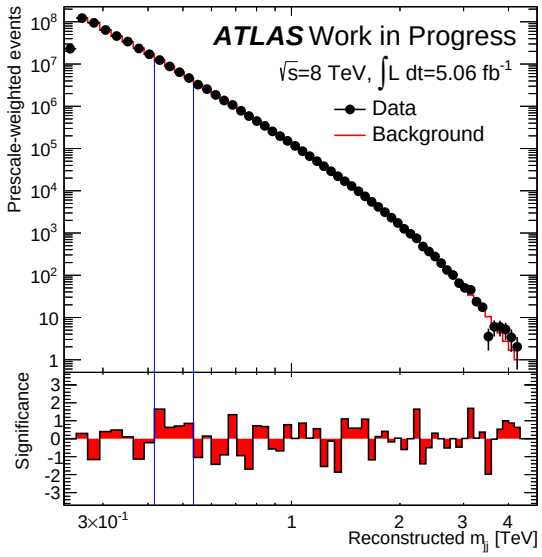
The information from a fit to the data alone is not sufficient to make any statement about which function is the most appropriate choice for describing a dataset. If the chosen parameterisation has more degrees of freedom than does the underlying distribution, then the function may be overconstrained so that some of the features to which it is adapting are



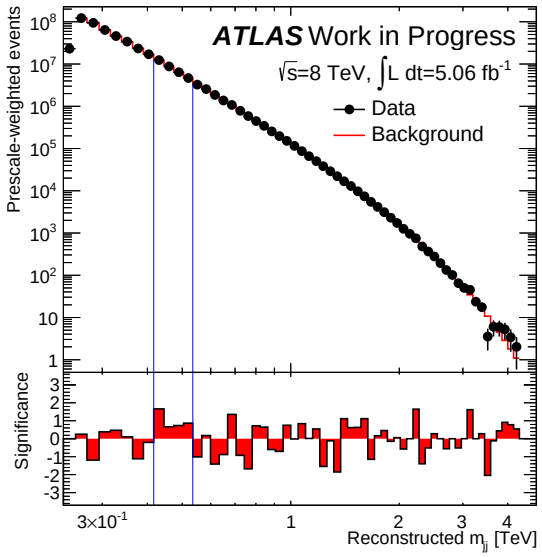
(e) Function V



(f) Function VI



(g) Function VII



(h) Function VIII

Figure 7.6: Fit, residual, and BumpHunter selected interval using the functions whose forms are given in Table 7.1. All results are shown in partial data (1/4 of the 2012 dataset).

statistical fluctuations instead of real physical features. This is tested for using *k-fold cross-validation*: the dataset T is split into k equal subsets T_k and each set T/T_k (the full data less one subset, or the k^{th} *training set*) is individually fit by the function being tested. The result of the fit is compared to T_k, which serves as the *validation set*. By repeating this using each of the k subsets for validation and averaging the resulting p-values, the ability of the function to predict the data is found. In the case that the function being tested has more degrees of freedom than the data, it will fit out statistical fluctuations in the training sets

and produce results which is less predictive of the validation sets. Where there are too few degrees of freedom, the real underlying physical features will not be adequately described. Thus the function which produces the best results in the testing data on average is the most appropriate choice for creating the background estimate [206].

A range of k values are used during the cross-validation to enable balancing of two factors. An overly high k will ensure a very small validation set dominated by statistical fluctuations and therefore without enough detail about underlying physical processes to distinguish between functions. An overly small k, in contrast, limits the number of separate training sets and thereby the stability of the prediction [217]. The most common selection to achieve this compromise is $k = 10$, which is also treated as the default value in this analysis.

A cross-validation was performed according to the following algorithm.

1. The full 2012 dataset was divided into k independent sections and the trigger streams combined according to a scale of $1/k$ times the full luminosity.
2. $k - 1$ of the k sub-datasets are added into a single training sample, then fit with a selected function to obtain a background estimate.
3. The background estimate is scaled by $1/(k - 1)$ to provide a prediction of the test sample.
4. The prediction is directly compared to the test sample without re-fitting. The χ^2/NDF and p-values of the comparison are recorded.
5. The above is repeated leaving out each of the other $k - 1$ subsets in turn. k different p -values are thus obtained, one for each validation set.
6. An average of the p-values from the k tests provides a measure of the predictive capacity of the function.

This process was repeated for each of the functions II, IV, V, VI, VII, and VIII; that is, the four-parameter function of choice, the two best-performing three-parameter functions, and the five-parameter and six-parameter functions first explored in this analysis. While the k-fold procedure is meant to determine the number of degrees of freedom appropriate in describing a dataset by determining the point at which it transitions from being poorly described to overfitted, in this context it was found that the number of parameters in the function did not directly correlate to the freedom of the description. For instance, the two

different variations on the standard dijet function labeled VI and VII vary in their predictive capacity despite sharing five fit parameters because allowing the centre-of-mass energy to float provides less flexibility in the overall function shape than permitting an additional $\ln(x)^2$ term in the function exponent. Thus in practice the test provides more information about individual functions than about the number of parameters they contain.

The full cross-validation results for $k = 10$ and the above list of functions are given in the following tables. Tables 7.3 through 7.6 list the results for each training set with functions II, IV, VII and VIII, showing the variability in prediction results between validation sets and giving one example for each number of function parameters tested. Table 7.7 provides the averaged values across validation sets for each of the six functions tested. A comparison between the fit to a test sample and the prediction for the corresponding validation sample is given for Function IV in Figure 7.7.

Sample	χ^2/NDF	LogL	LogL p-value	χ^2	χ^2 p-value	BH stat.	BH p-value
1	1.79	300.5	0.002	102.1	0.001	7.23	0.101
2	1.30	290.8	0.036	74.3	0.041	10.06	0.006
3	1.33	291.4	0.041	75.9	0.039	6.72	0.124
4	1.47	295.4	0.007	83.7	0.014	9.58	0.009
5	1.21	286.9	0.144	69.2	0.127	9.89	0.008
6	1.53	298.8	0.001	87.3	0.006	5.15	0.395
7	1.54	295.6	0.009	87.8	0.007	11.11	0.003
8	1.30	287.1	0.131	73.9	0.075	6.25	0.183
9	1.36	291.3	0.042	77.4	0.034	12.64	0.0
10	0.86	277.0	0.771	49.1	0.688	10.18	0.004

Table 7.3: Test statistics and their p-values for validation samples 1 to 10 using Function II, the three-parameter CDF function of 1995. $\chi^2/NDF > 1.2$ in 9 out of 10 cases.

The results of the k-fold cross-validation demonstrate that the data is under-described by both three-parameter functions. The six-parameter function offers no advantage over the four- and five-parameter functions, and by permitting an additional degree of freedom runs the risk of overfitting by adapting to statistical fluctuations in the training sets which do not represent the true distribution of the validation data. Of the three remaining functions, the results indicate that Functions IV and VI better describe the data by a narrow margin over Function V. Of the two five-parameter functions, this suggests that the additional factor

Sample	χ^2/NDF	LogL	LogL p-value	χ^2	χ^2 p-value	BH stat.	BH p-value
1	1.36	292.9	0.041	76.3	0.037	5.38	0.331
2	0.93	281.5	0.598	52.1	0.593	6.05	0.219
3	0.82	277.8	0.831	46.0	0.792	3.81	0.782
4	0.92	280.6	0.667	51.5	0.613	5.37	0.346
5	0.84	278.1	0.802	46.9	0.797	4.72	0.531
6	1.24	291.5	0.069	69.5	0.102	4.40	0.611
7	1.14	286.5	0.237	63.7	0.203	7.90	0.055
8	1.02	284.1	0.408	57.1	0.416	3.98	0.738
9	0.97	283.5	0.451	54.3	0.505	8.44	0.031
10	0.70	275.4	0.953	39.0	0.954	4.99	0.441

Table 7.4: Test statistics and their p-values for validation samples 1 to 10 using Function IV, the standard dijet function. χ^2/NDF varies between 0.82 and 1.36.

Sample	χ^2/NDF	LogL	LogL p-value	χ^2	χ^2 p-value	BH stat.	BH p-value
1	1.46	291.4	0.037	80.2	0.015	6.57	0.148
2	0.91	279.1	0.675	49.9	0.679	5.30	0.376
3	0.72	273.7	0.935	39.7	0.932	3.20	0.914
4	0.88	278.1	0.742	48.3	0.731	3.24	0.925
5	0.83	276.0	0.84	45.6	0.817	4.17	0.702
6	1.18	288.0	0.14	64.7	0.201	3.40	0.876
7	1.11	283.4	0.376	60.8	0.3	5.53	0.298
8	1.24	286.0	0.221	68.4	0.113	5.20	0.411
9	1.05	282.8	0.38	57.6	0.377	7.48	0.072
10	0.751	274.5	0.938	41.3	0.903	5.03	0.403

Table 7.5: Test statistics and their p-values for validation samples 1 to 10 using Function VII, the five-parameter variation on the standard function which introduces an $e^{[p]\ln(x)^2}$ term. χ^2/NDF has a wider range than in the four-parameter case, varying between 0.75 and 1.45.

fitting the tail is better adapted to the needs of the data than replacing the centre-of-mass energy with a free parameter. However, between Functions IV and VI no distinct preference was evident. The χ^2/NDF has a wider distribution around 1 in the five-parameter case, but the 10 samples have too low a statistical significance for this to be a substantial piece of evidence in favour of the standard fit. Both functions are considered good candidates for the background description.

Sample	χ^2/NDF	LogL	LogL p-value	χ^2	χ^2 p-value	BH stat.	BH p-value
1	1.48	291.4	0.034	80.0	0.015	6.54	0.141
2	0.92	279.2	0.675	49.7	0.685	5.36	0.357
3	0.73	273.6	0.934	39.5	0.933	3.23	0.927
4	0.89	278.1	0.739	48.2	0.723	3.22	0.929
5	0.84	276.0	0.827	45.5	0.804	4.19	0.669
6	1.20	288.0	0.136	64.6	0.192	3.40	0.87
7	1.13	283.4	0.375	61.3	0.298	5.58	0.299
8	1.30	286.1	0.193	70.1	0.097	5.40	0.321
9	1.06	282.9	0.400	57.4	0.399	7.52	0.088
10	0.76	274.5	0.935	41.2	0.912	5.01	0.446

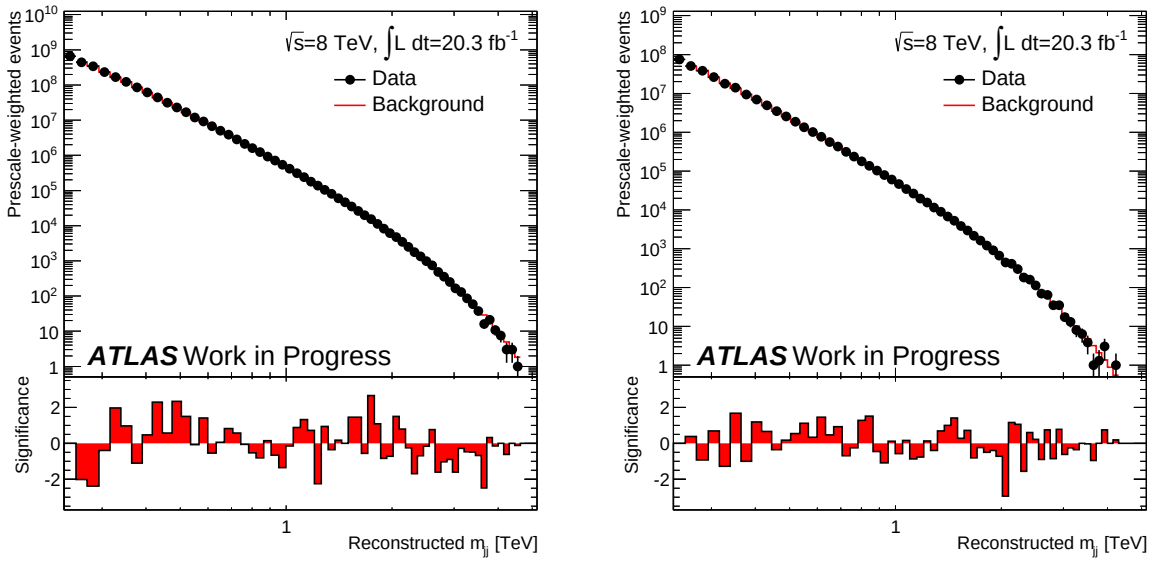
Table 7.6: Test statistics and their p-values for validation samples 1 to 10 using Function VIII, the six-parameter variation on the standard function. χ^2/NDF varies between 0.73 and 1.48 and the individual results are very similar to those obtained with Function VII.

Function	χ^2/NDF	$\lvert\chi^2/NDF - 1\rvert$	Log L p-value	χ^2 p-value	BH p-value
II	1.36984	0.397494	0.1184	0.1032	0.0833
IV	0.993373	0.158935	0.5057	0.5012	0.4085
V	1.52833	0.528334	0.0577	0.049	0.0486
VI	1.08621	0.243198	0.4929	0.4704	0.4567
VII	1.01156	0.194501	0.5284	0.5068	0.5125
VIII	1.03236	0.202306	0.5248	0.5058	0.5047

Table 7.7: Measures of goodness of fit averaged over the 10 validation samples for each of the six functions. Both the mean χ^2/NDF and its mean deviation from 1 (third column, $\lvert\chi^2/NDF - 1\rvert$) are shown. Where the χ^2/NDF falls close to 1 for several functions, the width of its distribution about that provides a further measure of its stability. The following three columns hold the mean p-value for each test statistic across the 10 sets.

7.5 Calculating the error on the background estimate

Deriving a background estimate by fitting the data distribution introduces two independent uncertainties. The first is on the fit result itself and is related to the confidence intervals on the parameter values estimated by the fitting process. The second uncertainty corresponds to the choice of parameterisation. With a range of plausible parameterisations of the data, a variety of different background estimates are all acceptable descriptions of the data. The size of the error corresponding to each source was calculated as described below.



(a) Training set (b) Validation set

Figure 7.7: Fit to a training set with the standard dijet function (Function IV) and the resulting prediction compared to the validation set. The χ^2/NDF for this prediction (validation sample 5) is 0.84.

7.5.1 Uncertainty on the fit result

When the best fit quantity of a parameter is calculated there is always a corresponding variance associated with the statistical uncertainty on the measurement. A confidence interval is typically chosen as the form of uncertainty to report, assigning a range of values within which the best-fit value of the parameter would be expected to lie in 68.3% of repeated experiments with different statistical fluctuations. In some cases it is possible to compute such an interval analytically, but for the fit parameters of the dijet function this approach is unfeasible. Instead, a Monte Carlo approach or an approximate computation using the likelihood function is used.

MINUIT calculates an uncertainty on each fitted parameter using the Rao-Cramér-Fréchet inequality, which sets a minimum bound on the variance of a parameter. Assuming an unbiased estimator and assuming the variances are at the minimum boundary, the inverse of the covariance matrix elements is given by:

$$(V^{-1})_{ij} = E[-\frac{\partial^2 \log L}{\partial\theta_i\theta_j}]\,. \tag{7.24}$$

$E[-\frac{\partial^2 \log L}{\partial\theta_i\theta_j}]$ is the expectation value of the second derivative of the log likelihood in the parameter of interest. For $i = j$, the matrix element represents the covariance of θ_i. Off-diagonal terms give the covariance, a dimensionful quantity related to the correlation, of the two parameters. An approximation of the derivatives of the likelihood function thus provides a measure of all parameter variances [210].

However, since MINUIT cannot calculate the second derivative of the likelyhood function analytically, it makes an approximation using the measured data and the maximum likelihood estimators of the parameters $\hat{\theta}$:

$$(\hat{V^{-1}})_{ij} = -\left.\frac{\partial^2 \log L}{\partial\theta_i\theta_j}\right|_{\theta=\hat{\theta}}. \tag{7.25}$$

In practice, the second derivatives around the maximum likelihood values are estimated using the finite difference method [209]. That is, the derivatives around the ML estimators are approximated using a Taylor series expansion and the values values of the function (here the log likelihood) at neighbouring points. As long as the log-likelihood is smoothly varying such that this approximation holds within appropriate uncertainties, the covariance matrix elements are well-calculated. A complication is however introduced since MINUIT is incapable of assessing the reliability of its error estimates. There are few ways for an analyst to determine whether or not the estimate is reliable in a given case, other than comparing several forms of function uncertainty estimation to ensure they are compatible.

Error estimation using the covariance matrix

Propagation of the uncertainty on the parameters to an uncertainty on the fit itself follows the equation:

$$\sigma(f) = \sum_{i=0}^{N} |\frac{\delta f}{\delta\theta_i}|^2 \sigma_{\theta i}^2 + 2\sum_{i!=j} \sigma_i\sigma_j \mathrm{cov}_{i,j} \tag{7.26}$$

This simple equation provides a fast and mathematical value for the uncertainty band surrounding a fitted function if the covariance matrix is well known.

An uncertainty determination was implemented using this approach and was found to work well in the background fit using the standard dijet function, where the log likelihood has a broad and smooth maximum. Tests of this method compared to the pseudoexperiment method below showed similar estimates. However, when an additional parameter is added

to the fit in order to account for the normalisation of a possible signal (as in the limit setting phase, Chapter 9), the range of the new parameter has a sharp cutoff at 0 and an ML estimator usually very close in parameter space. The result is a non-smooth region around the minimum and a poorly defined covariance matrix, leading to an unreliable uncertainty estimate. In order to ensure a stable uncertainty measurement for any shape of likelihood function, the covariance matrix estimation was abandoned in favour of an estimation from pseudoexperiments.

Error estimation from pseudoexperiments

An alternate method for uncertainty computation is evident from the physical interpretation of the variance. Since σ_θ is the amount by which $\hat{\theta}$ is expected to vary in repeated experiments, its propagation to the function produces error bands representing the range within which the function would be expected to lie in 68% of cases.

This result can easily be computed using repeated pseudoexperiments. The fit to the true data is used to generate pseudodata, and each pseudospectrum is fitted by the same function with the same starting parameters and conditions. For each bin, the set of values taken by the function in a fit to a pseudoexperiment is recorded. With a sufficiently high number of toys this provides a spectrum of fit values, of which the central 68.3% are taken to define the uncertainty on the function in that bin.

Tests in cases with a well-behaved MINUIT error calculation indicate that this process does accurately reproduce the statistical uncertainty on the fit result. All fit uncertainties used in this analysis are results of the pseudoexperiment method and are thus reliable for any shape of likelihood function.

7.5.2 Uncertainty on the choice of function

An additional uncertainty arises not from the fit of the function to the data from the choice of fit function itself. As discussed in Section 7.4.1, a wide range of possible functions can converge on the dijet spectrum, and the resulting background estimates can vary substantially.

No standard procedure exists for assigning an uncertainty to this situation. The solution

determined for the dijet analysis was to select two parameterisations, one *nominal function* to create the official background estimate (the standard dijet function, with a long publication history) and one *alternate function* to provide an second background estimate for determining an uncertainty in the limit-setting phase of the analysis. The definition of the uncertainty will be discussed in the dedicated systematics chapter. The selection of the alternate function was made from the k-fold cross-validation results and the observed fits of all considered functions to the blinded dataset. Since it performed second-best after the nominal function, the five-parameter function containing an exponential $\ln(x)^2$ term (Function VII) was selected for this purpose.

7.6 Background stability against signal contamination

To demonstrate that the background estimate is reliable it must be shown that it is not biased by the presence of signal. In the case of a sufficiently large bump, that is, a bump which causes the initial fit compared to data to have $p < 0.01$, a set of bins is selected and excluded as discussed in Subsection 7.2.2. If the background estimate is properly described by the errors assigned to the fit, then the fit to a background containing signal ought to remain within a distance of no more than a few times the error bars away from this prediction when a window is excluded or not. Since the window exclusion is defined by removing as many bins as necessary to return the remainder of the fit to a p–value of at least 0.01, the fit in this case tends to return to its initial value. There are two identifiable regimes in which the greatest discrepancy can be expected between a fit in the presence of signal and a fit to a background-only distribution:

1. A very, very large signal is present, such that the low-mass and high-mass tails not removed by the window exclusion are large enough to change the background estimate.
2. A signal is present which is almost large enough to cause a window exclusion, but not quite. At this point, the fit should be at its most biased by the presence of a bump, because once an exclusion in the fit is begun,the most discrepant region is ignored for all larger signals.

Since the blinding strategy consisted of looking at partial data and no excess was observed at that point, the first condition is less likely than the second. Both, however, have been

considered here in an attempt to validate the background estimate procedure as thoroughly as possible.

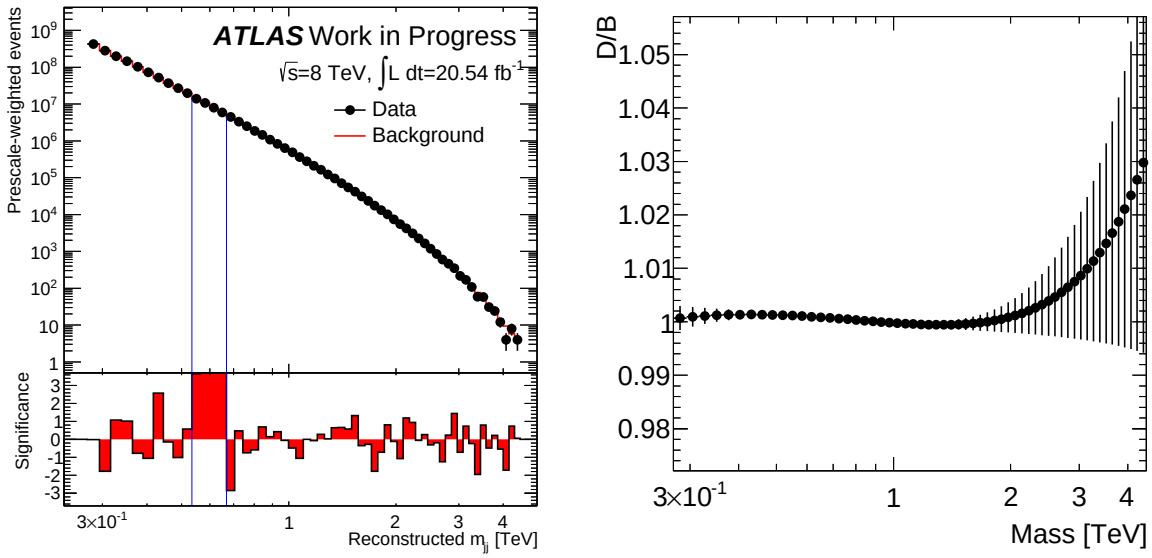
7.6.1 Signal contamination: large

Studies of the stability of the background prediction in the presence of very large signals used Monte Carlo signal templates for three different models superimposed on a pseudodata background. The q^*, W', and $s8$ signals were selected due to their varying widths. For each signal at each MC sample mass available, pseudodata was thrown using the template as a guide up to the number of events dictated by the nominal cross-section of the signal. This was superimposed on a background spectrum generated using a fit to PYTHIA MC as an ideal distribution and drawing pseudodata equivalent to 20.5 fb^{-1} from its shape.

The search phase algorithm was used to fit the combination of pseudo-background and pseudo-signal, leading in each case to the exclusion of a range of bins surrounding the injected signal peak. For each case, the background estimate without signal present was also computed by a fit to the pseudo-background alone. The ratio between the two background fits then provided a measure of the deviation undergone by the fit despite the window removal procedure.

In the ideal case, the search algorithm would remove enough bins encompassing excesses due to signal that the background estimate would not be biased by their presence, remaining at the same height. A counterbalancing requirement forces the exclusion of as few bins as possible: the background estimate will certainly change due to the loss of information on some large region of the spectrum. As some information must always be excluded with the window, a compromise must be enforced between the two conditions. Fits to the background distribution with three different signal models and masses are shown in Figure 7.8. In the very high-mass tails, the nominal cross-section of most signals is at the barely discoverable level, and is therefore more appropriately covered by the tests in the following section. Therefore the demonstration masses selected here are all such that a window is excluded from the background fit to accommodate the signals. The ratio plots on the right of each figure illustrate the ratio between the fits to background with and without injected pseudo-signal.

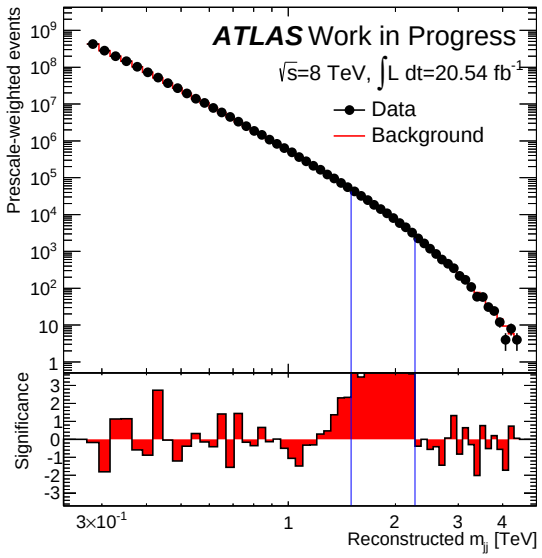
For each of the three sample signals at nominal cross section, the deviation in the background fit for every available mass is overlaid in Figure 7.9. Several examples are found



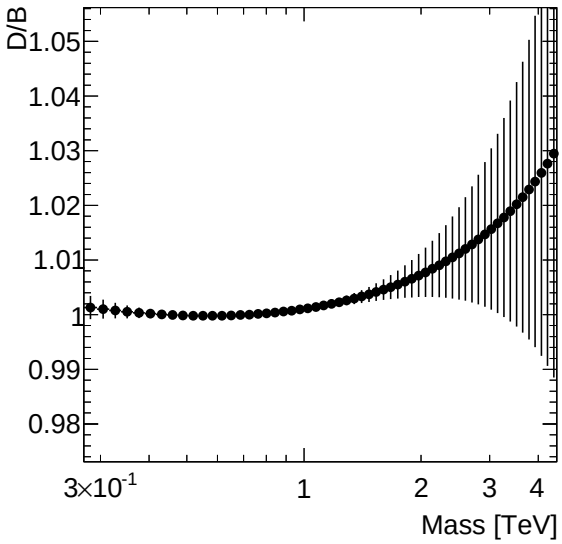
(a) Fit to pseudo-background with 600 GeV W' signal (b) Ratio of fit with and without W' signal

Figure 7.8: Effect on the background prediction of signals generated from MC templates with their nominal cross sections. Three representative masses and signals are given: a W' sample at 600 GeV, an $s8$ sample at 2000 GeV, and a q^* sample at 3000 GeV. The underlying distribution is pseudodata generated from a fit to MC. In the left-hand plots, the comparison between pseudodata and fit is shown with the largest bump outlined in blue. Right-hand plots show the ratio of the fit to background plus signal over the fit to background alone.

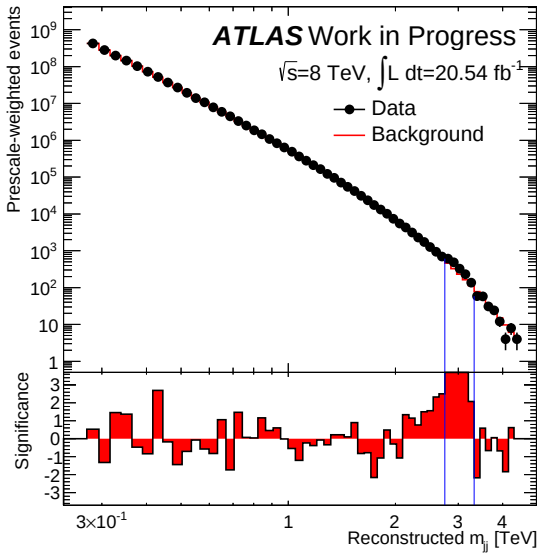
where the search phase algorithm is insufficient to account for the addition of an immense signal (q^* 400 GeV and 1000 GeV, three lowest-mass $s8$ samples), all of which correspond to the same physical conditions. Here, the signal is low enough in the mass range to have an enormous cross-section relative to QCD dijet production, but enough bins remain below the signal that instead of the low-mass tail of the function being biased upwards the centre of the function is pulled upwards and the tails forced to deviate sufficiently far from the data that the region of most significant discrepancy is in the first few bins of the spectrum rather than near the actual signal mass. No such scenario is possible in the current search, where generic resonant signals have already been excluded by the 5.06 fb^{-1} used in the blinded dataset to higher masses than these. However, it raises a flag for the upcoming 13 TeV analysis. No iteration of the dijet analysis to date has employed a search phase capable of producing an unbiased background estimate when the signal is sufficiently large that the region of greatest difference between the fit and data does not align with the new physics. Since this scenario is possible in the first analysis of Run II, the possibility must be seriously considered when selecting a search procedure. In every case where the region of greatest discrepancy between data and fit corresponds to the injected signal, the ratio between the



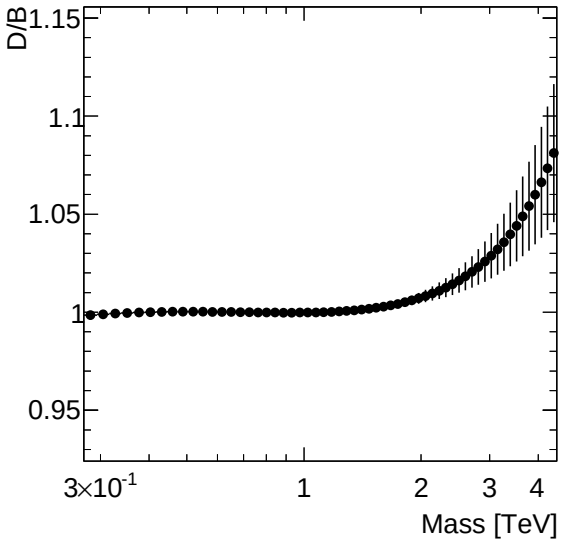
(c) Fit to pseudo-background with 2000 GeV $s8$ signal



(d) Ratio of fit with and without $s8$ signal



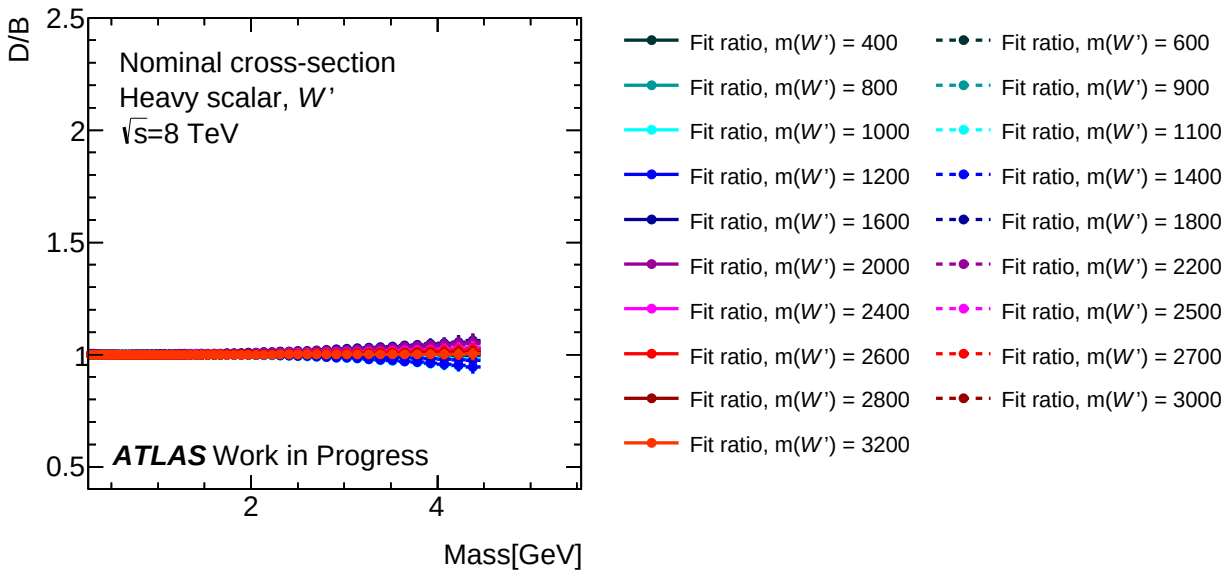
(e) Fit to pseudo-background with 4500 GeV q^* signal



(f) Ratio of fit with and without q^* signal

Figure 7.8: Effect on the background prediction of signals generated from MC templates with their nominal cross sections. Three representative masses and signals are given: a W' sample at 600 GeV, an $s8$ sample at 2000 GeV, and a q^* sample at 3000 GeV. The underlying distribution is pseudodata generated from a fit to MC. In the left-hand plots, the comparison between pseudodata and fit is shown with the largest bump outlined in blue. Right-hand plots show the ratio of the fit to background plus signal over the fit to background alone.

fits does not vary beyond 0.9 – 1.2.



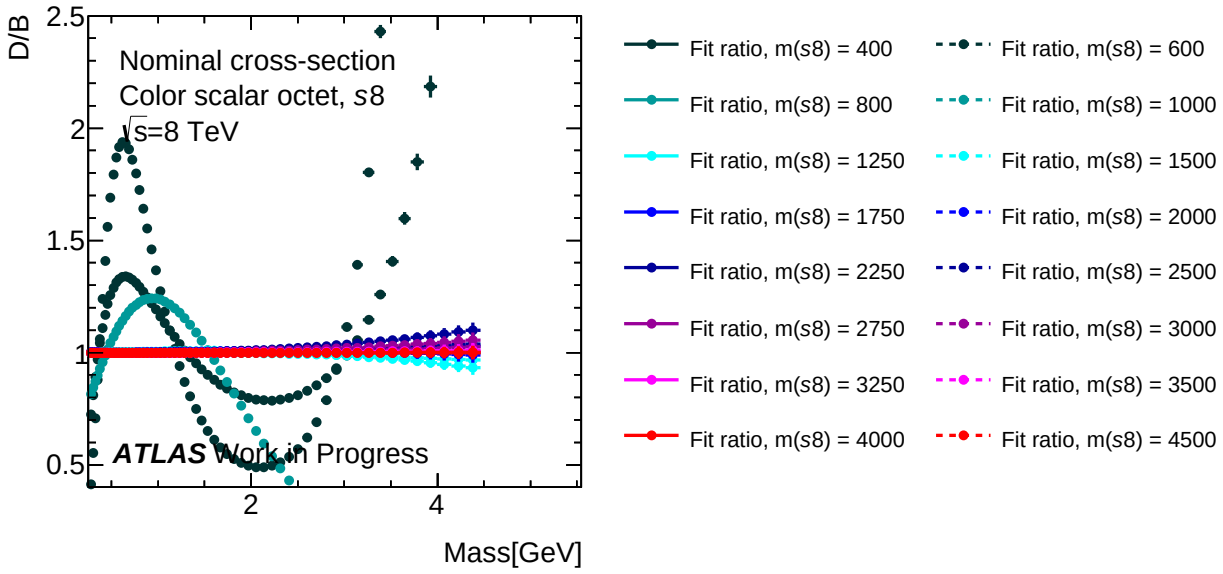
(a) W' signal at nominal cross-section

Figure 7.9: Ratios of fits to pseudo-background plus signal over pseudo-background alone for a large range of masses. The three plots represent W', $s8$, and q^* signals respectively.

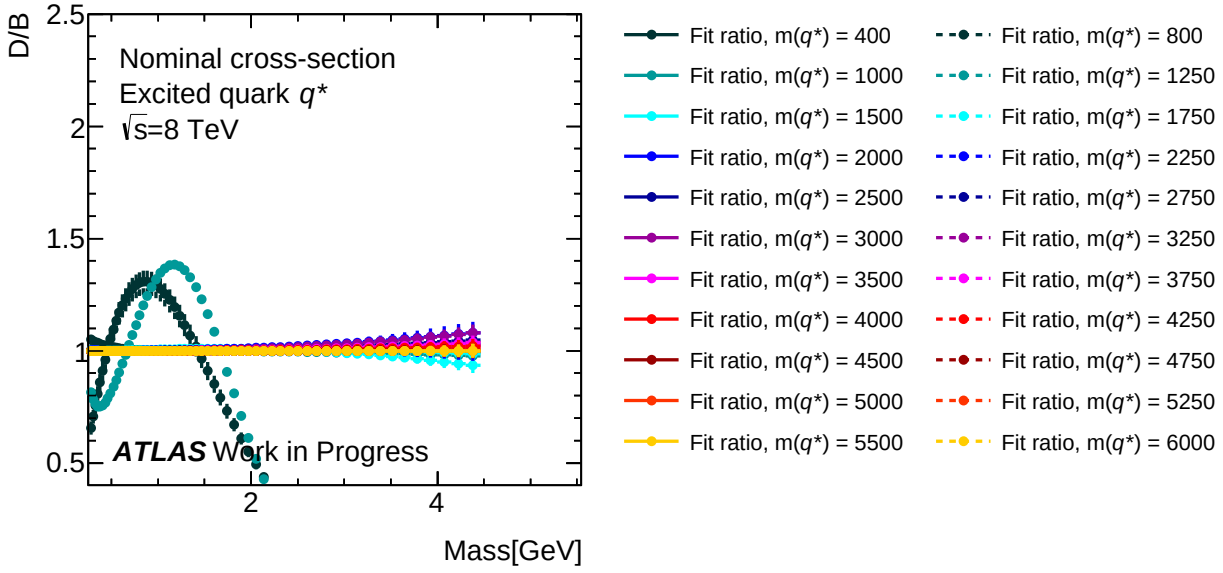
7.6.2 Signal contamination: small

The stability of the background prediction in the presence of nearly-detectable signal quantities was tested using generic resonance signal shapes and simulated background. The same background spectra generated for the previous test were used, consisting of pseudodata thrown using a fit to Pythia MC as a guide. For the simulated signal, Breit-Wigner functions at various masses and of various width-to-mass ratios ranging from 0.07 to 0.20 were generated and used as theoretical predictions from which to draw pseudodata in the desired amount.

For each mass point and width of signal, an iterative procedure was used to calculate a normalisation of the signal sample which, when added to the background, was fit with a p -value falling between 0.010 and 0.012. In certain cases, for example very high masses, the requirement that the number of signal events also be an integer value made it impossible to pinpoint this value so precisely; in that case, the normalisation producing the nearest p -value to 0.01 was used. In each case, the final fit to the pseudodata plus signal sample was compared to the same fit to a background-only sample in order to determine the degree to which the background prediction was affected. Figure 7.10 shows the fit results to the pseudo-background with injected signal and the ratio of the fits with and without signal for



(b) $s8$ signal at nominal cross-section

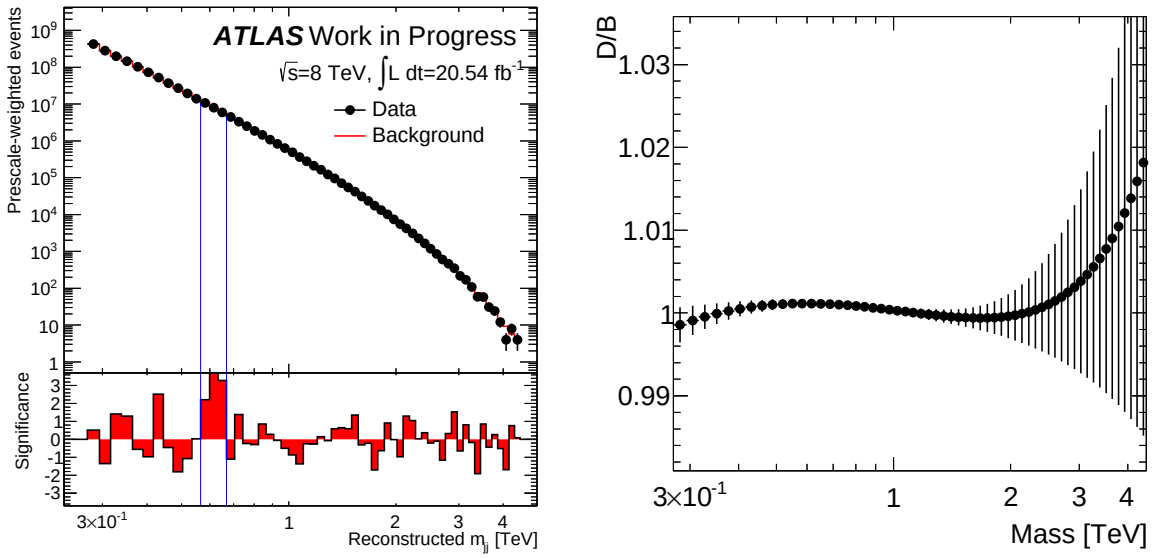


(c) q^* signal at nominal cross-section

Figure 7.9: Ratios of fits to pseudo-background plus signal over pseudo-background alone for a large range of masses. The three plots represent W', $s8$, and q^* signals respectively.

three representative masses.

From the preceding plots it can be seen that the signal has the greatest effect on the fit function in the tails, where the background prediction can increase or decrease by a substantial fraction as a result of the added events. The size of the error bands represents a change of 1σ in the fit. For every case shown, the fit deviates no farther than $\sim 3\sigma$ from its ideal value. In fact, an overlay of the fit ratio for every mass tested shows that in no



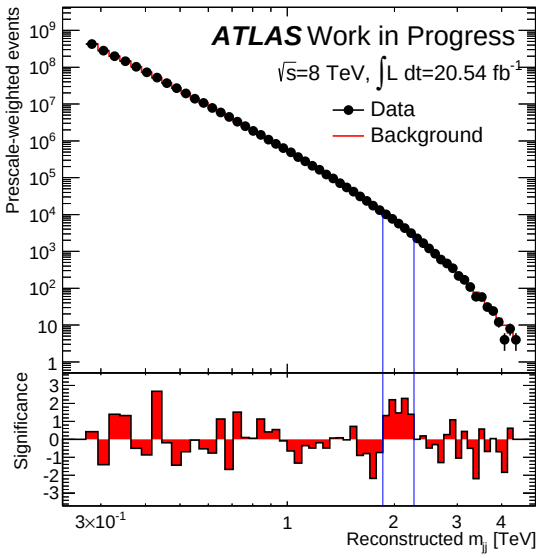
(a) Fit to pseudo-background with 600 GeV Breit-Wigner signal

(b) Ratio of fit with and without 600 GeV signal

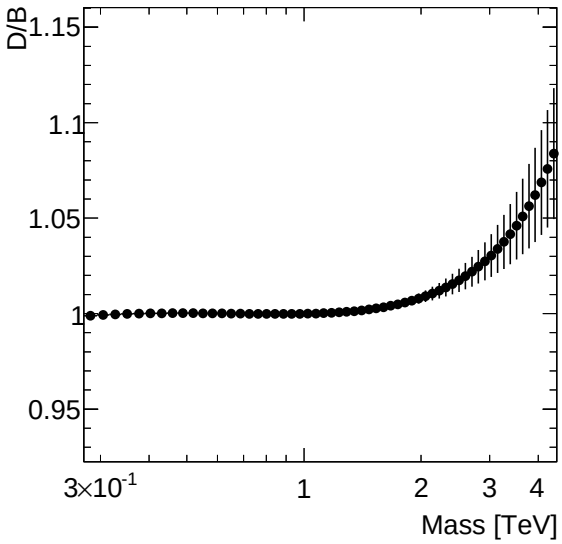
Figure 7.10: Effect on the background prediction of a Breit-Wigner signal with mass-to-width ratio 0.10 at three representative masses and the highest undetectable normalisation. The underlying distribution is pseudodata generated from a fit to MC. In the left-hand plots, the comparison between pseudodata and fit is shown with the largest bump outlined in blue. Right-hand plots show the ratio of the fit to background plus signal over the fit to background alone.

case does the underlying value of the fit change beyond this range. Figure 7.11 shows the fit ratios at all tested masses for four different signal widths. As demonstrated, the deviations increase with increasing signal width, but the ratios remain within a value of 1.3 in all cases.

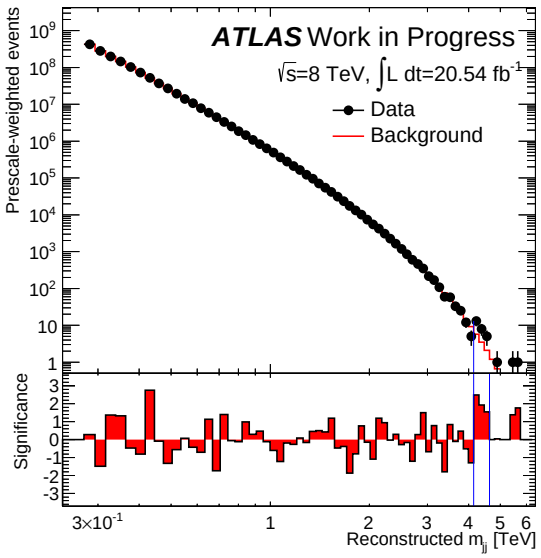
Although the plots in Figure 7.11 clearly indicate locations where the ratio of the functions changes dramatically, they do little to illuminate the regions where the central value of the function may change only a little, but the uncertainties are sufficiently small that they may not cover the change. To illuminate differences which are small on the scale of the fit normalisation but large compared to the errors, the plots in Figure 7.12 were created. These show the significance of the difference between the two fits as represented by their difference divided by the size of the total errors in each bin. Since the errors on each function were calculated to represent a 1σ deviation, the factor on the $y-$ axis of the plots corresponds to the multiple the number of standard deviations in the combined errors by which the fits are separated. The plots show that the discrepancy defined thus is largest for low masses but never exceeds a value of 5σ (in most cases much smaller), an appropriate value considering that every case tested here is an extreme situation designed to produce the largest deviations



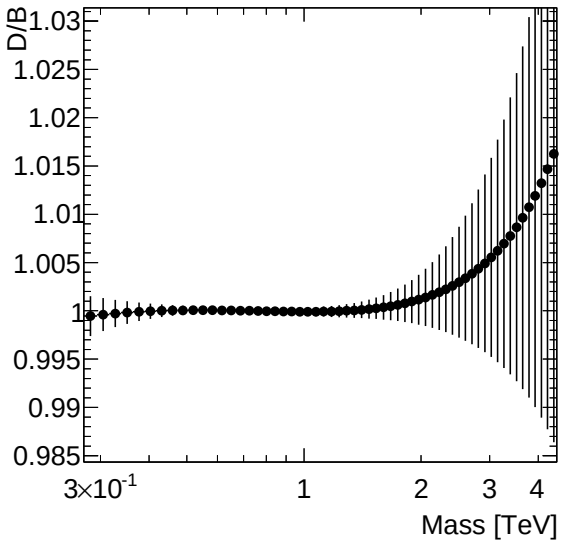
(c) Fit to pseudo-background with 2000 GeV Breit-Wigner signal



(d) Ratio of fit with and without 2000 GeV signal



(e) Fit to pseudo-background with 4400 GeV Breit-Wigner signal

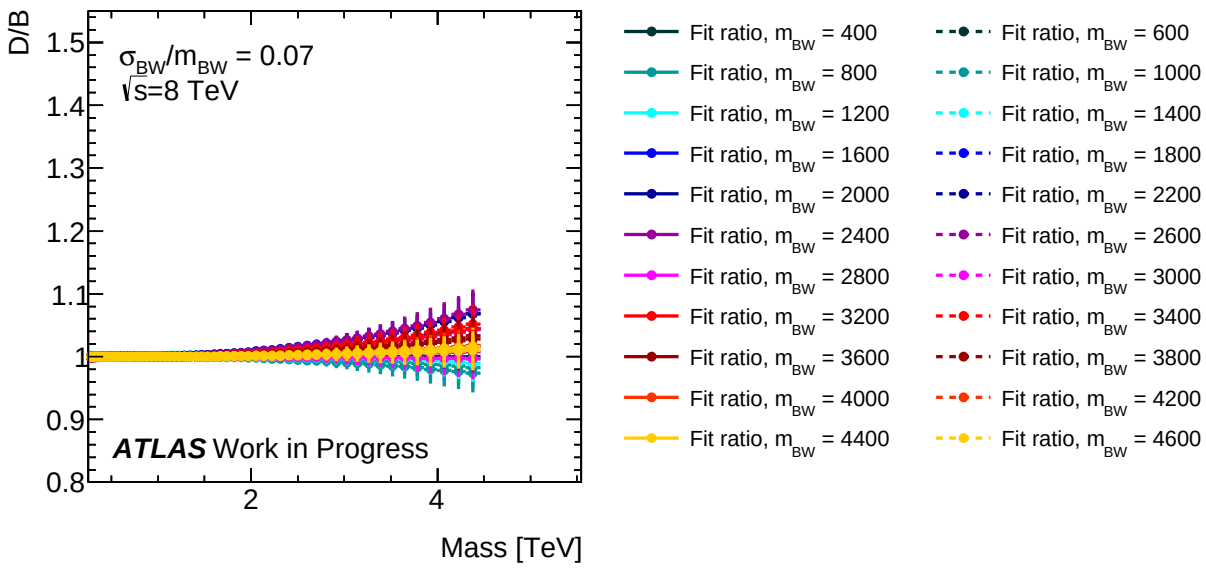


(f) Ratio of fit with and without 4400 GeV signal

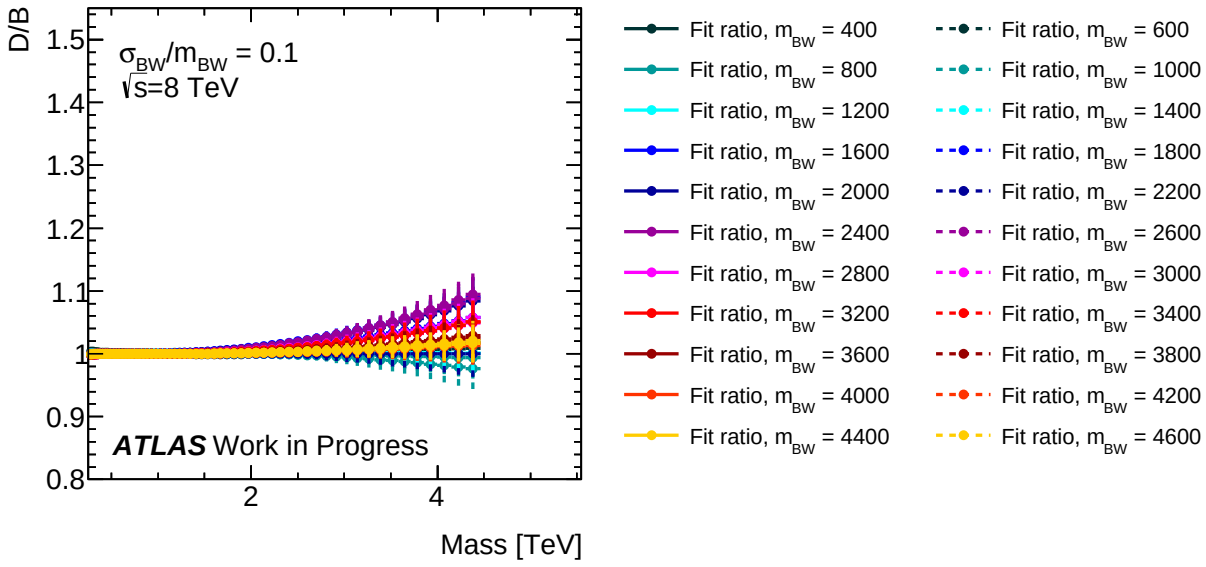
Figure 7.10: Effect on the background prediction of a Breit-Wigner signal with mass-to-width ratio 0.10 at three representative masses and the highest undetectable normalisation. The underlying distribution is pseudodata generated from a fit to MC. In the left-hand plots, the comparison between pseudodata and fit is shown with the largest bump outlined in blue. Right-hand plots show the ratio of the fit to background plus signal over the fit to background alone.

possible.

These tests can also be used to quantify the minimum signal size to which the search phase is sensitive. For each mass point above, once the signal normalisation providing a



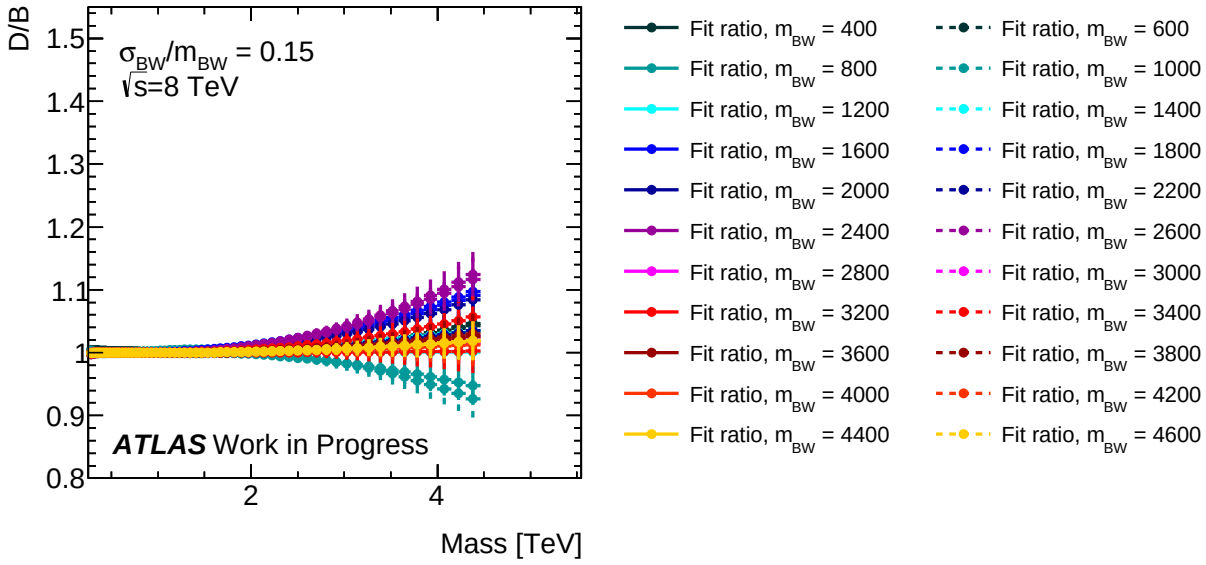
(a) Signal σ/m_{jj} = 0.07



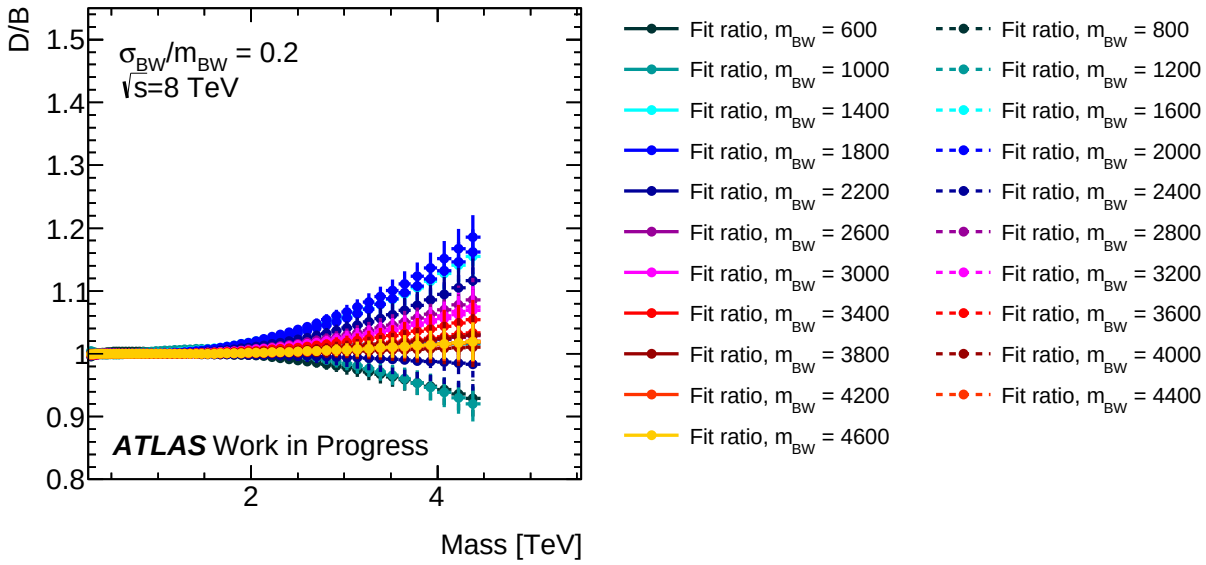
(b) Signal σ/m_{jj} = 0.10

Figure 7.11: Ratios of fits to pseudo-background plus signal over pseudo-background alone for a large range of masses. The four plots represent Breit-Wigner signals of different width to mass ratios.

p–value of 0.01 was found the normalisation of the central 68% of the signal was calculated and rounded to the nearest integer to provide a number of signal events. The same procedure was used to calculate the number of background events in the same range of bins. Thus the minimum S/B ratio was calculated to which the analysis is sensitive for each mass and width of signals tested. These values are listed in Table 7.8 for signals of width 0.10.



(c) Signal $\sigma/m_{jj} = 0.15$

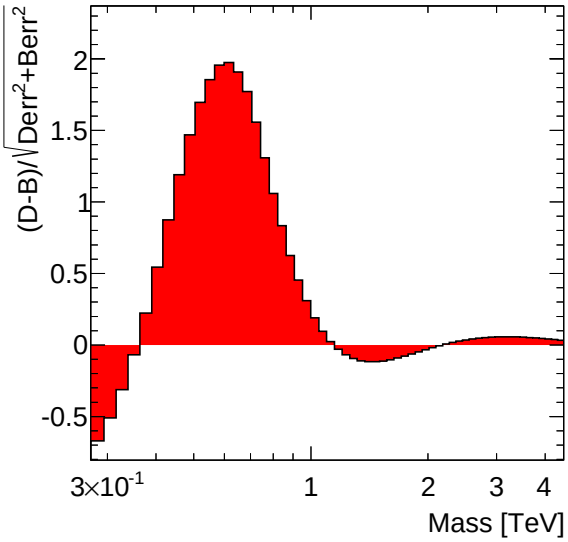


(d) Signal $\sigma/m_{jj} = 0.20$

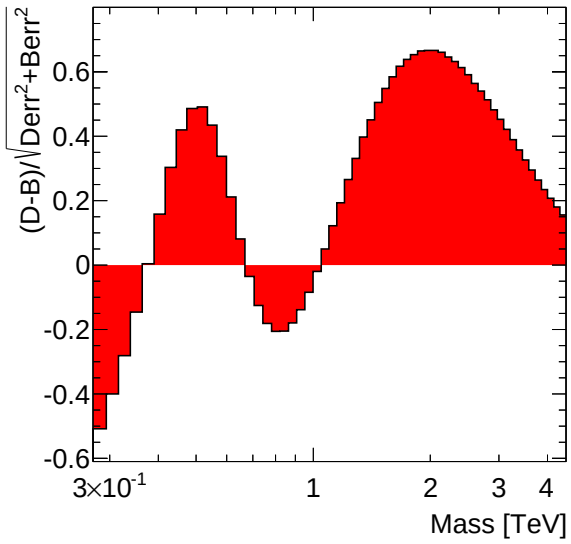
Figure 7.11: Ratios of fits to pseudo-background plus signal over pseudo-background alone for a large range of masses. The four plots represent Breit-Wigner signals of different width to mass ratios.

7.7 Search results in 8 TeV data

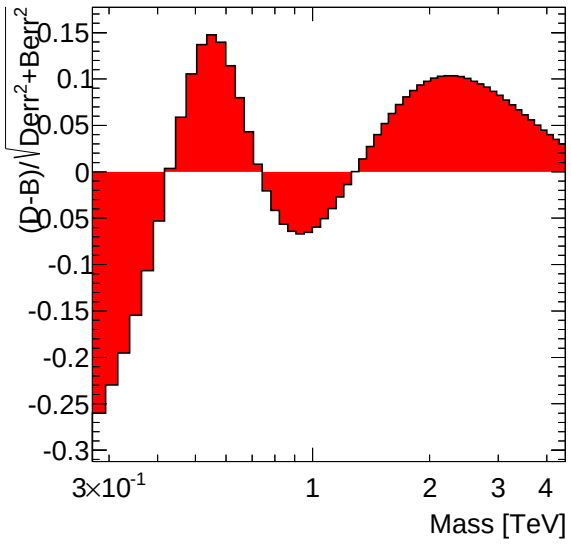
The observed dijet mass distribution after all selection cuts is shown in Figure 7.13. Also shown in the figure are the predictions for an excited quark for three different mass hypotheses [37, 38]. The χ^2-value of the fit is 79 for 56 degrees of freedom, and the reduced χ^2 is 1.41. The central section of the figure shows the relative difference while the lower part of the



(a) 600 GeV mass



(b) 2000 GeV mass



(c) 4400 GeV mass

Figure 7.12: Significance of the difference between fits to pseudo-background with and without signal for three different Breit-Wigner masses of $\sigma_{\mathrm{BW}}/m_{\mathrm{BW}}$ = 0.10. The significance is calculated as the difference in the two fits in each bin divided by the sum of squares of the errors.

figure is the residual and represents the significance, in standard deviations, of the difference between the data and the fit in each bin.

In Figure 7.14, the width and location in x of each horizontal line indicate one possible window tested by the BumpHunter algorithm, while the location in y of the line is the value of the BumpHunter test statistic obtained for that window. The most significant discrepancy identified by the BumpHunter algorithm in the observed dijet mass distri-

Mass	S	B	Minimum S/B
400	615797	56345975	0.0109
600	35446	12944776	0.00274
800	7655	2272363	0.00337
1000	3496	577622	0.00605
1200	1457	253953	0.00574
1400	984	85507	0.0115
1600	677	37252	0.0182
1800	441	15952	0.0277
2000	276	9004	0.0307
2200	265	3717	0.0713
2400	189	1929	0.0980
2600	116	1563	0.0742
2800	77	534	0.144
3000	43	208	0.207
3200	26	226	0.115
3400	18	100	0.180
3600	14	79	0.177
3800	15	38	0.395
4000	13	12	1.08
4200	10	23	0.435
4400	9	12	0.750
4600	6	2	3.00

Table 7.8: Minimum signal to background ratio to which the search phase is sensitive, for a range of masses at $\sigma/m_{jj} = 0.10$.

bution is a 7-bin excess in the interval 390 GeV to 599 GeV. Figure 7.15 compares the BumpHunter test statistic obtained for each pseudo-experiment (blue) to that observed in real data (red arrow). The percentage of events falling below the arrow is indicates the probability of obtaining a bump at least as significant as the observed excess. This p-value is 0.075, corresponding to a significance of 1.44 σ. It is concluded that no significant excess, indicative of new physics, is observed in the 2012 dataset.

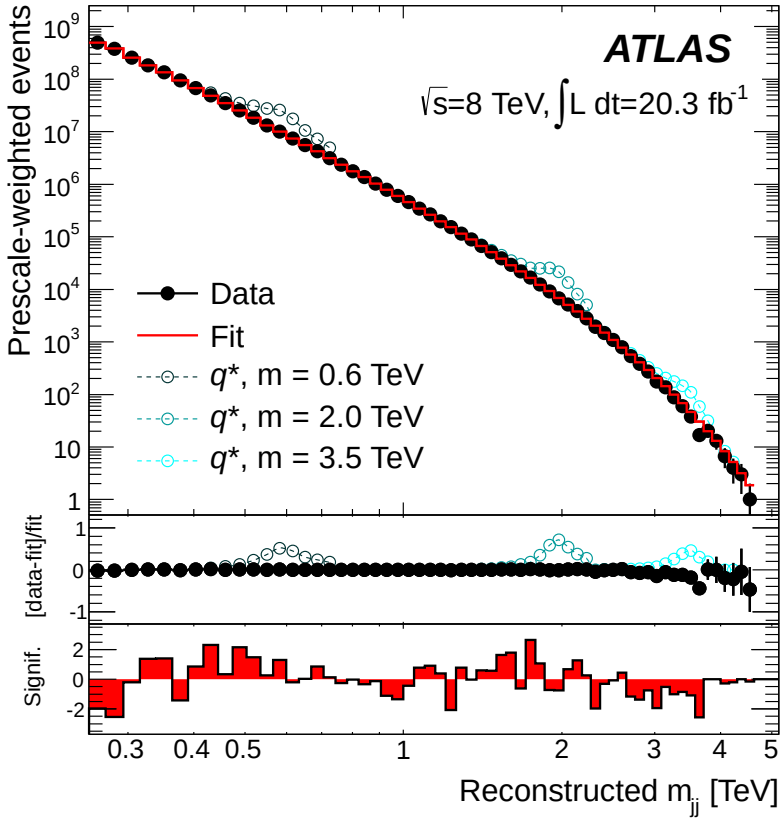


Figure 7.13: The reconstructed dijet mass distribution (filled points) fitted with a smooth functional form (solid line). Predictions for three q^* masses are shown above the background. The central panel shows the relative difference between the data and the background fit with overlaid predictions for the same q^* masses. The bin-by-bin significance of the data–background difference is shown in the bottom panel.

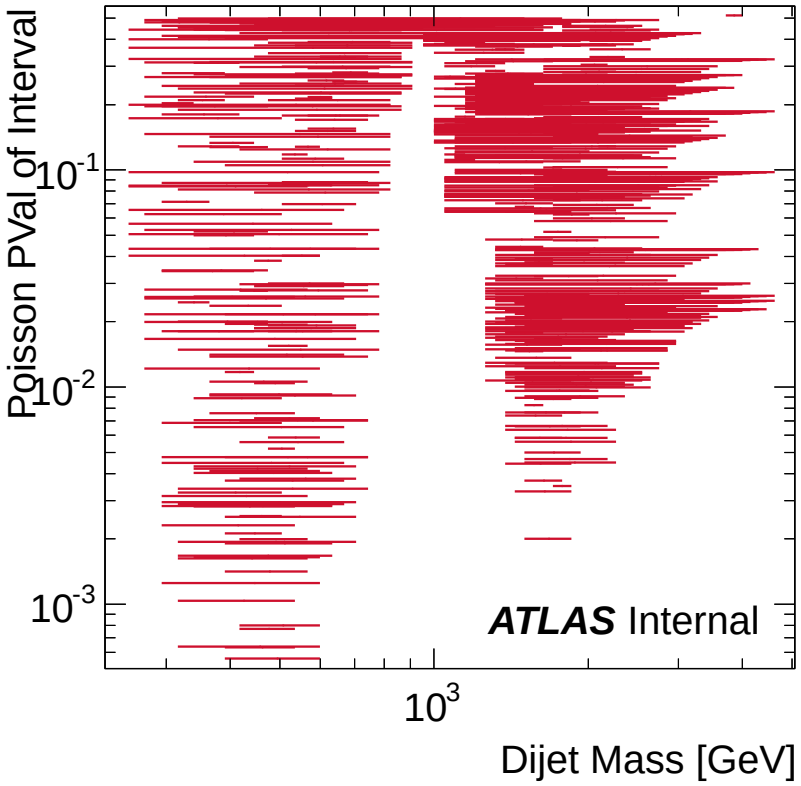


Figure 7.14: This plot indicates the Poisson probability associated with each possible set of bins in the distribution. As the BumpHunter searches for the region with the greatest excess, it tries every consecutive set of bins and calculates the p-value associated with the observed excess or deficit, assuming Poisson statistics in the combined bins. Each red line horizontally spans the set of bins under consideration and the vertical location of the line marks the probability associated with the observed data in that set of bins. The most discrepant interval is the one occurring lowest on the y-axis.

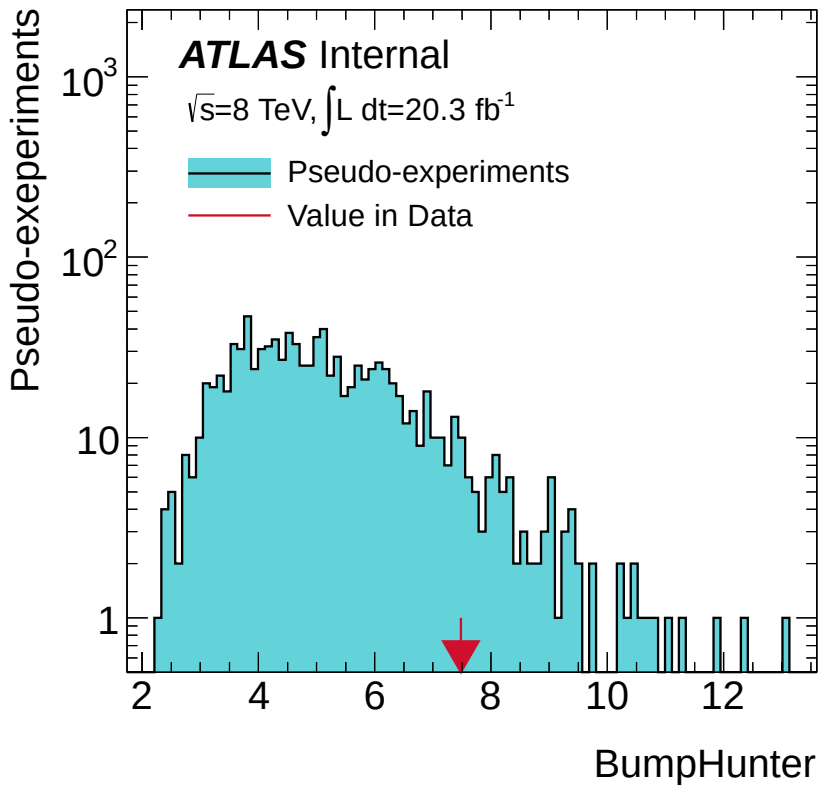


Figure 7.15: Comparison of the observed BumpHunter test statistic to the test statistics resulting from the pseudoexperiments. The blue histogram represents the values from pseudoexperiments while the red arrow indicates the measured value in data. The p-value of the measurement is defined to be the fraction of pseudoexperiments resulting in a more extreme BumpHunter test statistic than was observed in data.

Chapter 8

Systematic Uncertainties

'We demand rigidly defined areas of doubt and uncertainty!'
— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The dijet analysis is fortunate enough to be affected by relatively few systematic uncertainties. Those which do affect it are taken into account during the calculation of limits on the various signal models, which requires that those systematics which can affect the measurement be identified and the form their effects would take be carefully defined. Each independent source of systematic uncertainties is associated with a unique nuisance parameter to be discussed in the following chapter.

8.1 Luminosity uncertainty

In the dijet analysis, the luminosity of the dataset determines the relative normalisation of the dataset and a given signal MC sample in the limit setting procedure. Measurement of the luminosity recorded by ATLAS is a complex procedure involving numerous algorithms and centrally relying on a test only performed once or twice per year. The uncertainty can be calculated from the individual tests [218, 219].

Luminosity, defined in Equation 4.1 in terms of the beam parameters, can be specified from detector quantities as well:

$$\mathcal{L} = \frac{f_r n_b \mu_{\text{vis}}}{\sigma_{\text{vis}}} . \tag{8.1}$$

Here, f_r and n_b are revolution frequency and number of bunches. σ_{vis} is the visible cross

section: the product of the total inelastic cross section with the efficiency of the detector subsystem and algorithm used for the measurement. μ_{vis} is the visible bunch rate per crossing, and so both σ_{vis} and μ_{vis} are measurable quantities.

Measurements of σ_{vis} are done using van der Meer scans (vdM). Two low-intensity beams are passed through each other in steps, providing a measurement of μ_{vis} versus beam separation and parameters for the shape of the beam cross-section. These provide an independent measurement of $\mathcal{L}$ which combined with Equation 8.1 allows a definition of σ_{vis} entirely in terms of μ_{vis}^{peak} and the beam parameters. One set of vdM scans was performed in 2012 and the calculated σ_{vis} was assumed to be constant throughout the rest of the year.

Van der Meer scans are used to calibrate values of σ_{vis} for each measurement system used during regular data taking [220]. By measuring the instantaneous delivered luminosity with the vdM and determining the peak observed bunch rate per crossing for each detector system and algorithm, σ_{vis} for that measurement can then be solved for. In regular data collection periods the delivered luminosity is not known, but the measured value of σ_{vis} is assumed to be the same as during the vdM scans and thus the ongoing measurement of during data taking can be used to find $\mathcal{L}$ at all times. Two sources of systematics clearly appear: those affecting the vdM scans and those affecting the measurements of μ_{vis}. With the visible cross-section calibrated, the luminosity is determined by monitoring μ_{vis} using various detector systems during data-taking.

A variety of systematic uncertainties affect these measurements. The μ_{vis} calibration is affected by "afterglow", residual activity in a bunch crossing after a valid collision, and the assumption that the calibrations are valid throughout the data-taking period introduces an additional uncertainty. A third significant systematic is pileup dependence of the μ_{vis}-measuring algorithms. The effects of all three systematics can be estimated and constrained by comparing their different effects across the available algorithms and subdetectors. The total uncertainty on the luminosity measurement was found to be 2.8% in 2012 data.

8.1.1 Incorporation in dijets

The uncertainty in the luminosity measurement is incorporated in the dijet limit setting as a systematic on the signal samples. When, during the limit calculation, the signal sample must at each step be normalised to a number of events specified by the parameter of interest, the

relationship between that number of events and the luminosity of the sample is considered to be known. However, if the data sample constitutes in reality 1% more luminosity than was supposed, this same equivalent number of events must produce a sample with a proportionally smaller normalisation compared to the data. Thus the uncertainty in luminosity is treated as a scaling factor applied to the signal template: a nominal normalisation of the signal template is calculated but can be increased or decreased in accordance with the luminosity uncertainty. An increase or decrease in luminosity of 2.8% is treated as one standard deviation in the nuisance parameter.

8.2 Beam energy uncertainty

A detailed measurement of the momentum of the LHC beams was carried out during *Pb-p* running in January and February 2013, as detailed in [221]. The nominal momentum of the proton beam is defined by the dipole field, a quantity which can be calculated from the measured current drawn by the magnets and which had an uncertainty of $\sim 7 \times 10^{-4}$ in the 4 TeV beam. However, the quadrupole magnets also affect the true beam momentum. Their effect is proportional to the fractional difference between the RF frequency of the accelerator f_{RF} and the "central" RF frequency for which the beam is centred on average in all quadrupoles.

A measurement of the RF frequency of the beams is quite challenging. In ATLAS, this is done by exploiting the difference in revolution frequency between protons and lead ions due to their differing charge-to-mass ratios. The difference $\Delta f_{RF} = f^p_{RF} - f^i_{RF}$ gives a direct measurement of the beam momentum and was calculated in 2012 during a dedicated *Pb-p* collision period. The measurement is performed by forcing both the lead and proton beams onto a common frequency. With a frequency too high for the ion beam and too low for the proton beam, the former drifts to smaller radii and the second to larger, with the offset between them proportional to the change in beam momentum. The dominant uncertainty on the measurement is due to the beam position monitors (BPM) used to track the dispersion and location of the two beams. An additional systematic results from the drift in the frequency differences over time. The combination of all systematics on the measurement of Δf_{RF} for beams nominally at 4 TeV is shown in Table 8.1.

Systematic source	Size of error (Hz)
BPM selection	0.25
Bunch intensity effects	0.25
BPM scale uncertainty (0.5%)	0.30
Frequency difference drift	0.60
Total	0.76

Table 8.1: Various contributions to the systematic error on Δf_{RF}, which determine the uncertainty on the beam energy measurement.

The measurement of the frequency difference between beams is thus:

$$\Delta f_{RF,4\ \mathrm{TeV}} = 59.21 \pm 0.15(\mathrm{stat}) \pm 0.76(\mathrm{syst})\ \mathrm{Hz}\,. \tag{8.2}$$

Using this result to solve for the beam momenta, it is found that for a nominal beam energy of 14 TeV the actual momentum is:

$$P_{4\ \mathrm{TeV}} = 3988 \pm 5(\mathrm{stat}) \pm 26(\mathrm{syst})\ \mathrm{GeV}\,. \tag{8.3}$$

Note that the value is consistent with 4 TeV but that the central value is in fact lower by a full 12 GeV.

8.2.1 Effects of beam energy on signal MC

The signal Monte Carlo for all models tested was generated corresponding to an individual beam energy of 4000 GeV. In order to determine the effects of a change in beam energy, the samples are reweighted using a shifted PDF. To determine the MC shape for a beam energy $0.5\sqrt{s}_{\mathrm{new}}$ the Bjorken x factor of the PDF is scaled by a factor $\sqrt{s}_{\mathrm{new}}/\sqrt{s}_0$ where $0.5\sqrt{s}_0$ is the energy at which the sample was generated (4000 GeV). Templates corresponding to $\pm 1\sigma$ were generated by shifting the energy to values of $0.5\sqrt{s}_{\mathrm{new}} = 3962$ GeV and 4014 GeV, corresponding to the actual error around the measured real beam energy of 3988 GeV. A systematic uncertainty specified by a parameter θ was defined using these templates and was given a Gaussian prior. The combination of this prior with the asymmetric difference between the nominal template and the $\pm 1\sigma$ shifted templates meant that the probability of obtaining a shift $\delta\theta$ in each direction was the same (normally distributed) but the effect of

a positive shift was smaller than the effect of a negative one to account for the non-central location of the nominal template.

Figure 8.1 illustrates the effects of a shift in beam energy on q^* and $s8$ signal samples over a range of masses. The effects are most dramatic at high mass where the PDF changes sharply, dropping to zero at the centre-of-mass energy. Downward shifts in θ result in a greater shift away from the nominal distribution in keeping with the skewed location of the nominal energy compared to the real energy.

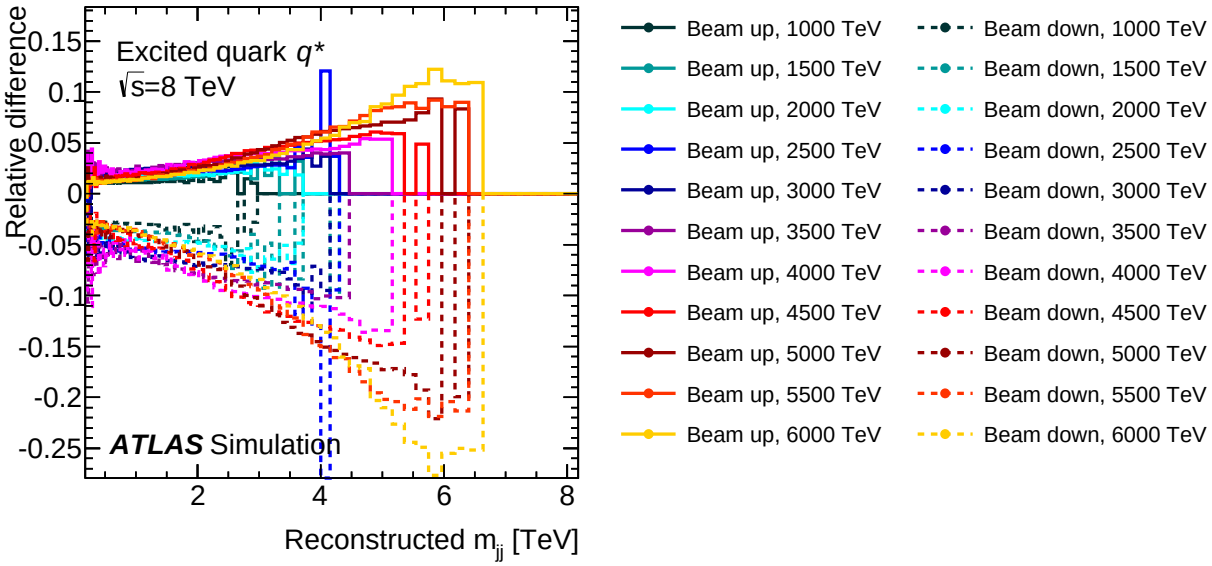
8.3 Background uncertainty

The uncertainty on the background fit from the selected function was calculated in Section 7.2.3 by determining the range of deviation of the fit in the central 68% of cases. Thus a movement of the fit function to either end of this range constitutes a change in the background estimate of $\pm 1\sigma$. A systematic uncertainty representing the background estimate was thus devised by assuming that the fit results varied linearly with a parameter θ representing the distance from the nominal fit: thus a shift in the fit by twice the size of the error band ought to correspond to a probability of 2σ. The parameter describing the background estimate was given a Gaussian prior in accordance with this calculation and allowed to range within $\theta \in [-3\sigma, +3\sigma]$. Thus if the value of the nominal fit in each bin was f_i with uncertainty δf_i, the background estimate in that bin for a given value of the systematic uncertainty was $f_i + \theta\,\delta f_i$.

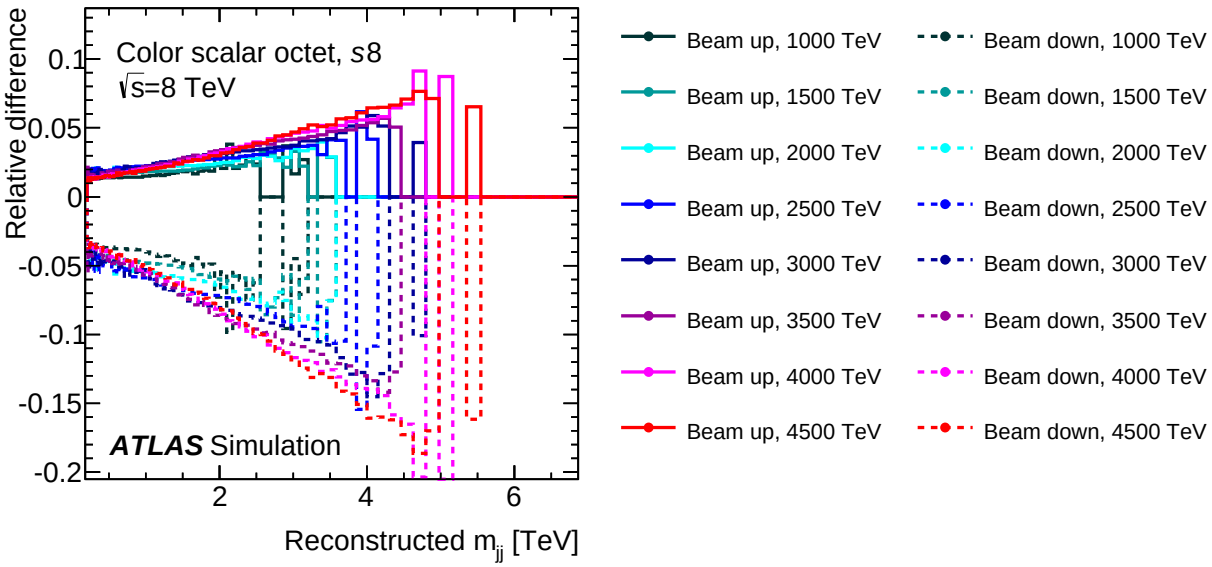
Figure 8.2 plots the nominal value of the fit function with a shaded band denoting its $\pm 1\sigma$ range compared to the data. The fit is tightly constrained by the data, as indicated by the narrowness of the plotted error band.

8.4 Background parameterisation

The full 8 TeV dijet paper is the first iteration of the analysis to include a systematic uncertainty representing the choice of parameterisation used to describe the background. The standard dijet function used in the search phase was selected as the nominal background estimate while Function V in Table 7.1 was the chosen alternate parameterisation. Tests performed on the two functions were described in Chapter 7. Although the two functions



(a) Excited quark



(b) Color octet scalar

Figure 8.1: Relative bin-by-bin difference between nominal signal MC and samples with the beam energy shifted by $\pm 1\sigma$. The y-axes show the shifted bin content divided by nominal bin content.

performed nearly identically in blinded data and the available Monte Carlo, they exhibited surprisingly different results in the full unblinded dataset: a p-value of 0.86 was found compared to the nominal fit p-value of 0.075. As a result, the systematic uncertainty assigned to the choice of function became the dominant uncertainty in the analysis. Figure 8.3 illustrates the difference between the two background parameterisations. On the left, the ratio of the two fits demonstrates that the biggest relative difference is in the high-mass tail,

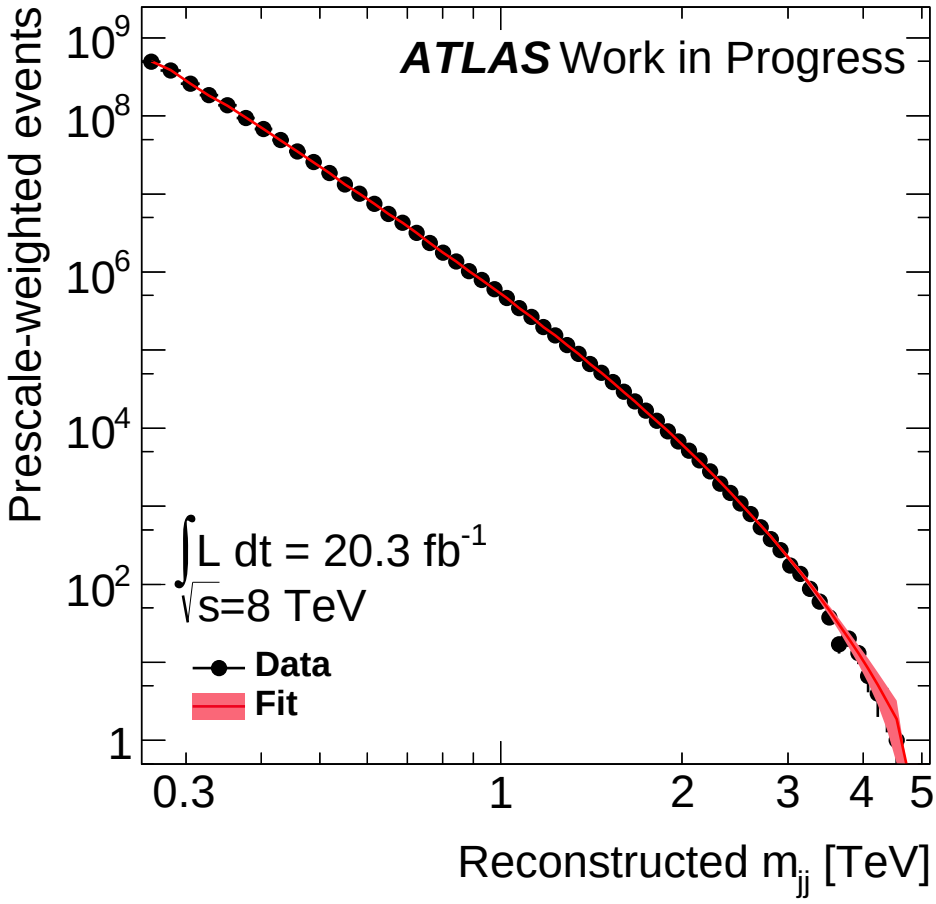
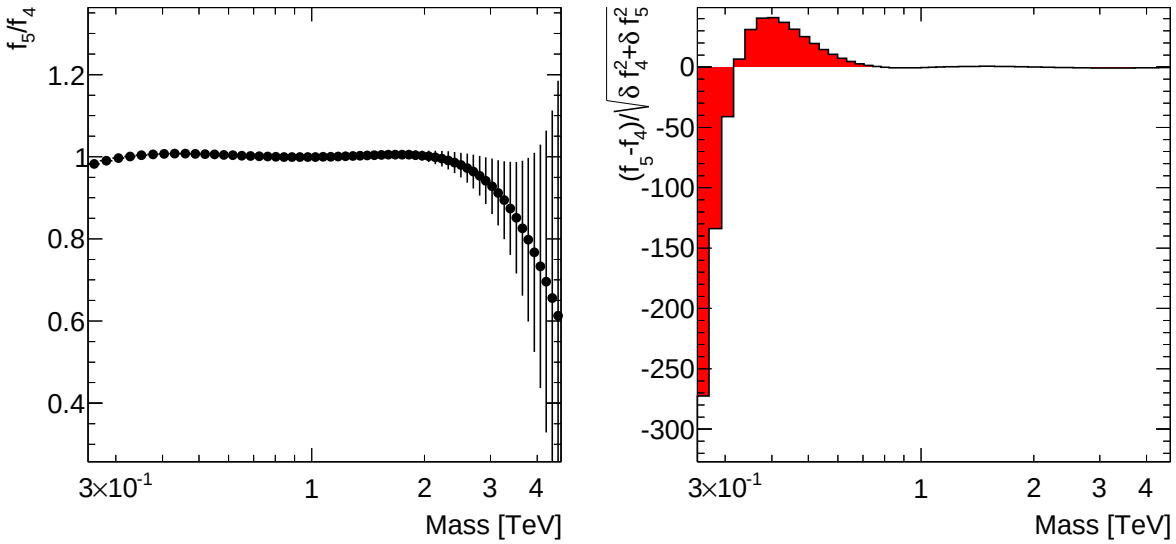


Figure 8.2: The nominal function fitted to data is shown as a smooth curve (red) within its uncertainties due to the variation of the function with data statistics (solid pink band). The black points illustrate the full data distribution. A variation of the background estimate uncertainty by $+1\sigma$ (-1σ) is used to move the background estimate to the top (bottom) of the uncertainty band.

where the alternate function takes a much lower value than the nominal function. On the right, the difference in the two estimates divided by the statistical fit uncertainty on each shows why this becomes the largest systematic uncertainty: although both functions fall within 2σ of the data with its statistical uncertainty in the first bin, both functions are also highly restricted in their low-mass tail, and do not vary far from their nominal value even with many repeated pseudoexperiments. Thus the difference between two valid descriptions of the data is far larger than the range over which either of the descriptions can vary, and the function choice uncertainty becomes more significant than the statistical uncertainty on one fit.

The parameter θ assigned to describing this uncertainty was treated as a two-point sys-



(a) Ratio of five- to four-parameter estimates

(b) Difference divided by sum of errors

Figure 8.3: Relative difference and significance of difference relative to the error bands of the nominal and alternate background parameterisations. A nuisance parameter in the limit setting procedure accounts for this difference by a varying within the range between the two estimates.

tematic, ranging between the background estimate from the two functions but not allowed to vary outside of this range, as any extrapolation of the difference between functions beyond this range was not physically motivated. The prior form and range were selected to indicate a slight but not overwhelming preference for the nominal function, so the parameter was given a Gaussian prior but constrained to the range $\theta \in [0, +1\sigma]$ where a value of 0 indicated the nominal background estimate f_{nom}, a value of 1σ indicated the alternate background estimate f_{alt}, and values in between would select a linearly scaled intermediate $f_{\text{nom}}(1+\theta\frac{f_{\text{alt}}-f_{\text{nom}}}{f_{\text{nom}}})$. The shape of the prior indicated that the alternate function was considered slightly less probable than the nominal function, but the posterior distribution would not be overly constrained by the choice. The restriction in the range of θ prevented the background prediction from moving outside of the range considered by either parameterisation, maintaining its physical meaning.

8.5 Jet Energy Scale

The 54 components of the jet energy scale uncertainty are not all treated independently in the analysis. The addition of so many degrees of freedom to the limit setting, so many of

which have almost no physical effect on the dijet spectrum, would not only add tremendous computational demands but would introduce the possibility of non-physical profiling without improving the description of actual variations. Instead, the treatment begins from a sets of reduced jet energy scale uncertainties for which the sets of components diagonalised into smaller sets of independent contributions have been kept grouped by type: *detector*, *modelling*, *statistical*, *pileup*, *mixed*, and special uncertainties which have not been recombined. Within these categories the new eigenvectors no longer have direct physical connotations; they are simply `Detector 1`, `Detector 2`, etc.

Of this reduced set of 23 orthogonal components, the impact of each on the MC was computed for all signals and mass points. For each signal MC mass point, all jets were shifted by 1σ in the selected JES uncertainty component. The m_{jj} spectrum was remade with shifted jets and the significance of the deviation between this and the nominal spectrum was calculated bin by bin. Considering only bins within a 68% window around the mass peak (thus suppressing large contributions from areas with limited statistics), the bin with the most significant deviation was selected and that value used as the effect of the systematic on that mass point. The procedure is repeated for every mass point and plotted. All systematics for which this maximum deviation remained below 2% in all mass points were considered small and did not need to be treated independently, while those which fluctuated above 2% at any point were kept separate. This cutoff was deemed appropriate because tests where components with effects below 2% were handled separately proved to have no observable impact on the limits. An example corresponding to the q^* signal sample is shown in Figure 8.4.

After repeating the computation for all uncertainties and signals, it was determined to keep the following components as independent systematic uncertainties:

- Single particle uncertainty (dominant JES component at high p_{T})
- Modelling, component 1
- Detector, component 1
- Relative non-closure (for AtlFast II samples only)
- η-intercalibration, modelling component
- Flavour composition
- Flavour response.

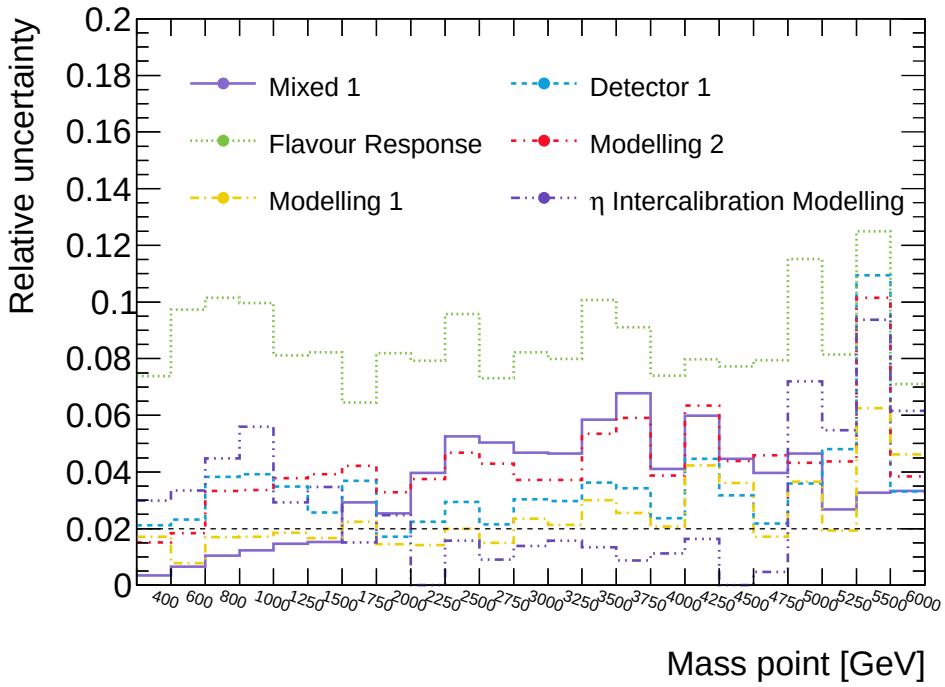


Figure 8.4: The relative uncertainty in each of six JES systematic uncertainty components are plotted for each mass sample in the $s8$ signal. Those which produce a shift > 2% will be treated as independent systematic uncertainties in the limit setting.

The remaining components, all with a very small impact on the dijet analysis, are combined in quadrature into a single parameter. They are:

- All statistical components
- All pile-up components
- Modelling, components 2-4
- Detector, components 2 and 3
- All mixed components
- η-intercalibration, statistical and method components.

8.6 Jet Energy Resolution

The energy resolution of the reconstructed and calibrated jets can potentially affect the dijet analysis if reconstructed jets fall in a different bin than their true value would indicate, and thus any imprecision in the knowledge of this resolution can contribute a systematic uncertainty to the measurement. The experimental origin and method of estimating the jet energy resolution uncertainty were explored in Section 5.3.2.

Effects of the uncertainty on the jet energy resolution are tested in the dijet analysis by comparing the spectrum after smearing by the JER uncertainty to the nominal spectrum. For each jet in an event, a random smearing factor is drawn from a Gaussian with width equal to the JER uncertainty for a jet of the nominal p_{T} and η. The jet's E and p_{T} are scaled by the smearing factor, the jets are reordered according to scaled energy to account for any differences in the selected leading and subleading jet, and the standard analysis selection is continued using the rescaled jets. An m_{jj} spectrum for the smeared jets which pass event selection is then compared to the m_{jj} spectrum with nominal jets. The ratio between the two provides a bin-by-bin estimate of the effects of the JER uncertainty. This quantity was found to be negligible everywhere, so the uncertainty is not used in the limit setting.

8.7 Angular resolution uncertainties

The jet angular resolution is a measure of the distance in η, ϕ by which reconstructed jets can vary from their true direction. Since the location of each jet in η, ϕ space affects its acceptance, a large angular uncertainty could mean certain desirable jets are being rejected by the selection or that jets which fall into a poorly measured region of the detector could nonetheless pass the analysis cuts. Standard jets formed from calorimeter information (*calo jets*), although they have excellent energy resolution, make no use of tracking information and therefore have limited angular resolution. In order to determine a measure for it, both MC truth jets and jets reconstructed from reclustering tracking information (*track jets*) are used [222].

The angular resolution is defined as the RMS of the distribution of distances between reconstructed jets and their matching truth jets. Calo jets and track jets in MC are matched to the nearest truth jet where the distance is defined by a measure $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. After some simple jet quality selections, the values of $\Delta\eta$ and $\Delta\phi$ are determined for all available matched jets and are used to form a set of distributions binned in truth jet p_{T} and η. For each bin the RMS of the $\Delta\eta$ and $\Delta\phi$ distributions is calculated, resulting in an MC-driven measure of the angular resolution in each dimension for both calo and track jets across the full p_{T} and η range. A data-driven alternative uses the same binned setup but measures the RMS of the angular distance between matched calorimeter jets and truth jets. Plots demonstrating the results for all cases in 2012 data can be found in Ref. [223]. The angular

resolution is smaller than 7% everywhere for both data and MC driven calculations, falling to below 4% for jets with $p_{\mathrm{T}} > 100$ GeV.

8.7.1 The effects on dijets: Negligible

As for the jet energy resolution uncertainty, the effects on the dijet distribution can be calculated by shifting every jet in η and ϕ by a random amount drawn from a Gaussian whose width is the angular resolution at the jet's nominal p_{T} and η. The difference between the nominal spectrum and the resulting one smeared by the angular resolution then gives a measure of the uncertainty due to the resolution. The results were found to be negligible in the dijet analysis.

8.8 Theoretical uncertainties

Theoretical uncertainties on the new signal samples are considered where they are expected to affect the shape or acceptance of the final MC template. There are three possible locations for such an effect to manifest: the choice of renormalisation scale, the choice of factorisation scale, and the choice of PDF.

8.8.1 Scale uncertainties: Negligible

Chapter 3 introduced the normalisation scale μ_R and factorisation scale μ_F upon which the calculated dijet cross-section depends. A value must be fixed for each scale during the Monte Carlo generation process, and the choice of value will affect the calculated cross-section. Since all signal samples are at leading order the effect could be substantial.

The renormalisation scale uncertainty was evaluated only for signals with renormalisable couplings where the quantity has a physical meaning, in this case the W' and $s8$. Variations of μ_R by factors of 0.5 and 2.0 were used to regenerate the MC samples and the resulting templates and cross sections were compared. The effects were found to be negligible in both cases. Comparisons of the factorisation scale were made by reweighting the existing samples to probe a different scale Q, where the scale was varied by factors of 0.5 and 2.0. Once again, the variation in signal acceptance and shape was negligible. Therefore no scale uncertainties were included in the limit setting.

8.8.2 PDF choice uncertainty

The effect of a different pdf choice on each MC sample is simulated by reweighting events in the existing sample using the `PDFReweightingTool` [224]. For the two interacting partons a weight can be calculated comparing the new pdf to that with which the sample was generated:

$$w = \frac{\mathrm{pdf}_1^{\mathrm{new}} \cdot \mathrm{pdf}_2^{\mathrm{new}}}{\mathrm{pdf}_1^{\mathrm{original}} \cdot \mathrm{pdf}_2^{\mathrm{original}}} \,. \tag{8.4}$$

In this analysis two alternate pdfs, MSTW2008LO and NNPDF2.1LO, were considered and the effects were defined by the larger of the two at each point.

There are two components to the effects of the pdf choice. First, a change in pdf can change the shape of the signal. Second, it could change its acceptance, which would be a scale-type uncertainty and be handled separately. These possibilities were treated independently. Although the shape alterations in the signal templates due to pdf choice were found in some high-mass cases to be substantial, they are also poorly defined from the perspective of a single nuisance parameter: there is no standard procedure for incorporating an uncertainty which is essentially the envelope of two different two-point systematics. Any attempt to define such an uncertainty would also enable the limit setting procedure to select a marginalised best-fit value of the nuisance parameter that would warp the signal shape into its best agreement with data, a profiling result which would not be physically justified. Due to the above considerations and to the added complication for theorist reinterpretation when signal shape is allowed to vary in the limit setting, it is not standard procedure in ATLAS to include shape-changing effects on signal MC templates but only those affecting acceptance and any shape change which may result from the acceptance change in different regions. Therefore after some deliberation it was decided to not include pdf shape uncertainties in the limit setting treatment. The uncertainty on the acceptance, although small in most cases, reached 2% for several samples and was extremely high for the W^* models, whose acceptance is compromised by the y^* cut. It was therefore included in the uncertainties as a scaling factor. Figure 8.5 illustrates the relative change in PDF acceptance due to this effect.

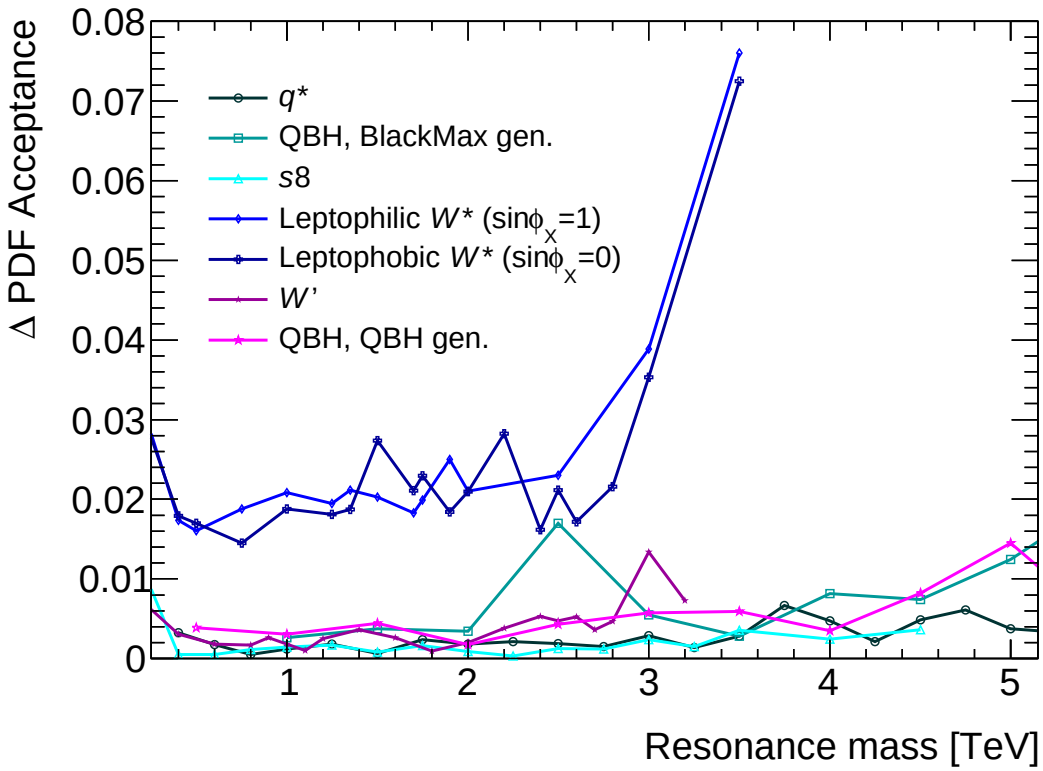


Figure 8.5: Relative change in the acceptance as a result of pdf choice for all signals and mass points. Each point on the graph is the size of a simple scale factor systematic applied to that signal sample during limit setting.

Chapter 9

Limit Setting

'Your dunce who can't do his sums always has a taste for the infinite.'
— George Eliot, *Felix Holt*

9.1 Theory of limit setting

At the end of Chapter 7, it was determined that there was no statistically significant evidence for new physics visible in the dijet spectrum. Having established this fact, it is next possible to use the observed data to set *limits* on the possible presence of any such new physics. For each of a set of models and shapes representing possible new resonances, one asks the question, "To what mass or cross section does the data exclude the existence of this process?"

This question can be broken down into smaller steps by dividing the possible range of masses into sections and asking for each mass considered, "What is the probability that a particle with this mass and behaviour exists, given the observed data?" A limit on the rate of production is then defined by the maximum number of new physics events which could be contained in the spectrum while still keeping this probability below the defined threshold of statistical significance.

It is possible to phrase the specific question in different terms; for instance, "What is the likelihood of observing the data collected given that model x is true?" But the question takes a more natural form for the observer when asking the comparative probabilities of different theories, leading to a Bayesian phrasing of the limit setting problem. The alternative, a maximum likelihood approach, would ultimately address the same issues but differs from

a Bayesian approach in one more critical form: the handling of systematic uncertainties. The frequentist maximum likelihood method for limit calculation finds values of all nuisance parameters which give the best description of the data overall; the Bayesian approach considers all permitted values on all nuisance parameters during a multidimensional integral (the marginalisation). This will be detailed momentarily. For now, suffice it to say that this process is slower but that in a case of complex interrelations between systematic uncertainties it can be relied upon to produce an answer which is representative of all the possible combinations.

Thus within each new physics model under consideration, for each of a range of masses the probability of the theory given the data is computed and an upper limit on the number of possible events present is determined. This can be converted into a maximum allowed cross-section for the model at each mass. For many models, the theory dictates a specific nominal cross-section. The intersection between this nominal cross-section and the observed maximum possible cross-section is thus the highest mass to which the signal in question can be excluded.

The remainder of the chapter details the definition and execution of the above process and provides the resulting cross-section and mass limits for a range of model-dependent and generic signal shapes.

9.1.1 Bayes' Theorem and its use in signal quantification

The theorem which relates conditional probabilities of two outcomes via *a priori* knowledge of the probability of each was first laid out by Thomas Bayes in the early 18th century, and now bears his name [225]. It was independently rediscovered by Pierre-Simon Laplace who, after encountering Bayes's additional work on non-uniform prior probabilities, gave the theorem its first modern recognisable form in his 1812 work *Théorie analytique des probabilités* [226]. Bayes' theorem is phrased thus in Laplace's 1814 summary:

> Si l'on calcule *à priori*, la probabilité de l'évènement arrivé, et la probabilité d'un evènement composé de celui-ci et d'un autre qu'on attend ; la seconde probabilité, divisée par la premieère, sera la probabilité de l'évènement attendu, tirée de l'évènement observé.

In mathematical form,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \ : \tag{9.1}$$

that is, the conditional probability of some event A given some event B is equal to the conditional probability of B given A multiplied by the ratio of the probability of event A alone to event B alone [227].

At this point, it is helpful to define several statistical terms and specify a distinct notation for each [210, 228].

Probability In a Bayesian context, probability takes on a subjective interpretation as plausibility of a proposal. Prior probability is the inherent plausibility assigned to an outcome or hypothesis with no external information. Posterior probability is the plausibility of the same outcome or hypothesis given other information (such an an experimental outcome), and is a conditional probability. The probability of item A is denoted $P(A)$.

Probability distribution function (p.d.f.) The probability of obtaining each particular outcome x in set of discrete possible outcomes may be specified by a function $p(x)$. This is normalised over the full space of possible outcomes S such that $\sum_S p(x) = 1$.

Probability density function (p.d.f.) When x is a continuous variable, the probability of any one value is not well defined. Instead a probability density function $p(x)$ is defined such that the probability of an outcome in $[x, x+dx] = p(x)\,dx$. This is normalised over the full space of possible outcomes S such that $\int_S p(x)\,dx = 1$. Following mainstream practice, both the notation and the abbreviation chosen for this are the same as for probability distribution functions, since the difference will either be irrelevant or clear from the context.

Likelihood A likelihood defines the conditional probability of a (fixed) observed experimental outcome as a function of a variable model. The likelihood of an observation x as a function of a model A would be written $\mathcal{L}(x|A)$.

Bayes' formula, presented above in terms of probability, takes an equivalent form stated in terms of probability densities. Consider then the probability of a continuous hypothesis H given a measured dataset x. The prior probability of the hypothesis $\pi(H)$ represents a degree of belief which can be more or less informative and is an explicit choice on the part

of the researcher. The selection of a specific function for $\pi(H)$ will be discussed in subsection 9.1.2. The prior probability of the dataset $\pi(x)$ can be understood as the integral of $p(x|H)\pi(H)$ over all possible hypotheses. Since this value is independent of H, it is essentially a normalisation coefficient and can therefore occasionally be ignored. The conditional probability of the dataset given the hypothesis is the likelihood $\mathcal{L}(H|x)$ (see Equation 7.8). Thus the posterior probability density of the hypothesis can be described as:

$$p(H|x) = \frac{\mathcal{L}(H|x)\,\pi(H)}{\int \mathcal{L}(H'|x)\,\pi(H')\,dH'}\,. \tag{9.2}$$

In the majority of physically interesting cases, the hypothesis to be compared to the data is not a single parameter but instead depends on a range of unknown parameters. Most of these will be experimental uncertainties which would affect the comparison between a model and the observed data while only one, such as the amount of a signal present, will be of interest to the observer. From equation 9.2 then, let $H \to \nu, \theta_1, \theta_2, \ldots = \nu, \boldsymbol{\theta}$ where ν is the *parameter of interest* and the θ_i are *nuisance parameters*. To proceed from a multidimensional probability density function $p(\nu, \boldsymbol{\theta})$ to one depending on a single variable, the uninteresting variables must be integrated out: $p(\nu) = \int p(\nu, \boldsymbol{\theta})\,d\theta$ [229]. Thus it follows that:

$$\begin{aligned} p(\nu|x) &= \int p(\nu, \boldsymbol{\theta}|x)\,d\boldsymbol{\theta} \\ &= \frac{1}{C}\int \mathcal{L}(x|\nu, \boldsymbol{\theta})\,\pi(\nu, \boldsymbol{\theta})\,d\boldsymbol{\theta} \\ &\propto \int \mathcal{L}(x|\nu, \boldsymbol{\theta})\,\pi(\nu, \boldsymbol{\theta})\,d\boldsymbol{\theta} \end{aligned} \tag{9.3}$$

where C is a normalisation constant equal to the denominator of Equation 9.2, which depended only on the dataset.

This equation can now at last be used for practical purposes of establishing limits on the presence of a signal in the data spectrum. Let a model be selected which is fixed in mass and shape: for instance, one dictated by a histogram of MC generated according to a specific theory. In order to determine to what extent this model could be present in the data, let the parameter of interest be the amount of this new physics present. Then Equation 9.3 defines the probability density function of the amount of signal, independent of any other

variable, a function which can be used to state a definite limit on the amount of signal which may be present to a certain level of probability: for example, the value of the signal strength parameter below which 90% of the p.d.f. integral is located (the 90% *quantile* of the distribution) marks the amount of signal which we say is excluded by the data with 90% probability.

In a high-energy physics analysis any prior information on the probability of a certain quantity of new physics events will likely be independent of the nuisance parameters, and this is certainly true in our case: the prior probability of having any given number of signal events does not depend on the jet energy scale, luminosity, etc. In the case of the dijet analysis, the systematic uncertainties also have no prior relationship to one another, sometimes by virtue of their physical meaning and sometimes by construction (as in the case of the jet energy scale sub-components). Thus for this analysis all of the priors can be separated and the final equation to be used becomes:

$$p(\nu|x) \propto \int \mathcal{L}(x|\nu,\boldsymbol{\theta})\,\pi(\nu)\prod_i \pi(\theta_i)\,d\boldsymbol{\theta} \tag{9.4}$$

Integration over the nuisance parameters $\boldsymbol{\theta}$ is referred to as *marginalisation* and is the computationally challenging step involved in obtaining a final probability density function for the parameter of interest.

Elaboration on the choice of a Bayesian approach

The frequentist approach to determining limits on signal presence is more popular in ATLAS and results in the calculation of a different probability: while a Bayesian reports an upper limit on signal strength as the cross-section above which it is 95% certain that no new physics lies, a frequentist reports instead the upper edge of a 95% CL confidence interval whose lower edge is at $-\infty$ [196]. As discussed in Chapter 7, both Bayesian and frequentist approaches to limit setting have been supported by ATLAS during Run I data taking, but there is some indication that the Bayesian support will be dropped for Run II in an attempt to consolidate the limit tools in use in the community. One of the primary arguments in favour of a frequentist approach to limit-setting is convenience: the calculations converge more quickly and time can be saved at a later step due to the existence of equations for predicting

the distributions of the posterior quantiles [230]. But there are more philosophical strains to the discussion as well, which will be briefly mentioned here.

One of the primary complaints leveled at Bayesian methodology is the incorporation of subjectivity during the selection of prior probability distributions. Indeed, there are many possible choices, but as long as a reasoned physical argument is given for the option selected and it is clearly conveyed to the end user of the analysis, the actual choice made is often insignificant. When the data powerfully supports a certain conclusion, it can "overwhelm" the prior and produce the same end result regardless of the chosen initial description: here, a posterior with a roughly truncated Gaussian shape is dictated by data regardless of a completely flat prior. Several simple choices have become standards in the community, including flat and Jeffreys priors. The choice of a prior will be discussed in detail in the following section. The critical point is that the analysis contains many subjective choices – cut values, systematic descriptions, mass ranges – which, while physically motivated, are ultimately an arbitrary decision, and as long as the choice made is clearly conveyed and no physical behaviours are obscured by the decision, this is a non-issue.

The second substantial stumbling block for Bayesian analysis is the difference in convergence of the regions of certainty. The frequentist confidence interval is constructed such that it will contain the true unknown value of some parameter in 68% of repeated trials, a property defined as *coverage*. More importantly, the concept applies to real experiments: should many experiments measure the same property and each construct a confidence interval for it, then with sufficient experiments 68% of their reported confidence intervals should contain the true value of the parameter. Although Bayesian credible intervals hold a very different meaning and are not constructed around coverage, their coverage in a Frequentist context can be calculated. Depending on circumstances they can over-cover, under-cover, or exactly cover the Frequentist result. So long as they do not undercover, Bayesian credible intervals will generally be acceptable even to frequentists. It has also been shown that Bayesian intervals cover *on average*; that is, the coverage of a confidence interval weighted by the prior in the parameter of interest s and integrated over ds will equal the frequentist coverage [231]. Bayesian credible intervals are also able to produce meaningful results in cases where frequentist ones are unphysical. This complex subject leaves neither approach in the wrong. Instead, it is worth considering both ranges and their interpretations in different contexts.

9.1.2 Selecting a function for the signal prior

All the components of Equation 9.4 are determined by the experiment and its results with the exception of $\pi(\nu)$. For the dijet analysis, the selection of a functional form to use was motivated by previous experience, practicality, and ease of interpretation for the user.

Bayesian prior distributions can be informative or noninformative. An informative prior is one which incorporates previous knowledge about the parameter, so in this case one could use as a prior the posterior probability distribution for a given signal sample from an independent dataset such as ATLAS 2010 data or the output of a CMS dijet analysis. There are, however many technical difficulties inherent in this approach: for instance, how can it be certain that the dataset used to calculate the input prior was in actuality completely independent from the one being currently used? Within ATLAS, the systematic uncertainties at least will be correlated between older searches and the current one. If CMS data were to be used, the lack of access for an external researcher makes it extremely difficult to check the exact model used to generate a sample and match it to the current input. Furthermore, if a model under examination was not used in an older search, no informative prior might be available.

The alternative is a noninformative prior, which makes no statement about the relative probabilities of the signal values. Theoretical arguments are often made to support the use of noninformative priors, which can be seen to be as near to objective as is possible in a Bayesian methodology. With the use of a noninformative prior, the resulting posterior probability is dictated entirely by the data and there is no danger of a strongly incorrect prior biasing the outcome of the measurement. Both Bayes and Laplace favoured the use of a *uniform prior* which is flat for all permitted values of the parameter, Bayes appreciating its phrasing of the problem entirely in terms of observables while Laplace believed it to be the most appropriate way of specifying complete ignorance of a parameter [229]. Several other options have been put forward in the intervening years, such as the Jeffreys prior which is invariant under transformations of the parameter of interest [232]. The reference prior, introduced in its simplest form in 1979 and developed in multivariate forms since, matches the Jeffreys prior in its one-dimensional form but eliminates concerns about dependence among parameters in multidimensional systems [233, 234]. An effort is now being made within ATLAS to encourage the use of these more complex noninformative priors [235].

The dijet analysis has elected to use a uniform prior for several reasons. First, since the parameter of interest here is simply a number of signal events (translating into the cross-section of the model under examination), a prior definition which makes any cross-section equally likely is easy to understand and to explain to users for reinterpretation. Second, with a uniform prior the posteriors returned by this method are identical to the results of the Frequentist CLs method, an ATLAS standard. Third, all previous versions of the dijet analysis as well as many other ATLAS analyses have employed flat prior distributions in the signal, meaning that its use here facilitates the comparison of results. Third, it is easy to state for theorists who may need to know the method of derivation of the limits for reinterpretation. For each posterior probability to be calculated for a signal MC sample, a range of acceptable values of ν was found by calculating the likelihood of the simple background with superimposed signal normalised to ν compared to the data distribution and increasing ν until a point was reached where the likelihood was 10^5 times smaller than its maximum value. The prior distribution was then defined up to this maximum relevant point:

$$\pi(\nu) = \begin{cases} 1/\nu_{\max} & \text{for } x \text{ in } [0, \nu_{\max}] \\ 0 & \text{otherwise,} \end{cases} \tag{9.5}$$

where $\mathcal{L}(x|\nu_{\max}) = 10^{-5}\,\mathcal{L}(x|\hat{\nu})$.

9.1.3 Upper limits and their distributions

A Bayesian credible interval defines a region within which the true value of a parameter is expected to lie with a specified probability. Note that this differs from the classical interpretation, where a confidence interval is defined such that in many repeated experiments, the interval would contain the true parameter value in a specified fraction of cases. The Bayesian interpretation is the intuitive one for most users, matching the instinctive interpretation of such statements[1].

Since a Bayesian credible interval simply states that the true value of ν lies within

[1]From D'Agostini, "I conducted an extensive poll in July 1998, personally and by electronic mail. The result [...] is that for the large majority of people the above statement ["95% confidence level lower bound of 77.5 GeV/c^2 is obtained for the mass of the Standard Model Higgs boson"] means that 'assuming the Higgs boson exists, we are 95% confident that the Higgs mass is above that limit, i.e. the Higgs boson has 95% chance (or probability) of being on the upper side, and 5% chance of being on the lower side,' which is not what the operational definition of that limit implies." [236]

the interval to a probability α, then the region can be found quite simply by determining an interval in ν within which a fraction α of the posterior probability density $p(\nu|\alpha)$ is contained [237]. For a normalised posterior density,

$$\alpha = \int_a^b p(\nu|x)\, d\nu\,. \tag{9.6}$$

Note that this does not specify a unique a, b but allows for an infinite number of different intervals to be selected [196]. In general however there are a few meaningful choices of credible interval available [238, 239]:

- The narrowest possible region in ν: this is that region which contains as much of the peak of the posterior as possible and excludes the tails as much as possible
- The central region. This region excludes the same fraction of the integral above as below: $\int_{-\infty}^{a} p(\nu|x)\, d\nu = \int_b^{\infty} p(\nu|x)\, d\nu$
- The region beginning at $-\infty$: the upper limit. The nonfixed edge of the confidence interval gives a probable upper bound on the true value of ν

In the case of setting a boundary on the presence of signal events, the second choice provides the appropriate information. A standard probability to use in reporting a limit on the existence of signal is 95%. So the *confidence limit* on the amount of a certain signal present in the observed data is found by integrating the posterior density from 0 to the desired *confidence level*, the point ν_{up} at which 95% of the area is contained. It can now be stated that the true quantity of signal present in the data is less than ν_{up} with 95% probability.

Expected upper limits

In addition to reporting an observed upper limit on the signal presence, it is important also to provide an expected upper limit and its uncertainties. This provides an answer to the question of how much the upper limit could be expected to move due to statistical fluctuations alone, given the best null hypothesis model available, and what the central value of this spread of possible results is. When reported together with the observed limit, the expected limits provide a context for interpreting the measured results by giving an idea of the number of standard deviations between the observed limit and what would be expected

in the null case. The standard method for doing so in ATLAS is frequentist: an ensemble of pseudo-datasets are generated from a chosen background prediction, the upper limit of the signal is calculated for each using an identical Bayesian marginalisation to that which determined the observed value, and the resulting distribution of quantiles defines a central value and 1σ and 2σ error bands. The background estimate used here takes the values of all nuisance parameters which give the maximum likelihood when the signal strength is fixed at $\nu = 0$, since this is the best available background-only (null) hypothesis which accounts for all systematic uncertainties.

Given how closely the final output now resembles a frequentist result, users must take care to distinguish the meaning of reported confidence bands on an observed limit in the two different methodologies. Both methods quote an expected value which is the median upper limit on the signal at 95% CL, but in the classical case this is the value of ν for which the median frequentist p-value is 0.05 while in the Bayesian case it is the median ν below which 95% of the signal posterior is found [230].

9.2 Implementation of Bayesian limits

With the full procedure in view for obtaining observed and expected limits on signal strength, the challenge is to make it computable. To find the desired posterior probability density, the integral of the likelihood and prior distributions must be performed over all nuisance parameters. There are a number of possible options for performing this numerical integration.

One option, used by the earlier iterations of the dijet analysis, is a simple grid calculation. Each of the n dimensions in the parameter space is divided into some number of intervals m and for each node in each dimension, a prediction for the background and signal are calculated. Then the likelihood at each node is weighted with the values of the priors for all parameters at that point and the results combined. This calculation has several drawbacks: first, the calculation time grows like m^n. Second, it focuses equal attention on unlikely nuisance parameter combinations as on likely ones, leading to an under-exploration of the most important regions of phase space. Third, it is challenging to combine the effects of different nuisance parameters using only the information on each of them independently, since the combination of their effects is not always clear to describe. This point will reappear shortly.

An option which solves several of these issues is to perform the numerical integration using points in the parameter space which are then propagated through the entire analysis to generate a different final spectrum for every choice of $\boldsymbol{\theta}$. This eliminates the third issue above by allowing the parameter to interact and produce an entirely accurate final description, instead of needing to attempt a *post hoc* combination. By selecting the random points in parameter space using weights from their priors, the spread of the points can also be controlled to address the issue of the focus. But for an analysis like dijets with an enormous number of signal MC events, the computational requirements of redoing the entire analysis for each of some 10^6 parameter combinations are absolutely crippling.

A third option, and the one selected by the dijet analysis in the full 8 TeV paper, uses a more sophisticated Monte Carlo method to perform the numerical integration. This retains one of the challenges of the first method, that of defining a likelihood function which can adequately combine the effects of all systematics after the data analysis has been completed. But it fixes the other issues faced by the previous methods: using a Markov Chain Monte Carlo method ensures that the calculation focuses on regions of interest and that the computational complexity is polynomial [240].

9.2.1 Markov Chain Monte Carlo and the Bayesian Analysis Toolkit

Markov chain Monte Carlo (MCMC) is a method for calculating the posterior distribution based on sequentially selecting values of ν from an approximate distribution which is updated on each draw to better reflect the results. Since the approximate distributions of the parameters are improved at each step this process quickly converges on the correct target distribution. The term "Markov chain" refers to a sequence of variables $\theta_0, \theta_1, \theta_2, \ldots$, for which each new value θ_i depends only on the most recent value θ_{i-1} and has no memory of the steps preceding that (*memoryless*) [241]. Generally, several such chains are seeded at independent points in the parameter space and the process runs until the chains have converged to the same final distribution [229].

The dijet limit setting employs the Markov chain defined by the Metropolis algorithm [242, 243]. This is a stochastic algorithm which proceeds between iterations by proposing a possible next point θ^* using a *jumping algorithm* $J_t(\theta^*|\theta)$ dependent on the iteration number t and current location θ, then always accepts the new point if it improves the probabil-

ity density $p(\theta|x)$ and accepts or rejects it with probability $r = p(\theta^*|x)/p(\theta|x)$ otherwise. The resulting chain can be shown to converge under a few simple conditions to the correct posterior distribution: see for example [229].

ATLAS analyses requiring an MCMC implementation are encouraged to use the Bayesian Analysis Toolkit (BAT), a package designed for physics use which interfaces easily with HEP analysis code [244, 245]. BAT takes as input a user-defined likelihood function and prior distributions for each nuisance parameter and for the parameter of interest, performs the marginalisation over all parameters, and returns the posterior for each parameter marginalised over the others. It furthermore provides several tools for quantifying the results of the marginalisation, such as comparing the posterior to prior distributions of all parameters, with most likely values and their quantiles, to ensure every parameter is well-behaved.

The challenge for the user is to provide a physically meaningful definition of the likelihood. The likelihood function must return a single number representing the comparison between data and an expectation e for the signal plus background hypothesis for any combination of parameters $(\nu, \boldsymbol{\theta})$. Thus there are several requirements this function describing the expectation must satisfy:

- It must be continuously defined throughout the allowed parameter space. Since the Markov chain may return any real value of the parameters, the expectation $e(\nu, \boldsymbol{\theta})$ must be calculable everywhere. This can be achieved by extrapolation between closely spaced templates demonstrating the parameter effects.
- It must be easily extensible with the addition of new nuisance parameters.
- It must be fast to calculate. The number of steps in each Markov chain required to provide an accurate output is well over 10^6, and each step requires a computation of the expectation.
- It must provide an accurate physical description of the system for all combinations of nuisance parameter values. This is the most challenging requirement and will be discussed in the following section.

9.2.2 Constructing a likelihood for the MCMC

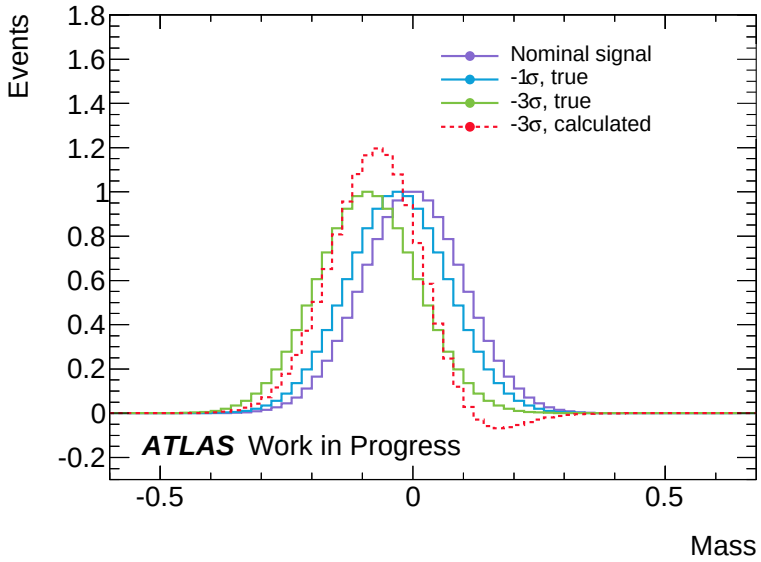
The naïve approach to constructing such a likelihood, and the one implemented by the majority of analyses using BAT, defines the difference in a signal or background data histogram (*nominal histogram*) using pairs of histograms to describe the effects of those systematics on the nominal, and then combining the effects of all systematics. One pair of histograms is generated for each nuisance parameter by rerunning the analysis code with the values of that parameter shifted by $+1\sigma$ for the first histogram and -1σ for the second. Then for a selected shift $\delta\theta$ in that nuisance parameter away from its central value, an effect on the nominal histogram is computed by shifting the content of each bin by an amount $\delta\theta \cdot (b_{\pm 1\sigma} - b_{\text{nominal}})$. This effect is calculated for each systematic uncertainty at its specified $\delta\theta$ and the effects are summed bin-by-bin to produce a final histogram, meant to represent the shape of the signal or background distribution for parameter values $\boldsymbol{\delta\theta}$. Where b_f is the final content in a single histogram bin, b_{nom} is the content of that bin in the nominal histogram, and $b^i_{\pm 1\sigma}$ is the bin content in the template histogram representing a shift of $\pm 1\sigma$ in nuisance parameter i, then [245, 246]:

$$b_f = b_{\text{nom}} + \sum_i \delta\theta_i (b^i_{\pm 1\sigma} - b_{\text{nom}}) \,. \tag{9.7}$$

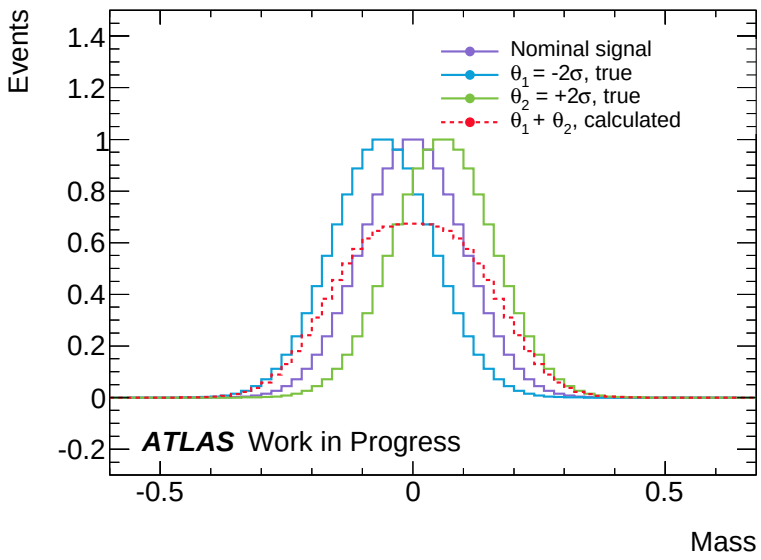
There are several problems with this approach, all of which take effect if the systematic uncertainty being considered does not represent a change in efficiency (more or fewer events kept in a bin) but rather a shifting of events between bins. Take for instance the jet energy scale. A change in this uncertainty could increase jet energy and move some set of events to a higher mass bin. No events are added or removed from the spectrum. Instead the systematic changes the shape or location of the signal peak. These "shifts" are not properly accounted for by the above method:

1. The assumption that a 2σ or 3σ systematic effect can extrapolated from a 1σ template does not hold in the case of a lateral shift as described above. Within $[-1\sigma, +1\sigma]$ the effect would be smooth and physically motivated but beyond that point the signal would be distorted vertically instead of displaced. The figure below illustrates this scenario. The purple line is a nominal signal sample and the blue line is the signal after a shift of -1σ in this nuisance parameter; these are used to find the template for generating shifts. The true value of the signal shifted by -3σ in the nuisance parameter

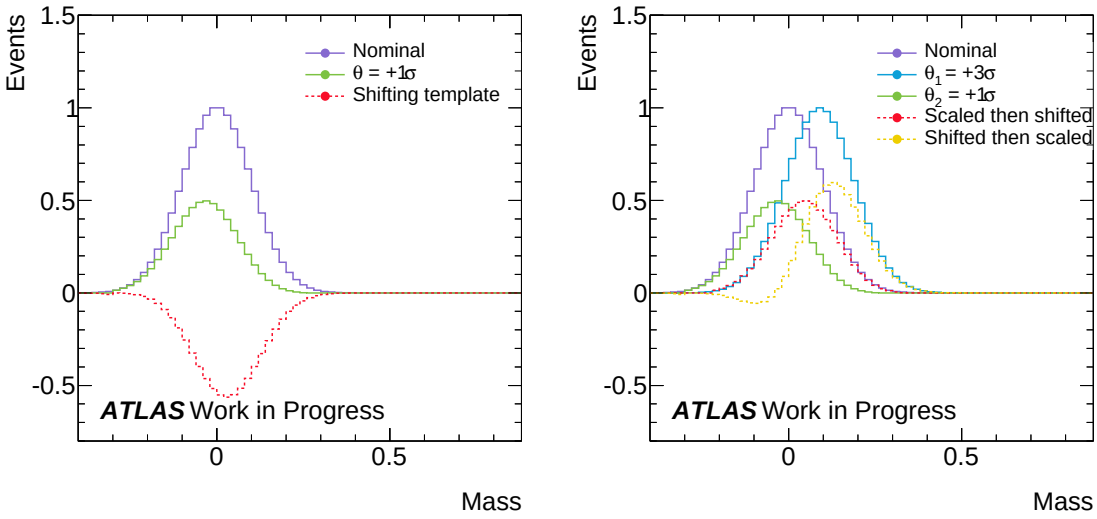
is the green curve. The red histogram results from using the 1σ template and scaling the difference by 3: it fails to reproduce the true signal shape or normalisation.



2. The combination of several systematics which all produce a lateral shift in the signal fails using this system. To illustrate, suppose two systematic uncertainties θ_1 and θ_2 both shift events between bins proportionally to the value of $\delta\theta$. In this case a set of parameter values $\delta\theta_1 = -1\sigma$, $\delta\theta_2 = +1\sigma$ would in theory shift the signal peak first to lower masses, then to higher masses, with a final result looking much like the original signal. If instead the application of the first systematic bin-by-bin removes events from the mass region where the peak centre originally stood and adds events to a lower mass region, while the application of the second systematic removes events from the same original location and moves them to a higher mass region, then the combined effect of the two systematics is rather to flatten the peak than to reproduce its original shape. An example is the case of two separate jet energy scale components. The figure below illustrates the result of two nuisance parameters, both of which shift the peak of the signal towards higher or lower masses, when they are moved by 2σ in opposing directions. The result of displacing a peak in both directions should be a return to the original spectrum; instead, a stretched version of the signal shape at lower normalisation is produced.



3. A nuisance parameter whose effects were calculated on a specific signal template at a bin-by-bin level are correct when applied using templates like those described by the naïve method, but would be incorrect if applied after a systematic which had shifted the location of events between bins. With a peak successfully shifted to a different mass location, applying a bin-by-bin difference calculated to effect a vertical shift in the original template would no longer alter the signal peak but rather would shift some location in the tail of the distribution (the original peak location) up and down instead. Therefore in addition to the need for a mechanism to describe shifting events, it is also necessary to apply systematics in a certain order to preserve their meaning. On the left of the figure below, a systematic uncertainty distorts the signal according to the dashed template shown. On the right, this effect is combined with a 3σ shift in a systematic which moves the signal peak laterally. The result is dramatically different if the template is applied first (red dashed line) than if the lateral shift is applied first (yellow dashed line).



In order to address these issues, all systematics used in the dijet analysis were divided into three types: those which affect the template in a mass-dependent manner independent of the signal form (such as beam energy effects) and are thus correctly described using template histograms, those which move events from one location to another within the spectrum, and those which affect the spectrum proportionally to its height and can therefore be described as simple scaling effects. The details of the three systematic types and their combination follow.

Template-based systematics

For systematic uncertainties which are best described by histograms defined bin-by-bin and applied to the nominal template, the second problem does not apply and the third is eliminated by ensuring that such effects are always applied first, before any shifts. The problem of linearity may not apply in all cases, but is addressed by using more than two shifted templates and ensuring that a sample template exits for the endpoints of the parameter range. Thus if the nuisance parameter is constrained such that $\theta_i \in [-3\sigma, +3\sigma]$ then one template must exist in the range $\leq -3\sigma$ and one in $\geq +3\sigma$, and more templates should be used in the intervening locations to increase the accuracy by only assuming linear behaviour between the nearest values of θ for which samples are available. To calculate the adjustment for a nuisance parameter, the nominal template is normalised to 1 and all the variation templates are normalised such that their areas are in the correct proportion to that of the nominal histogram. Then an adjustment to be applied to each bin is calculated. Let $b_{\theta+}$ be the bin value in the template thus normalised corresponding to a nuisance parameter value θ^+ and

$b_{\theta-}$ the same for parameter value θ^-, where these are the nearest available templates to the drawn parameter θ: $\theta^- < \theta < \theta^+$. The adjustment which must be added to the bin of the normalised template in order to account for this systematic is:

$$b_{adj} = b_{\theta-} + \left(\frac{\theta - \theta^-}{\theta^+ - \theta^-}\right)(b_{\theta+} - b_{\theta-})\,. \tag{9.8}$$

Then for a system with multiple template-based systematics, the final value in each bin is given by:

$$b_f = b_{\text{init}} + \sum_i b^i_{adj}\,. \tag{9.9}$$

For the dijet limit setting procedure, one template was defined for every 0.5σ step in the nuisance parameter value, so the systematic uncertainties were only assumed to behave linearly over that range.

Shape-changing systematics

A variety of options were explored to determine how best to describe nuisance parameters which move events between bins. If a single such parameter exists, it can be well described using the template-based method above where templates exist over the full range of the nuisance parameter. In this way the horizontal shifting behaviour is properly modelled over the full parameter range. However, this still fails to address the second problem above: it cannot accurately describe more than one such parameter at a time.

The first choice method for handling combinations of shifts was the ROOT template morphing methods, which interpolate from existing forms (usually functions) to find template shapes at points in between. Initial tests indicated, however, that the computation involved was too slow to make this a practical choice for a process which must be iterated over many times. Another option would involve parameterising the signal shape with a functional form and assessing how it changed with the systematics, but no functional forms were found which could adequately describe the signals.

A more computationally lightweight solution was to define *transfer matrices* to relate the location of events in systematic-shifted histograms to their locations in the nominal template to derive intermediate matrices representing a shift for any given θ, and then use the product

of these matrices to describe the shifting back and forward of events under the application of many such uncertainties. A range of matrices was defined identical to the range of templates in the above uncertainty: one transfer matrix per 0.5σ was calculated, with bin content equal to the fraction of events in bin x_i of the spectrum which migrated to bin y_i after applying the systematic shift. The bin values were extrapolated linearly between neighbouring matrices, similarly to the treatment of the template-based systematics:

$$M_\theta = M_{\theta-} + \left(\frac{\theta - \theta^-}{\theta^+ - \theta^-}\right)(M_{\theta+} - M_{\theta-})\,. \tag{9.10}$$

Then for several of these *shape-changing* systematic uncertainties, with no other types of systematic present, the final bin contents can be found by applying the product of all transfer matrices to a vector of the bin contents. Let v be such a vector. Then,

$$v_f = \prod_i M_{\theta i}\, v_{\text{init}}\,. \tag{9.11}$$

The matrix algebra is implemented using the ROOT `TMatrixTSparse` class in order to save all possible computation time due to the large number of empty bins in the transfer matrices.

A few conditions apply in order for this technique to work entirely reliably. First, the matrices must commute or be close to this behaviour in practice. Otherwise the order of application of the systematic uncertainties will affect the outcome. Furthermore, the fraction of events shifted between bins should be similar or only slowly changing across the full mass range, such that a shift in the peak does not change the manner in which it is affected by its systematic.

Scale-changing systematics

The simplest systematic uncertainties to handle are those which change only the scale of the spectrum with no dependence on the kinematic variable. These can be specified by a single number: this represents the percentage by which the value in each bin should change under the application of a $+1\sigma$ change in θ. Using a system which takes as input a single value in place of a template to represent such effects decreases computation time and increases numerical accuracy in the handling of such systematic uncertainties. With each *scale-changing* systematic defined by a parameter a, the final bin content in a spectrum after

their application is:

$$b_f = \prod_i a_i b_{\text{init}} \,. \tag{9.12}$$

The full implementation

The full likelihood creates an expected signal plus background value by combining all systematic uncertainties on each of the signal and background processes individually and adding them together. As mentioned above, in order to preserve the physical meaning of the systematics, they are applied in a fixed order: first all template-bases systematic uncertainties, then all shape-changing ones, then the scale uncertainties at the end. In order to ensure that the correct statistical weight is given to the data, the effective number of entries in each bin of the spectrum is used. This requires the expected value in that bin to be divided by the effective weight w_{eff} of the data so that they can be directly compared. The data is assumed to be Poisson distributed around the effective number of events N_{eff} in the bin. The final calculable likelihood after assembling all components is thus defined by:

$$\mathcal{L}(x|\nu,\boldsymbol{\theta}) = \prod_i \text{Pois}(N_{eff}, e(\nu,\boldsymbol{\theta})/w_{eff}) \tag{9.13}$$

where $\text{Pois}(a,b)$ is the probability of observing a number of events $\geq b$ given a Poisson distribution with parameter $\lambda = a$. The expected value e is:

$$e(\nu,\boldsymbol{\theta}) = \sum_{\phi=\text{sig, bkg}} \prod_k a_k \prod_j M_{\theta j}(v^{\phi}_{\text{init}} + \sum_i v^{\theta i}_{\text{adj}}) \tag{9.14}$$

where i are template-based systematic uncertainties, j are shape-changing, and k are scale-changing. If the resulting sum is < 0 the result is reset to 0 since negative events are never expected.

The calculation speed of this process is slower than the template-only approach released in BAT by less than a factor of two, although the exact value varies with the proportion of systematics of each of the three types. It is the user's responsibility to consider which type of systematic to employ when creating a new nuisance parameter. The decision is a physically motivated one:

- A systematic which changes the efficiency should be template-based. These systematic

uncertainties were template-based in the dijet limit setting calculation:

1. Fit quality
2. Function choice
3. Beam energy

- A systematic which changes the bins in which events fall without changing (substantially) the number of events present should be matrix-based. These systematic uncertainties were matrix-based:

  1. All jet energy scale components (up to 7)

- A systematic which changes each bin by an amount proportional to its content should be scale-changing. These systematic uncertainties were scale-changing in the limit calculation:

  1. Luminosity uncertainty
  2. PDF scale acceptance uncertainty

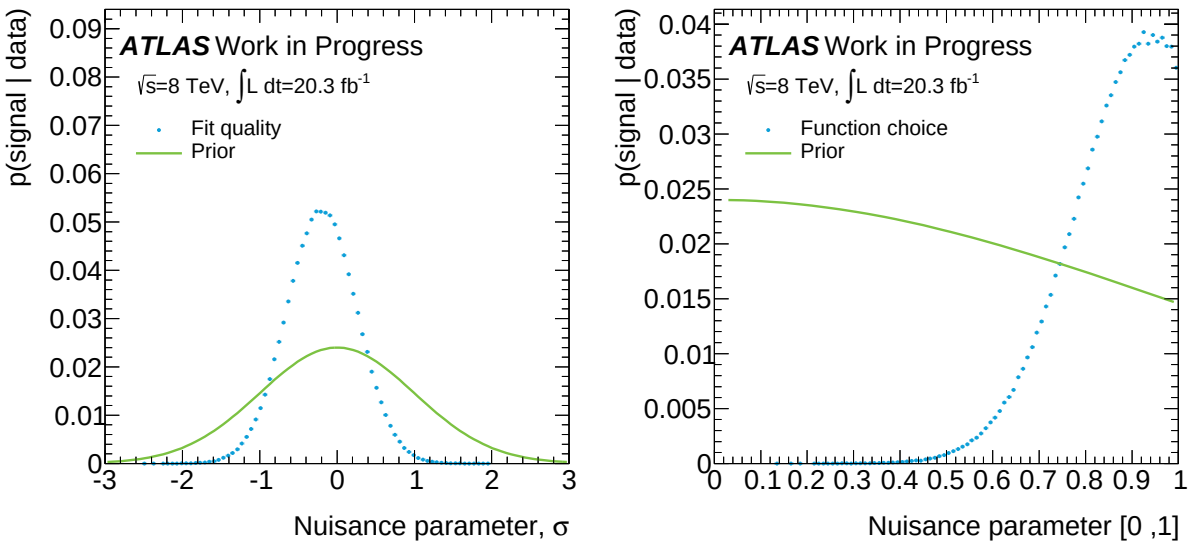
Impact of the detailed systematic treatment

A single objection has been levelled against the method of systematic uncertainty handling outlined above: when considering incorporating the dijet results into a combined limit effort, this was claimed to be unnecessarily complicated as the difference between the template-only and full treatment is negligible for small systematic effects. Indeed, the effect is small in this case, as the systematic uncertainties are all less critical to the final result than the statistical fluctuations in the data.

This does not, however, alter the fact that a treatment of the uncertainties using a template-driven method alone would be not be physically meaningful. If the difference between the two treatments is small, it is because the uncertainties themselves are small and should not be taken as a reason to use a fundamentally meaningless implementation of their effects. Doing so is tantamount to simply creating noise in association with the parameters rather than using them to describe a physical effect. As such the additional time investment involved in both developing this treatment and in performing the marginalisations which use it have been important additions to the integrity of the analysis and should be considered carefully in future iterations.

9.2.3 The complete dijet calculation

For every mass point in every signal sample for which Monte Carlo was available, the marginalisation using the above likelihood was performed. BAT provides in its output the marginalised posterior distribution for every parameter treating all other parameters as nuisances. Below are sample output plots for a randomly selected mass point with a single JES nuisance parameter: the q^* m_{q^*} = 2 TeV sample. Figure 9.1 presents the posterior distributions for all nuisance parameters. Each nuisance parameter shown here was given a Gaussian prior as denoted by the green line. The posteriors indicate that the preferred background estimate for the selected fit was constrained to a narrower range than the prior anticipated and that the probability drops sharply for values beyond the error bands calculated in the search phase. It also indicates that instead of the presupposed weak preference for the nominal fit function, the observed data showed a strong preference for the alternate (five-parameter) function. The remainder of the nuisance parameters exhibit very similar prior and posterior distributions, indicating that the nuisance parameters behave as expected and prefer central values to the predicted degree.

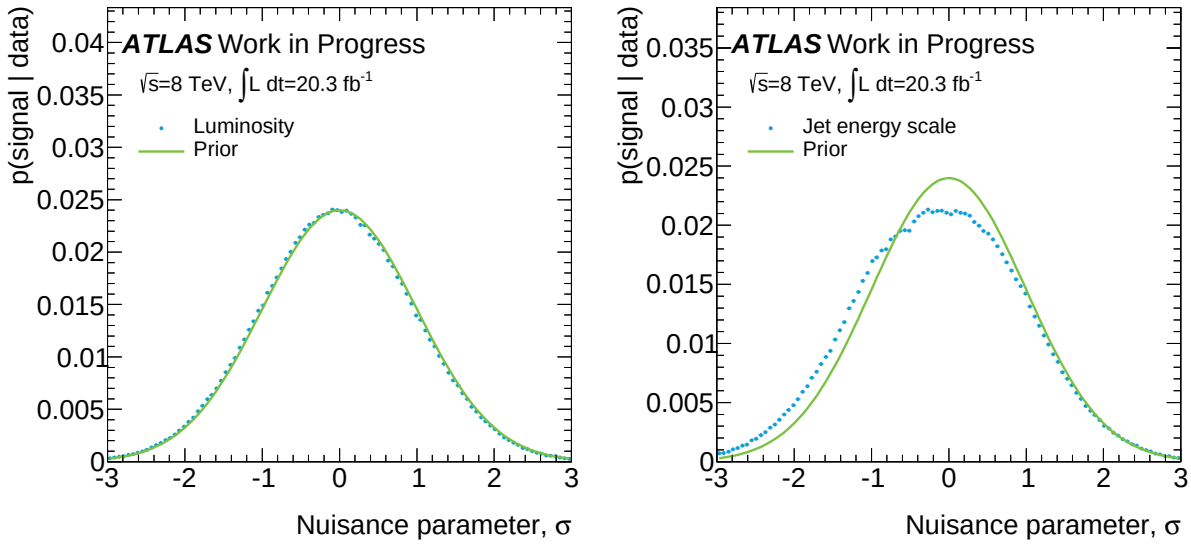


(a) Background fit

(b) Function choice

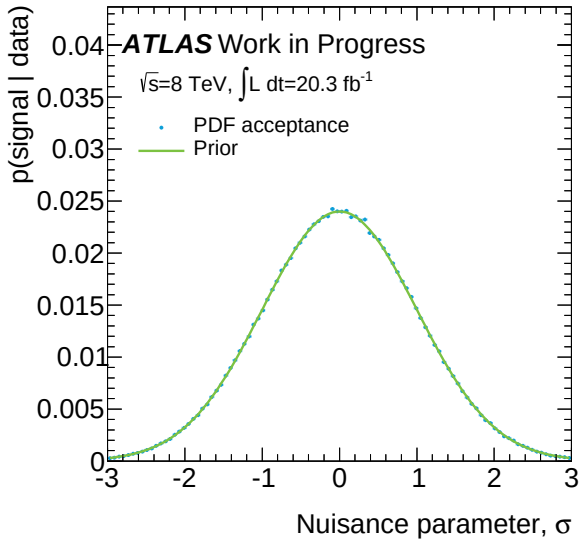
Figure 9.1: Posterior probability densities for all nuisance parameters after integrating over all other parameters (including signal). The signal sample used was a 2 TeV q^*.

Figure 9.2 presents the posterior probability density for the parameter of interest, the signal normalisation. The vertical line on this plot marks the 95% quantile of the distribution

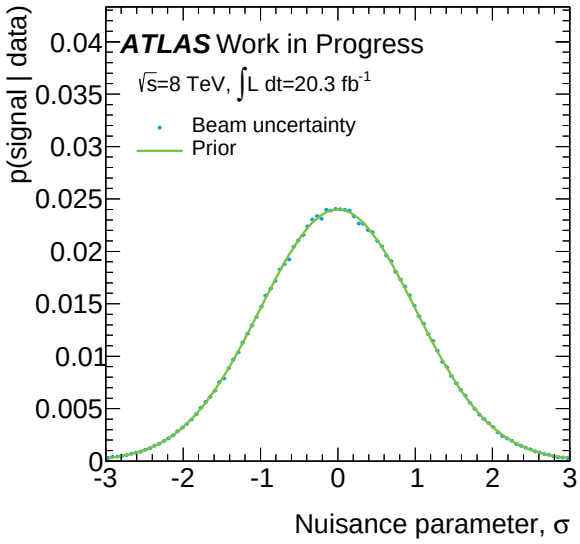


(c) Luminosity

(d) Jet energy scale



(e) PDF acceptance



(f) Beam energy

Figure 9.1: Posterior probability densities for all nuisance parameters after integrating over all other parameters (including signal). The signal sample used was a 2 TeV q^*.

and thus the observed upper limit on the number of 2 TeV q^* events present. Finally, Figure 9.3 provides the correlations between all parameters in a single image together with the posterior distributions of each.

The marginalisation began by running 5 Markov chains until they converged, which required order 14,000 iterations for seven parameters. The chains are subsequently re-run for 100,000 iterations and the result of this second set used to determine the posteriors.

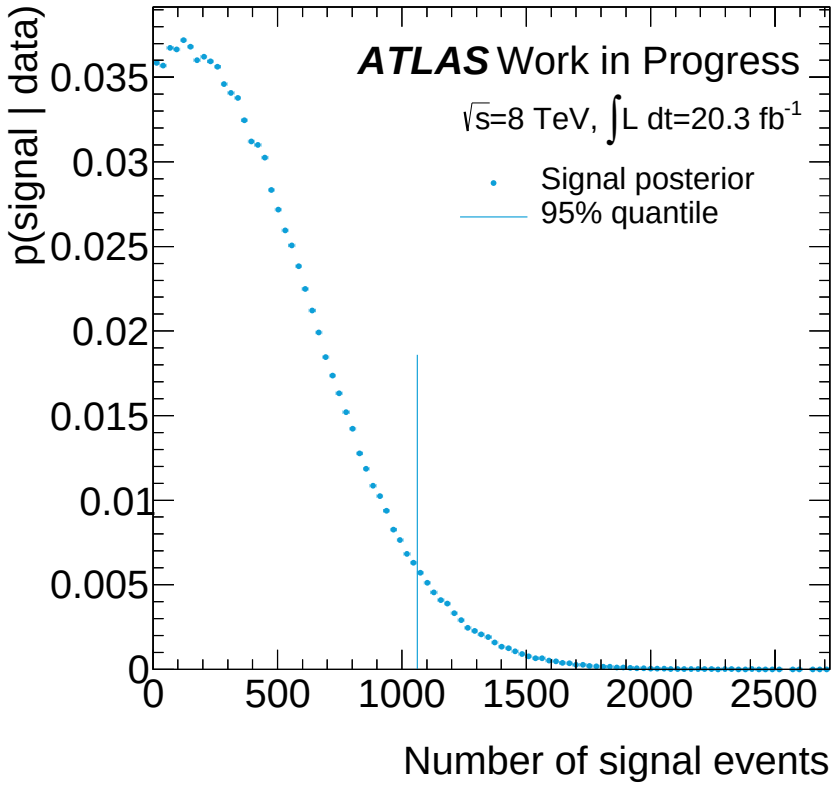


Figure 9.2: The posterior probability density $p(\nu, x)$ for a q^* sample at 2 TeVobtained from BAT after marginalising over all systematic uncertainties. The vertical line marks the 95% quantile of the distribution.

100 pseudoexperiments were generated from the best background-only model using Poisson fluctuations around the effective number of events. The marginalisation was rerun from the same initial state for each and the 95% quantile recorded in each case. The distribution of measured quantiles over the pseudoexperiments is shown in Figure 9.4 with the observed value in real data signified by the arrow.

9.3 Limits on benchmark models at 8 TeV

Once the data summarised in Figure 9.4 had been computed for every mass point in a given signal sample, it was converted into an observed limit on the cross section. The limit-setting procedure produces a measure of the maximum number of events of a given signal shape and type which could be present in the measured data. This number of events, divided by the luminosity of the dataset considered, is the recorded cross-section of the process.

The recorded cross-section will not necessarily equal the true cross-section σ of the

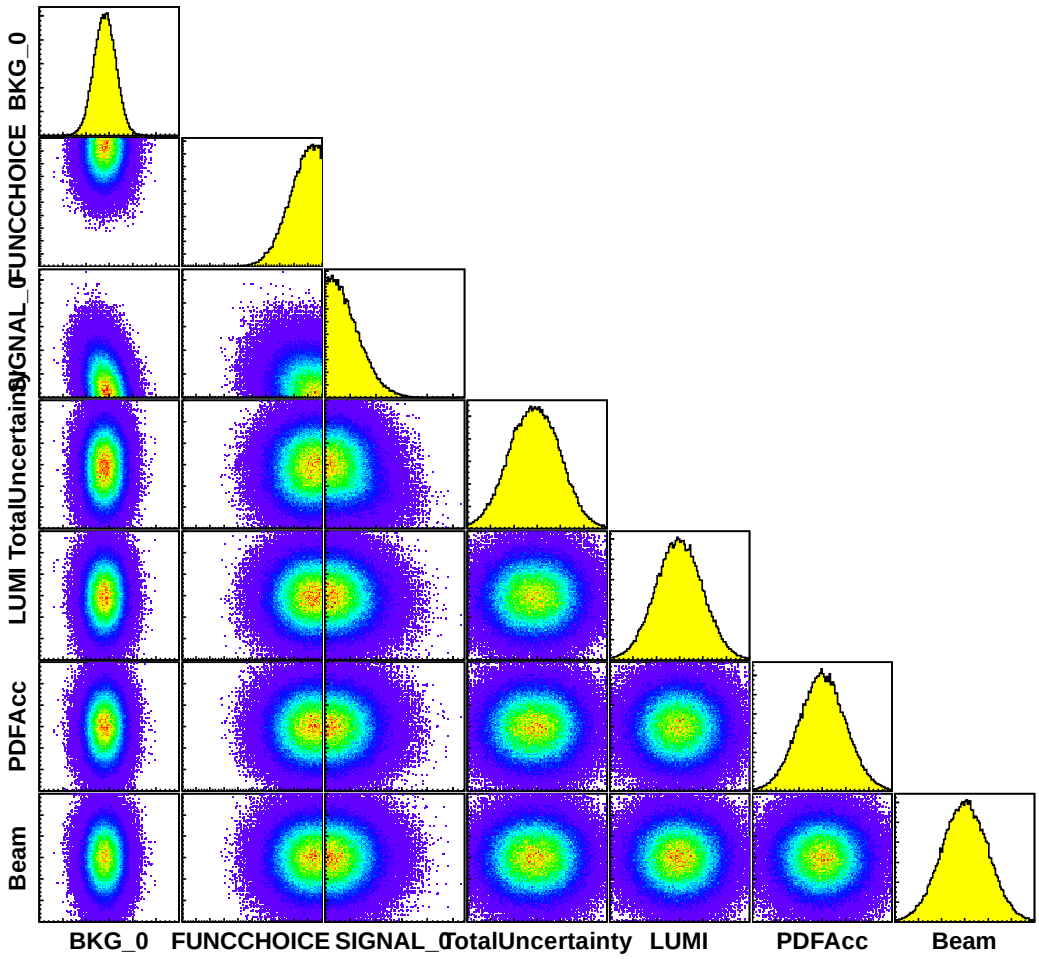


Figure 9.3: Correlations between and marginalised posteriors for all parameters. From left to right (top to bottom): background estimation, function choice, signal normalisation, jet energy scale, luminosity uncertainty, PDF acceptance, and beam energy uncertainty. The diagonal displays the posterior for each parameter while the remaining portions show the correlation between the parameters listed on the x and y axes.

particle. It is always scaled by the acceptance of the signal: where some events fell outside the analysis acceptance the number of events calculated by the limit setting procedure would be low, and thus the limits plotted and reported are quoted in terms of $\sigma \times \mathcal{A}$. An additional complicating factor is the branching ratio of the signal into dijets. Where the signal Monte Carlo was generated permitting all possible decay channels, only that fraction producing quarks and gluons detectable by the dijet analysis would have been recorded and the reported cross section would be scaled down by the branching ratio to dijets. Many

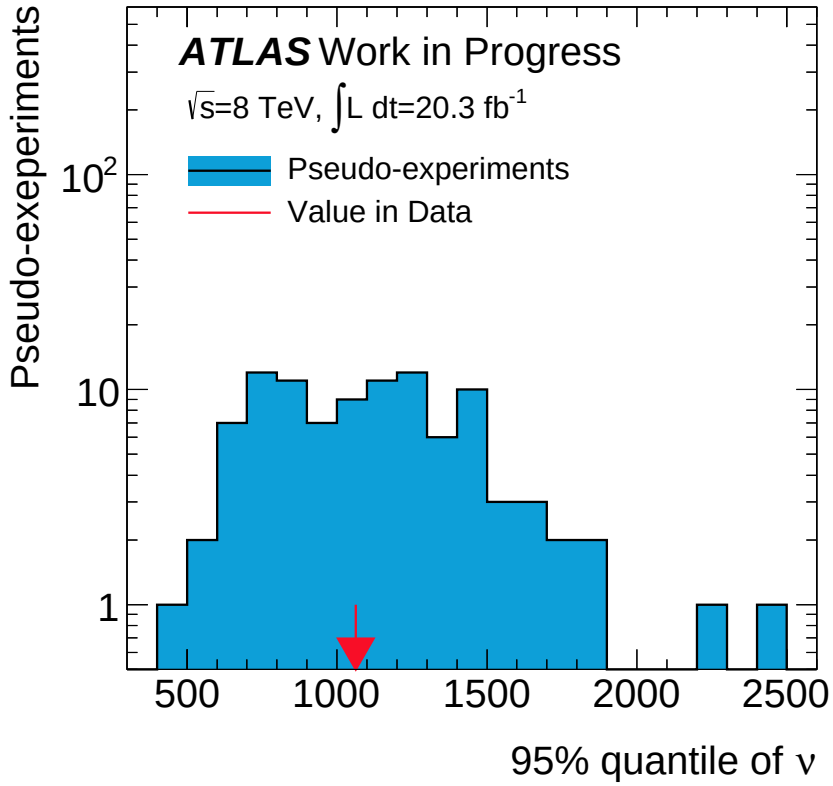


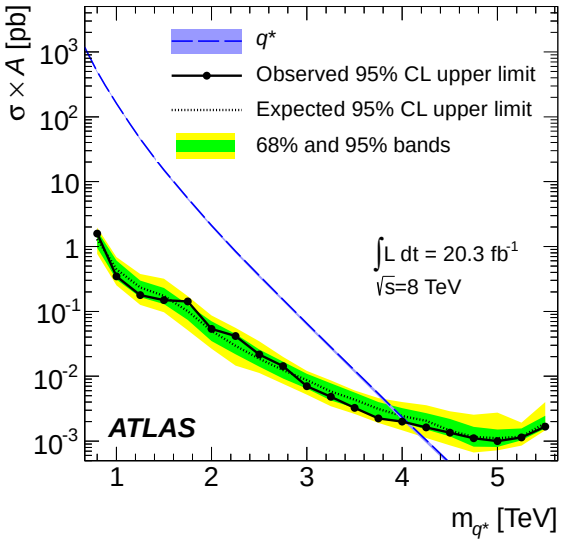
Figure 9.4: Distribution of 95% quantiles obtained from 100 pseudoexperimens for a sample of q^* MC at a mass of 2 TeV. Each quantile corresponds to the point below which 95% of the posterior probability distribution $p(\nu, x)$ is contained.

of the Monte Carlo samples used in this paper with only the dijet decay chain permitted: in these cases the calculated rate is exactly $\sigma \times \mathcal{A}$. For the signals with all decay channels active in the simulation (...), the results were rescaled by the branching ratio to provide plots consistent across all signals.

A theoretical prediction of the cross-section for each model is well-known. The model is considered excluded for all masses at which predicted cross-section is greater than the observed 95% upper limit on the same signal. The *observed limit curve* is extrapolated using an exponential function between neighbouring mass points while the nominal cross-section is similarly extrapolated between its known values at each simulated mass. The intersection of the two curves is taken as the upper mass limit on the model.

Figure 9.5 shows the observed limits and nominal cross-sections for all signals. The green and yellow curves are the 68% and 95% quantiles of the pseudoexperiments performed on the limit measurement, while the black dashed line is their mean. These provide a

measure of the *expected limit curve* with the 1σ and 2σ uncertainty bands which show its expected variance under statistical fluctuations. The beam energy uncertainty is illustrated as an error band around the signal cross-section curve; in the case of the W' an additional uncertainty on the next-to-next-to-leading order cross-section is added in quadrature to this error band. The limits for the two black hole generators are shown together in Figure 9.5e. Although the reported limits for these two models were calculated separately, their observed and expected limit curves were so similar by eye that they were combined into a single plot for publication purposes. The W^* limits are also shown combined in Figure 9.5d. Here, leptophobic W^* observed limit was calculated and displayed but the theoretical cross-section for the leptophilic model is also included since the two vary in acceptance by only 1%. The observed and expected limit values for all models are summarised in table 9.1.

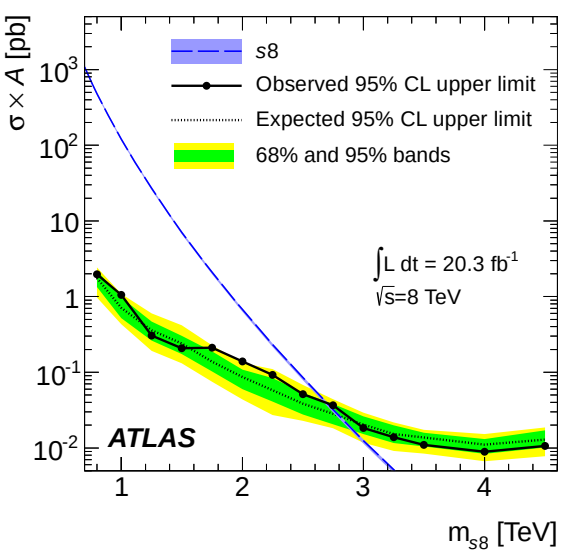


(a) Excited quark q^*

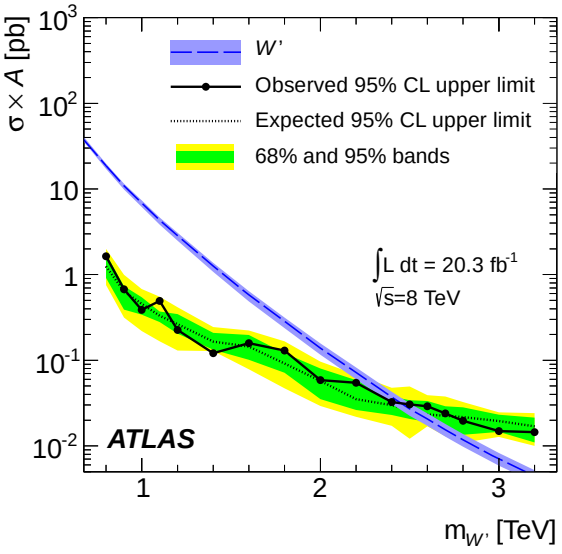
Figure 9.5: Observed and expected limits on benchmark models. The green and yellow bands correspond respectively to 1σ and 2σ variations in the expected limits.

9.4 Limits on generic resonance shapes

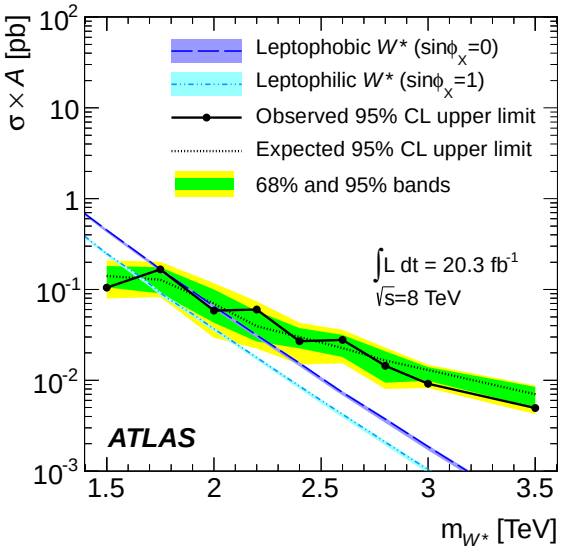
In addition to the model-dependent limits, several limits are calculated for generic template shapes designed to emulate the behaviour of an abstract resonance. These are provided for the benefit of the theory community where they are widely used for reinterpretations of the analysis in the context of other models. Gaussian shapes have been used for generic resonance limits since the earliest ATLAS dijet analyses and provide an element of continuity



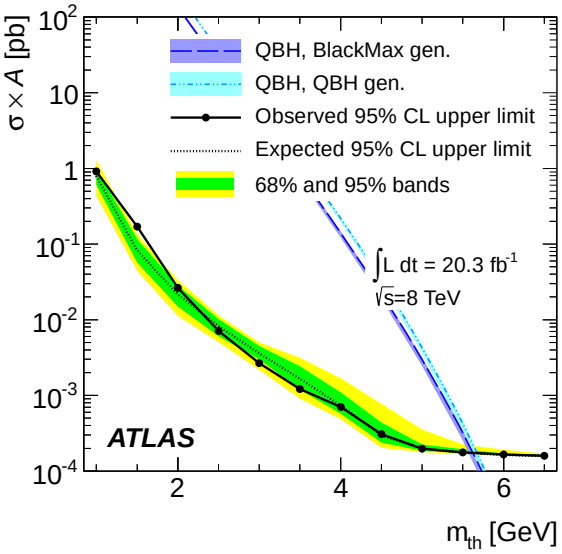
(b) Color scalar octet $s8$



(c) Heavy boson W'



(d) Excited chiral boson W^*



(e) Quantum black holes QBH,BLACKMAX

Figure 9.5: Observed and expected limits on benchmark models. The green and yellow bands correspond respectively to 1σ and 2σ variations in the expected limits.

comparable to past results. This analysis also introduced generic limits based on Breit-Wigner resonance shapes, taking into account pdf effects and detector effects, which have more likeness to a real physical signal.

Only a subset of the systematics affecting model-dependent limits are relevant here. The jet energy scale uncertainty is also simplified to a single component which for which a variation of 1σ effects a 3% shift in the Gaussian mean but does not alter its width. The systematics considered for generic resonance limits are:

Model and Final State	95% CL Limits [TeV]	
	Expected	Observed
$q^* \to qg$	3.98	4.06
$s8 \to gg$	2.80	2.70
$W' \to q\bar{q}'$	2.51	2.45
Leptophobic $W^* \to q\bar{q}'$	1.95	1.75
Leptophilic $W^* \to q\bar{q}'$	1.66	1.65
QBH black holes	5.66	5.66
(q and g decays only)		
BLACKMAX black holes	5.62	5.62
(all decays)		

Table 9.1: The 95% CL lower limits on the masses and energy scales of the models examined in this study. All limit analyses are Bayesian, with statistical and systematic uncertainties included.

- Luminosity
- Background fit statistical uncertainty
- Background parameterisation choice
- One JES uncertainty component.

9.4.1 Gaussian limits

The simplest generic resonance shape used in the limit setting is a Gaussian distribution. It is assumed that events from a signal of some arbitrary sort, having passed through the event selection procedure, would form a normal distribution with mean equal to the signal mass and some width related to the intrinsic width of the resonance combined with the detector effects. Several Gaussian widths are provided in the analysis (7%, 10%, and 15%), and the selection of the correct width for a given reinterpretation would be made by emulating the analysis cuts in a detector-level simulation and comparing the result to a Gaussian fit.

A smooth histogram representing the shape of a signal with no statistical fluctuations is created from a Gaussian function and trimmed to its central 95% of data to remove the infinitely extending tails. The histogram is passed into the limit setting code and the maximum number of events calculated by a limit-setting procedure identical to that using a Monte Carlo signal template. The resulting maximum number of events, divided by the dataset luminosity, can be interpreted as an upper limit on the $\sigma \times \mathcal{A} \times \mathrm{BR}$ of any model of

interest matching the Gaussian form. The observed limits on Gaussians with width-to-mass ratios of 7%, 10%, and 15% as well as with width equal to the experimental dijet resolution at each mass are shown in Figure 9.6. This last curve is the narrowest distinguishable signal and therefore the best limits accessible to the analysis.

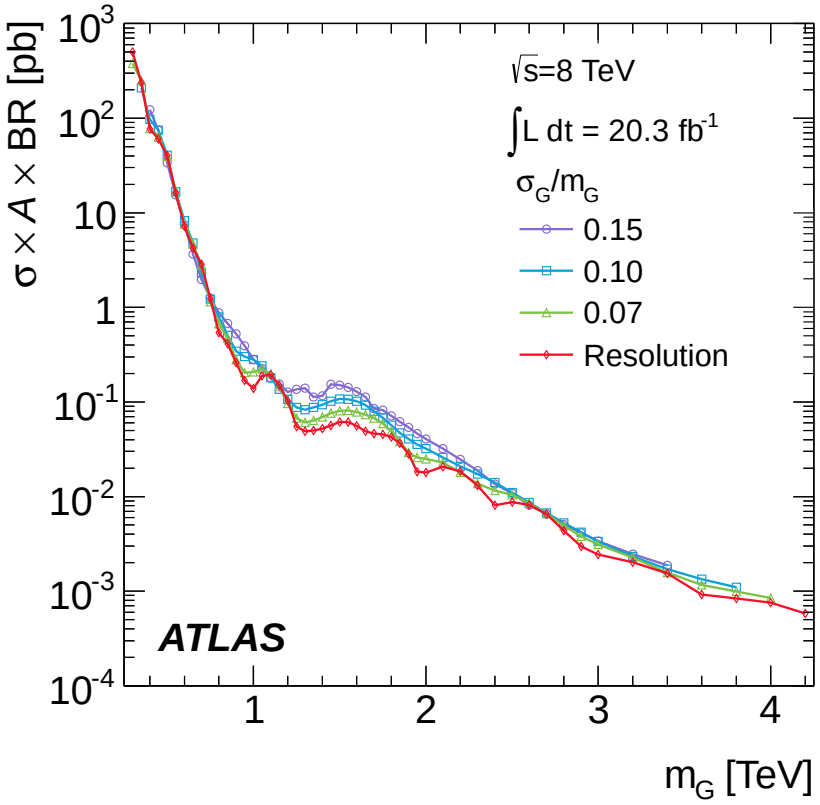


Figure 9.6: The 95% CL upper limits on $\sigma \times \mathcal{A} \times$BR for a simple Gaussian resonance decaying to dijets as a function of the mean mass, m_{G}, for four values of $\sigma_{\mathrm{G}}/m_{\mathrm{G}}$, taking into account both the statistical and systematic uncertainties.

9.4.2 Breit-Wigner limits

This analysis newly introduces a set of generic limits designed to more closely mimic the behaviour of a true physical resonance. These signal shapes are characterised by the mass and intrinsic width of the resonance to be considered, removing any obligation to undergo analysis simulation on the part of a user in the theory community. PDF and acceptance effects are accounted for in the signal shape generation, which is then convoluted with parton to particle showering and detector resolution effects to produce a final template shape.

The non-relativistic Breit-Wigner equation is given by:

$$f(x, m, \Gamma) = \frac{1}{2\pi} \frac{\Gamma}{(x-m)^2 + (\Gamma^2/4)} \,. \tag{9.15}$$

This, more than a Gaussian shape, is the expected distribution for the decay of a short-lived particle with mass m and intrinsic width Γ. For a collision with infinite centre-of-mass energy and a particle created at rest, this shape would well describe the distribution of invariant masses of the parton pair produced. However, two effects must be considered for real particle collisions at the LHC.

Let s be the centre-of-mass energy and x_1, x_2 be the fractions of the proton momenta carried by the two interacting partons. For a particle of mass m to be produced, it must hold that $m^2 = s \cdot x_1 \cdot x_2$. As m approaches s this constrains $x_{1,2}$ to larger values and thus less probable configurations. The probability of producing a pair of final-state particles at any specific invariant mass can be quantified through their *parton luminosity*, the rate of change of luminosity in that channel with respect to $d\hat{s}\, dy = s\, dx_1\, dx_2$ [247]. This parton luminosity is expressed in terms of the pdfs of the interacting initial state partons, such that $\mathcal{L}(qq \to \text{two jets}) \neq \mathcal{L}(gg \to \text{two jets})$: four parton luminosities are considered for each possible final state qq, $q\bar{q}$, qg, and gg. Although the exact shape changes between parton pairs, the trends are similar: parton luminosities are highest at low dijet invariant mass, drop slowly initially, and plunge steeply to zero as the dijet mass approaches s. This behaviour reflects the increasing improbability of producing a heavy resonance as its mass approaches $\sqrt{s}$. The Breit-Wigner distribution produced by a resonance in an ideal situation is thus suppressed at each point by the appropriate parton luminosity, producing substantially higher low-mass tails and a reduced peak compared to the bare function of Equation 9.15. Parton luminosities were modelled using the CT10 pdf set [140].

As the resonance mass changes, so too does the acceptance. A resonance with $m \sim \sqrt{s}$ would be produced at rest, but lighter particles can be produced with a substantial imbalance in x_1, x_2 leading to a boost along the beam axis. For sufficiently large boosts, at least one jet will always be produced outside the analysis acceptance. At intermediate boosts, a calculable fraction of events will be lost outside the analysis acceptance. The acceptance fraction for each mass is calculated by integrating the acceptance in rapidity over the distribution of boosts. Near 1 at high masses, it drops increasingly rapidly towards lower masses (larger

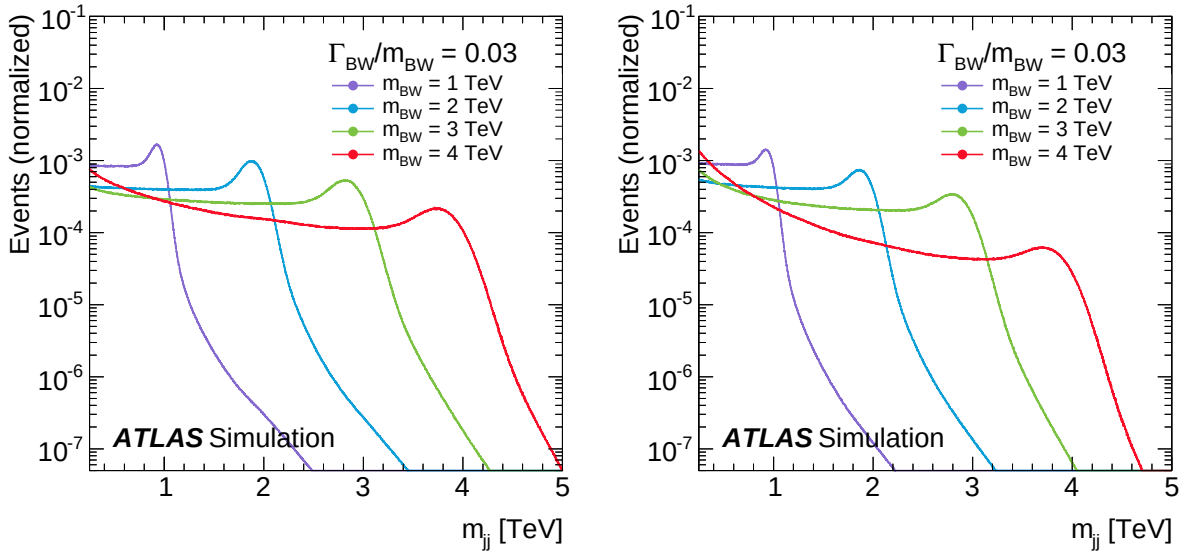
boosts) before plateauing at a flat acceptance of ~ 0.54 where the boost range becomes limited by the phase space of extreme boosts in which some events are always lost. The full "raw" Breit-Wigner generic signal shape before any detector effects is thus the product at each point in mass of the simple Breit-Wigner distribution, the acceptance fraction, and the parton luminosity.

Showering and non-perturbative effects introduce a slight dependence on Monte Carlo modelling, though not on a specific signal. The resolution due to parton showering and hadronisation was calculated by comparing the m_{jj} spectrum in Monte Carlo at particle and at parton level. The resolution $m_{jj}^{\text{particle}}/m_{jj}^{\text{parton}}$ was found using both PYTHIA and HERWIG++, and the latter was selected for the limits as it was found to produce the more conservative result. The raw Breit-Wigner was convolved with the resolution shape to produce a more smeared particle-level resonance distribution.

Detector effects similarly broaden the final result. These were simulated by convolving the generic signal shape with a Gaussian distribution whose width at each mass point was drawn from the dijet resolution distribution shown in Figure 6.11.

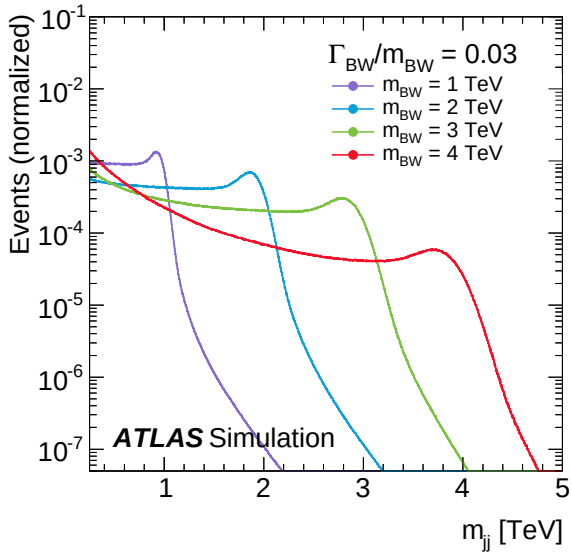
The final generic signal shapes, resulting from the product of a Breit-Wigner function with pdf and acceptance effects, then convolved with showering and detector resolutions, are shown in Figure 9.7 for a sample of mass points and at the fixed intrinsic width-to-mass ratio of 0.03. The resulting limits are given in Figure 9.8. Four intrinsic widths of BW are shown in each limit plot, and four plots are present, one for each of the parton luminosity distributions. The difference in limits, especially at high masses, between Breit-Wigner signals with different parton luminosities is evident in Figure 9.8. The most striking difference is in the gg pdf and is due to the substantial low-mass tails affecting gluon parton luminosities.

Reinterpretation of these generic limits in terms of a new resonance model to be tested is simpler for the theory community, since showering and detector effects are already accounted for by the analysis team and need not be simulated by a user. Only the mass, width, and production mode of a resonance needs to be known to obtain the correct limit on $\sigma \times \mathcal{A} \times \text{BR}$ from the plots above and the detailed tables of limit values included with the publication.

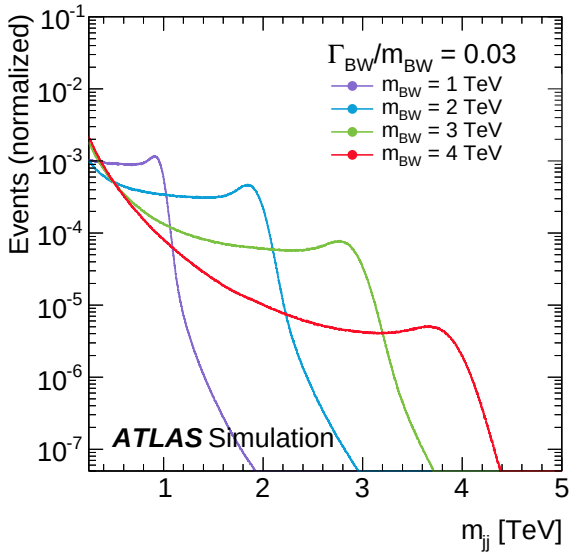


(a) qq PDF

(b) $q\bar{q}$ PDF

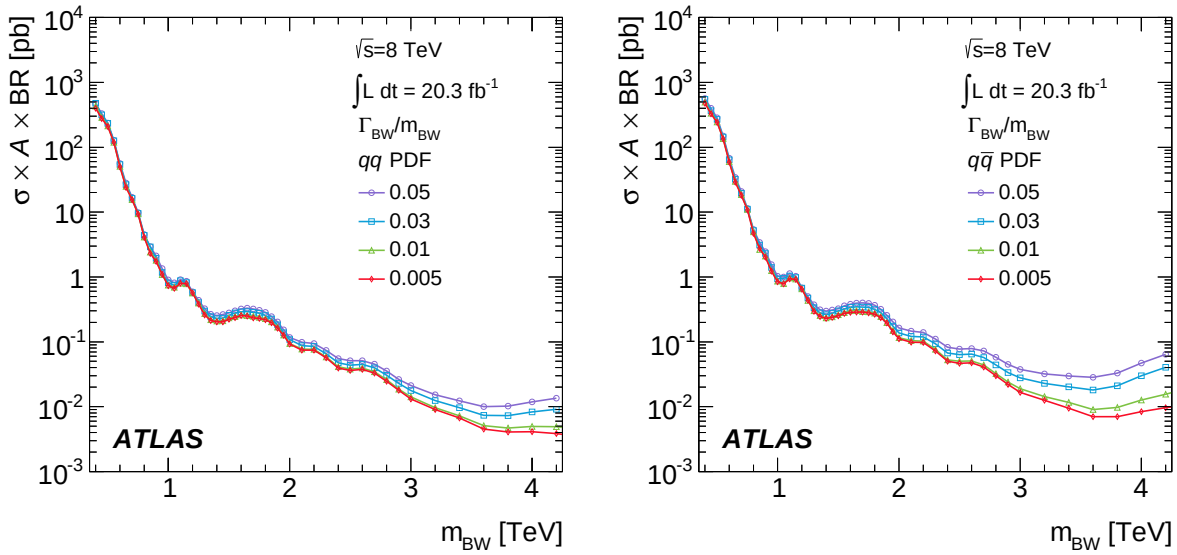


(c) qg PDF



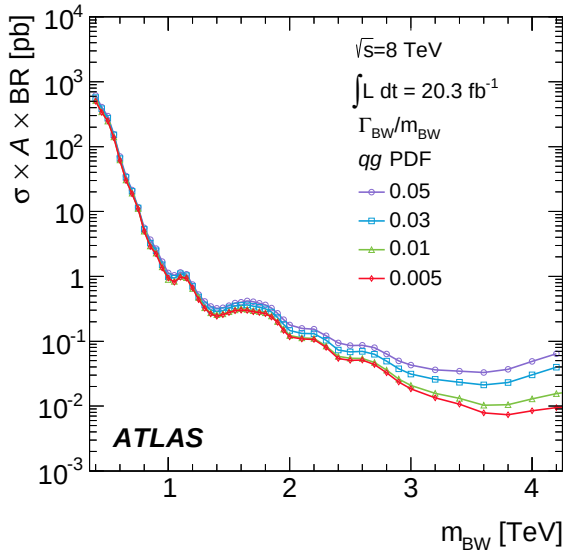
(d) gg PDF

Figure 9.7: Signal shapes for model-nonspecific resonances based on a Breit-Wigner functional form. The BW form was multiplied by a different parton distribution function in each subfigure and the results were convolved with two Gaussians whose widths represented showering effects and the detector resolution.

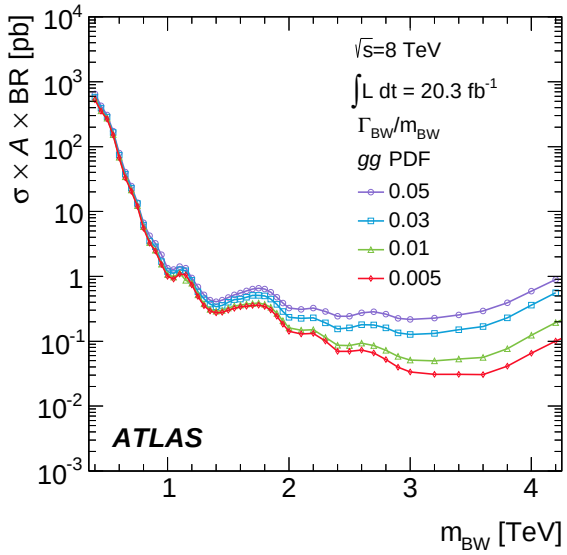


(a) qq PDF

(b) $q\bar{q}$ PDF



(c) qg PDF



(d) gg PDF

Figure 9.8: Model-nonspecific limits derived from the Breit-Wigner signal shapes. The 95% CL upper limits on $\sigma \times \mathcal{A}$ are shown as a function of the mean mass, m_{BW}, for different values of intrinsic width over mass ($\Gamma_{\mathrm{BW}}/m_{\mathrm{BW}}$), taking into account both the statistical and systematic uncertainties.

9.5 Evolution of limits

The limits from this analysis are put into context by comparing them with pre-existing limits. First, the evolution of Gaussian limits is shown over the three most recent ATLAS dijet publications (Figure 9.9). The current analysis shows, as expected, an improvement in limits at all mass points over the smaller datasets at $\sqrt{s}$ = 7 TeV. Furthermore, this plot illustrates the extension in limits to lower masses not covered by the previous results. All Gaussian limits shown in this plot correspond to a width-to-mass ratio of 10%.

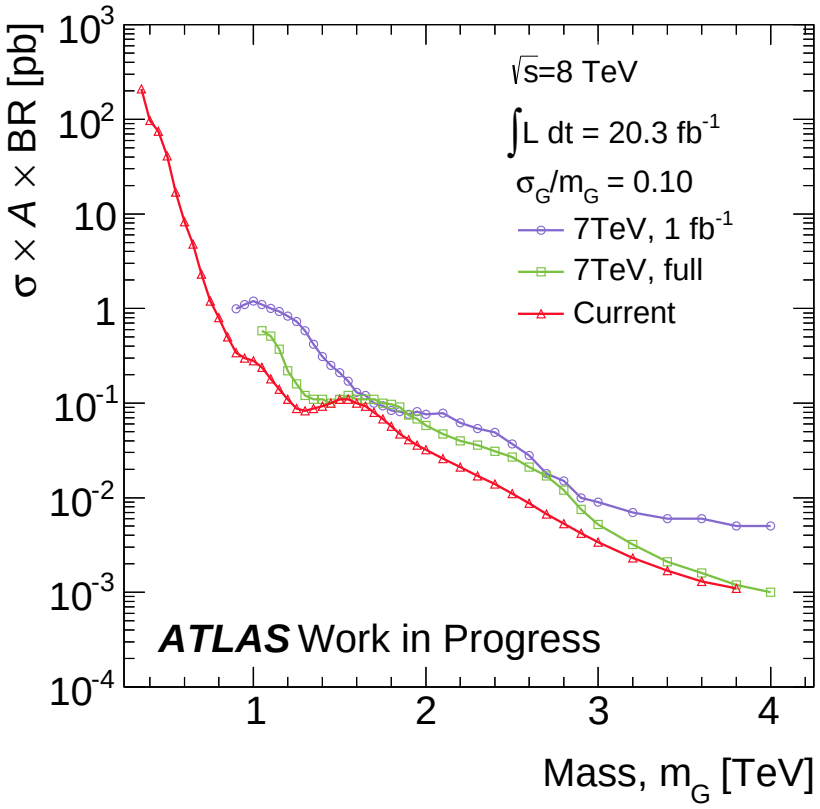


Figure 9.9: The $\sigma \times \mathcal{A}$ limits on Gaussians signal shapes with width-to-mass ratio of 10% are shown for the current analysis and for the previous two published ATLAS dijet search results [174, 175]. The Gaussian limits from the present analysis, shown in red, are the strongest limits and extend over the greatest mass range.

Table 9.2 compares the expected upper limits on the q^* and $s8$ models to previous limits from ATLAS dijet publications. The other models have either no or sparser published precedent in the analysis. The substantial increase in limits on both models from the previous to the current publication is due both to the larger luminosity and the increased centre-of-mass energy.

$\sqrt{s}$	$\mathcal{L}$	Citation	q^* [TeV]	$s8$ [TeV]
7 TeV	36 pb^{-1}	[173]	2.07	-
7 TeV	1.0 fb^{-1}	[174]	2.81	1.77
7 TeV	4.8 fb^{-1}	[175]	2.94	1.97
8 TeV	20.3 fb^{-1}	current	3.98	2.80

Table 9.2: ATLAS previous and current expected 95% CL upper limits [TeV] on excited quarks and color-octet scalars.

The usefulness of the new low-mass exclusions has been illustrated by a reinterpretation of the Gaussian generic resonance limits in the context of a Z' model. The analysis presented in Reference [248] calculates the exclusion curve in coupling and mass for a hypothetical charged heavy boson decaying into dijets for each of a wide range of dijet searches including those from CMS and the Tevatron experiments. The exclusion curves are computed using the published generic Gaussian limits from all analyses. This introduces a model-dependence into the comparison between analyses, but permits an equalisation of the results for centre-of-mass energy and luminosity across which limit comparisons are generally unclear. Their comparison thus shows the strength of dijet limits, subject to the stipulations of the simple model chosen but compatible across years, experiments, and collision energies. Since increasing backgrounds and trigger thresholds make searches in low masses more challenging, the strongest limits obtained on low-mass Z' particles in Dobrescu and Yu's paper result from UA2 and CDF, along with one early CMS search.

The analysis of Dobrescu and Yu employed a simple model with only two free parameters: the resonance mass $M_{Z'_B}$ and its coupling to dijets g_B. An on-shell Z' sample was generated for each mass point at each collider and passed through showering, detector simulation, and the appropriate jet reclustering. The cross-section times branching ratio for a fixed coupling $g_B = 0.2$ was obtained from the simulation as a function of mass for each collider. Acceptances were determined by applying the relevant analysis cuts to the simulations. The ratio between $\sigma_{g=0.2} \cdot \mathrm{BR} \cdot \mathcal{A}$ from simulations and the observed cross-section limits reported by the analysis thus provided an upper limit on the coupling as a function of mass.

The described analysis was rerun by the dijet team and the Gaussian results from the current analysis were added. Figure 9.10 shows the current reproduction of the Z' reinter-

pretation with the Gaussian results from this analysis added. The solid red line corresponds to the current result. It is compatible with the CMS result in full 8 TeV data across the latter's range, although the two fluctuate due to statistical effects. It also, however, extends downwards to masses which have not been excluded since the first LHC dataset or the Tevatron experiments. Between masses of 800-1300 GeV the current analysis provides by far the most stringent exclusions, filling several significant gaps. Below 600 GeV the CDF results become dominant since the high trigger prescales in the current analysis strongly counteract its high luminosity in this region. This figure provides a measure of the importance of the low-mass extension performed in the current analysis and motivates continued efforts to access the low mass region in Run II.

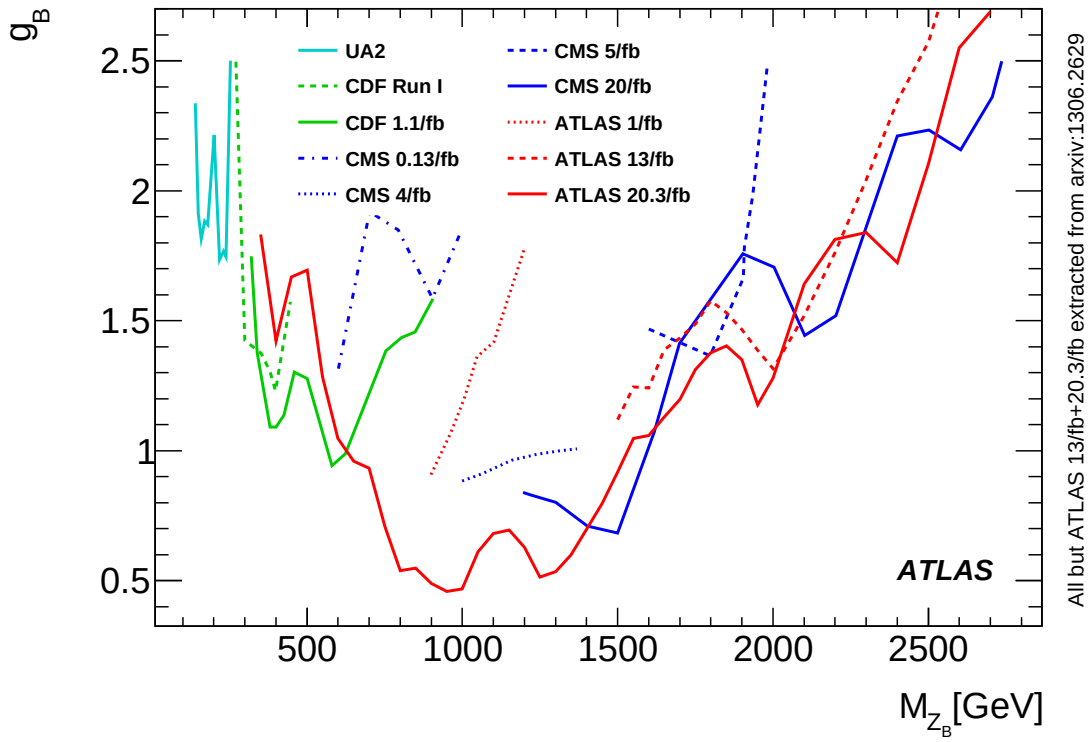


Figure 9.10: A plot comparing generic limits from ATLAS, CMS, CDF, and UA2 analyses reinterpreted as exclusion limits on a Z' model was published by Dobrescu and Yu in [248] and using as its most current ATLAS limit the CONF note result produced with 13 fb^{-1} of 8 TeV data. Following the technique outlined in the paper, the plot has been reproduced with the addition of the generic Gaussian limits from the present analysis (solid red line). The current limit is competitive with the CMS result in full 8 TeV data over the relevant mass range but also sets new, stringent limits in the 800-1300 GeV mass range which was previously only weakly constrained by early LHC searches.

Part III

The Future

Chapter 10

Dijet-based searches in Run II

'It's a magical world, Hobbes, ol' buddy – let's go exploring!'
— Bill Watterson, *Calvin and Hobbes*

10.1 Dijets in Run II

The dijet analysis will be one of the first public search results from the LHC in Run II. Preparations for the 13 TeV analysis are in full swing and the analysis team expects to be ready to proceed as soon as the LHC restarts. Current schedules anticipate a trial physics run at 50 ns spacing between bunches during June of 2015, amounting to a little under 1 fb^{-1} of data. From July, physics runs at 25 ns spacing will produce a dataset of ~ 12 fb^{-1} by the beginning of November. This dataset will provide an enormous increase in sensitivity for the dijet analysis.

10.1.1 Projected performance in early 13 TeV data

The change in centre-of-mass energy and increasing dijet production cross section will allow the first 13 TeV dijet search to surpass current sensitivities with very little integrated luminosity. Initial studies using 13 TeV Monte Carlo indicate that the current limits will be matched with the first 100 pb^{-1} of data. Discovery potential for QBH models will pass the current exclusion limits within 1 fb^{-1} of data, with other models close behind. Figure 10.1 illustrates this gain. By computing the increase in signal and background events using the parton luminosity ratio between 13 and 8 TeV, simple limit predictions can be

made as a function of luminosity. Solid lines mark the predictions; squares and circles the corresponding ATLAS and CMS analyses.

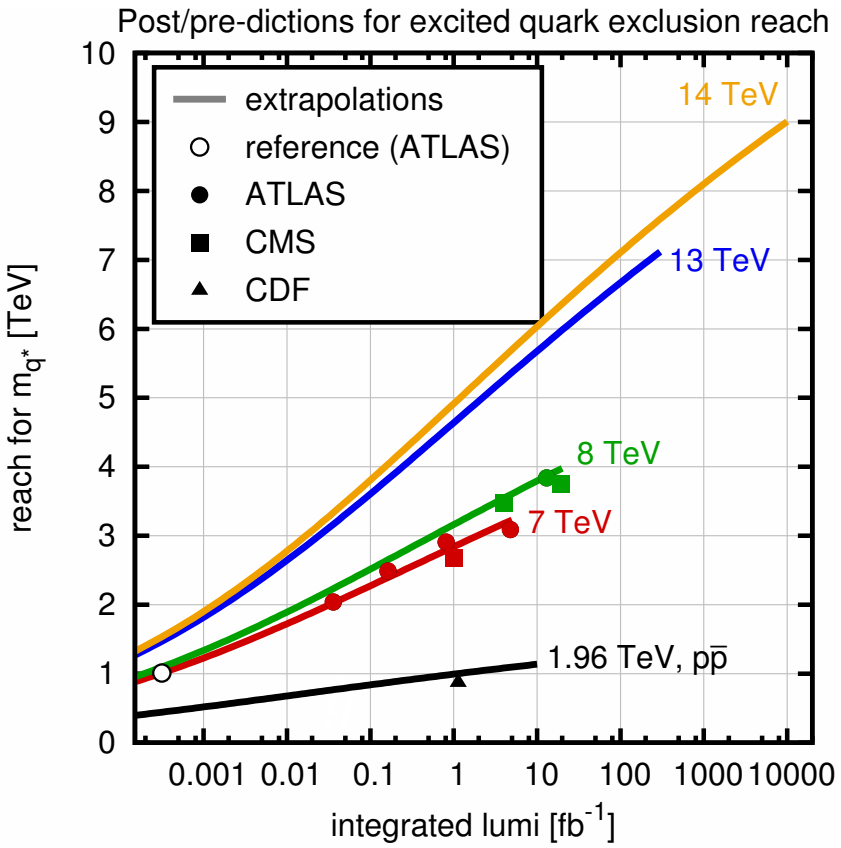


Figure 10.1: From G. Salam [249]. Predictions for the exclusion sensitivity of the dijet search for q^* samples are shown for a range of centre-of-mass energies and integrated luminosities. The predictions are compared to published results from CMS, ATLAS, and CDF. The same sensitivity is seen for 20 fb^{-1} of data at 8 TeV and for 0.2 fb^{-1} of data at 13 TeV.

Current plans anticipate an integrated luminosity rate of ~ 48 pb^{-1} per day for the 50 ns data, allowing the current limits to be surpassed with the first two days of data-taking. The biggest challenges to be posed by this nearly immediate leap into sensitivity are currently in discussion: how can the background estimation method be verified in advance, and how should the analysis be blinded, if at all?

The first analysis, intended for publication at the point where sensitivity surpasses current limits and therefore possibly using only the 50 ns data, is being designed for speed and reliability. Only the highest unprescaled single-jet trigger will be used and therefore only the high-mass regime explored. The intention is a rapid sweep of the newly available phase space for any large and obvious new physics signals, to be followed by analyses with greater mass ranges in later iterations.

10.1.2 Data scouting

A variant on the dijet analysis currently in development will increase sensitivity to new physics signals in the low-mass region by recording data at unprescaled rates from a dedicated L1 trigger. The bandwidth requirements which force jet triggers with low p_{T} thresholds to employ high prescales are based on both the event rate and the size of event data. By using a reduced data format which stores only the calorimeter information essential to jet reconstruction, the number of events recorded can be dramatically increased without surpassing bandwidth limitations.

Initial studies show that a disk size allocation of ~ 15 MB/s is sufficient to allow storage of all events passing an L1 single-jet trigger. Discussions have been ongoing with the trigger coordinators since early 2014 and permission has been granted for the new dedicated trigger chain and the required storage space. The most significant challenge this analysis faces is ensuring appropriate cleaning and calibration for trigger-level jet objects, but response studies so far have indicated that the online and offline objects behave very similarly, and discussions with the performance group responsible for jet recommendations are ongoing. The quantities required for fake jet identification are being recorded at the trigger level and initial cleaning tests with these are promising.

Three significant open questions remain. First, no definitive way has yet been found to prove that the online trigger jets can be used for the analysis. Second, if a resonance were to be discovered, the next steps are unclear. It would provide a focus point for detailed offline analysis (hence "scouting" in the name), but whether it would be publishable in its own right is still under debate. Finally, an appropriate blinding method has not yet been found. One current possibility is to fix the analysis method, decide on the dataset to be used for the first publication, and undergo the approval to unblind before data-taking starts, then record all data unblinded from the beginning. The projected increase in low-mass sensitivity is two orders of magnitude at m_{jj} = 400 GeV.

10.2 Related analysis plans

Three variations on the dijet analysis are also planned for early 13 TeV data.

Angular analysis This analysis has been performed repeatedly in ATLAS and takes advantage of the expected difference in angular distribution between new physics models and QCD backgrounds. This non-resonant search looks for turn-on curves visible in $\chi = e^{2|y^*|}$ and in the centrality ratio

$$F_\chi = \frac{dN/dm_{jj}(|y^*| < 0.6)}{dN/dm_{jj}(|y^*| < 1.7)}. \tag{10.1}$$

It will surpass current limits below 1 fb^{-1} of data and is a possible candidate for publication in the 50 ns dataset.

3-jet analysis A resonance search in the invariant mass spectrum of the subleading and sub-subleading jet is sensitive to signals produced via a $pp \to X(\to jj) + j$ mechanism. The leading jet is used for triggering, allowing an analysis reach to low invariant masses. Sensitive to dark matter mediators and q^* models produced in a contact interaction-like mechanism, 1 fb^{-1} of data will allow a q^* discovery up to 5 TeV.

Heavy flavour dijets Certain new physics theories, including some Z' and diquark models, are expected to couple preferentially to the third generation quarks. Requiring 1 or 2 b-tagged jets in each event will increase the sensitivity to these models provided the tagging efficiency is high and its uncertainties are sufficiently small. This analysis is expected to publish on a longer timescale than the others.

Initial studies have been undertaken for all three analyses, although manpower is limited for the second and third.

10.3 What if?

Despite ATLAS's efforts to check in every pocket, there is some chance that LHC Run II will not bring any discoveries of physics beyond the Standard Model. The next step for the physics community in this case is unclear. Some of the apparent discrepancies which incline theorists to believe that the current picture is incomplete, particularly the hierarchy problem, simultaneously suggest that the solution should be found on the TeV scale. Without a discovery at the LHC, the specific theoretical motivations for the construction of a more powerful collider are largely removed, as there would no longer be a coherent reason to suppose that the Standard Model could not remain intact until the Planck scale. Dark matter, if

discovered at the LHC or a direct detection experiment, could no doubt be appended to the Standard Model in a consistent, if inelegant, theory and the variety of apparently fine-tuned parameters in the model may need to be simply accepted.

This is not, however, the time to give up. The efforts of phenomenologists in the past three years have demonstrated how many ways there are to stretch the bounds of naturalness and even measurability for popular theories. But beyond that, there is always cause to hope that the universe may be more surprising than we ever expected: the solution may be something which has not yet occurred to anyone. For now, all that can be done is to look everywhere with as broad and unrestricted an analysis programme as possible, and wait to see what happens.

Part IV

Appendices

Bibliography

[1] C. Burgess and G. Moore, *The Standard Model: A Primer*. Cambridge University Press, 2007.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory (Frontiers in Physics)*. Westview Press, 1995.

[3] F. Halzen and A. Martin, *Quarks and leptons: an introductory course in modern particle physics*. Wiley, 1984.

[4] S. P. Martin, *A Supersymmetry primer*, Adv.Ser.Direct.High Energy Phys. **21** (2010) 1–153, arXiv:hep-ph/9709356 [hep-ph].

[5] J. Donoghue, E. Golowich, and B. Holstein, *Dynamics of the Standard Model*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2014. https://books.google.ca/books?id=QB9GAwAAQBAJ.

[6] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Physical Review Letters **13** (1964) no. 9, 321–323.

[7] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964) no. 2, 132–133.

[8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys.Rev.Lett. **13** (1964) 508–509.

[9] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys.Rev.Lett. **13** (1964) 585–587.

[10] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys.Rev. **145** (1966) 1156–1163.

[11] T. Kibble, *Symmetry breaking in nonAbelian gauge theories*, Phys.Rev. **155** (1967) 1554–1561.

[12] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[13] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[14] R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering*. Oxford University Press, 2004.

[15] S. Bethke, G. Dissertori, and G. P. Salam, *Quantum chromodynamics*, . http://pdg.lbl.gov/2014/reviews/rpp2014-rev-qcd.pdf. Particle Data Group Review.

[16] LHCb;collaboration, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys.Rev.Lett. **112** (2014) no. 22, 222002, arXiv:1404.1903 [hep-ex].

[17] G. Sterman et al., *Handbook of perturbative QCD*, Reviews of Modern Physics **67** (1995) no. 1, 157.

[18] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys.J. **159** (1970) 379–403.

[19] M. Markevitch, A. Gonzalez, D. Clowe, et al., *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, Astrophys.J. **606** (2004) 819–824, arXiv:astro-ph/0309303 [astro-ph].

[20] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys.Rept. **405** (2005) 279–390, arXiv:hep-ph/0404175 [hep-ph].

[21] Planck Collaboration Collaboration, P. Ade et al., *Planck 2013 results. XVI. Cosmological parameters*, Astron.Astrophys. **571** (2014) A16, arXiv:1303.5076 [astro-ph.CO].

[22] WMAP Collaboration, *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, Astrophys.J.Suppl. **192** (2011) 18, arXiv:1001.4538 [astro-ph.CO].

[23] U. Sarkar, *Particle and Astroparticle Physics*. Series in High Energy Physics, Cosmology and Gravitation. CRC Press, 2007. https://books.google.ch/books?id=t-qL07MSvxEC.

[24] E. W. Kolb and M. S. Turner, *The Early Universe*. Addison-Wesley, 1990. Frontiers in Physics, 69.

[25] M. Trodden, *Electroweak baryogenesis*, Rev.Mod.Phys. **71** (1999) 1463–1500, arXiv:hep-ph/9803479 [hep-ph].

[26] Y. Nagashima, *Beyond the Standard Model of Elementary Particle Physics*. Wiley-VCH Verlag GmbH & Co. KGaA, 2014.

[27] H. Georgi and S. Glashow, *Unity of All Elementary Particle Forces*, Phys.Rev.Lett. **32** (1974) 438–441.

[28] G. Altarelli, B. Mele, and M. Ruiz-Altaba, *Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders*, Z.Phys. **C45** (1989) 109.

[29] G. Altarelli, B. Mele, and M. Ruiz-Altaba, *Erratum: Searching for new heavy vector bosons in $p\bar{p}$ colliders*, Z. Phys. C **47** (1990) 676.

[30] ATLAS Collaboration, G. Aad et al., *Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **1409** (2014) 037, arXiv:1407.7494 [hep-ex].

[31] M. Chizhov and G. Dvali, *Origin and Phenomenology of Weak-Doublet Spin-1 Bosons*, Phys.Lett. **B703** (2011) 593–598, arXiv:0908.0924 [hep-ph].

[32] M. Chizhov, V. Bednyakov, and J. Budagov, *A unique signal of excited bosons in dijet data from pp-collisions*, Phys.Atom.Nucl. **75** (2012) 90–96, arXiv:1010.2648 [hep-ph].

[33] M. Chizhov, *A Reference Model for Anomalously Interacting Bosons*, Phys.Part.Nucl.Lett. **8** (2011) 512–516, arXiv:1005.4287 [hep-ph].

[34] M. Chizhov. private communication, 2014.

[35] M. Chizhov, *On the two kinds of vector particles*, arXiv:hep-ph/0008187 [hep-ph].

[36] M. Chizhov, V. Bednyakov, and J. Budagov, *Proposal for chiral bosons search at LHC via their unique new signature*, Phys.Atom.Nucl. **71** (2008) 2096–2100, arXiv:0801.4235 [hep-ph].

[37] U. Baur, I. Hinchliffe, and D. Zeppenfeld, *Excited Quark Production at Hadron Colliders*, Int. J. Mod. Phys. **A2** (1987) 1285.

[38] U. Baur, M. Spira, and P. M. Zerwas, *Excited-Quark and -Lepton Production at Hadron Colliders*, Phys. Rev. **D42** (1990) 815.

[39] T. Han, I. Lewis, and Z. Liu, *Colored Resonant Signals at the LHC: Largest Rate and Simplest Topology*, JHEP **1012** (2010) 085, arXiv:1010.4309 [hep-ph].

[40] C. T. Hill and E. H. Simmons, *Strong dynamics and electroweak symmetry breaking*, Phys.Rept. **381** (2003) 235–402, arXiv:hep-ph/0203079 [hep-ph].

[41] E. Eichten, I. Hinchliffe, K. Lane, and C. Quigg, *Signatures for Technicolor*, Phys.Rev. **D34** (1986) 1547.

[42] R. S. Chivukula, P. Ittisamai, J. Ren, and E. H. Simmons, *Technicolor in the LHC Era*, arXiv:1202.1505 [hep-ph].

[43] B. A. Dobrescu, K. Kong, and R. Mahbubani, *Leptons and photons at the LHC: Cascades through spinless adjoints*, JHEP **0707** (2007) 006, arXiv:hep-ph/0703231 [HEP-PH].

[44] B. A. Dobrescu, K. Kong, and R. Mahbubani, *Massive color-octet bosons and pairs of resonances at hadron colliders*, Phys.Lett. **B670** (2008) 119–123, arXiv:0709.2378 [hep-ph].

[45] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. **B429** (1998) 263–272, arXiv:hep-ph/9803315 [hep-ph].

[46] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity*, Phys.Rev. **D59** (1999) 086004, arXiv:hep-ph/9807344 [hep-ph].

[47] L. A. Anchordoqui, J. L. Feng, H. Goldberg, and A. D. Shapere, *Inelastic black hole production and large extra dimensions*, Phys.Lett. **B594** (2004) 363–367, arXiv:hep-ph/0311365 [hep-ph].

[48] S. Dimopoulos and G. L. Landsberg, *Black holes at the LHC*, Phys. Rev. Lett. **87** (2001) 161602, arXiv:hep-ph/0106295 [hep-ph].

[49] R. C. Myers and M. Perry, *Black Holes in Higher Dimensional Space-Times*, Annals Phys. **172** (1986) 304.

[50] X. Calmet, W. Gong, and S. D. Hsu, *Colorful quantum black holes at the LHC*, Phys. Lett. B **668** (2008) 20–23, arXiv:0806.4605 [hep-ph].

[51] P. Meade and L. Randall, *Black Holes and Quantum Gravity at the LHC*, JHEP **0805** (2008) 003, arXiv:0708.3017 [hep-ph].

[52] M. Cavaglia, *Black hole multiplicity at particle colliders: Do black holes radiate mainly on the brane?*, Phys.Lett. **B569** (2003) 7–13, arXiv:hep-ph/0305256 [hep-ph].

[53] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Ann.Rev.Nucl.Part.Sci. **37** (1987) 383–409.

[54] J. M. Campbell, J. Huston, and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept.Prog.Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

[55] S. Höche, *Matching fixed-order Matrix Elements and Parton Showers*, in *MCnet-LPCC School on Event Generators for LHC*. July, 2012.

[56] R. K. Ellis, W. J. Stirling, and B. Webber, *QCD and collider physics*, Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol. **8** (1996) 1–435.

[57] L. Lipatov, *The parton model and perturbation theory*, Sov.J.Nucl.Phys. **20** (1975) 94–102.

[58] V. Gribov and L. Lipatov, *Deep inelastic ep-scattering in perturbation theory*, Sov.J.Nucl.Phys. **15** (1972) 438–450.

[59] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl.Phys. **B126** (1977) 298.

[60] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov.Phys.JETP **46** (1977) 641–653.

[61] B. Foster, A. D. Martin, and M. G. Vincter, *Structure functions*, . pdg.lbl.gov/2014/reviews/rpp2014-rev-structure-functions.pdf. Revised September 2013.

[62] R. D. Ball, V. Bertone, S. Carrazza, et al., *Parton distributions with LHC data*, Nucl.Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[63] C. Meyer and M. Oreglia, *Measurement of Dijet Cross Sections in Proton-Proton Collisions at 7 TeV Center-of-Mass Energy Using the ATLAS Detector*. PhD thesis, Chicago U., 2013. Presented 2013.

[64] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J.Math.Phys. **3** (1962) 650–677.

[65] T. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, Phys.Rev. **133** (1964) B1549–B1562.

[66] R. Akhoury, M. Sotiropoulos, and V. I. Zakharov, *The KLN theorem and soft radiation in gauge theories: Abelian case*, Phys.Rev. **D56** (1997) 377–387, arXiv:hep-ph/9702270 [hep-ph].

[67] F. Bloch and A. Nordsieck, *Note on the Radiation Field of the electron*, Phys.Rev. **52** (1937) 54–59.

[68] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[69] B. Webber, *Fragmentation and hadronization*, Int.J.Mod.Phys. **A15S1** (2000) 577–606, arXiv:hep-ph/9912292 [hep-ph].

[70] B. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl.Phys. **B238** (1984) 492.

[71] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys.Rept. **97** (1983) 31–145.

[72] G. F. Sterman and S. Weinberg, *Jets from Quantum Chromodynamics*, Phys.Rev.Lett. **39** (1977) 1436.

[73] G. F. Sterman, *QCD and jets*, arXiv:hep-ph/0412013 [hep-ph].

[74] H. Kirschenmann. http://cms.web.cern.ch/news/jets-cms-and-determination-their-energy-scale, July, 2012.

[75] C. Buttar, J. D'Hondt, M. Kramer, et al., *Standard Model Handles and Candles Working Group: Tools and Jets Summary Report*, arXiv:0803.0678 [hep-ph].

[76] J. E. Huth, N. Wainer, K. Meier, et al., *Toward a standardization of jet definitions*, .

[77] G. P. Salam, *Towards Jetography*, Eur.Phys.J. **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].

[78] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at* $\sqrt{s} = 7$ *tev*, arXiv:1308.6064 [hep-ex].

[79] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys.Rev. **D48** (1993) 3160–3166, arXiv:hep-ph/9305266 [hep-ph].

[80] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, arXiv:hep-ph/9907280 [hep-ph].

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

[82] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys.Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].

[83] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) no. 08, S08001. http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[84] A. Team, *The four main LHC experiments*, Jun, 1999.

[85] F. Soler, *Large Hadron Collider Phenomenology*. Scottish Graduate Series. Taylor & Francis, 2004. https://books.google.ca/books?id=onLbH3WDgiYC.

[86] O. S. Brüning et al., *LHC Design Report*. CERN, Geneva, 2004.

[87] LHC Study Group Collaboration, T. S. Pettersson and P. Lefèvre, *The Large Hadron Collider: conceptual design*, Tech. Rep. CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995.

[88] A. Team, *Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC*, Jun, 1999.

[89] M. Lamont, *Status of the LHC*, in *Journal of Physics: Conference Series*, vol. 455, p. 012001, IOP Publishing. 2013.

[90] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[91] ATLAS Luminosity Group, *Preliminary Luminosity Determination in pp Collisions at sqrt(s) = 8 TeV using the ATLAS Detector in 2012*, Tech. Rep. ATL-COM-LUM-2012-013, CERN, Geneva, Nov, 2012.

[92] ATLAS Collaboration, *Luminosity public results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults, 2014.

[93] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013) 2518, arXiv:1302.4393 [hep-ex].

[94] *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999.

[95] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar, 2008.

[96] *ATLAS magnet system: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[97] P. Bussey, L. Chytka, R. Staszewski, T. Sykora, and M. Trzebinski, *LHC and ATLAS Coordinate Conventions*, Tech. Rep. ATL-COM-FWD-2013-002, CERN, Geneva, Sep, 2013.

[98] *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[99] J. Pequenao, *Computer generated image of the ATLAS inner detector*, Mar, 2008.

[100] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[101] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Geneva, Sep, 2010.

[102] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report Addendum*, Tech. Rep. CERN-LHCC-2012-009. ATLAS-TDR-19-ADD-1, CERN, Geneva, May, 2012.

[103] N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1998.

[104] K. Lantzsch, *Status and performance of the ATLAS Pixel Detector at the LHC after three years of operation*, Tech. Rep. ATL-COM-INDET-2014-003, CERN, Geneva, Jan, 2014.

[105] A. Abdesselam and T. Akimoto, *The Barrel Modules of the ATLAS SemiConductor Tracker*, Tech. Rep. ATL-INDET-PUB-2006-005. ATL-COM-INDET-2006-009. CERN-ATL-COM-INDET-2006-009, CERN, Geneva, Jul, 2006.

[106] A. Abdesselam and F. Anghinolfi, *The ATLAS semiconductor tracker end-cap module*, Nucl. Instrum. Methods Phys. Res., A **575** (2007) no. 3, 353–389.

[107] C. Alpigiani, *Performance of the ATLAS semiconductor tracker*, Tech. Rep. ATL-COM-INDET-2014-025, CERN, Geneva, Jun, 2014.

[108] E. Stanecka, *Performance and operation experience of the Atlas Semiconductor Tracker and Pixel Detector at the LHC.*, Tech. Rep. ATL-INDET-PROC-2013-023, CERN, Geneva, Dec, 2013.

[109] ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013.

[110] T. A. T. collaboration, *The ATLAS TRT Barrel Detector*, Journal of Instrumentation **3** (2008) no. 02, P02014. http://stacks.iop.org/1748-0221/3/i=02/a=P02014.

[111] T. A. T. collaboration, *The ATLAS TRT end-cap detectors*, Journal of Instrumentation **3** (2008) no. 10, P10003. http://stacks.iop.org/1748-0221/3/i=10/a=P10003.

[112] *Basic ATLAS TRT performance studies of Run 1*, Tech. Rep. ATL-INDET-PUB-2014-001, CERN, Geneva, Mar, 2014.

[113] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[114] ATLAS Electromagnetic Barrel Liquid Argon Calorimeter Group, *Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter*, Nucl.Instrum.Meth. **A558** (2006) 388–418.

[115] M. Aleksa et al., *Construction, assembly and tests of the ATLAS electromagnetic end-cap calorimeters*, J. Instrum. **3** (2008) P06002.

[116] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics*. CERN, Geneva, 2009.

[117] M. Andrieux et al., *Construction and test of the first two sectors of the ATLAS barrel liquid argon presampler*, Nucl.Instrum.Meth. **A479** (2002) 316–333.

[118] *ATLAS*, Aug, 1993.

[119] A. Bazan, *The ATLAS liquid argon calorimeter read-out system*, IEEE Trans. Nucl. Sci. **53** (2006) 735–740.

[120] ATLAS Collaboration, *Monitoring and data quality assessment of the ATLAS liquid argon calorimeter*, Tech. Rep. arXiv:1405.3768. CERN-PH-EP-2014-045, CERN, Geneva, May, 2014.

[121] A. E. B. C. Collaboration, *Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam*, Nucl.Instrum.Meth. **A568** (2006) 601–623, arXiv:physics/0608012 [physics].

[122] C. Leroy and P.-G. Rancoita, *Principles Of Radiation Interaction in Matter and Detection*, .

[123] D. M. Gingrich et al., *Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter*, Tech. Rep. ATL-LARG-PUB-2007-009. CERN-ATL-COM-LARG-2007-006, CERN, Geneva, Apr, 2007.

[124] P. Puzo, *ATLAS calorimetry*, Nucl.Instrum.Meth. **A494** (2002) 340–345.

[125] *ATLAS tile calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.

[126] ATLAS TileCal Collaboration, *Design, Construction and Installation of the ATLAS Hadronic Barrel Scintillator-Tile Calorimeter*, Tech. Rep. ATL-TILECAL-PUB-2008-001. ATL-COM-TILECAL-2007-019, CERN, Geneva, Nov, 2007.

[127] K. Anderson et al., *Design of the front-end analog electronics for the ATLAS tile calorimeter*, Nucl.Instrum.Meth. **A551** (2005) 469–476.

[128] G. Usai, *Signal Reconstruction of the ATLAS Hadronic Tile Calorimeter: implementation and performance*, Tech. Rep. ATL-TILECAL-PROC-2010-008, CERN, Geneva, Aug, 2010.

[129] B. S.-M. Peralva, *Calibration and Performance of the ATLAS Tile Calorimeter*, Tech. Rep. arXiv:1305.0550, May, 2013.

[130] P. Adragna et al., *Testbeam studies of production modules of the ATLAS tile calorimeter*, Nucl.Instrum.Meth. **A606** (2009) 362–394.

[131] B. Burghgrave and E. Tiouchichine, *ATLAS detector performance in Run1: Calorimeters*, .

[132] L. Heelan, *Run 1 Performance of the ATLAS Tile Calorimeter*, .

[133] A. Belloni, *Integration and commissioning of the ATLAS Muon spectrometer*, Tech. Rep. arXiv:0810.2869, Oct, 2008.

[134] A. Salvucci, *Measurement of muon momentum resolution of the ATLAS detector*, Tech. Rep. arXiv:1201.4704, Jan, 2012.

[135] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur.Phys.J. **C74** (2014) no. 11, 3130, arXiv:1407.3935 [hep-ex].

[136] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 2003.

[137] ATLAS Collaboration, *The performance of the jet trigger for the ATLAS detector at the LHC during 2011 data taking*, Tech. Rep. ATL-COM-DAQ-2014-133, CERN, Geneva, Dec, 2014.

[138] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[139] *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul, 2011.

[140] H.-L. Lai et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[141] Z. Nagy, *Three jet cross-sections in hadron hadron collisions at next-to-leading order*, Phys.Rev.Lett. **88** (2002) 122003, arXiv:hep-ph/0110315 [hep-ph].

[142] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys.Rev. **D68** (2003) 094002, arXiv:hep-ph/0307268 [hep-ph].

[143] S. Catani and M. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, Nucl.Phys. **B485** (1997) 291–419, arXiv:hep-ph/9605323 [hep-ph].

[144] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[145] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.

[146] C. Eck et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical Design Report LCG. CERN, Geneva, 2005.

[147] ATLAS Collaboration, *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005.

[148] M. Baak et al., *Data Quality Status Flags and Good Run Lists for Physics Analysis in ATLAS*, Tech. Rep. ATL-COM-GEN-2009-015, CERN, Geneva, Mar, 2009.

[149] ATLAS Collaboration, *Calorimeter clustering algorithms, description and performance.*, CERN (2010) . ATL-LARG-PUB-2008-002.

[150] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur.Phys.J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[151] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur.Phys.J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[152] ATLAS Collaboration, *Local hadronic calibration*, .

[153] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex].

[154] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys.Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[155] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, CERN (2013) . ATLAS-CONF-2013-083.

[156] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*, Tech. Rep. ATLAS-CONF-2013-004, CERN, Geneva, Jan, 2013.

[157] ATLAS Collaboration, *Probing the measurement of jet energies with the ATLAS detector using Z+jet events from proton-proton collisions at sqrt(s) = 7 TeV*, Tech. Rep. ATLAS-CONF-2012-053, CERN, Geneva, May, 2012.

[158] ATLAS Collaboration, *Probing the measurement of jet energies with the ATLAS detector using photon+jet events in proton-proton collisions at sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2012-063, CERN, Geneva, Jul, 2012.

[159] ATLAS Collaboration, *TeV-scale jet energy calibration using multijet events including close-by jet effects at the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2013-003, CERN, Geneva, Jan, 2013.

[160] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013) 2305, arXiv:1203.1302 [hep-ex].

[161] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at* $\sqrt{s}$ = 7 *TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].

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

[163] UA1 Collaboration, *Angular distributions and structure functions from two-jet events at the CERN SPS* $p\bar{p}$ *collider*, Phys. Lett. **B136** (1984) 294.

[164] UA2 Collaboration, *Measurement of jet production properties at the CERN* $\bar{p}p$ *Collider*, Phys. Lett. **B144** (1984) 283.

[165] CDF Collaboration, *Search for new particles decaying into dijets in* $p\bar{p}$ *collisions at* $\sqrt{s}$ = 1.96 *TeV*, Phys. Rev. **D79** (2009) 112002, arXiv:0812.4036 [hep-ex].

[166] D0 Collaboration, *Measurement of Dijet Angular Distributions at* $\sqrt{s}$ = *1.96 TeV and Searches for Quark Compositeness and Extra Spatial Dimensions*, Phys. Rev. Lett. **103** (2009) 191803, arXiv:0906.4819 [hep-ex].

[167] ATLAS Collaboration, *Search for New Particles in Two-Jet Final States in 7 TeV Proton-Proton Collisions with the ATLAS Detector at the LHC*, Phys.Rev.Lett. **105** (2010) 161801, arXiv:1008.2461 [hep-ex].

[168] ATLAS Collaboration, *Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at* $\sqrt{s}$ = 7 *TeV Measured with the ATLAS Detector*, Phys. Lett. **B694** (2011) 327, arXiv:1009.5069 [hep-ex].

[169] CMS Collaboration, *Search for Dijet Resonances in 7 TeV pp Collisions at CMS*, Phys. Rev. Lett. **105** (2010) 211801, arXiv:1107.4771 [hep-ex].

[170] CMS Collaboration, *Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Rev. Lett. **105** (2010) 262001, arXiv:1010.4439 [hep-ex].

[171] CMS Collaboration, *Search for resonances in the dijet mass spectrum from 7 TeV pp Collisions at CMS*, Phys. Lett. **B704** (2011) 123, arXiv:1107.4771 [hep-ex].

[172] CMS Collaboration, *Measurement of Dijet Angular Distributions and Search for Quark Compositeness in pp Collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Rev. Lett. **106** (2011) 201804, arXiv:1102.2020 [hep-ex].

[173] ATLAS Collaboration, *Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at* $\sqrt{s} = 7$ *TeV Measured with the ATLAS Detector*, New J.Phys. **13** (2011) 053044, arXiv:1103.3864 [hep-ex].

[174] ATLAS Collaboration, *Search for new physics in the dijet mass distribution using 1* fb^{-1} *of pp collision data at* $\sqrt{s} = 7$ *TeV collected by the ATLAS detector*, Phys. Lett. **B708** (2012) 37, arXiv:1108.6311 [hep-ex].

[175] ATLAS Collaboration, *ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at* $\sqrt{s} = 7$ *TeV*, Journal of High Energy Physics **2013** (2013) no. 1, . http://dx.doi.org/10.1007/JHEP01%282013%29029.

[176] CMS Collaboration, *Search for narrow resonances and quantum black holes in inclusive and b-tagged dijet mass spectra from pp collisions at* $\sqrt{s} = 7$ *TeV*, JHEP **1301** (2013) 013, arXiv:1210.2387 [hep-ex].

[177] CMS Collaboration, *Search for narrow resonances using the dijet mass spectrum in pp collisions at* $\sqrt{s}$=8*TeV*, Phys.Rev. **D87** (2013) no. 11, 114015, arXiv:1302.4794 [hep-ex].

[178] ATLAS Collaboration, *Search for new phenomena in the dijet mass distribution using pp collision data at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, arXiv:1407.1376 [hep-ex].

[179] *ATLAS Computing, CoolLumiCalc*, https://twiki.cern.ch/twiki/bin/view/AtlasComputing/CoolLumiCalc.

[180] *ATLAS Computing, CoolLumiCalc tutorial*, https://twiki.cern.ch/twiki/bin/view/AtlasComputing/CoolLumiCalcTutorial.

[181] V. Lendermann, J. Haller, M. Herbst, et al., *Combining Triggers in HEP Data Analysis*, Nucl.Instrum.Meth. **A604** (2009) 707–718, arXiv:0901.4118 [hep-ex].

[182] G. Bohm and G. Zech, *Introduction to statistics and data analysis for physicists*. DESY, 2010. http://books.google.ca/books?id=aQBHcgAACAAJ.

[183] *Data preparation recommendations for physics analyses, 2012 data*, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DataPreparationCheckListForPhysicsAnalysis.

[184] *EGamma Group, LAr cleaning and object quality recommendations*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/LArCleaningAndObjectQuality.

[185] *Timing, Trigger and control (TTC) Systems for the LHC*, http://ttc.web.cern.ch/TTC/intro.html.

[186] *TileCal Group, `TileTripReader` package*, https://twiki.cern.ch/twiki/bin/view/Atlas/TileTripReader.

[187] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, Tech. Rep. ATLAS-CONF-2012-020, CERN, Geneva, Mar, 2012.

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

[189] *JetEtmiss Group, jet cleaning recommendations, 2011 data*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011.

[190] *JetEtmiss Group, jet cleaning recommendations, 2012 data*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2012.

[191] J. Lacey, *MC Weights and Quality Cuts in 2012*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetStudies2012, July, 2012.

[192] *Cell Reconstruction: From RawChannels to Cells*, https://twiki.cern.ch/twiki/bin/view/AtlasComputing/CellReconstruction.

[193] *JetEtmiss Group, `BCHCleaningTool` package*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BCHCleaningTool.

[194] *Tile calorimeter group, `TileTripReader` package*, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TileTripReader.

[195] L. Lyons, *Bayes and Frequentism: a Particle Physicist's perspective*, arXiv:1301.1273 [physics.data-an].

[196] R. D. Cousins, *Why isn't every physicist a Bayesian?*, Am.J.Phys. **63** (1995) 398.

[197] G. Cowan, *Statistics for Searches at the LHC*, arXiv:1307.2487 [hep-ex].

[198] L. Demortier, *P values and nuisance parameters*, .

[199] K. Pearson, *X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, Philosophical Magazine Series 5 **50** (1900) no. 302, 157–175, http://dx.doi.org/10.1080/14786440009463897. http://dx.doi.org/10.1080/14786440009463897.

[200] I. Narsky and F. C. Porter, *Statistical Analysis Techniques in Particle Physics*. Wiley-VCH Verlag GmbH & Co. KGaA, 2013.

[201] J. Aldrich et al., *RA Fisher and the making of maximum likelihood 1912-1922*, Statistical Science **12** (1997) no. 3, 162–176.

[202] R. A. Fisher, *On the mathematical foundations of theoretical statistics*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character (1922) 309–368.

[203] G. Choudalakis, *On hypothesis testing, trials factor, hypertests, and the BumpHunter*, PHYSTAT2011 (2011) , arXiv:1101.0390v2 [data-an].

[204] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists*. Academic Press, 6 ed., July, 2005.

[205] G. Choudalakis and D. Casadei, *Plotting the Differences Between Data and Expectation*, Eur. Phys. J. Plus **127** (2011) no. 25, , arXiv:1111.2062 [data-an].

[206] O. Behnke, K. Kröninger, G. Schott, and T. Schörner-Sadenius, *Data Analysis in High Energy Physics*. Wiley-VCH Verlag GmbH & Co. KGaA, 2013.

[207] CDF Collaboration Collaboration, T. Aaltonen et al., *Search for new particles decaying into dijets in proton-antiproton collisions at s**(1/2) = 1.96-TeV*, Phys.Rev. **D79** (2009) 112002, arXiv:0812.4036 [hep-ex].

[208] R. M. Harris and K. Kousouris, *Searches for Dijet Resonances at Hadron Colliders*, Int.J.Mod.Phys. **A26** (2011) 5005–5055, arXiv:1110.5302 [hep-ex].

[209] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput.Phys.Commun. **10** (1975) 343–367.

[210] G. Cowan, *Statistical data analysis*. Oxford University Press, Oxford, 1998.

[211] ATLAS Collaboration, *Search for New Phenomena in the Dijet Mass Distribution updated using 13.0 fb^{-1} of pp Collisions at sqrts = 8 TeV collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-148, CERN, Geneva, Nov, 2012.

[212] F. Ruehr. Private communication, November, 2013.

[213] UA2 Collaboration Collaboration, J. Alitti et al., *A Measurement of two jet decays of the W and Z bosons at the CERN $\bar{p}p$ collider*, Z.Phys. **C49** (1991) 17–28.

[214] CDF Collaboration Collaboration, F. Abe et al., *Search for new particles decaying to dijets in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys.Rev.Lett. **74** (1995) 3538–3543, arXiv:hep-ex/9501001 [hep-ex].

[215] CDF Collaboration Collaboration, F. Abe et al., *Search for new particles decaying to dijets at CDF*, Phys.Rev. **D55** (1997) 5263–5268, arXiv:hep-ex/9702004 [hep-ex].

[216] ATLAS Collaboration Collaboration, G. Aad et al., *Search for microscopic black holes and string balls in final states with leptons and jets with the ATLAS detector at sqrt(s) = 8 TeV*, JHEP **1408** (2014) 103, arXiv:1405.4254 [hep-ex].

[217] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, *The "K" in K-fold Cross Validation*, in *ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. Bruges, Belgium, April, 2012.

[218] ATLAS Collaboration Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur.Phys.J. **C71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[219] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013) 2518, arXiv:1302.4393 [hep-ex].

[220] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[221] J. Wenninger, *Energy Calibration of the LHC Beams at 4 TeV*, Tech. Rep. CERN-ATS-2013-040, CERN, Geneva, May, 2013.

[222] B. Butler and A. Schwartzman, *Track-Jet Reconstruction and Performance*, Tech. Rep. ATL-PHYS-INT-2010-040, CERN, Geneva, Apr, 2010.

[223] F. Guescini, *Measurement of Jet Angular Resolution*, Tech. Rep. ATL-COM-PHYS-2013-1467, CERN, Geneva, Oct, 2013.

[224] *SUSY working group, `PDFTool` package*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PDFReweight.

[225] M. Bayes and M. Price, *An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.*, Philosophical Transactions **53** (1763) 370–418. http://rstl.royalsocietypublishing.org/content/53/370.full.pdf+html.

[226] Pierre-Simon marquis de Laplace, *Théorie analytique des probabilités*. Courcier, Paris, 1812.

[227] Pierre-Simon marquis de Laplace, *Essai philosophique sur les probabilités*. Courcier, Paris, 1814.

[228] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge University Press, New York, NY, USA, 2005.

[229] B. Carlin and T. Louis, *Bayesian Methods for Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2011. http://books.google.co.uk/books?id=GTJUt8fcFx8C.

[230] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011) no. 2, . http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[231] J. Heinrich, C. Blocker, J. Conway, L. Demortier, L. Lyons, G. Punzi, and P. K. Sinervo, *Interval estimation in the presence of nuisance parameters. 1. Bayesian approach*, arXiv:physics/0409129 [data-an].

[232] H. Jeffreys, *An Invariant Form for the Prior Probability in Estimation Problems*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **186** (1946) no. 1007, pp. 453–461. http://www.jstor.org/stable/97883.

[233] J.-M. Bernardo, *Reference posterior distributions for Bayesian inference*, J. Roy. Statist. Soc. Ser. B **41** (1979) no. 2, 113–147. http://links.jstor.org/sici?sici=0035-9246(1979)41:2<113:RPDFBI>2.0.CO;2-P&origin=MSN. With discussion.

[234] J. O. Berger, J. M. Bernardo, and D. Sun, *The Formal Definition of Reference Priors*, The Annals of Statistics **37** (2009) no. 2, . http://www.jstor.org/stable/30243652.

[235] *ATLAS Statistics Forum, Recommendations on the choice of a prior, 2011*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ChoosingPriors.

[236] G. D'Agostini, *Confidence limits: What is the problem? Is there the solution?*, arXiv:hep-ex/0002055 [hep-ex].

[237] Particle Data Group Collaboration, K. A. Olive et al., *2014 Review of Particle Physics*, Chin. Phys. C. **38** (2014) 090001.

[238] L. Demortier, *Objective Bayesian Upper Limits for Poisson Processes*, CDF5928 (2005) .

[239] L. E. Eberly and G. Casella, *Estimating Bayesian credible intervals*, Journal of statistical planning and inference **112** (2003) no. 1, 115–132.

[240] A. Beskos and A. Stuart, *Computational complexity of Metropolis-Hastings methods in high dimensions*, in *Monte Carlo and Quasi-Monte Carlo Methods 2008*, pp. 61–71. Springer, 2009.

[241] A. A. Markov, *Rasprostranenie zakona bolshih chisel na velichiny, zavisyaschie drug ot druga*, Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete **15** (1906) no. 135-156, 18.

[242] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of state calculations by fast computing machines*, The journal of chemical physics **21** (1953) no. 6, 1087–1092.

[243] N. Metropolis and S. Ulam, *The Monte Carlo Method*, Journal of the American Statistical Association **44** (1949) no. 247, pp. 335–341. http://www.jstor.org/stable/2280232.

[244] *ATLAS Statistics Forum, Bayesian limit recommendation implementation, 2014* , https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BayesianLimitRecommendationImplementation.

[245] A. Caldwell, D. Kollár, and K. Kröninger, *BAT–The Bayesian analysis toolkit*, Computer Physics Communications **180** (2009) no. 11, 2197–2209, arXiv:0808.2552 [data-an].

[246] A. Caldwell, D. Kollár, and K. Kröninger, *An introduction to BAT*, tech. rep., September, 2013.

[247] C. Quigg, *LHC Physics Potential versus Energy*, arXiv:0908.3660 [hep-ph].

[248] B. A. Dobrescu and F. Yu, *Coupling-mass mapping of dijet peak searches*, Phys.Rev. **D88** (2013) no. 3, 035021, arXiv:1306.2629 [hep-ph].

[249] G. Salam, *General aspects of QCD at future colliders*, SLAC Workshop on Physics at a 100 TeV Collider, 2014.