A study of the ATLAS detector's sensitivity to quark compositeness

submitted to

University College London

Oliver Harris

February 6, 2009

I, Oliver Harris, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The sensitivity of ATLAS to processes characterised by high-energy hadronic jets has been investigated, with analysis of both theoretical and experimental sources of uncertainty. This analysis has been applied in particular to a study of a quark compositeness model.

Contents

List of Figures

Chapter 1

Introduction

A number of possible new physics scenarios at the LHC have signatures including high-$p_\perp$ jets, that is, high-momentum groups of strongly-interacting particles (hadrons) given off at a large angle relative to the axis of the colliding beams. One of these scenarios is the interaction between constituent particles making up composite quarks (which are assumed to be elementary in the Standard Model). Two main sub-analyses are presented: firstly, a study of the sensitivity of ATLAS to quark compositeness, with consideration given to both theoretical and experimental sources of error; and a study of jet cross-sections calculated at next-to-leading order, with the emphasis here on two kinds of theoretical uncertainty.

Chapter 2

Theoretical overview and motivation

2.1 The Standard Model of particle physics

In the early 1970s the Standard Model of particle physics began to emerge as a coherent picture of the relationship between the known elementary 'matter' particles and the quanta of the various fields by which they interact. The so-called matter particles are more properly known as *fermions* (although only three of the them make up 'matter' as we normally understand the term), while those that mediate forces between the fermions are called *bosons*.

The Standard Model describes 12 fermions, that is, 12 particles with a spin (a sort of intrinsic angular momentum) of $\frac{1}{2}$, in units of Planck's Constant [1]. Six of these are *quarks*, which take part in all three kinds of interaction described by the Standard Model; the strong, weak and electromagnetic forces. There are three so-called 'up-type' quarks, namely the u ('up') quark itself and its heavier relatives, the c ('charm') and t ('top') quarks: all have an electric charge of $+\frac{2}{3}$ in electronic charge units. The others are the 'down-type' quarks; these are the d ('down'), s ('strange') and b

[1] Strictly speaking, the fundamental quantum unit of angular momentum is the reduced Planck's Constant, $\hbar$, signifying Planck's Constant h divided by 2π.

('bottom') quarks; they have a charge of $-\frac{1}{3}$.

Quarks are never observed as free particles as they are subject to *confinement*, meaning they are only ever found inside composite particles called hadrons. These come in two varieties: mesons, consisting of a quark bound to an antiquark ($q\overline{q}$), and baryons/antibaryons, consisting of three quarks/antiquarks ($qqq/\overline{qqq}$). All quarks possess a property known as *colour*, which is a form of charge analogous to electric charge (see Section 2.1.2). Hadrons are overall colour-neutral or 'white'[2], so nuclear forces between them are a residual phenomenon analogous to the van der Waals force between two overall neutral atoms. The two lightest baryons are the familiar proton and neutron of normal nuclear matter.

There are three charged leptons, which feel the electromagnetic and weak forces but not the strong force; these are the electron, e^-, and its heavier cousins, the μ^- and τ^-. Finally there are three neutrinos, ν_l, each associated with a particular charged lepton ($l = e$, μ or τ), which feel only the weak force and are consequently extremely hard to detect: at the LHC experiments, their presence will be inferred from 'missing' energy and momentum, since they will exit the detectors without interacting.

The other ingredients of the Standard Model are called *gauge bosons*, which mediate exchanges of momentum between the fermions. They all have one unit of spin. The electromagnetic force is mediated by a single particle, the photon, denoted γ; this is the quantum of electromagnetic radiation, such as visible light. It has neither mass nor electric charge, imparting an infinite range to the electromagnetic force. There are three so-called massive vector bosons, the neutral Z^0 and the charged W^+ and W^-; these three bosons together mediate the weak force, which is responsible for certain kinds of radioactive decay and is vital to the thermonuclear processes that power stars. The theories of the electromagnetic and weak forces were unified in the early 1960s into a theory of *electroweak* interactions. Finally there are eight massless gluons - identical except for their colour

[2]Note that mesons exist in a superposition of colour-anticolour states ($\frac{1}{\sqrt{3}}(r\overline{r} + b\overline{b} + g\overline{g})$). While a simple combination like $r\overline{r}$ would be colour-neutral, it is not invariant under transformations in the relevant symmetry group and is therefore not a colour singlet; this illustrates the shortcomings of the 'colour mixing' analogy to charge in SU(3).

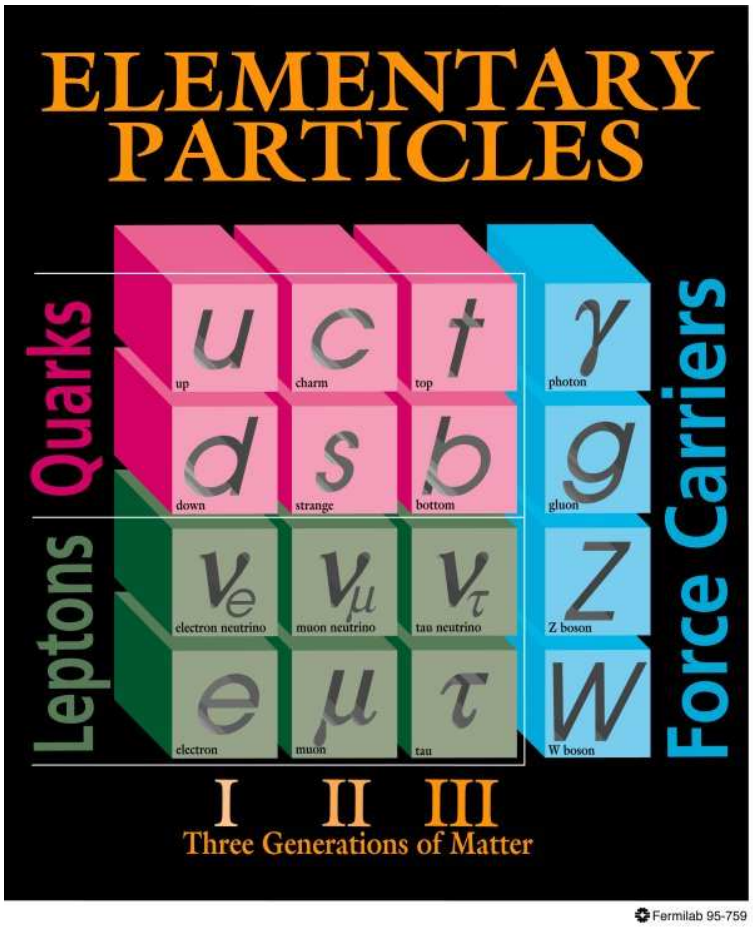

Figure 2.1: The Standard Model of particle physics.

charge - which mediate the strong force that binds quarks together inside hadrons. The fermions and gauge bosons of the Standard Model are tabulated in figure 2.1.

In addition to the fermions named above, the Standard Model includes antiparticle partners to all of them, which have identical masses and lifetimes but opposite quantum charges. The gauge bosons too have antipartners, but these are not listed separately; the photon and Z^0 are their own antiparticles, the W^+ and W^- are antiparticles of each other and the antiparticle of any given gluon will be another gluon in the octet.

The final missing piece of the Standard Model is the Higgs boson, a scalar (i.e. spin-0) particle envisaged as the quantum of the field proposed to give rise to spontaneous breaking of the electroweak symmetry via the so-called Higgs Mechanism. This was originally proposed to explain the

masses of the Z^0 and $W^{\pm}$ (in contrast to the masslessness of the photon) and is also thought to give rise to all the other 'bare' masses in the Standard Model [3]. The search for the Higgs boson - or bosons - is one of the main motivations for the LHC.

Note that one form of interaction is not described by the Standard Model at all: gravity. The search for a theory that unifies gravity with the quantum field theories of the Standard Model is the single greatest challenge facing theoretical physicists, and any theory that does achieve this will be a good candidate for the long-sought 'theory of everything'.

2.1.1 Gauge Symmetries

Central to the Standard Model is the concept of *gauge symmetry*.[4] A symmetry in this sense is a transformation on a field that leaves the Lagrangian of the field invariant. Symmetries can be *global*, meaning they are applied identically to the field at all points in space, or *local*, meaning the transformation is varied throughout space; gauge invariance is an example of local symmetry. When this is applied to a matter field the derivative term in the original Lagrangian must be modified to include a *gauge field*; this field may then be formally identified as the interaction between the particles of the original field. In the Standard Model, the field of the original Lagrangian is fermionic in nature, while the interaction field is bosonic. There is a dimensionless number called the *coupling constant* (though, as we shall see, they are not strictly constant) that quantifies the strength of the coupling between the fermions and bosons; in the electromagnetic interaction, it is called α (sometimes α_{EM}) and is related to the electronic charge by:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}, \tag{2.1}$$

in which e is the electronic charge, c the speed of light *in vacuo* and ϵ_0 the permittivity of free space[5]; the expression is simply $e^2/4\pi$ in natural units[6]. The equivalent constant for the strong

[3] There are other sources of particle mass; for instance, almost all the mass of protons and neutrons comes from the energy of the gluonic field binding the (practically massless) quarks together.

[4] This discussion is based largely on [1].

[5] A fundamental property of the vacuum, measured empirically to be about $8.85 \times 10^{-12}\mathrm{Fm}^{-1}$

[6] In this convention fundamental constants such as c and $\hbar$ are set equal to 1, so that mass and momentum both have the dimension of energy and units of eV (or MeV etc.) and angular momentum is dimensionless.

force is denoted α_s, and its behaviour with varying distance scales is of vital importance in the phenomenology of the strong interaction; see section 2.1.3.

The form taken by the gauge symmetry will depend on the number of components to the fermionic field. For example, in electromagnetism there is just a single type of charge, so the field describing the electron (as the 'prototypical' charged particle in Dirac's original formulation) has a single component; thus the gauge symmetry in quantum electrodynamics (QED) is a multiplication by a single (complex) scale factor of magnitude 1, which forms the group U(1). This is called the gauge group of the theory. In quantum chromodynamics (QCD), the theory that describes the strong interaction, the fermionic field has three components, so its gauge group is SU(3); this is the group of dimension-3 *special unitary* matrices, 'special' indicating that they have a determinant of $+1$.

As may be expected from the simple form of its gauge group, QED is the simplest gauge theory, although its development served as a prototype for more complex theories such as QCD. In natural units, the Lagrangian of the electron field, based on the Dirac equation, is

$$\mathcal{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.2}$$

where m is the electron mass, ψ is its wave-function (the electron field) and $\overline{\psi}$ is the wave-function's conjugate. Repeated indices in a single term are implicitly summed over.

The symmetry for this system is $\psi \mapsto e^{i\theta}\psi$ (the U(1) symmetry), where θ is just a phase change of the wave-function; if $\theta = \theta(x_\mu)$, the symmetry is local, so that $\partial_\mu \theta$ is nonzero. This implies that $\partial_\mu \psi$ in the Lagrangian must be replaced with something a little more complex; the required operator is called the *covariant derivative*, D_μ, and its form is

$$D_\mu = \partial_\mu - ieA_\mu \tag{2.3}$$

where e is the electronic charge and A_μ is the gauge field, sometimes written $A_\mu(x)$ to explicitly show its locality. In the case of QED, it is identified with the electromagnetic four-vector potential. This leads to the so-called *interaction Lagrangian*:

$$\mathcal{L}_{int} = e\overline{\psi}\gamma^\mu \psi A_\mu = J^\mu A_\mu \tag{2.4}$$

where J^μ is the four-vector electric current density, related to the electromagnetic *field strength tensor* by $4\pi J^\nu = \partial_\mu F^{\mu\nu}$. This tensor is in turn derived from the gauge field: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This object can then be inserted into the original expression to obtain the full Lagrangian for QED:

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.5}$$

A somewhat similar prescription may be used to gauge the three quark fields associated with the three colour charges in QCD.

2.1.2 The theory of strong interactions

The interaction of quarks and gluons is described in the Standard Model by the theory of quantum chromodynamics (QCD), named in an analogous way to quantum electrodynamics. Although there are some similarities between QCD and QED, such as massless gauge bosons, there are two very important differences; whereas in QED there is only one type of charge, QCD has three; and as a result, the gauge bosons of QCD carry charge as well, whereas in QED the photon is electrically neutral. The reason for this is explained below.

In the language of gauge theory, QCD is based on the non-Abelian symmetry group[7] SU(3), so the particles that form the group's *fundamental representation* - the quarks - have three kinds of charge. These were named *red*, *green* and *blue* in analogy to the primary colours of light-mixing theory; however, it should be noted that this has nothing to do with colour in the literal sense. The antiparticle of a quark with a certain colour simply has the 'opposite' colour; for example, the antiparticle of a red u quark is an anti-red $\overline{u}$.

Gauge theories described by an SU(n) symmetry group have $n^2 - 1$ degrees of freedom, so the *adjoint representation* of SU(3) is an octet; namely, the eight gluons. They have colour charges of a different sort from the quarks; for example, a quark with colour r could give off an $r\overline{g}$ gluon, becoming a g-coloured quark in the process; or a g-coloured quark could annihilate with an $\overline{r}$ antiquark to produce a $g\overline{r}$ gluon. The triple gluon vertex, by which two gluons fuse to form a single gluon or a gluon splits into two, has no analogue in QED since the latter theory has only one

[7] Meaning transformations under the symmetry group do not commute, as they do in the case of U(1).

degree of freedom, which is 'used up' by the one type of charge (i.e. positive or negative on a single axis of charge) so there is no way for a photon to have a *different kind* of charge: no orthogonal axis for the photon's charge to lie on.

The Lagrangian describing QCD is as follows:

$$\mathcal{L}_{\text{QCD}} = \overline{q}(i\gamma^\mu \partial_\mu - m)q - g(\overline{q}\gamma^\mu T_a q)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{2.6}$$

in which q is the quark field, g is interaction strength, γ^μ are the Dirac matrices, G^a_μ are the eight gauge fields (gluons) and T_a are the Gell-Mann matrices, related to the structure of SU(3). In this Lagrangian, the covariant derivative takes the form $D_\mu = \partial_\mu + igT_aG^a_\mu$.

The term $G^a_{\mu\nu}$ is the gluonic field tensor, analogous to the electromagnetic field tensor $F_{\mu\nu}$ in QED, but with an extra term that represents the self-interaction of the field. It is given by:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f_{abc} G^b_\mu G^c_\nu \tag{2.7}$$

where the $G^b_\mu G^c_\nu$ is the self-interaction term and f_{abc} are the structure constants, derived from the commutation relations (Lie algebra) of the generators of SU(3):

$$[g^a, g^b] \equiv g^a g^b - g^b g^a = i f_{abc} g^c. \tag{2.8}$$

The Gell-Mann matrices T^a are a particular conventional choice for g^a.[8]

2.1.3 The running coupling constant

The strong coupling 'constant' is not actually constant, but varies with the distance between strongly-interacting particles. This is usually expressed as a dependence on some energy scale (often, in practice, an energy scale squared) rather than on a distance; the two quantities may be thought of as inversely proportional, as high energies correspond to short distances.

Since gluons also carry strong colour charges, they can interact with each other in a way that has no analogue in QED. One result of this is that the strong attraction between any two quarks gets weaker as the quarks are brought together (a phenomenon known as asymptotic freedom)

[8]The generators g^a are not to be confused with the interaction strength g, which is related to the strong coupling constant α_s by $\alpha_s = g^2/4\pi$.

and, conversely, very strong as they are pulled apart. Eventually enough latent potential energy is stored in the 'stretched' colour field lines (self-interacting gluons) that new quark-antiquark pairs are produced, immediately forming new hadrons with the two original quarks.

The scaling of the coupling constant with an energy scale μ is given by:

$$\alpha_s(\mu^2) \approx \frac{1}{\beta_0 ln(\mu^2/\Lambda^2_{\mathrm{QCD}})} \tag{2.9}$$

where β_0 is a constant computed by Wilczek, Gross and Politzer [2] and Λ_{QCD} is the QCD scale, measured experimentally to be approximately 217 MeV. One consequence of this is that many kinds of calculation involving the strong interaction may be made using the approximation that the three lightest quarks (u, d and s) are effectively massless, as their masses are below Λ_{QCD}, while the masses of the c, b and t quarks must be taken into account.

The behaviour of $\alpha_s(\mu^2)$ with energy is a result of *vacuum polarisation*, and is responsible for both asymptotic freedom and confinement. In QED, this has the effect of increasing the strength of the electromagnetic coupling constant α_{EM} at short distances, as virtual e^-e^+ pairs latent in the vacuum, which normally serve to shield the 'bare' electric charges of pointlike particles, have less effect on the interaction between the charges of any two real particles; there is less shielding at shorter distances. In QCD, there is a similar shielding effect due to virtual quark-antiquark pairs but a stronger 'anti-shielding' effect due to virtual gluons; this effect decreases as any two real quarks are brought together, leading to the strong coupling constant α_s asymptotically approaching zero as the distance approaches zero: see figure 2.2. Conversely, the potential increases as the quarks are pulled apart. For this reason perturbative calculations in QCD ('pQCD' for short) are possible in high energy regimes but not at low energies, or equivalently large distances; the opposite is the case in QED.

2.2 Compositeness models

Since the discovery of atoms in the late 19th century, matter has repeatedly been revealed to consist of smaller and more elementary particles that those previously thought to be fundamental - first

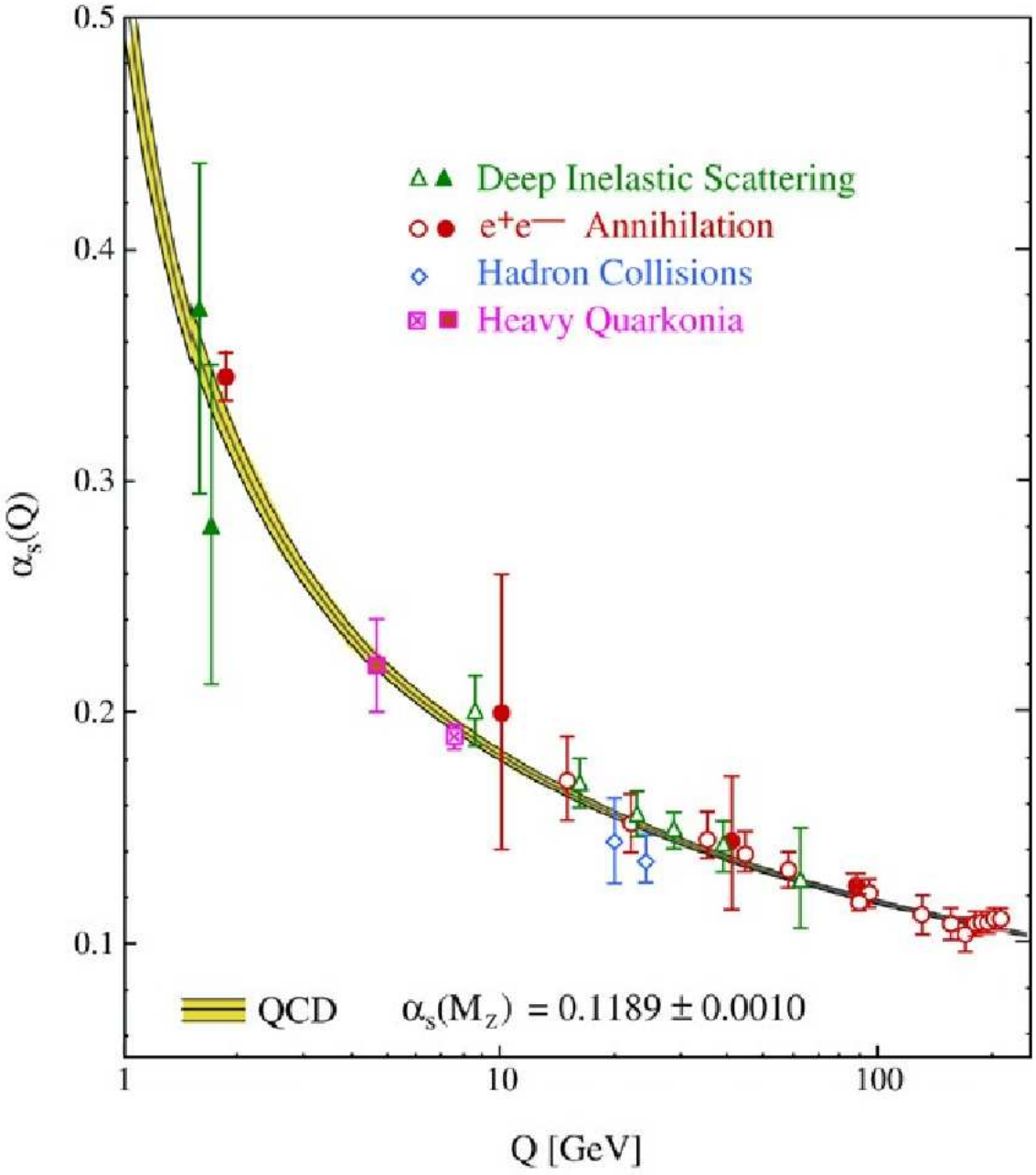

Figure 2.2: Asymptotic freedom and confinement in the strong interaction. $\alpha_s(\mathrm{M_Z})$ is the value of the strong coupling constant at energies equal to the mass of the Z^0 boson, 91.1876 GeV [15]; Q is the energy scale of the interaction.

with Rutherford's discovery of the nucleus around 1910, followed by the discovery of the proton and neutron between the World Wars, and most recently unequivocal evidence for the existence of quarks, which began to emerge in the late 1960s. It is therefore reasonable and entirely natural to ask whether quarks themselves are truly fundamental, or whether they too possess substructure.[9]

The concept of particles existing on a smaller scale than electrons first appeared in 1948, in a reprint of E.E. Smith's science fiction novel of 1930, *Skylark Three*, which mentions *"sub-electrons of the first and second type"*, the latter of which apparently had properties that in some sense anticipated the graviton [12]. However, the earliest scientific investigation of hypothetical particles existing on a scale smaller than those of the Standard Model dates from 1974, when Abdus Salam and Jogesh Pati [13] proposed the existence of particles they named 'preons', which were postulated to make up the fermions of the Standard Model (most of which were known at the time). It was hoped that a relatively small number of types of preon could explain the large number of supposedly elementary particles, in much the same way that the proton and neutron make up the 3,000 or so known nuclides (elements and isotopes), and that just a few elements may be combined to make countless organic compounds. It was also hoped that fermion substructure could explain a number of other phenomena inexplicable in the Standard Model, such as the huge range of particle masses - some 11 orders of magnitude between the top quark and the current best value for the lightest neutrino [10] - and the apparent 'redundancy' of three generations of fermions, identical except for mass and width (inverse lifetime). Additionally, attempts were made to used the preon model as a starting point for a Higgs-less theory of electroweak symmetry breaking, a Grand Unified Theory to unify the electroweak and strong interactions and even a theory of quantum gravity.

One interesting feature of compositeness models is that they generally have experimental features which are largely or wholly independent of the models themselves, depending instead only on

[9]It should be noted that while some theorists believe superstrings to exist on a much smaller length scale than that normally associated with quarks and other Standard Model particles, superstrings (in these models) do not 'make up' quarks in the same way that quarks make up hadrons; each particle is believed to be a single string, it's just that the strings exist on a scale far smaller than that currently or forseeably probeable with conventional accelerator experiments.

[10]Top quark mass: 171.2±2.1 GeV [14]; ν_e mass: < 2 eV [15]

the scale at which the Standard Model particles start to exhibit substructure. The word 'model' is used rather than 'theory', since these hypotheses are not generally advanced enough to include a Lagrangian from which dynamical equations could be derived. For completeness, a brief overview of the main models of compositeness are given below:

Harari's 'Rishon' Model [16] postulates a level of sub-Standard Model[11] elementary particle called *rishons*, which come in just two varieties; T (for 'Third') and V (for 'Vanishes') and their antiparticles, denoted t[12] and v. Ts have a charge of $+\frac{1}{3}$ and Vs are uncharged, while both kinds of rishons have a spin of $\frac{1}{2}$, so the first-generation SM fermions may be constructed from triplets of rishons as follows:

- TTT - e^+
- VVV - ν_e
- $TVV/TVT/VTT$ - u-type quarks (three permutations provide three QCD colour charges)
- $TTV/VTV/VVT$ - d-type quarks

Antiparticles of the above fermions may be obtained simply by substituting t for T and v for V in the above combinations. The second and third generations of fermions are assumed to be excited states of the first generation, just as the ρ mesons may be considered excited states of the π mesons. The vector bosons consist of six rishons; for example, the W^+ (W^-) is made up of $TTTVVV$ ($tttvvv$); this allows the violation of lepton and baryon number (L and B) by processes such as $u + u \rightarrow \overline{d} + e^+$ via an intermediate state with a charge of $+\frac{4}{3}$; however, the quantity $B - L$ is conserved. [13] Rishons are assumed to be bound to each other by a 'hyperstrong' force acting on their 'hypercolour' charge; the existence of gauge bosons responsible for this force is not discussed. Also, particle masses are not explained or predicted.

[11] Henceforth 'SM'.

[12] Not to be confused with the t, or 'top' quark.

[13] A phenomenologically identical process is predicted in some GUTs and has been the subject of extensive searches - all negative so far - for proton decay.

Bilson-Thompson's Topological Model [17] considers preons not as pointlike elementary particles but as topological features of a quantised spacetime; this approach show some promise in linking the phenomenology of particle physics to the theory of loop quantum gravity, which starts from the premise that spacetime itself may be quantised as a gauge field theory, just as particle fields are quantised and gauged in 'standard' field theories. So far, Bilson-Thompson has been able to model the first generation of fermions with correct parity properties by treating them as composites of preons which are in some sense 'braids' of spacetime. However, this model too lacks a Lagrangian to describe interactions, and it has not yet been possible to prove that the SM particles modelled are indeed fermions.

Bilson-Thompson's model does, however, provide a neat way to sidestep a problem that has plagued earlier preon models, namely the very large masses they tend to predict for composite particles, including those known to be very light, such as electrons and neutrinos. This comes about as a result of the Heisenberg Uncertainty Principle (HUP), which implies that particles confined inside a small volume (meaning they have a small uncertainty in their position) must necessarily have a large uncertainty in momentum. As a consequence of this, the expectation value of the particles' absolute momentum is large, meaning they have a large kinetic energy and, by mass-energy equivalence, the volume in which they are confined must have a large mass density. Consider, as an example of a simple generic preon model, a bound state of three massless preons held together by some 'hypergauge' field to form an electron; using the currently accepted lower bound on the compositeness scale (denoted Λ_{comp}) of the electron of around 10 TeV [15], this gives a maximum electron 'radius' of

$$r_e \leq \frac{\hbar c}{10\text{TeV}} \approx 2 \times 10^{-20}\text{m} \tag{2.10}$$

which may be thought of as the largest allowed uncertainty in position (Δx) for the preons. From the HUP, the preons therefore have an uncertainty in momentum given by

$$\Delta p = \frac{\hbar}{2\Delta x} = \frac{\hbar}{2r_e} = \frac{\hbar}{2(\hbar c/\Lambda_{\text{comp}})} \approx 5\text{TeV}/c. \tag{2.11}$$

If this value is multiplied by 9 for the three space dimensions and the three preons, the estimated mass for the composite electron seems to be some eight orders of magnitude greater than the known mass of the electron, 511keV/c^2 [15]. While previous preon theories have attempted to cancel this anomalously high mass with large negative binding energies, Bilson-Thompson's preons are not bound by a gauge interaction but exist on the Planck scale, far below the scales of effective field theories [17], and are in fact bound by quantum topology, according to which topological invariants correspond to conserved quantum numbers. However, if this is the case, it seems difficult to imagine any way of directly probing this structure in the standard collider paradigm due to the disparity between the energies of colliders and the Planck energy.

2.2.1 Compositeness interactions

Feynman diagrams for two possible kinds of event involving new physics at the quark or sub-quark level are shown in figure 2.3. In the upper diagram, quark substructure is revealed in an s-channel exchange involving preons (left), and a new interaction is revealed by t-channel exchange of a novel force carrier (right). Here the invariant mass of the hard interaction, M, is of the same order as Λ_{comp}. In the lower diagram, the internal propagator essentially disappears as M is much larger than Λ_{comp} and both types of interaction appear as a 'contact interaction' (quartic coupling). The concept of the contact interaction is not limited to searches for fermion substructure but is also used in searches for other new physics processes, such as additional heavy gauge bosons or extra-dimensional graviton exchange, where experimental signatures are sought despite the energy exchange of the interaction being well below the mass scale of the new particle, resonance or dimension under consideration. It should be noted that the failure so far of any experiment to find evidence in favour of quark substructure does not rule out such theories, as it may be that energies so far achieved by colliders are much lower than the characteristic binding energy of the preons within the apparently elementary known particles; after all, if a beam of electrons is incident on a proton target with energy lower than the binding energy of a proton, the latter will appear elementary as the quark/gluon substructure is inaccessible and a simple elastic scatter

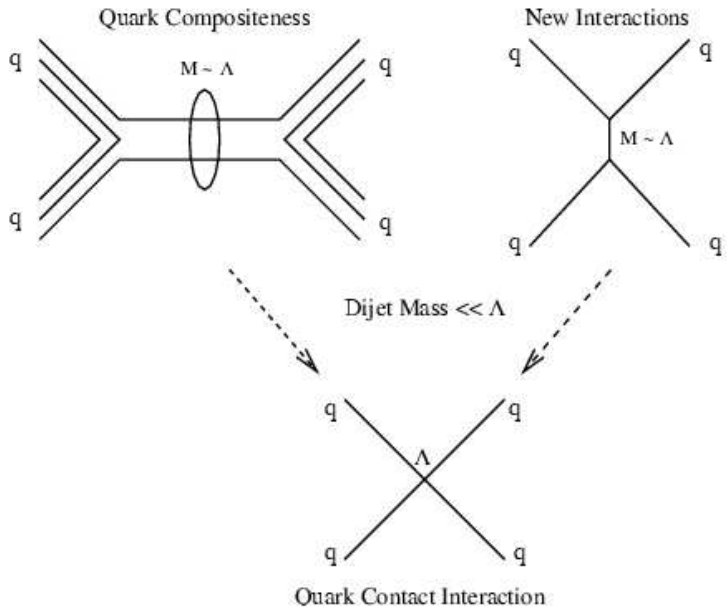

Figure 2.3: Preon-level interactions observed as an apparent contact interation.

occurs. Thus experimental searches for compositeness may be thought of as setting upper bounds on the distance scale of the substructure; to date, the current experimental lower bound on the energy scale of quark compositeness of the sort considered in this analysis is 2.7 TeV [15].

2.3 Experimental signatures of compositeness

If hadrons collide with a partonic kinetic energy in the centre-of-mass frame (henceforth 'CoM') approaching the binding energy of the preons, sub-quark interactions will occur, whereby 'naked' preons come into contact with each other either by co-annihilation or exchange of a preon-level force mediator, allowing quark substructure to become apparent. An important experimental signature of this will be an excess of high-$p_\perp$ jets (over the rate predicted for pointlike quarks) in the central pseudorapidity region[14], with *less* of an excess at higher absolute rapidity. The reason for this may be explained by analogy to the experiment by which Rutherford discovered the nucleus. The 'plum pudding' model of the atom prevalent at the time, in which electrons

[14]See section 4.4 for a definition of this term

and positive charges were uniformly distributed throughout the atom, suggested that a stream of charged particles passing through a thin layer of material should have been deflected through small angles as they negotiated the 'lumpy' potential landscape in the atoms. What was instead discovered was that the vast majority of α particles incident on thin ($\sim$200 atoms thick) gold foil passed through without apparent interaction, but that a few were deflected through very large angles, some even being backscattered in the direction of the source. Rutherford described this result as astonishing "...as if you fired a 15-inch shell at a piece of tissue paper and it came back to hit you". [18] The only possible explanation was that a few of the massive, high-momentum α particles were being deflected by an even more massive, positively charged and *very small* particle within the atom; this was soon established as the nucleus.

By analogy, a 'target' proton may be considered to contain a cloud of partons, so that 'test' quarks in an approaching proton are most likely to be involved in a glancing collision, that is, a low-μ^2 t-channel gluon exchange, with only a few taking part in head-on collisions with high μ^2 (the square of the momentum exchanged in the process). If the quarks are not considered pointlike but have substructure, head-on collisions between *preons* are more likely, simply because each quark is composed of several of them. These high-μ^2 interactions produce jets with large energy at low $|\eta|$, and therefore high $p_\perp$.

Chapter 3

The Large Hadron Collider

The Large Hadron Collider (figure 3.1) currently undergoing commissioning at the European Organisation for Nuclear Research (CERN) on the Swiss/French border will be the world's most powerful particle accelerator/collider when it starts operating at design performance in 2009. Two beams of protons will circulate in opposite directions, each with a kinetic energy of 7 TeV in the laboratory rest frame, giving a CoM energy of 14 TeV. This is an order-of-magnitude increase from the 1.96 TeV CoM energy achieved by the current most powerful collider, FNAL's Tevatron $p - \overline{p}$ machine in the United States. Just as significantly, the LHC at design performance will have an instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, an increase of two orders of magnitude over the Tevatron[1]. The machine is being constructed in the 27 km-circumference tunnel dug for its predecessor as CERN's main accelerator, the Large Electron-Positron collider (LEP), which was decommissioned in 2000 after a long and scientifically fruitful working life.

The protons, travelling in 2,808 bunches of approximately 10^{11}, will be injected into the LHC from the SPS (Super Proton Synchrotron) at an energy of 450 GeV. They will be accelerated to 7 TeV by radio-frequency cavities and steered by superconducting magnets, then made to collide head-on at four points around the LHC ring. Each of these interaction points lies at the centre of an array of high-precision detector components, collectively known as an experiment.

[1] A 'Super-LHC' upgrade planned for 2015 will increase this by a further order of magnitude.

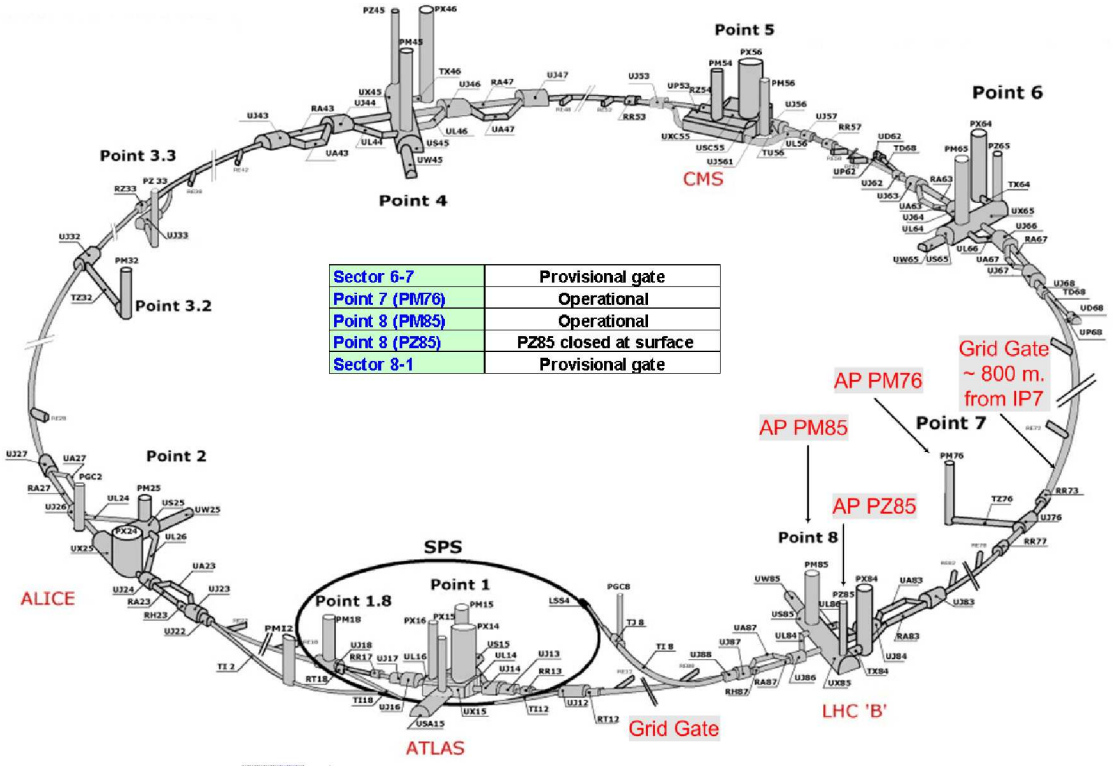

Figure 3.1: The Large Hadron Collider.

Two of these are specialised experiments; ALICE, which is dedicated to studying the results of heavy ion collisions after the LHC's proposed conversion to a Pb-Pb collider, and LHCb, whose main purpose is to make precision measurements of the phenomenon of charge-parity violation in B mesons. The other two experiments are more 'general purpose' in scope; these are CMS and ATLAS. Amongst their intended physics goals are the discovery of the Higgs boson, required in the Standard Model to impart mass to other particles; heavy partners of the known SM particles predicted by supersymmetry, which may account for the so-called 'dark matter' thought to make up most of the matter in the Universe; and the investigation of a host of 'exotic' scenarios, such as 'large' extra dimensions, leptoquarks, quark compositeness, extra fermion families and so on. A fifth, much smaller experiment - TOTEM - shares an interaction point with CMS and will be used to measure, amongst other things, the total cross section of the pp collisions, which will be an important quantity in the analysis of data from the other experiments.

3.1 Hadron collider physics

The physics of hadron colliders is inherently more complex than that of electron-positron colliders such as LEP, in that the particles undergoing collision are not themselves elementary, but are composed of several types of (apparently) elementary constituents, collectively known as *partons*. These are the quarks and gluons that make up (in the LHC's case) the protons in each beam. In a collision, the two protons have the following momentum 4-vectors:

$$\mathbf{P}_1^\mu = (E, 0, 0, p_z); \qquad \mathbf{P}_2^\mu = (E, 0, 0, -p_z), \tag{3.1}$$

in which the z-axis is defined as the direction of one of the beams and the μ superscript runs from 0 to 3 for the four space-time components. Note than, since $E^2 = \mathbf{p}^2 + \mathrm{m}^2$, in which $\mathbf{p}$ is the momentum 3-vector, $E \approx |p|$ is a very good approximation, as the CoM energy of each proton is far larger than its invariant mass-energy ('rest mass'). In practice, it is often convenient to use the so-called Mandelstam variable s, which is the squared total energy:

$$s = \mathbf{P}_1^\mu \mathbf{P}_{2\mu}, \tag{3.2}$$

which of course has the value of $(14\mathrm{TeV})^2$ at the LHC.

However, the momentum quantities involved in the actual collision (the 'hard subprocess') are not $\pm p_z$, but the quantities $x_1 p_z$ and $x_2(-p_z)$, in which x_1 and x_2 are the fractions of the incoming protons' momenta carried by the partons involved in the hard scatter. If each proton has kinetic energy $E = |p_z|$ (using the above approximation), the total squared energy involved in the hard subprocess is $\hat{s} = x_1 x_2 s$. The probability of a parton having a particular value of x is described by the probability density function $f(x)$, called a *parton density function* (PDF).

3.2 Parton densities

Parton densities are an important source of theoretical uncertainty in hadron collider experiments, and as such major efforts are continually being made to improve our knowledge of them. One of the main difficulties is that the functional form of a PDF cannot be be calculated analytically

using our current knowledge of QCD; instead, an empirically-motivated function is selected as an informed guess, and then fitted to data points extracted experimentally. The following formula gives the hadronic cross-section in terms of partonic cross-section and parton density functions:

$$\sigma(\mu^2) = \sum_{i,j=1}^{n_f} \int_{x_1=0}^{1} \int_{x_2=0}^{1} \hat{\sigma}_{1,2\to X}(p_1, p_2, \alpha_s(\mu^2), \mu^2/\mu_F^2) f_1^i(x_1, \mu^2) f_2^j(x_2, \mu^2) dx_1 dx_2, \qquad (3.3)$$

in which σ is the hadronic cross-section, $\hat{\sigma}_{1,2\to X}$ is the partonic cross-section (for partons in hadrons 1 and 2 going to some final state X), $f_{1,2}^i(x_{1,2}, \mu^2)$ is the PDF for partons 1 and 2, μ_F^2 is the factorisation scale (see 7.2.4) and n_f is the number of flavours being considered (typically four or five). The PDF sets used in this analysis are the ZEUS-JETS 2005 [3] and CTEQ6.1M [4] sets. More details on the techniques used to extract and fit PDFs can be found in Appendix A.

Although the proton has been known to science for close to a century, its properties have been far from exhaustively investigated. It became clear in the early 1960s that protons were not elementary particles, but were made up of as-then unknown constituents, which Richard Feynman named 'partons'. After Murray Gell-Mann's quark hypothesis (which was independently put forward by George Zweig) became increasingly accepted in the physics community through the late 1960s and early '70s, partons came to be identified with quarks and the gluons that hold them together. Experimental verification of quarks and gluons - at SLAC in 1968 and DESY in 1979, respectively - confirmed this identity.

The observation that first suggested hadrons are not elementary particles was dubbed 'scaling'. This is the phenomenon of the structure function F_2 (a coefficient in the amplitude for $e^- p$ scattering) depending, to a very good approximation, only on x and not on the energy scale of the interaction, μ. (see Appendix A). This could be explained by postulating pointlike elementary particles inside a proton of finite physical extent; the non-dependence on the energy scale indicated that, no matter how deeply the proton was probed, the constituents continued to appear pointlike, with no structure of their own. Further evidence came from the fact that, in collisions between protons and electrons, only a fraction of the proton's momentum was observed to be involved in the collision. This fraction, in the CoM frame, is called 'Björken x', after James Björken, who first described the phenomenon.

A parton density function, at a given value of μ^2, is the probability that a parton of flavour i will have a particular value of x; that is to say, $f_i(x)$, defined as

$$P_i(x_a < x < x_b; \mu^2) = \int_{x_a}^{x_b} f_i(x, \mu^2) dx \tag{3.4}$$

which is the probability that, at a given value of μ^2 (the square of the energy scale of the interaction), a parton of flavour i will have a value of x between limits x_a and x_b; or, differentially,

$$f_i(x) = \frac{dP_i}{dx}. \tag{3.5}$$

A graphical representation of a PDF will typically have different lines representing u and d valence quarks, sea quarks and gluons, plotted on the same axes. In practice, the y-axis variable is usually plotted as $xf_i(x)$, so that it represents the probability-weighted fractional momentum carried by each type of parton, so

$$p_i(\mu^2) = p \int_0^1 x f_i(x, \mu^2) dx \tag{3.6}$$

is the momentum contribution from a given parton, in a hadron with a total momentum p involved in some process at an energy scale μ. As PDFs represent probability distributions, they are naturally normalised, in that the integrated momentum contributions from all parton types sum to unity:

$$\sum_i \int_0^1 x f_i(x, \mu^2) dx = 1. \tag{3.7}$$

This is the *momentum sum rule*, and is of great utility in compiling PDFs since it allows the probability density of a parton type that is hard to measure in a certain $\{x, \mu^2\}$ regime to be constrained using the more precisely known probability densities of the other parton types. Further information on the PDF sets used in this analysis may be found in section 7.1. Note that in later discussion the symbol Q^2 is used instead of μ^2, in accordance with the publishers' convention,

Chapter 4

The ATLAS detector

The ATLAS detector - named for *A Large Toroidal LHC ApparatuS* - is the largest detector ever constructed for a collider experiment, measuring 42 m in length and 11 m in radius, and weighing some 7,000 tonnes. Its detector components are arranged in three main layers: from the beam pipe outwards, the Inner Detector (subdivided into the Pixel Detector, the Semiconductor Tracker and the Transition Radiation Tracker), the Calorimetry (subdivided into electromagnetic and hadronic calorimeters) and finally the Muon Spectrometer. A simplified cutaway diagram is shown in figure 4.1.

Since this analysis is concerned mainly with high-$p_\perp$ hadronic jets, particular emphasis will be placed on ATLAS's hadronic calorimetry.

4.1 The Inner Detector

The purpose of the inner detector components [5] is to measure the tracks of charged particles to a high degree of accuracy. The curvature of the tracks, caused by the solenoidal magnetic field of 2 T, allows the momentum and charge of the particles to be reconstructed. These components are designed to absorb as little energy as possible from the particles so this quantity can be accurately measured by the calorimeters. The inner detector is 7 m long and a metre in radius, giving coverage of the pseudorapidity region $|\eta| < 2.5$.

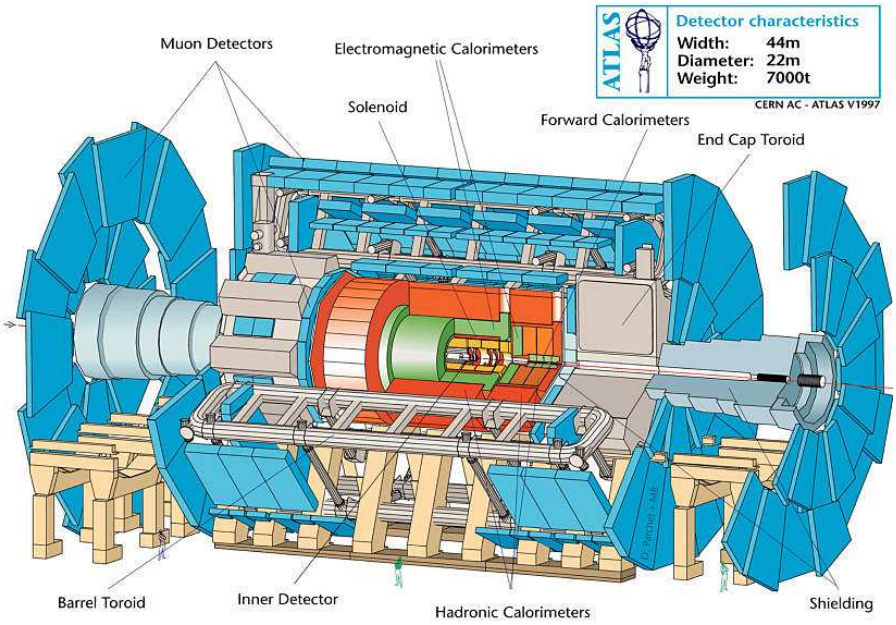

Figure 4.1: ATLAS.

4.1.1 The Pixel Detector

The Pixel Detector consists of three concentric layers of silicon wafers, which have a spatial resolution of 12 μm in $r - \phi$ and 60 μm in z. Each individual sensor measures 16.4 mm by 60.8 mm, and has 46,080 pixels, each pixel measuring $50\mu m$ by 400 μm. There are 1,744 sensor modules in the entire Pixel Detector, giving close to 80 million readout channels. The innermost of these layers is just 4cm from the beam pipe, allowing excellent three-dimensional track reconstruction; this is important for the determination of impact parameters in tracks that have come from a secondary vertex, such as the decay products of B mesons that have travelled several millimetres from the primary vertex (at the interaction point) before decaying. It is designed to withstand the extremely harsh radiation environment that will exist this close to the interaction point, but even so, the innermost layer is expected to need replacing every few years once the LHC enters its high-luminosity running phase.

4.1.2 The Semiconductor Tracker

This component, know as the SCT, is also based on modules of silicon wafers. In the toroidal 'barrel' region, it is composed of four concentric layers at radii of 30 cm, 37.7 cm, 44.7, cm and 52 cm; in addition there are nine sets of endcap wheels. The silicon wafers measure 6.4 cm by 6.4 cm; they are connected in pairs to form rectangular strips, and these strips are attached back-to-back at a small crossing angle to form modules. The small crossing angle is to allow resolution of tracks in the z-axis, i.e. parallel to the beams.

The SCT modules have a spatial resolution of 16 μm in $r - \phi$ and 580 μm in z. Each wafer has 768 readout channels and the component as a whole contains a total of 61 m^2 of silicon wafers.

4.1.3 The Transition Radiation Tracker

The Transition Radiation Tracker (TRT) consists of 370,000 4 mm-thick aluminium tubes, each filled with a $Xe/CO_2/CF_4$ gas mixture and having a thin gold-plated W-Re wire in the centre. Charged particles traversing these tubes give off transition radiation as they travel through materials of different dielectric properties; this radiation creates electron-ion pairs that can allow currents to flow, and these can be picked up from the wires. The straws each have a spatial resolution of 170 μm, providing an overall spatial resolution of 50 μm for the TRT as a whole.

The barrel section contains 52,544 axial straws of about 150 cm in length, at radii between 56 cm and 107 cm. The end-caps contain 245,760 radial straws at radii between 64 cm and 103 cm.

4.2 Calorimetry

Calorimeters, in contrast to the tracking components, are designed to interact with particles as much as possible, with the intention of efficiently turning their energy into a measurable form, typically charge separation or light. Calorimeters are usually specialised to measure energy deposited either by electromagnetic or nuclear (i.e. strong) interactions. See figure 4.2 for the layout of ATLAS's various calorimeter systems.

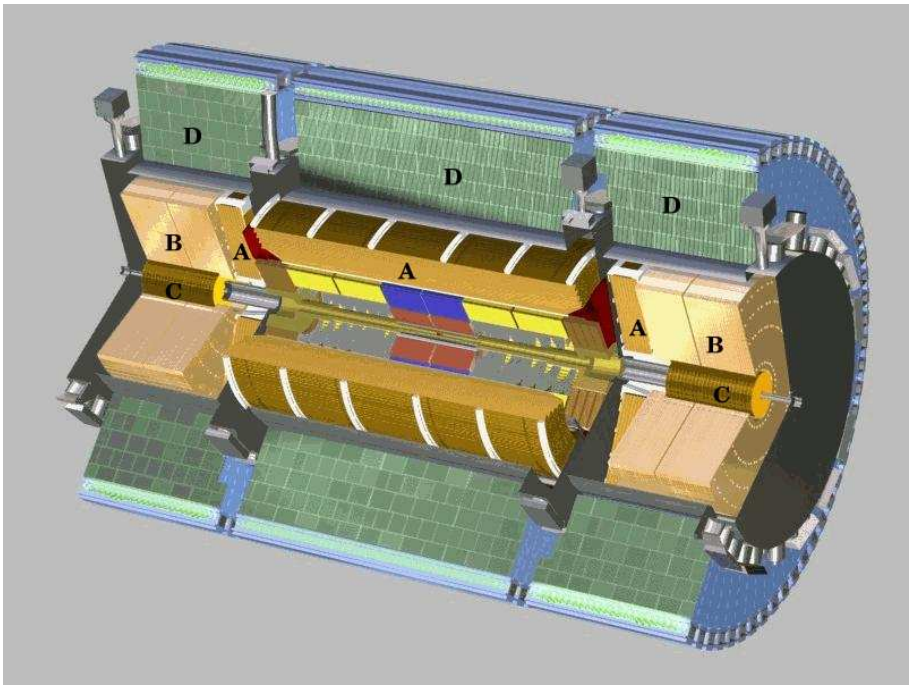

Figure 4.2: The ATLAS calorimetry, showing principal components: electromagnetic accordion calorimeters (A), hadronic endcap LAr calorimeters (B), forward LAr calorimeters (C) and hadronic tile calorimeters (D).

4.2.1 Electromagnetic calorimeter

ATLAS's electromagnetic calorimeter covers a pseudorapidity range of $|\eta| < 2.5$ and is made of alternating layers of lead and liquid argon arranged in an accordion pattern. When 'electromagnetic' particles (mainly photons and electrons/positrons) pass through the lead, they initiate showering, that is, the production of a large number of secondary particles. This mainly happens by ee pair production ($\gamma \rightarrow e^- e^+$) or *bremsstrahlung* ($e^{\pm} \rightarrow e^{\pm}\gamma$). The number of particles in the shower rapidly increases and their mean energy decreases until the remaining particles have too little energy to produce new ones, and all the energy has been dumped into the calorimeter. Electromagnetic showers penetrate to a depth parameterised by the *radiation length*, X_0 (the mean free path of electromagnetically interacting particles), and have a narrow transverse profile.

4.2.2 Hadronic calorimeter

The hadronic calorimeter relies on hadronic showering, which is an analogous process to electromagnetic showering; hadrons collide inelastically with nuclei in the target material to produce showers of secondary hadrons, mainly pions and kaons. The absorption of hadrons by material is parameterised by the *interaction length*, λ, analogous to X_0 in the EM calorimeter. This is the parameter that describes the average rate of absorption of hadrons with depth of material traversed:

$$P_{abs}(x) = e^{-\lambda/x}, \tag{4.1}$$

$P_{abs}(x)$ being the probability that a particle has been absorbed after travelling a distance x through the material.

The hadronic calorimeter [6] consists of three main sub-component types: in the barrel and extended barrel regions ($|\eta| < 1.7$), iron is used as a target and plastic strips are used as a scintillator (a material that emits flashes of light when charged particles pass through it); in the $1.5 < |\eta| < 3.2$ region, overlapping somewhat with the extended barrel, is a Cu/LAr (liquid argon) sampling calorimeter; in $3.2 < |\eta| < 4.9$ Cu/W/LAr is used as a combined hadronic/electromagnetic calorimeter. At $\eta = 0$, the hadronic calorimeter has a thickness of eleven interaction lengths,

making 'punchthrough' (hadrons passing through the calorimeter and making it as far as the muon chambers, where they can cause false signals) highly unlikely.

The design performance energy resolution is

$$\frac{\sigma_E}{E} = [\frac{58\% \pm 3\%}{\sqrt{E}} + (2.5\% \pm 0.3\%)] \oplus \frac{(1.7 \pm 0.2)}{E}. \quad [6] \qquad (4.2)$$

in which energies are measured in GeV.

The absolute energy resolution, σ_E, varies principally with $\sqrt{E}$ as the number of secondary particles produced in the shower is proportional to the energy of the incoming hadron, or hadronic jet; it then follows that the relative statistical uncertainty (σ_E/E) on the measured energy deposit is inversely proportional to the square root of the number of shower particles, and therefore inversely proportional to the square root of the energy deposit. There is also a term representing electronic noise, which in absolute terms is not a function of energy and is therefore inversely proportional to energy when evaluated as a relative uncertainty.

4.3 Muon spectrometers

Although not used in this analysis, an brief description of the muon spectrometers [7] will be given for the sake of completeness.

Muons are classed as Minimum Ionising Particles, or MIPs, as their relatively large mass (compared to electrons) allows them to pass through large distances of dense matter with very little interaction with atomic electrons; in addition, being leptons, they have no strong interaction. This, together with the fact muons are 'stable' (they have a lifetime far longer than the time typically taken for them to traverse the detector) dictates that muon detectors are typically the outermost component of a high-energy physics experiment, and ATLAS is no exception; its muon systems are 20m in diameter and 26m long. Having the muon spectrometer located here is also useful for determining the sign of muons (i.e. whether they're μ^+ or μ^-) by looking at the curvature of their paths due to the toroidal field, since the sign of high-energy muons is often difficult to tell from their tracks in the inner detector as these can be almost straight, making misidentification

likely.

In the barrel, muon chambers are mounted in cylindrical layers at 4 m, 7.5 m and 10 m from the beampipe; in the endcaps, they are located in discs at 7 m, 10 m and 14 m from the interaction point. Monitored drift tubes cover most of the η range, while cathode strip chambers are located closer to the interaction point.

4.4 Detector coordinates

The ATLAS detector coordinates are based on cylindrical polar coordinates, with the z axis defined as the direction of the beams and the pseudorapidity by

$$\eta = -\log_e(\tan(\theta/2)), \tag{4.3}$$

where θ is the polar angle. Pseudorapidity is used as an approximation to true rapidity:

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

(in which E is the particle's energy and p_z its longitudinal momentum), differences in which are invariant under longitudinal Lorentz boosts[1]. This is a prerequisite for quantities to be meaningful in the analysis of data from particle collisions, although rapidity calculations require knowledge of both momentum and energy, which cannot be distinguished if the mass of the particle is unknown. However, the approximation $y \approx \eta$ is good for relativistic particles, that is, particles with a kinetic energy much greater than their invariant mass-energy.

The radial coordinate r is the perpendicular distance from the z axis, and ϕ is the azimuthal angle within this plane. In addition to η and ϕ, another quantity used to describe particle trajectories is transverse momentum, the component of momentum in the $r-\phi$ (or $x-y$) plane: $p_\perp \equiv \sqrt{p_x^2+p_y^2}$. Related to this is transverse energy, $E_\perp \equiv \sqrt{p_\perp^2+m^2}$. When Cartesian coordinates are used, the positive x-axis points towards the centre of the LHC ring, the positive y-axis points straight up and the positive z-axis points down the beam-pipe, roughly in the direction of nearby Geneva.

[1]Transformations to a reference frame moving at some constant speed in the z-axis

4.5 Trigger and online software

ATLAS has a three-level trigger [8, 9] to sift the vast quantity of raw data for the far smaller number of interesting events that will occur. The Level 1 trigger (LVL1) reduces data from the initial interaction rate of ∼40 MHz to ∼100 kHz; the Level 2 trigger (LVL2) reduces this further to around ∼1 kHz, and the data is then passed through the final level, the Event Filter (EF), to be recorded at a rate of ∼200 Hz (that is, 200 events per second).

The Level 1 trigger makes event selections based on reduced-granularity data from the calorimeters (0.1×0.1 in $\eta - \phi$ space) and data from the muon trigger chambers. This trigger searches for objects with high $p_\perp$; muons from the muon chambers, electrons, photons and jets from the calorimeters and large $E_\perp^{miss}$ inferred from these quantities. The decision whether or not to pass on data to the next level is made within 2 μs.

At Level 2, full detector granularity is used, but only in the Region of Interest (RoI) identified by LVL1. Tracks are reconstructed using information from the inner detector and muon systems. This level of decision-making takes 1-10 ms per event.

Finally, the Event Filter has access to full event data and takes around a second to make something approaching a complete event reconstruction. Events passing a menu of trigger options (cuts on the $E_\perp$ of jets, b-jets, photons and leptons, and on inferred $E_\perp^{miss}$) are then stored on disk for full offline analysis. A schematic representation of the trigger is given in figure 4.3.

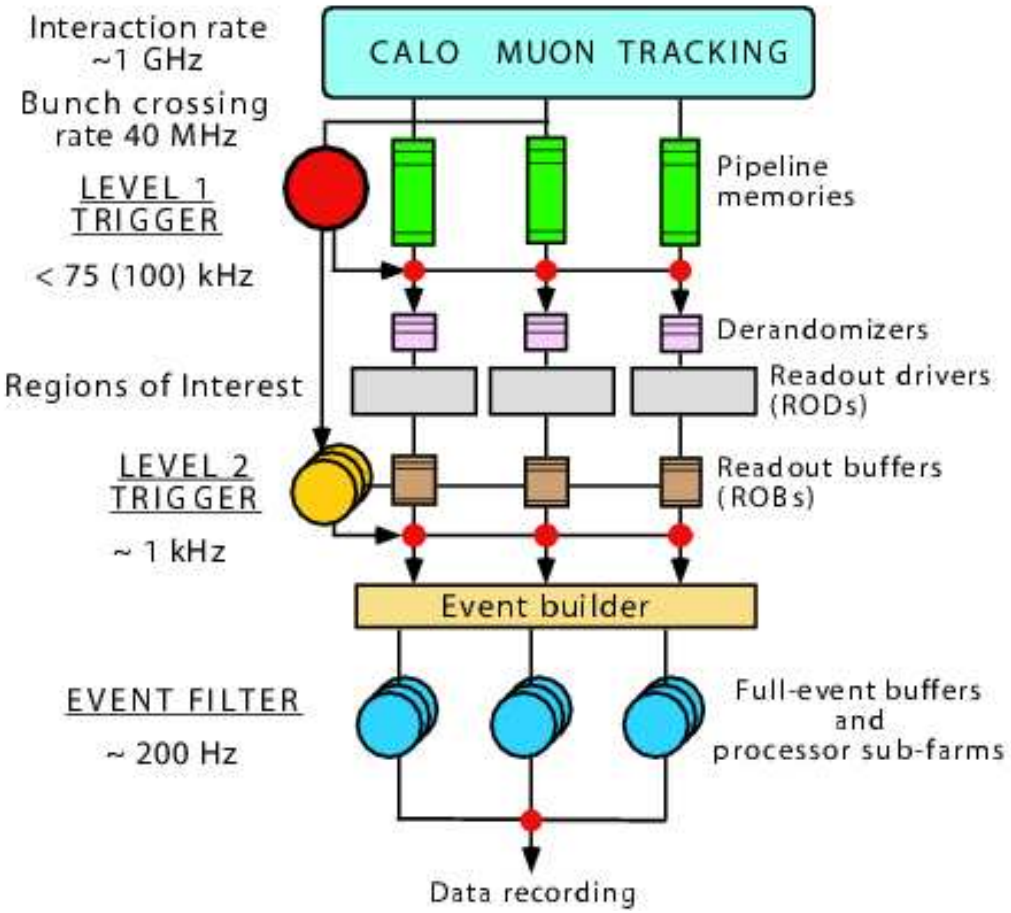

Figure 4.3: Schematic diagram of the ATLAS trigger chain.

Chapter 5

The ATLAS offline software

The offline software currently used for the generation of 'truth' data (pseudodata modelling a particle interaction), simulation of the ATLAS detector and analysis of results is written within an overarching control framework called Athena. This is based on a software architecture called Gaudi, which was originally developed by the LHCb collaboration.

The purpose of a software framework is to provide the 'skeleton' of an application into which users plug their own code; it provides all the common libraries and algorithms that most users are likely to need, to save them having to write their own code from scratch. In addition, it encourages a common approach to writing analysis software that makes it easier to understand and use code written by someone else.

There are two paths a user may take from truth data (the output of a class of programs such as `PYTHIA`, collectively known as Monte Carlo generators), to a form that can be used as the input to the analysis. Firstly, one can use the **full chain** of generation/reconstruction, which has four stages:

generation: a 'Monte Carlo' program (event generator) is used to create *truth data*, i.e. 4-vectors of final-state particles resulting from a collision; this is fed through a software

simulation of ATLAS (an extremely resource-intensive process), using the GEANT4

program [10] to simulate the passage of particles through matter; the *simulated data* then undergoes

digitisation, to produce data in the same format as that which will be read out from the onboard electronics when the machine is running; finally, the *digits* are used as input to a

reconstruction program, which infers physics objects (hadronic jets, τ jets, photons, electrons, muons, $E_{\perp}^{miss}$) from detector objects (inner detector tracks, calorimeter deposits and muon hits).

5.1 Atlfast

The other path is to use the ATLAS Fast Simulation, or Atlfast for short [11]. This is a program that effectively combines the four steps of generation, simulation, digitisation and reconstruction into one, so it has great advantages over the full chain in terms of runtime (four or five orders of magnitude faster), intermediate file storage space and user convenience. This is achieved by a process called 'smearing', whereby four-vectors generated by the Monte Carlo have random Gaussian errors convoluted with them to simulate the performance of the simulation, digitisation and reconstruction steps done separately by the full chain. The disadvantage is that the user is not able to tune the many free parameters that may be adjusted in the various stages of the full chain reconstruction; therefore, it is customary for users interested in tuning and modifying the software to use the full chain, and users who are willing to trust the officially validated Atlfast releases to use that instead. It is intended to compliment, rather than replace, the full simulation chain, which is still used to produce officially validated Computer Systems Commissioning (CSC) samples.

Atlfast simulates three calorimeters; a central region ($|\eta| < 3.2$) and two forward regions ($3.2 < |\eta| < 5$). The cells have a granularity of 0.1×0.1 (in $\eta - \phi$ space) in the central region, and 0.2×0.2 in the forward regions. The final output is in the form of an Analysis Object Data (AOD) file that 'distills' useful physics information into a format of convenient size and user accessibility.

Physics quantities that may be chosen by the user in the Atlfast job options include the processes allowed to take place (from a library of hundreds provided by `PYTHIA`) and various parameters relating to the treatment of the incoming protons (such as initial state radiation, treatment of α_s, choice of PDF). User-defined variables relating to the reconstruction process include choice of jet reconstruction algorithm, various (fairly rudimentary) acceptance cuts on kinematic variables and a choice of what kinds of information to include in the output AOD.

5.2 Analysis in Athena

Once data from the reconstruction program is available in an analysable format - most typically AOD - one can then run an Athena analysis algorithm on it. In the versions of Athena currently used, the output is a format called Athena-Aware NTuple, or AANT, which replaces the older format CBNT (ComBined physics NTuple). AANT allows the user to navigate back from objects in the analysis output to the objects they correspond to in the input AOD; furthermore, it includes data in the form of both ntuples, which may be used for further analysis either in Athena or using stand-alone code, and histograms, which may be analysed directly in ROOT, as the user requires.

5.3 Jet-finding algorithms

All physics processes that will occur at the LHC will include jets in the final state. A jet is one or more hadrons in a tightly confined beam which is treated as a single physics object from the point of view of the analysis. Just how close together the trajectories of any two hadrons have to be in order to be considered to belong to the same jet depends on certain parameters set by the user; specifically, the type of 'clustering' or 'jet-finding' algorithm used to define jets from hadronic calorimeter deposits, and the value of the parameter used to define the maximum size of a jet before it must be split up into several jets.

In this analysis the cone algorithm is used, whereby two hadrons are considered to belong to

the same jet if

$$\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq R, \tag{5.1}$$

where $\Delta\eta$ and $\Delta\phi$ are the distances between the two hadrons in $\eta - \phi$ space (see section 4.4) and R is the cone size parameter, which in this analysis is assigned the value 0.9 (typical sizes may range from 0.1 to 1.0 or more). The algorithm used in this analysis is a *seeded* cone algorithm, in which high-$E_\perp$ particles in the event are used as 'seeds' around which putative cones are built; lower-$E_\perp$ particles and jets are then merged with the seed if they fall within the cone. (In alternative *seedless* schemes, either all calorimeter cells are used to define cones or points on a pre-defined $\eta - \phi$ grid are used as seeds.) An iterative procedure is followed until all jet candidates have either been made into jets in their own right or absorbed into neighbouring jets. Jets are assumed to be massless, since even high-mass particles have an invariant mass negligible (to within detector resolution) next to their kinetic energy.

Chapter 6

Experimental uncertainties

This analysis was performed using Atlfast and the Athena analysis framework in release 11.0.5, with the detailed dijet analysis being performed separately with stand-alone code. In generating the data, only left-left spin quark doublets have been considered, with the interference term between the contact interaction and the SM Lagrangian assumed to be positive (out of a choice of +1 and -1), implying destructive interference [19]. The `PYTHIA` routines used to generate events with quark compositeness are based on [32] and [33]. The generic contact interactions generated for this analysis have a Lagrangian of the following form:[1]

$$\mathcal{L}_{comp} = \frac{g_0^2}{\Lambda^2_{\text{comp}}}\{\eta_{LL}(\overline{q}_L\gamma^\mu q_L)(\overline{f}_L\gamma_\mu f_L)\} \tag{6.1}$$

where

$$q_L = \begin{bmatrix} u \\ d \end{bmatrix}_L, \tag{6.2}$$

with u and d signifying up-type and down-type quarks respectively, and f and $\overline{f}$ being any fermion-antifermion pair (assumed to be quarks in this analysis, since we are concerned with hadronic jets). This Lagrangian may represent a number of distinct physical processes:

[1] It should be noted that this is based on the Lagrangian for effective contact interactions theories, and is not derived from deep symmetry arguments as are $\mathcal{L}_{\text{QED}}$ and $\mathcal{L}_{\text{QCD}}$.

- an s-channel annihilation process;
- a $q\overline{q} \rightarrow q\overline{q}$ t-channel momentum exchange;
- by reversing the direction of one quark line in the t-channel scatter, one achieves a $qq \rightarrow qq$ or $\overline{qq} \rightarrow \overline{qq}$ scatter and by exchanging the two outgoing quarks one achieves a
- u-channel exchange.

Quantum interference takes place, as usual, between any two (or more) diagrams with the same initial and final state particles; in this study, interference between Standard Model and sub-quark processes may occur with diagrams that have qq, $q\overline{q}$ or $\overline{qq}$ in the initial and final states, since gluons are assumed to be elementary.

Multiple files of 1,000 events were generated and reconstructed in six non-overlapping $p_\perp$ regions, with the $p_\perp^{min}$ and $p_\perp^{max}$ for each region defined by the minimum and maximum $p_\perp$ of the two outgoing partons at the generator level (of the order of 100,000 events per $p_\perp$ region). This is to ensure sufficient statistics at higher $p_\perp$, as the cross-section falls by roughly an order of magnitude with each doubling of $p_\perp$. These regions are known as 'J-regions', following the prescription used in the preparation of AODs for the CSC study, in which the cross-section distributions may then added together with each region weighted by its total cross-section to produce a continuous spectrum. Region J0 corresponds to events in which the partons exiting the hard scatter have a $p_\perp$ between 8 GeV and 17 GeV; in J1 from 17 GeV to 35 GeV, and from then on a doubling with each successive J-region. An example of such a $p_\perp$ spectrum is given in figure 6.1.

Another important quantity in jet analyses is the invariant mass of the final dijet system; this is the combined invariant mass of the two hardest (highest $p_\perp$) jets in the final state. It is calculated as follows:

$$m_{jj} = \sqrt{(\mathbf{P}_1^\mu + \mathbf{P}_2^\mu)(\mathbf{P}_{1\mu} + \mathbf{P}_{2\mu})} = \sqrt{(E_1 + E_2)^2 - (p_{1x} + p_{2x})^2 - (p_{1y} + p_{2y})^2 - (p_{1z} + p_{2z})^2} \tag{6.3}$$

in which $\mathbf{P}_i^\mu$ is the four-momentum vector of the ith jet ($i = 1$ or 2), with energy E_i and momentum components p_{ix} etc. A dijet mass distribution is given in figure 6.2.

Following from the discussion in section 2, it seems that a good way to search for signatures

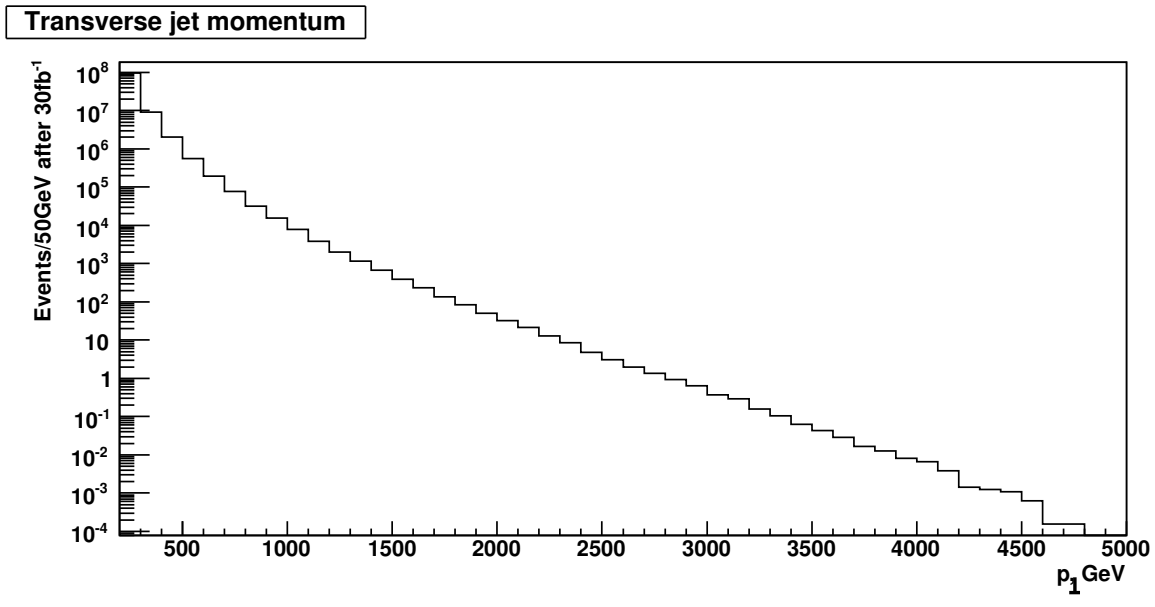

Figure 6.1: The differential transverse momentum spectrum for Standard Model QCD dijet events after 30fb^{-1} integrated luminosity.

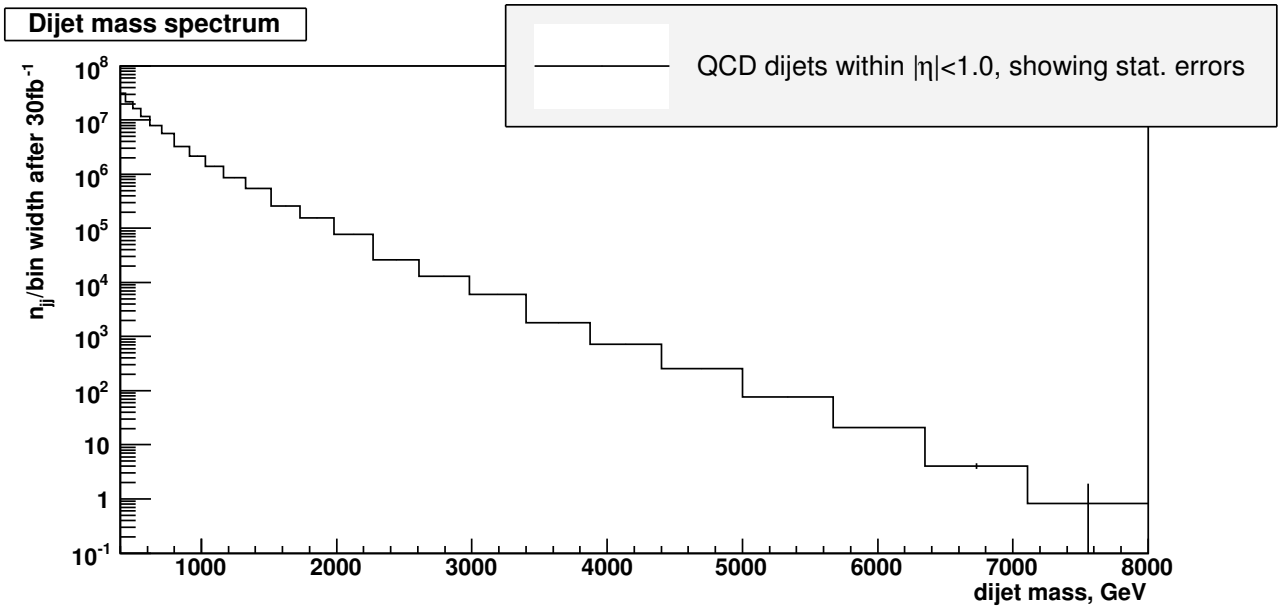

Figure 6.2: The mass spectrum for Standard Model QCD dijet events, showing statistical uncertainty after 30fb^{-1} integrated luminosity.

of compositeness may be to look at the distribution of hadronic jets in η; hadronic jets, because these are the main products of partons involved in hard collisions (whether interacting by Standard Model QCD processes or new physics in a contact interaction), and the distribution in η because this is the kinematic quantity sensitive to contact interactions due to compositeness, as explained qualitatively in 2.3.

Specifically, the kind of jet events used in this analysis are *dijets*, which is to say, the pair of jets having the highest and second-highest $p_\perp$ in the η range under consideration. The dijet invariant mass is then calculated according to the formula above. Dijets are used because the hard process is a $2 \to 2$ process, so the two hardest jets will come from the two outgoing partons involved in this process; by using only the 'leading' (i.e. highest-$p_\perp$) jet, useful information is thrown away, while using more than two jets would include semi-hard QCD radiation that is not of interest to this analysis and so 'muddy the waters'.

Figure 6.3 (upper plot) shows this distribution for two η regions: $|\eta| < 0.5$ and $0.5 < |\eta| < 1.0$, for three compositeness models and standard-model QCD. Figure 6.3 (lower plot) shows the dependence on m_{jj} of the ratio:

$$R = N(|\eta| < 0.5)/N(0.5 < |\eta| < 1.0), \tag{6.4}$$

in which N is the number of dijet events with the two leading jets in the same η acceptance region. The basic cuts carried out at the start of the analysis are $p_\perp > 100$GeV and $|\eta| < 1.0$ for all jets. A ratio is used as a discriminant between the different physics scenarios as certain experimental sources of systematic uncertainty, such as imperfect calorimeter calibration, will to some extent cancel out when one dijet rate is divided by another.

The three compositeness models mentioned above differ in the value of the characteristic energy scale of the preons, Λ_{comp}, which may be thought of as their 'binding energy'. Parton interactions at CoM energies greater than Λ_{comp} may probe quark substructure directly, just as interactions between hadrons at energies greater than Λ_{QCD} will probe hadron substructure. However, quark compositeness can still have measurable effects even if the interaction energy is below Λ_{comp}, due

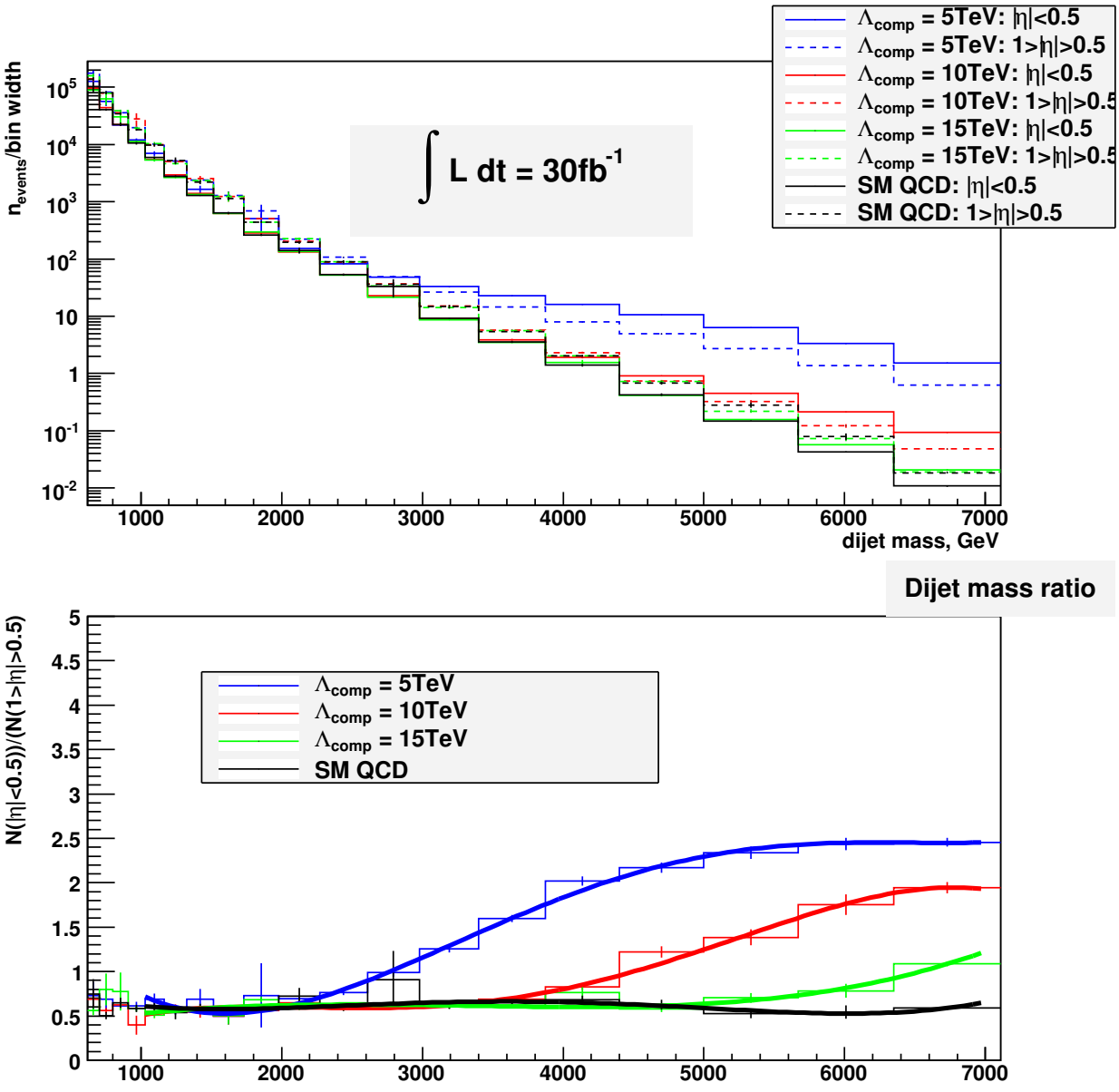

Figure 6.3: The cross-section and dijet ratio as a function of dijet invariant mass.

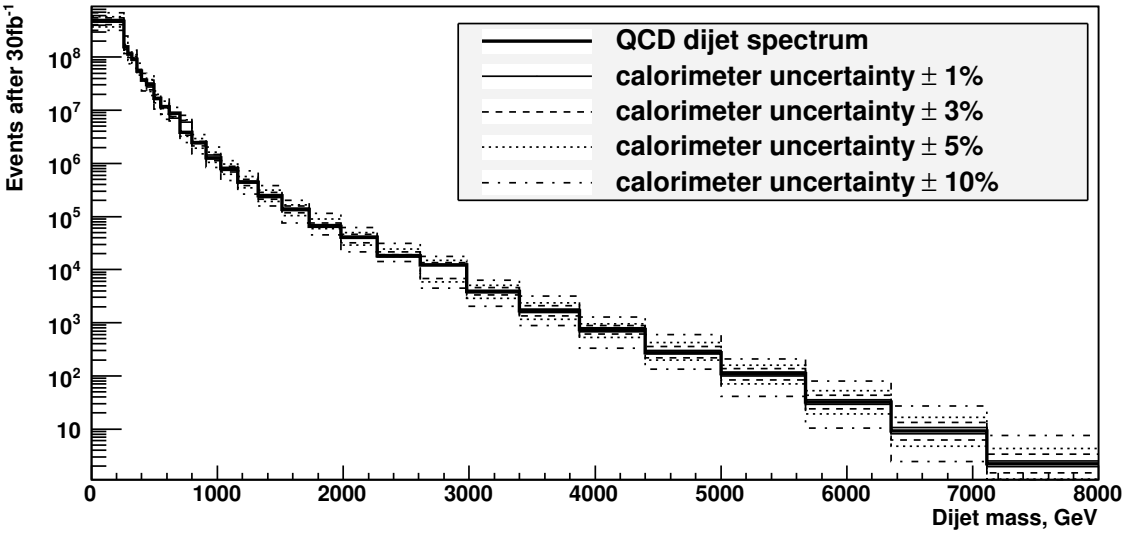

Figure 6.4: The dijet rate vs. dijet mass, showing the effect of four margins of uncertainty on the ATLAS calorimeter response.

to virtual propagator effects[2]. This may be seen in figure 6.3, where the dijet ratio R for $\Lambda_{\text{comp}} = 10$ TeV begins to be distinguishable from that for SM QCD at a dijet mass of around 4 TeV, less than half the compositeness energy scale.

Naturally, the higher value of Λ_{comp}, the more tightly bound the preons and the more difficult their experimental detection; the 5 TeV model has the greatest deviation from the expected SM result, while the 15 TeV model is indistinguishable from it up to dijet masses in excess of 5 TeV. The SM QCD scenario is equivalent to a compositeness model with an infinitely high energy scale. The finite resolution of ATLAS's hadronic calorimeters and tracking will lead to uncertainty in the mass of jets (which is derived from both energy and momentum measurements), and this will in turn lead to uncertainties in the number of jets passing kinematic cuts used in this analysis.

Two kinds of calorimeter uncertainty have been considered:

[2]This is due to interference between the SM gauge coupling and the contact interaction, which (to leading order in $\hat{s}/\Lambda^2_{\text{comp}}$) contributes an amplitude of $\hat{s}/\alpha_s\Lambda^2_{\text{comp}}$: see [20].

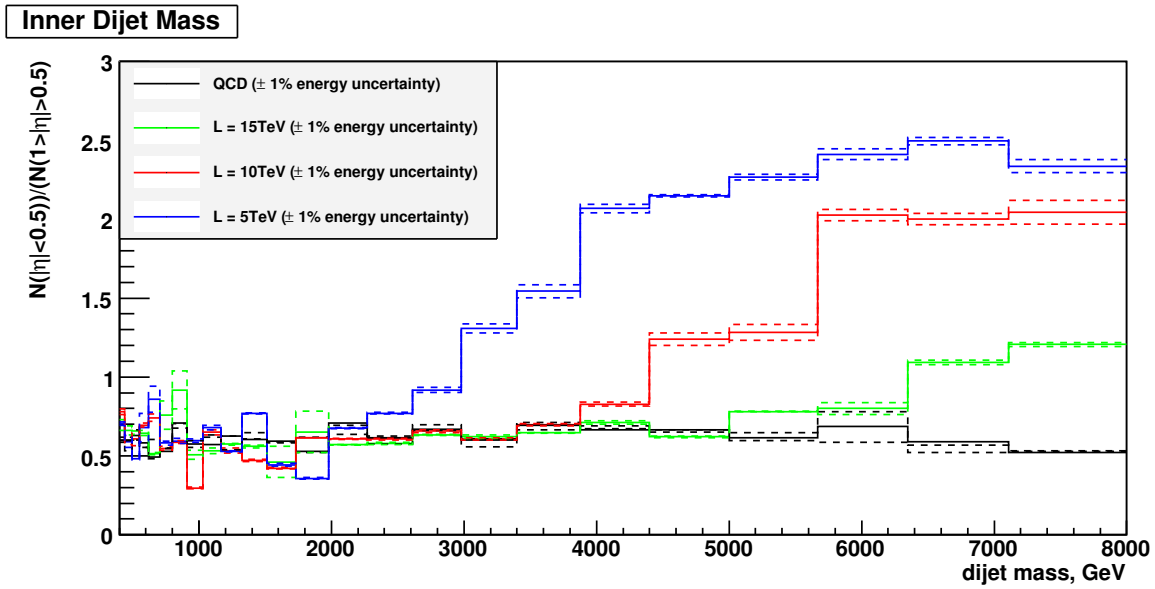

Figure 6.5: The effect of a 1% jet energy scale miscalibration, compared to the deviation from SM QCD for the three compositeness models.

6.1 Absolute energy uncertainty

Absolute uncertainty is assumed to be independent of jet energy for the sake of simplicity; in reality, there will be a dependence (see 4.2). It should be stressed that the uncertainty under consideration here is the *jet energy scale*, i.e. a consistent miscalibration of the detector response, so that energy readings are consistently above or below the true energy of the hadrons impacting on it. This is not to be confused with the detector resolution described in section 4.2.2. Four values of $\Delta E/E$ have been chosen: 1%, 3%, 5% and 10%. Figure 6.4 shows the effect of these uncertainties on the dijet event rate for the inner pseudorapidity region; figures 6.5 to 6.8 show the effect of these four uncertainty values on both the SM QCD dijet ratio and the ratio for the three compositeness models; it can be seen that larger uncertainties effectively push up the minimum dijet mass for which the four scenarios may be distinguished, especially the $\Lambda_{comp} = 15$TeV and SM QCD scenarios.

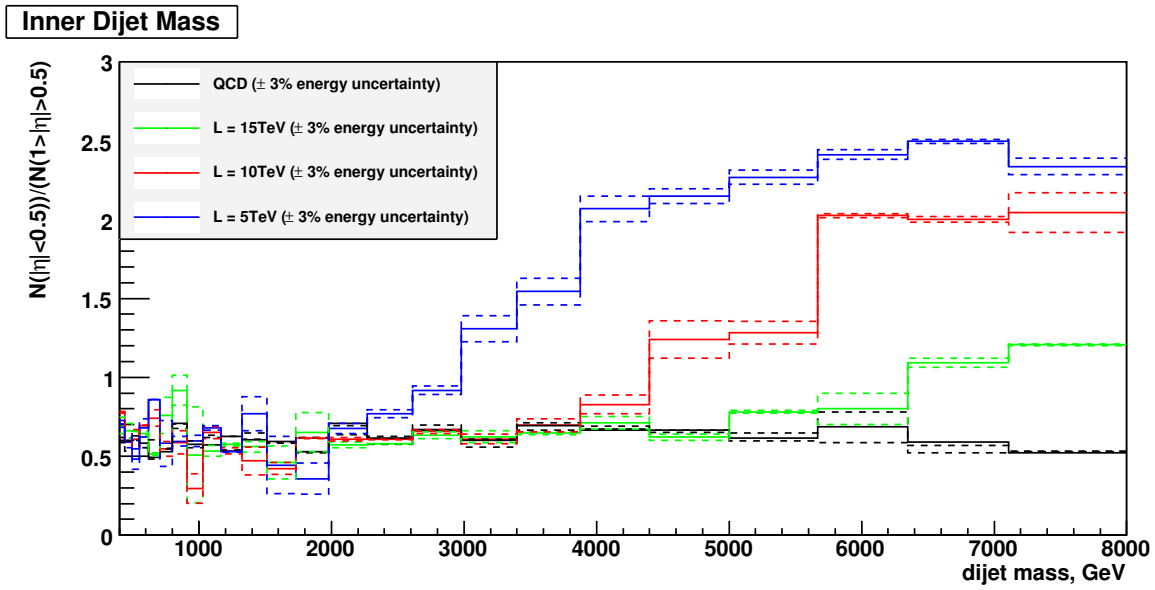

Figure 6.6: The effect of a 3% jet energy scale miscalibration, compared to the deviation from SM QCD for the three compositeness models.

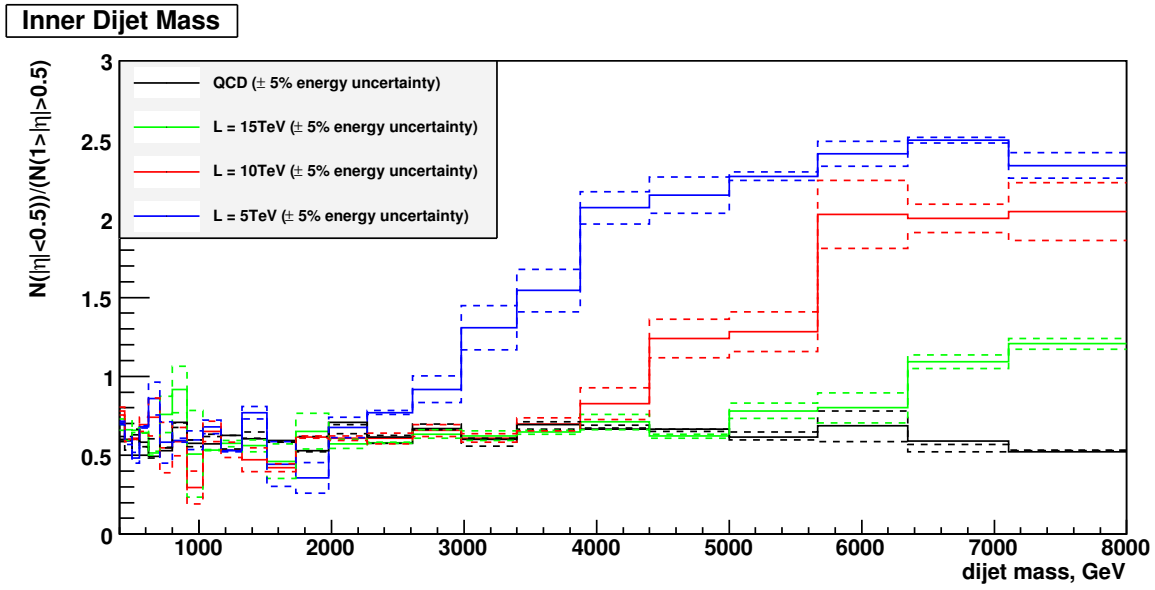

Figure 6.7: The effect of a 5% jet energy scale miscalibration, compared to the deviation from SM QCD for the three compositeness models.

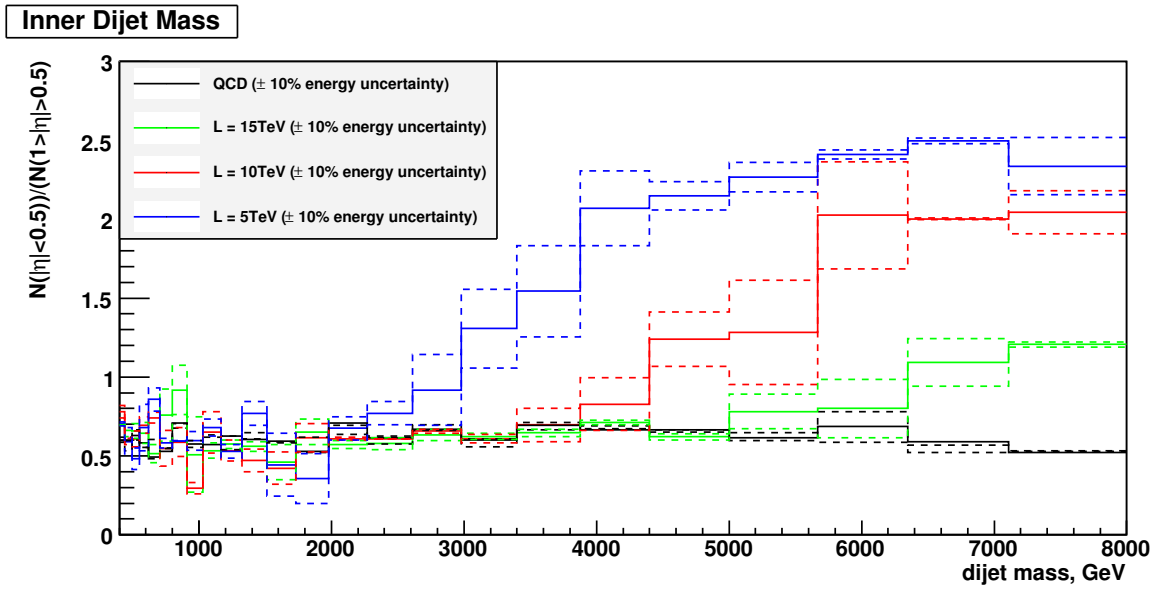

Figure 6.8: The effect of a 10% jet energy scale miscalibration, compared to the deviation from SM QCD for the three compositeness models.

6.2 Relative energy uncertainty between $|\eta|$ regions

It is possible that in addition to an overall, absolute uncertainty in energy and momentum, there will be a relative difference in detector response in the two $|\eta|$ regions considered in this analysis. Figures 6.9 and 6.10 show the effect of a 0.5% and 1.0% difference, respectively, between the inner ($|\eta| < 0.5$) and outer ($0.5 < |\eta| < 1.0$) detector regions used in this analysis. This is taken into account because it will obviously have a direct impact on the ratio of dijets accepted in the two η regions, unlike the absolute jet energy scale uncertainty where a degree of cancellation may be hoped for.

6.3 Energy resolution

A quantity of interest in an analysis of high-energy jets is the resolution of the reconstructed jet energy:

$$R_E = \frac{E_{truth} - E_{reco}}{E_{truth}}, \tag{6.5}$$

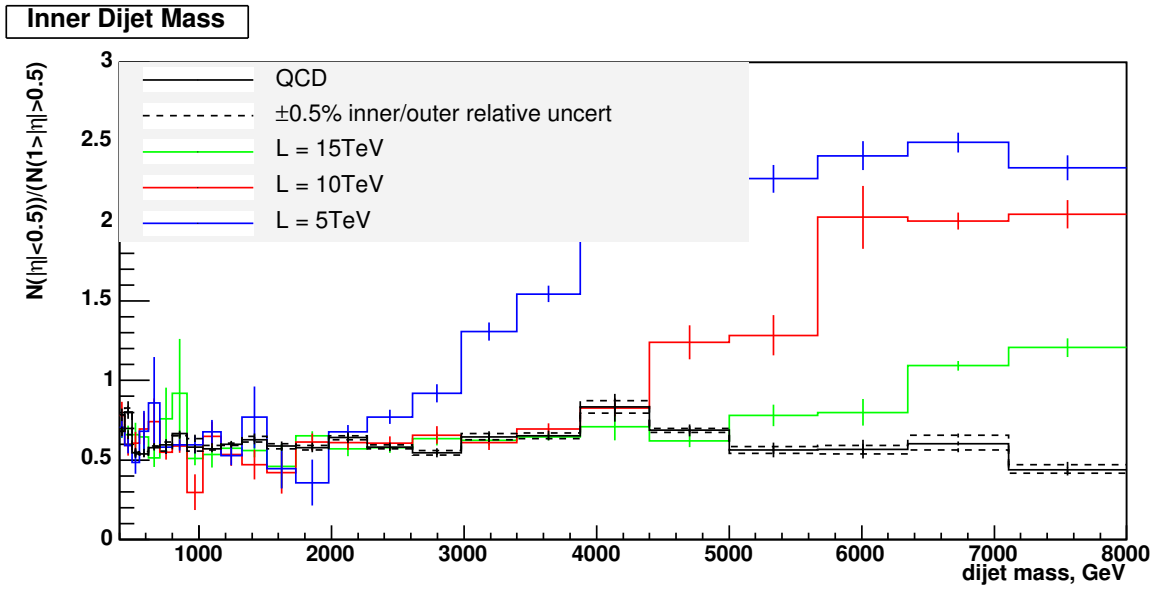

Figure 6.9: The effect of a 0.5% relative uncertainty in jet energy and jet momentumbetween $|\eta| < 0.5$ and $0.5 < |\eta| < 1.0$, compared to the deviation from SM QCD for the three compositeness models.

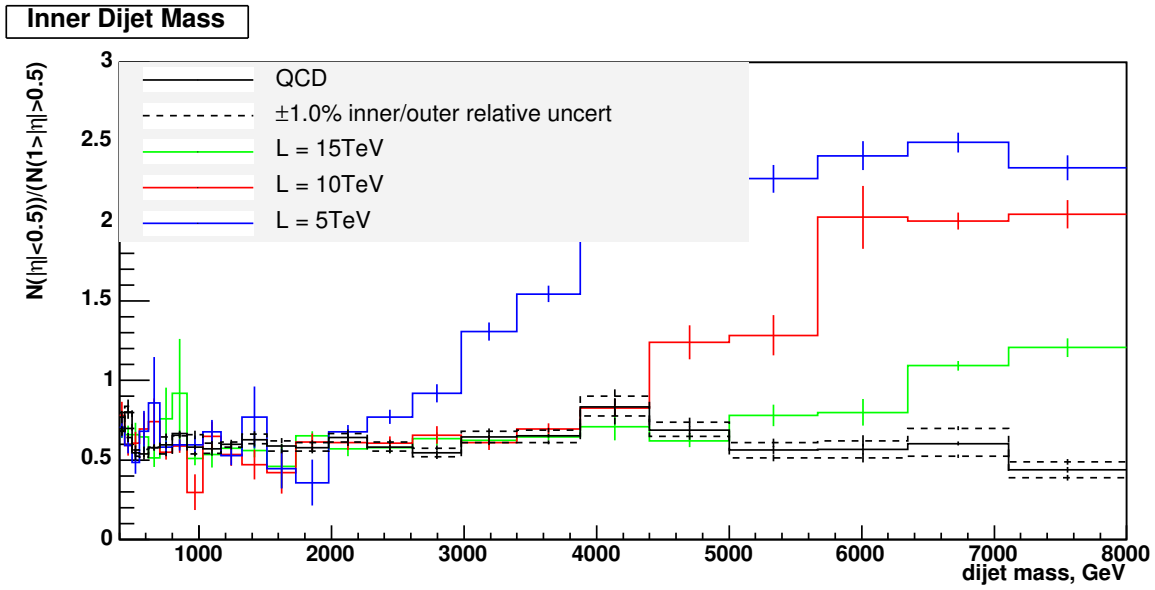

Figure 6.10: The effect of a 1.0% relative uncertainty in jet energy and jet momentumbetween $|\eta| < 0.5$ and $0.5 < |\eta| < 1.0$, compared to the deviation from SM QCD for the three compositeness models.

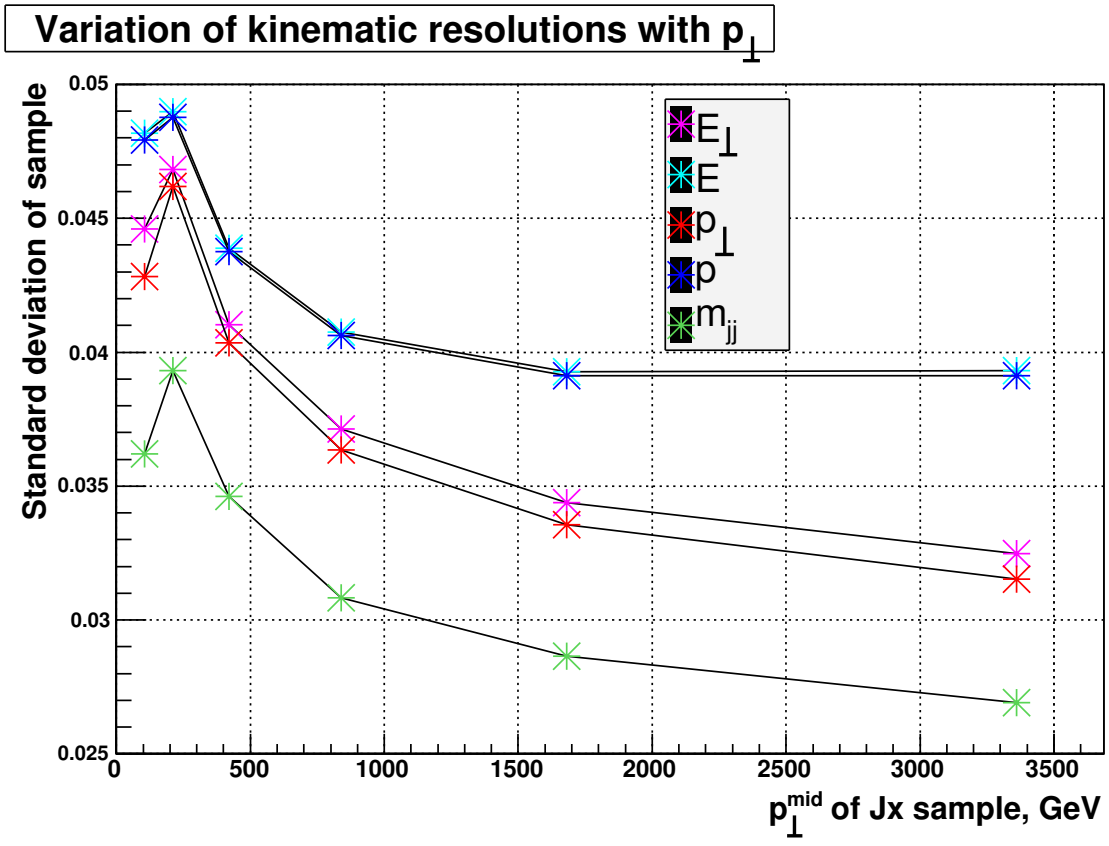

Figure 6.11: Fractional resolution (see text) of various basic kinematic quantities of interest to a jet analysis, plus dijet mass (a derived quantity), as a function of $p_\perp$.

in which E_{truth} is the true energy, accessed from the output of the Monte Carlo, and E_{reco} is the energy reconstructed by Atlfast. This impacts on the resolution of the dijet mass, which is defined similarly:

$$R_{M^{jj}} = \frac{M^{jj}_{truth} - M^{jj}_{reco}}{M^{jj}_{truth}}. \tag{6.6}$$

These resolutions are plotted, as a function of dijet mass, in figure 6.11. The abscissa of each datum is the middle of the 'J-region' used. For the dijet mass spectrum plots, the data have been rebinned so that each dijet mass bin is equal to twice the dijet mass resolution in that $p_\perp$ region multiplied by the mass of the lower bin edge:

$$M^{jj}_{n+1} - M^{jj}_{n} = 2M^{jj}_{n}\sigma(M^{jj}_{n}). \tag{6.7}$$

Chapter 7

Theoretical uncertainties

In addition to the experimental uncertainties discussed above, there also exist *theoretical uncertainties*, whereby uncertainties are propagated to the final computed quantities from uncertainties in the underlying theory. Two sources of theoretical error are considered here; uncertainties in the parton density functions used by the Monte Carlo, and uncertainties in two arbitrary energy parameters, called the *renormalisation scale* and the *factorisation scale*.

Theoretical uncertainties on LHC jet cross-section measurements have been calculated using two complimentary techniques. Firstly, a program performing QCD calculations to next-to-leading order (NLO) precision was used to generate parton-level histograms of kinematic quantities of interest; this technique was used to investigate the effects of uncertainty in two theoretical energy scales and also in a set of PDFs which were constructed using jet data to constrain the proton's gluon content (which is the largest source of PDF uncertainty in proton-proton collisions). Then parton-level information from the 'Monte Carlo' program `PYTHIA` was used to generate tables of weights using a different PDF set.

The first technique has the advantage that NLO calculations are more precise than those made only to leading order (LO), as they are in `PYTHIA`; however, the program used for these calculations only generates information at parton level, with no hadronization and no detector description. The second technique, although less theoretically precise, is more powerful since hadronization and soft

radiation are simulated and the output data may be passed through the ATLAS event simulation and reconstruction chain, allowing the resolution of the detector to be taken into account in addition to the theoretical uncertainty.

7.1 Parton Density Functions

This part of the analysis is concerned with predictions for jet cross-sections at the LHC with analysis on the uncertainties due to the PDF set used as a theoretical input to the calculation and also uncertainties in the renormalisation and factorisation scales. The program used to generate this data is stand-alone so no detector description is attempted.

7.1.1 The ZEUS-JETS PDF set

The ZEUS experiment at HERA is an ideal machine for probing the internal structure of the proton over a wide range in x (the fraction of the proton's momentum carried by a given parton in the laboratory rest frame) and Q^2 (the negative squared four-momentum, or virtuality, of the exchanged boson). A number of PDF sets have been published by the ZEUS collaboration, the most recent of these being the ZEUS-JETS 2005 set [3]. In previous ZEUS PDF sets, data from ZEUS were combined with those from other experiments; however, ZEUS-JETS is based solely on ZEUS data, in particular:

- neutral and charged current e^+p and e^-p DIS (deep inelastic scattering) inclusive cross-sections;
- neutral current e^+p inclusive jet cross-sections;
- direct photoproduction dijet cross-sections.

These data are then used as input to a next-to-leading order QCD DGLAP analysis to produce the ZEUS-JETS PDF set [3]. These PDFs are calculated in the general mass (Thorne-Roberts) variable flavour number, or TRVFN, scheme [29].

The two types of jet cross-section are especially useful in constraining the PDF of the gluon, particularly in the mid- to high-x region ($x \approx 0.01$ - 0.5). This region of the gluon PDF is the dominant source of uncertainty in jet cross-sections in high-Q^2 pp collisions, so constraining gluons in this region has a large impact on the overall cross-section uncertainty. ZEUS-JETS makes a large improvement in this area on previous PDFs because the inclusion of jet data makes the analysis very sensitive to the behaviour of $\alpha_s(Q^2)$ over a large range in Q^2, on which the gluon density strongly depends. (The same impact is not made on the quark distributions, as there are other complicating factors here, such as non-zero quark masses and fermionic statistics.)

The distribution of the outgoing jets is determined in part by the x-values of the incoming partons. In particular, the total energy of the jets will be equal to $\hat{s} = x_1 x_2 s$ (see section 3.1) as the partons may all be considered massless in collisions in this energy regime. (Even the b quark may be considered massless, to a rough approximation, for partons with x greater than around 10^{-3} as its invariant mass of $\sim$4 GeV is less than the $\sim$7 GeV such a parton would carry in kinetic energy.)

Consider first a simple dijet configuration in the final state. The two jets will always be back-to-back in a centre-of-mass frame; in terms of detector coordinates, they will be back-to-back in the $r - \phi$ plane in the laboratory rest frame, but only back-to-back in the $r - \eta$ plane if boosted in η to some suitable frame of reference. If $x_1 \gg x_2$, both jets may have high $|\eta|$ with η of the same sign; events of this type are naturally sensitive to high x and low x partons, and may be used to probe the proton PDF once the LHC begins taking data. The situation is essentially similar in events with three jets, with the jets balanced in momentum once boosted in η to the system's centre of mass.

Many of the principle physics programmes to be pursued at the LHC will involve the search for high-mass resonances; these will occur in a collision of two high-x partons interacting by an s-channel annihilation or a high-Q^2 boson exchange. This may then result in high-energy jets given off in a range of η values, with the possibility of jets with high $p_\perp$, i.e. large $|\vec{p}|$ and small $|\eta|$. Jets with large $|\vec{p}|$ and large $|\eta|$ will have relatively low $p_\perp$ and may represent partons involved

in a glancing collision, with little momentum exchanged between them. If both incoming partons have low x the outgoing jets will invariably have low $p_\perp$ and will not be of use for searching for high-mass new physics; however, such events will be of use to those studying low-x physics, which forms an important subfield of QCD phenomenology in itself.

Figure 7.1 shows proton PDFs, in a range of Q^2 values, generated by the author using QCDNUM with the ZEUS-JETS 2005 parameters as input. The u and d valence quarks and gluons only are shown. Note that the gluon distribution is scaled down by a factor of 20. The valence quarks exhibit *Björken scaling*, whereby they remain relatively constant as Q^2 increases, while the gluon starts to dominate the proton's structure at successively higher values of x. The u quark PDF is twice as big as that of the d quark for the obvious reason that the proton contains two u quarks and one d quark, and both PDFs may be seen to peak at an x close to $\frac{1}{3}$, although the gradient is actually zero at an x value somewhat below $\frac{1}{3}$ due to effects of the gluonic field on the dynamics of the valence quarks.

The 16 parameters of the fit (see Appendix B) are reduced, by certain assumptions and constraints (such as the number and momentum sum rules), to 11 *free* parameters, from which 11 orthogonal combinations are made. These combinations form the set of eigenvectors of the Hessian uncertainty matrix, which encodes the parameters and their uncertainties. The ZEUS-JETS PDF set therefore comprises 23 member PDFs: 11 pairs representing excursions up and down each of the 11 eigenvectors, plus a central value PDF representing a best fit of the other 22, obtained by a χ^2 minimisation. These PDFs can each be used as input to a Monte Carlo program to produce 23 sets of cross-sections; a value for the upper and lower limits on the total uncertainty (on the cross-section in each histogram bin, n) can then be calculated:

$$\sigma^-_{tot}(n) = \sigma_0(n) - \Big(\sum_{i=1}^{11}\big(\sigma^0(n) - \sigma^-_i(n)\big)^2\Big)^{1/2} \tag{7.1}$$

$$\sigma^+_{tot}(n) = \sigma_0(n) + \Big(\sum_{i=1}^{11}\big(\sigma^0(n) - \sigma^+_i(n)\big)^2\Big)^{1/2} \tag{7.2}$$

where $\sigma^-_{tot}(n)$ and $\sigma^+_{tot}(n)$ are lower and upper total uncertainties on cross-section central values $\sigma^0(n)$, and $\sigma^-_i(n)$ and $\sigma^+_i(n)$ are the minimal and maximal cross-sections due to eigenvector i.

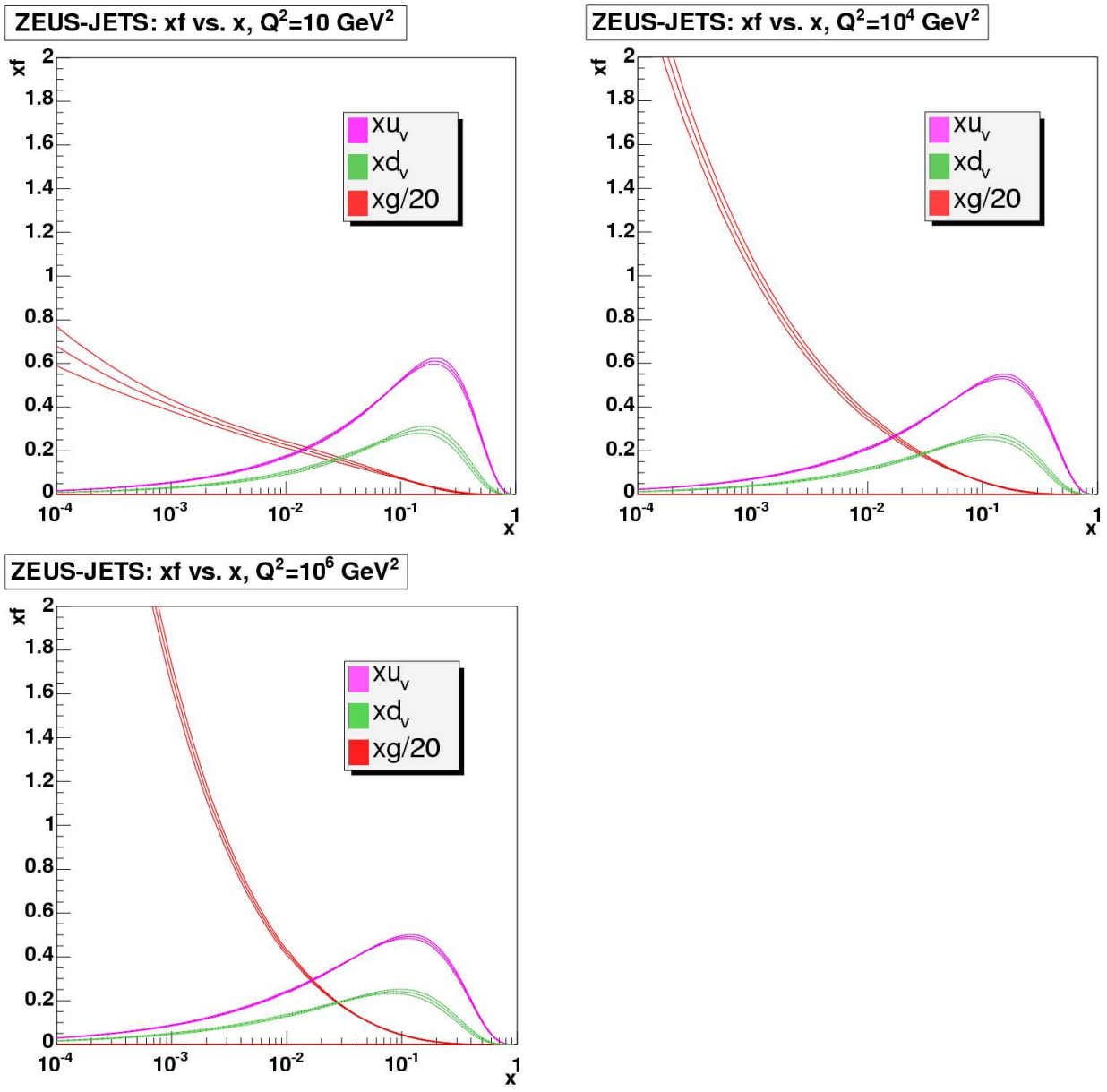

Figure 7.1: The ZEUS-JETS PDFs for u and d valence quarks and gluons over a range of Q^2.

7.1.2 Frixione's and Ridolfi's jet production package

The program used to generate the jet cross-sections was developed by Stefano Frixione and Giovani Ridolfi. It calculates QCD matrix elements at leading order (LO) and next to leading order (NLO) and then convolutes these with the user's chosen PDF to generate an event with a given weight. An inbuilt histogramming package then sums these weights to produce a histogram of cross-sections.

From a theoretical point of view, this represents an increase in the accuracy of the physical calculations compared to a leading order Monte Carlo program such as `PYTHIA`. While a lack of detector simulation makes it impossible to use such a program to estimate experimental uncertainties, it can be run very quickly and is useful for investigating effects of theoretical parameters, such as PDFs and calculational energy scales. These may then be used in more thorough analyses that include simulations of detector effects on the data.

The prototypical events calculated by the program up to order α_s^3 are:

$$
\begin{aligned}
q_i &+ q_j \rightarrow q_i + q_j; \quad i \neq j, \\
q_i &+ q_i \rightarrow q_i + q_i, \\
q_i &+ \overline{q}_i \rightarrow g + g, \\
g &+ g \rightarrow g + g
\end{aligned}
$$

for 2→2 processes, and

$$
\begin{aligned}
q_i &+ q_j \rightarrow q_i + q_j + g; \quad i \neq j, \\
q_i &+ q_i \rightarrow q_i + q_i + g, \\
q_i &+ \overline{q}_i \rightarrow g + g + g, \\
g &+ g \rightarrow g + g + g
\end{aligned}
$$

for 2→3 processes.

Here i and j label quark flavours. All other matrix elements for parton-parton scattering processes in $O(\alpha_s^3)$ can be obtained from the above by time reversal and crossing [22]; e.g. the diagram for

$$q_i + g \rightarrow q_i + g$$

is just the diagram for

$$q_i \quad + \quad \overline{q}_i \quad \rightarrow \quad g \quad + \quad g$$

with space and time axes exchanged.

7.1.3 The Les Houches Accord PDF interface

This is a collection of `FORTRAN` routines and data files that provides an interface between PDFs and the user's Monte Carlo or other program of choice. Once interfaced to the user's code, the desired PDF can be selected using an input parameter; then when the user compiles and runs the code for the first time, LHAPDF [23] uses the program QCDNUM [24, 25] to solve the DGLAP equations [26, 27, 28] using the relevant PDF set's parameters as input. The results from this calculation are stored in high-density grids in x and Q^2 (the standard Björken variables: see below) to generate parton densities that can be used by the Monte Carlo. The grid is saved so subsequent runs of the program don't have to recalculate it, as generating the grid is highly processor-intensive.

7.1.4 Contributions to the uncertainty due to PDF eigenvectors

The availability of the ZEUS-JETS PDF set in the form of separate eigenvector excursions allows the user to calculate a pair of cross-sections for each eigenvector - giving upper and lower uncertainty limits - and therefore look at the contribution from each one to the overall uncertainty. Although there is not a one-to-one correlation between eigenvectors and PDF parameters, the components of the former are given in terms of the latter in a table in the relevant LHAPDF data file, which can be viewed as plain text.

This analysis is based on differential cross-sections; $d\sigma/dp_\perp$ and $d\sigma/d\eta$. Since this analysis concerns highly relativistic jets, $p_\perp$ (the transverse momentum of the hardest jet) may be considered fully equivalent to $E_\perp$. Partons are then clustered using an implementation of the popular cone algorithm, with the cone size parameter R set to 1.0. This algorithm serves the purpose of defining the relevant phase-space used in calculating a (physical) hadronic-level cross-section from the (unphysical) partonic-level cross-section and PDF; here, the implementation of the algorithm

is used to decide whether two partons close together in $\eta - \phi$ space should be merged into one jet.

Figure 7.2 shows the $d\sigma/dp_\perp$ spectrum with uncertainties due to total PDF uncertainty compared to uncertainty due to eigenvectors 3, 8 and 11 only, which are the main eigenvectors encoding the gluon. Figures 7.4 to 7.7 are plots showing a central value differential cross-section, based on the central value PDF (number 0 in the ZEUS-JETS numbering scheme), and relative uncertainties for each eigenvector. These are simply the difference between the upper and lower differential cross-sections and the central value one, normalised with respect to the central value:

$$E_r^+ = \left(\frac{d\sigma^+}{dX} - \frac{d\sigma^0}{dX}\right) / \frac{d\sigma^0}{dX} \tag{7.3}$$

$$E_r^- = \left(\frac{d\sigma^-}{dX} - \frac{d\sigma^0}{dX}\right) / \frac{d\sigma^0}{dX} \tag{7.4}$$

where $E_r^\pm$ is the (positive or negative) relative uncertainty and X stands for either $p_\perp$ or η.

The cross-section $d\sigma/d\eta$, which is integrated over $p_\perp$, is dominated by jets with lower $p_\perp$ which have a higher cross-section. Therefore the statistical precision is higher than for the $d\sigma/dp_\perp$ cross-sections, especially at high $p_\perp$, which are shown in figures 7.6 and 7.7. Note that in both sets of plots the upper and lower uncertainty limits are not symmetrical about the central value for every eigenvector; for example, eigenvector 10, which encodes mainly the mid-x d_v quark, gives rise to an upper differential cross-section bound which is lower than the central value for both $d\sigma/d\eta$ and $d\sigma/dp_\perp$ spectra. However, this is a known feature of these error sets [30].

The results indicate that some eigenvectors contribute far more to the overall uncertainty than others. In particular, eigenvectors 8 and 11 make the dominant contribution to uncertainty on the $d\sigma/d\eta$ cross-section, as they parameterise primarily the behaviour of the gluon in the high- and mid-x regions respectively, and $g + g \to X$ and $g + q \to X$ diagrams dominate the cross-section for QCD processes in the LHC energy regime. This can clearly be seen in figures 7.2 and 7.3.

Some qualitative comments may be made regarding the distribution of uncertainty in $p_\perp$ and η by eigenvector. Firstly, it is apparent that eigenvectors 1 and 4, which encode primarily the low-x sea and overall sea normalisation respectively, make only very small contributions to the overall uncertainty. This is to be expected, as the sea distribution is strongly constrained by the momentum-sum rule, and has very small fractional uncertainty at low x [30]. In addition, heavy

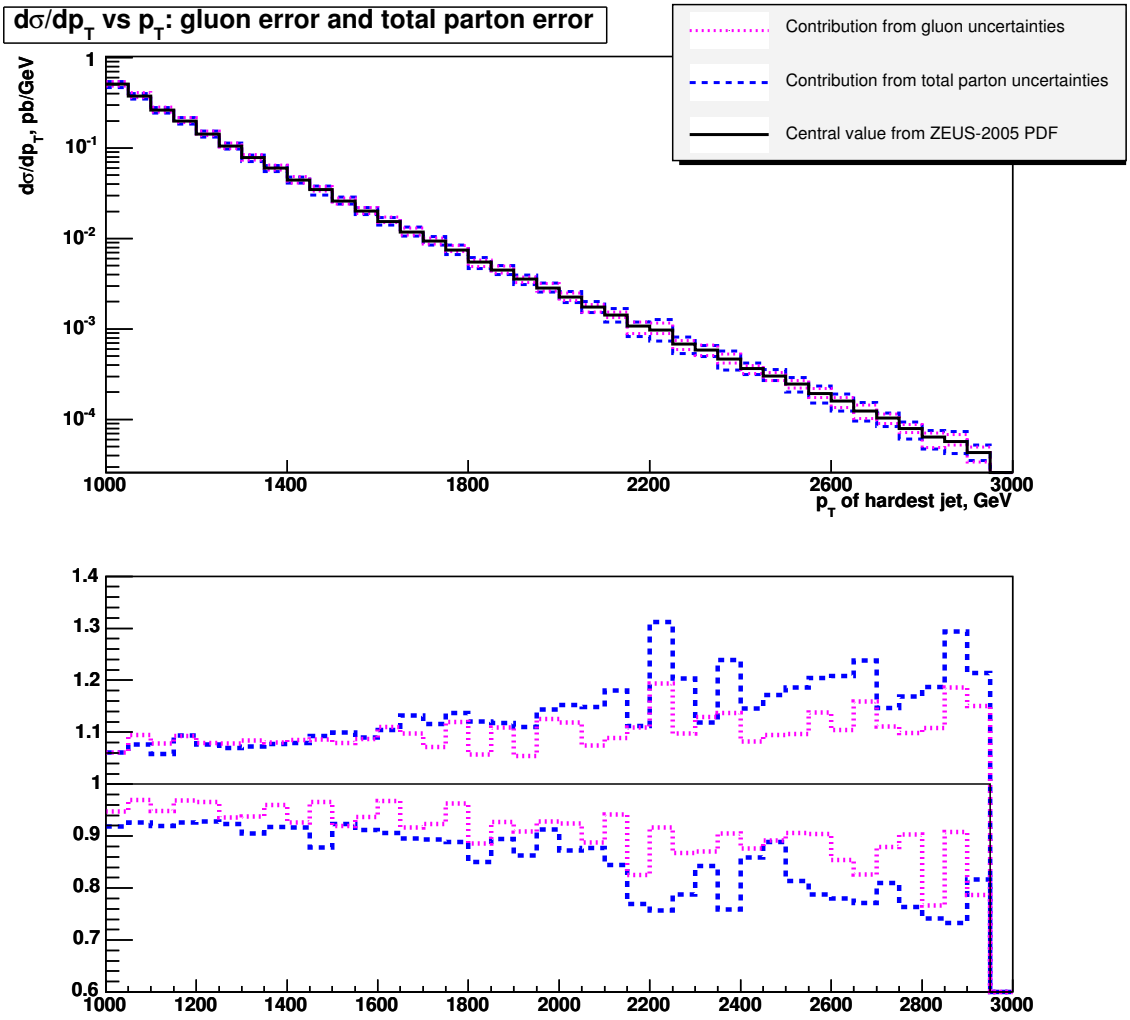

Figure 7.2: Differential cross-section of the hardest jet against $p_\perp$ of the hardest jet (top) and fractional uncertainty (bottom), showing contributions from gluonic and total partonic uncertainties.

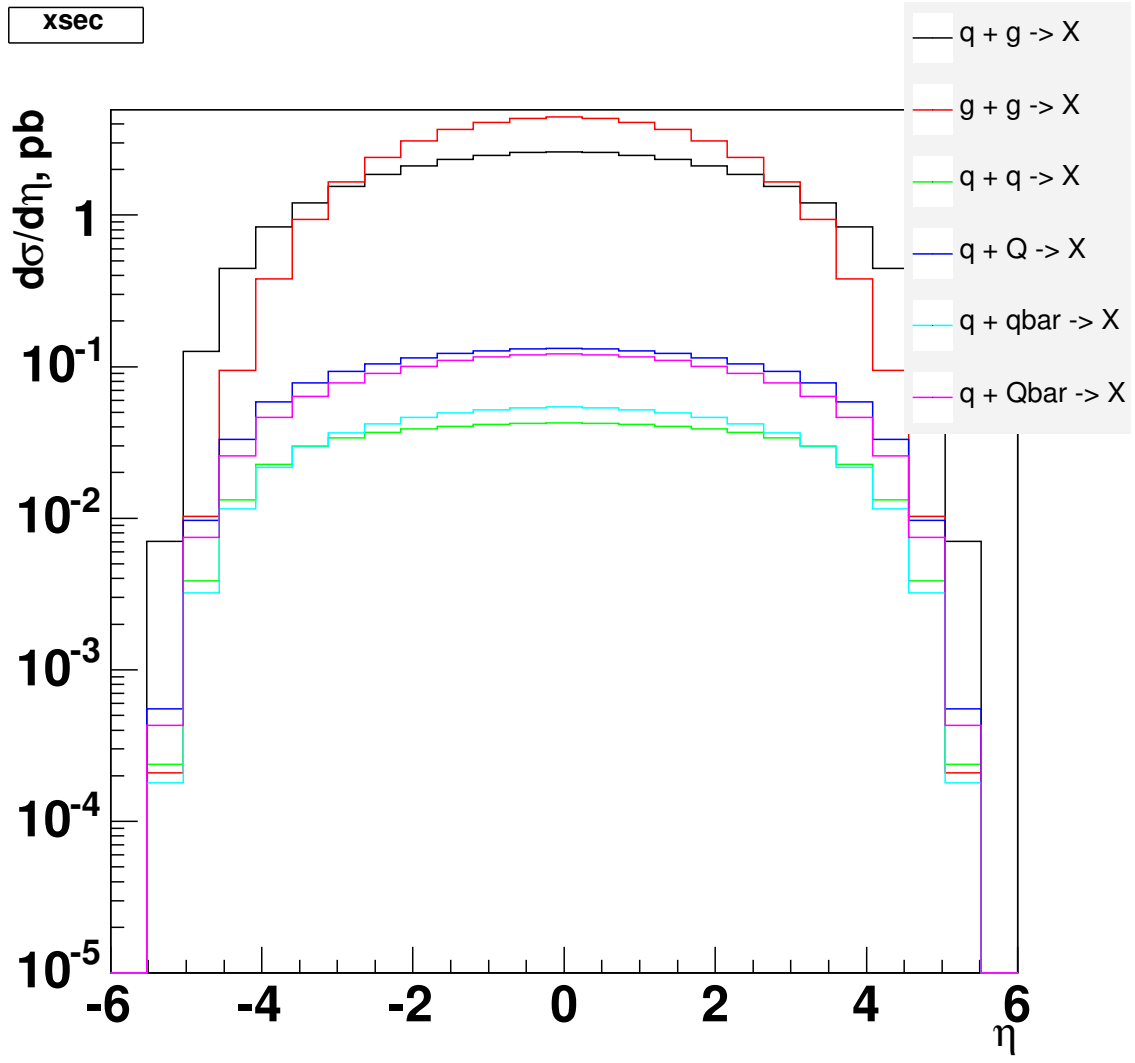

Figure 7.3: Differential cross-section of hardest jet against η for the various allowed partonic initial states; q and Q are used to indicate quarks of different flavour.

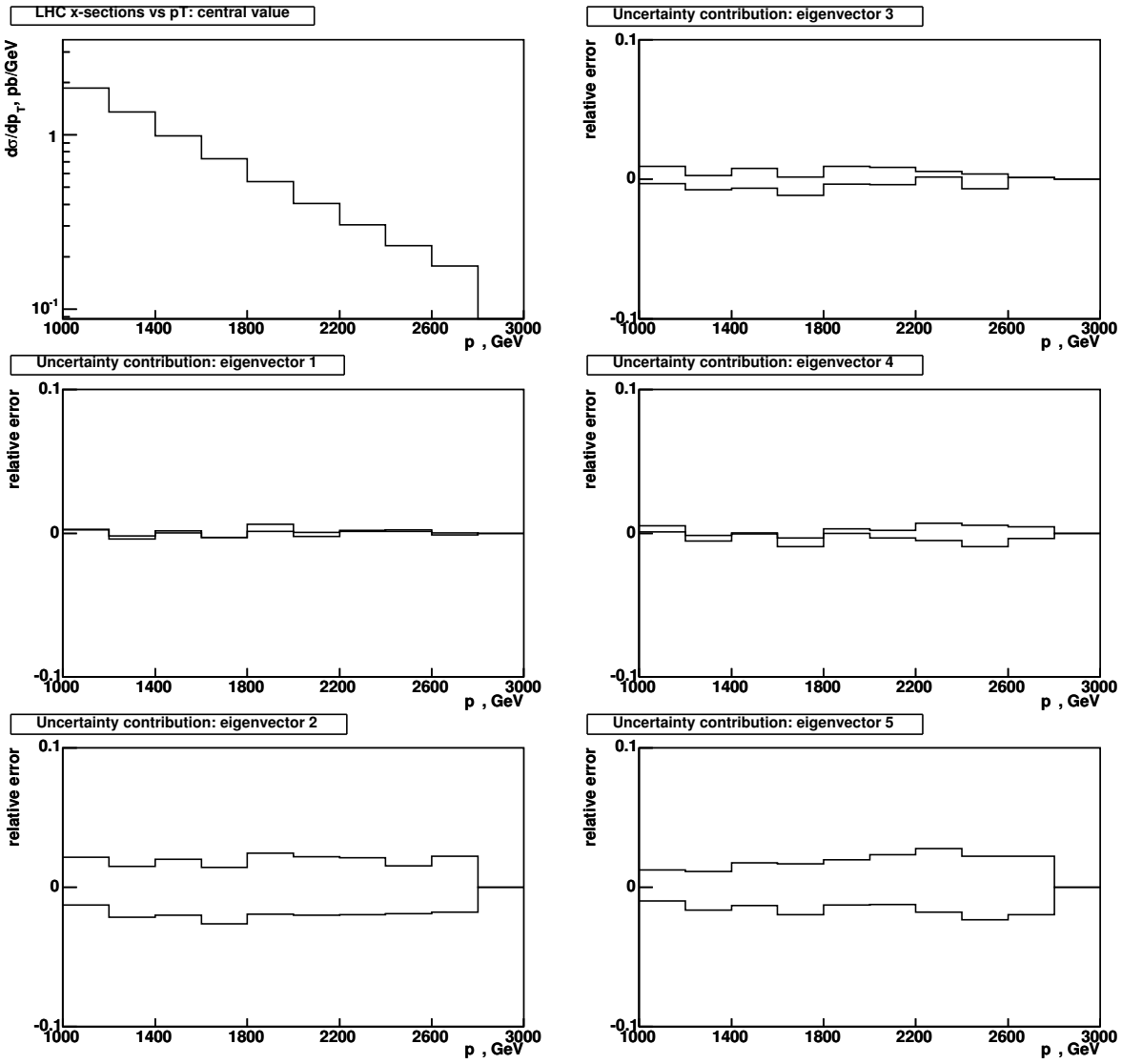

Figure 7.4: The $d\sigma/dp_\perp$ spectrum of the hardest jet, in the range -2 $< \eta <$ 2, and uncertainty contributions relative to central value cross-section from the eigenvectors 1 to 5.

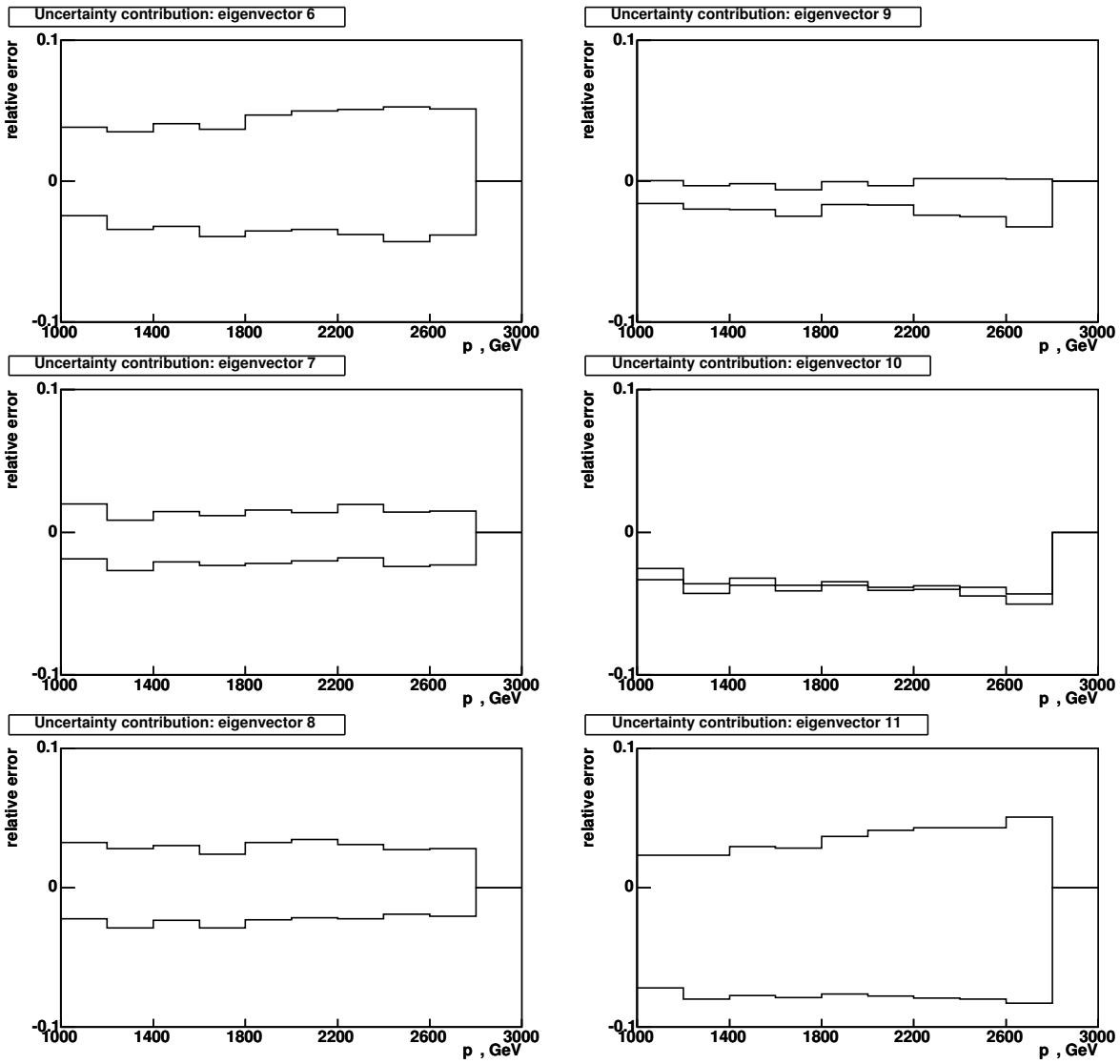

Figure 7.5: Uncertainty contributions relative to central value cross-section from the eigenvectors 6 to 11.

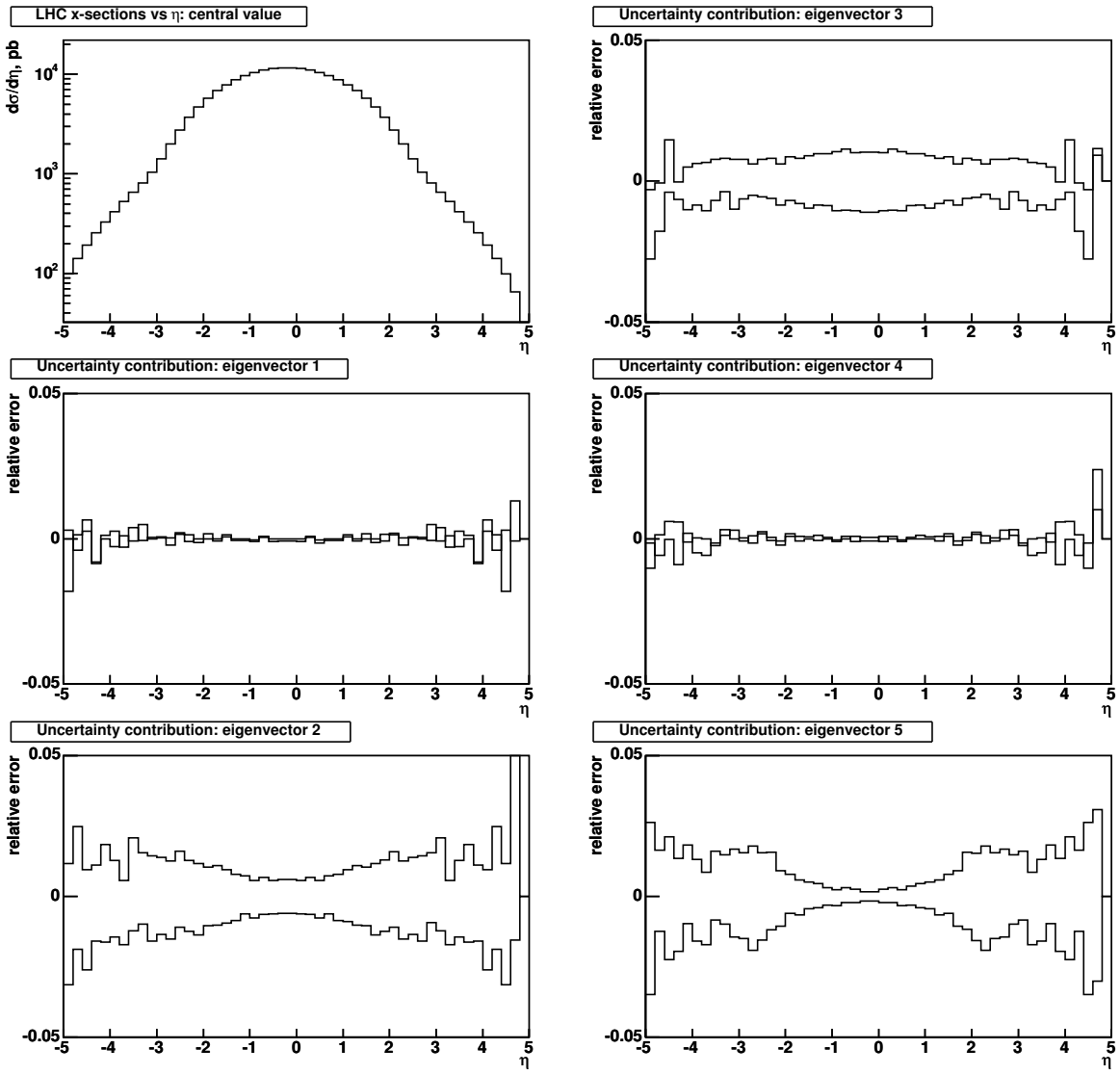

Figure 7.6: The $d\sigma/d\eta$ spectrum of the hardest jet, in the range 1000GeV $< p_\perp <$ 3000GeV, and uncertainty contributions relative to central value cross-section from the eigenvectors 1 to 5.

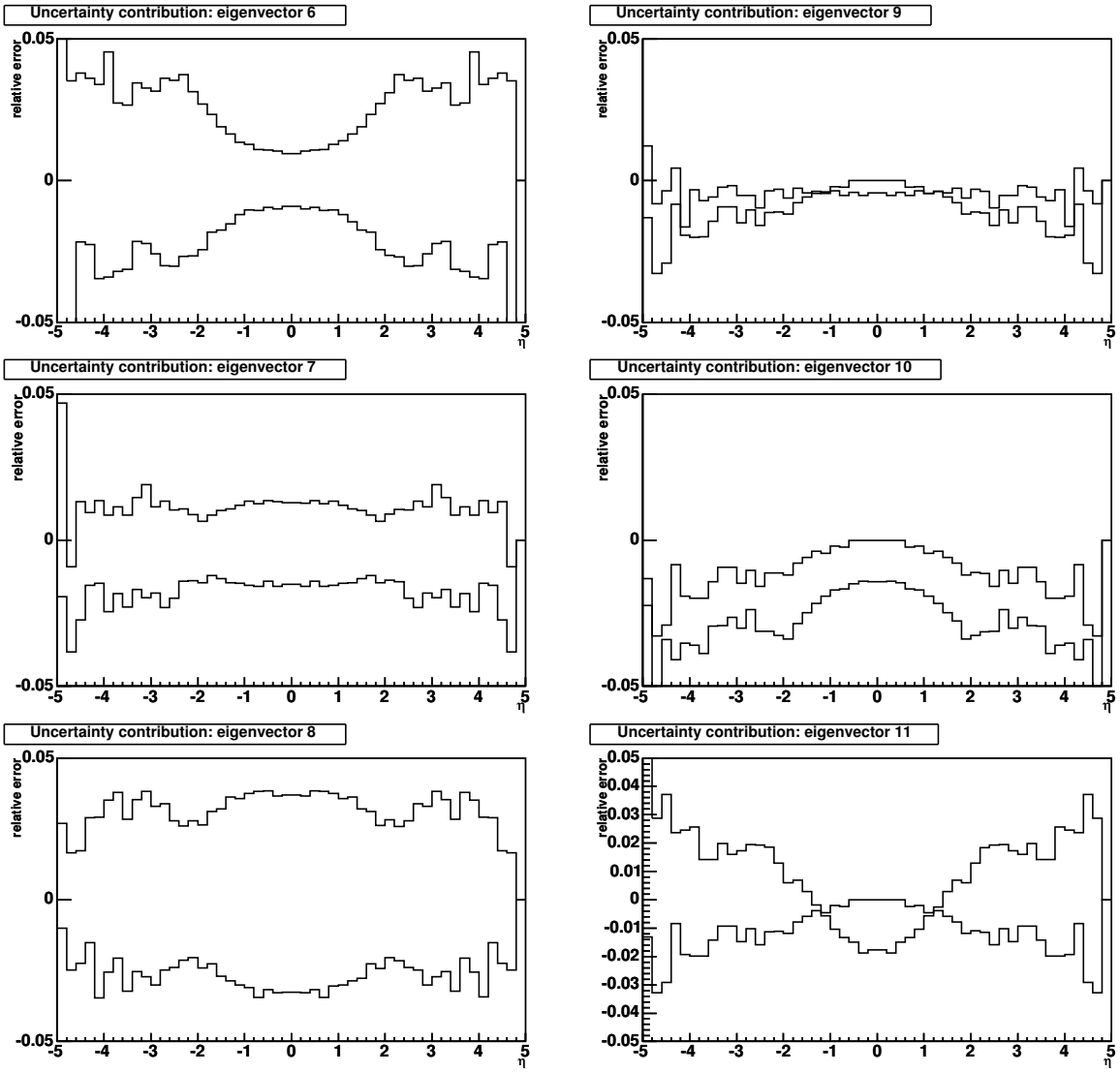

Figure 7.7: Uncertainty contributions relative to central value cross-section from the eigenvectors 6 to 11.

quarks in the sea (charm and bottom) are strongly suppressed by their large masses (top quarks are not considered in this analysis; their primary mode of production at the LHC will be through gluon-gluon fusion, rather than excitation of virtual top quarks in the proton [31]). Eigenvectors 2, 5 and 6, which encode the low-x and high-x u and high-x d quark respectively, all give rise to very small uncertainties at low $|\eta|$ which rapidly increase when $|\eta| > 2$ (the low-x d is not represented by one main eigenvector in the ZEUS-JETS parameterisation; the p_2^d parameter is set equal to p_2^u). This may be because the valence quarks are more important in either glancing collisions whereby a gluon is exchanged in the t-channel, such that the outgoing partons retain much of their initial momentum, or events in which the partons exchange almost all their momenta and are deflected at a small angle to their incoming trajectories; either sort of collision leads to two back-to-back jets at high $|\eta|$ with high $|\vec{p}|$ but low $p_\perp$.

In events involving one parton with high x and one with low x, the outgoing system is highly boosted in η. In such events it is likely that one parton has a much larger fractional uncertainty than the other, in which case the larger uncertainty will dominate in contributing to the uncertainty on the cross-section.

Eigenvector 10, which encodes mainly the mid-x d and a little of the high-x d, is anomalous in that it gives rise to a cross-section with a mean value significantly below the central value - whereas all the other eigenvectors give rise to pairs of differential cross-sections that are at least approximately symmetric about the central value. This may have something to do with the ZEUS-JETS PDF's poorer flavour separation compared to earlier ZEUS PDF sets that included fixed-target data, which included neutrons and therefore a higher density of d_v quarks.

The greatest contribution to uncertainty in the $d\sigma/d\eta$ spectrum is eigenvector 8, which encodes the high-x gluon. This is because the gluon dominates the proton PDF over the entire range of x except very high x (close to 1), as can be seen in the plots of the ZEUS-JETS PDF reproduced above. Eigenvector 11, which determines the mid-x (strictly, mid- to high-x) gives the largest contribution to uncertainty in $d\sigma/dp_\perp$, due to the dominance of high-x partons in producing high $p_\perp$ events and the overall dominance of the gluon in the partonic cross-section for QCD events.

Frixione's and Ridolfi's package allows the user to set his own value for the renormalisation and factorisation scales. These are theoretical quantities involved in the calculation of Feynman diagram amplitudes used to calculate cross-sections. These scales are defined, and their relevance to jet cross-sections discussed, in sections 7.2.2 and 7.2.3.

The renormalisation scale is the energy scale at which renormalisation procedures are carried out: these are calculations performed to cancel the divergences (i.e. infinities) in cross-sections and masses due to contributions from virtual loops in Feynman diagrams. Both of these energy scales are, to a certain extent, arbitrary and can be varied according to the precise prescription used by the person (or computer) doing the calculations. In Frixione's program, the default value for both of these quantities is half the sum of the $p_\perp$ of the outgoing partons (which may be calculated on the assumption that the event has occurred, then used in the calculation of the event's weight).

Figure 7.8 compares the total cross-section uncertainty due to PDF uncertainty to the uncertainty introduced by allowing both the factorisation and renormalisation scales to vary from 0.5 to 2.0 times their default values. It may be seen that the uncertainty due to scale factor uncertainty is greater than that due to PDF uncertainty at $p_\perp$ values of less than $\sim$2 TeV.

7.1.5 The CTEQ6.1M PDF set

A set of PDFs that incorporates uncertainties on the partons can be constructed using a Hessian matrix [21]. First of all, a χ^2 function is defined for each data set used in the fit, based on correlated and uncorrelated errors in those data. These functions are summed to make a total χ^2, and a Jacobian of this function is then taken with respect to a matrix of PDF parameters (i.e. the parameters determining the functional form of the distribution for each flavour concerned), $\{a_i a_j\}$; the number of parameters thus determines the dimension of the Hessian matrix. In the case of the CTEQ6.1M set, there are 20 parameters. The point $\{a_0\}$ in parameter-space that minimises χ^2 gives a value χ^2_0, and the corresponding PDF is the best fit or so-called 'central value', and is numbered 0, as is conventional. Then 20 eigenvectors of the Hessian matrix are defined, and along each of these the χ^2 is allowed to vary in such a way that $\chi^2 - \chi^2_0 < T^2$, where T is called the

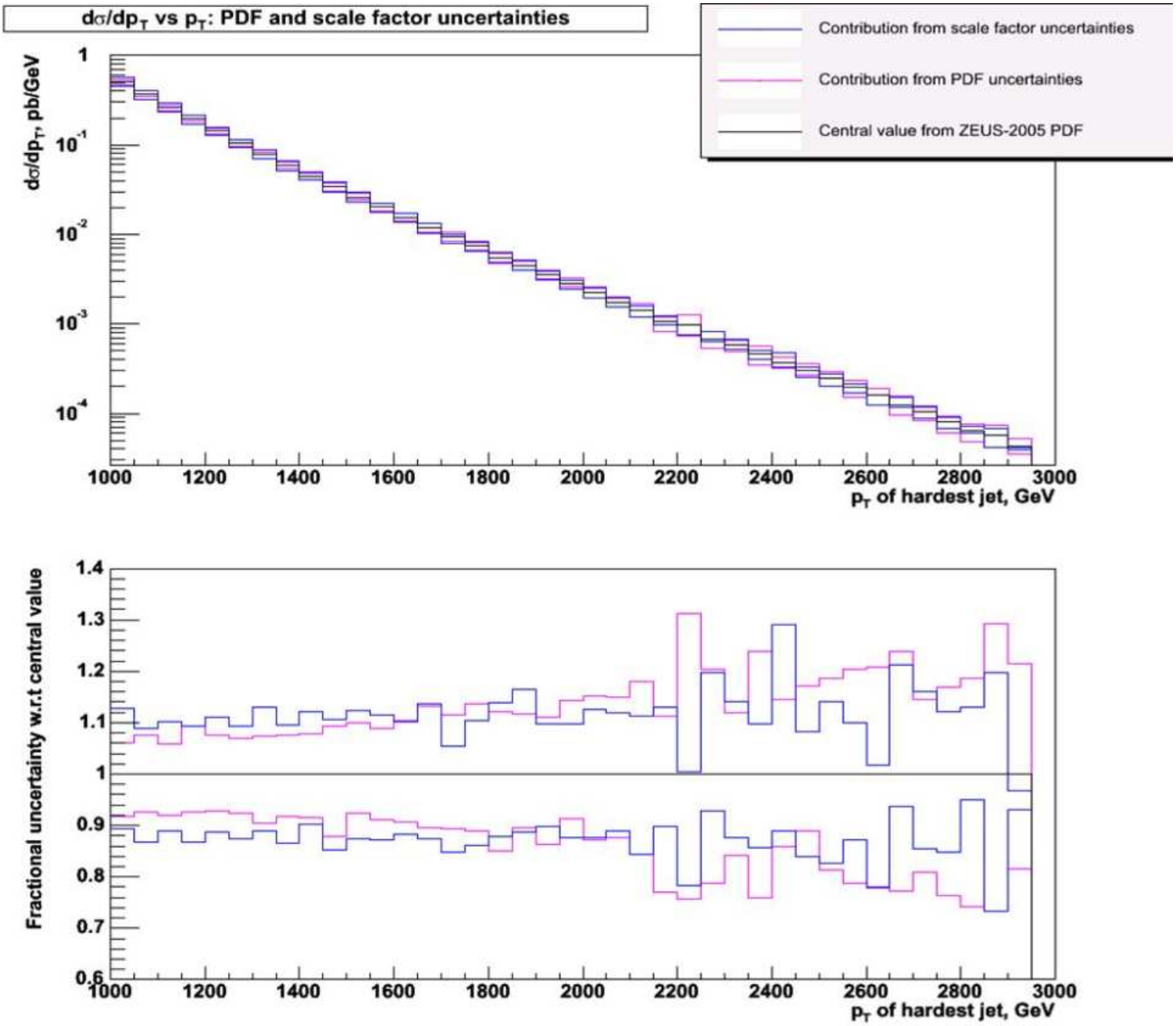

Figure 7.8: Uncertainty on the $d\sigma/dp_\perp$ distribution due to variation of the renormalisation and factorisation scales, which are set equal to each other in Frixione's code; an uncertainty range is introduced by varying both scales to 2 and 0.5 times their default value. Uncertainty due to PDF uncertainty is shown for comparison.

tolerance, which is a somewhat arbitrary value that may be chosen so the PDF set appropriately reflects the degree of uncertainty on the input data sets. Now 40 *eigenvector basis sets*, $\{a_i^{\pm}\}$ - two for each of the 20 eigenvectors - may be chosen to span the parameter-space in the neighbourhood of the minimum. These, then, are the 40 member PDFs which, together with the central value, make up the set.

In terms of this eigenvector basis set, any physical quantity X that has a predicted value $X_0 = X(a_0)$ and a range of uncertainty δX that has a linear approximation:

$$(\delta X)^2 = \frac{1}{4}\sum_{i=1}^{n}[X(a_i^+) - X(a_i^-)]^2; \tag{7.5}$$

however, this depends on the assumption that $\chi^2(a)$ is a quadratic function of the parameters $\{a\}$, and that $X(a)$ is linear. In actual fact this approximation is not strictly valid, so asymmetric error bounds are used instead:

$$(\delta X_{\pm})^2 = \sum_{i=1}^{n}[X(a_i^{\pm}) - X(a_0)]^2. \tag{7.6}$$

So there are 41 members to the CTEQ6.1M PDF set: the central value, labelled 0, and the 20 pairs of eigenvector basis sets, labelled 1-40 where 1 and 2 represent the pair of PDFs associated with eignenvector 1, 3 and 4 with eigenvector 2 and so on.

In technical terminology each pair of basis sets represents an 'excursion' up and down the relevant eigenvector of the Hessian matrix; however, since there is a 20-dimensional phase space of variables, 'up' and 'down' are ambiguous. With the ZEUS-JETS PDF set, the pairs of PDFs corresponding to 'up' and 'down' excursions do *generally* tend, when enfolded with the cross-section weightings computed by a Monte Carlo, to give cross-sections that are mostly higher and lower than that based on the central value. With CTEQ6.1M this is not the case: in the case of dijet invariant mass, it is often found that cross-sections due to a pair of basis sets cross over, or that in some regions *both* cross-sections are either higher or lower than the central value prediction. This suggests the following algorithm for computing an overall uncertainty from the uncertainties due to all 40 member PDFs:

- if $\sigma_i^a > \sigma_0$ AND $\sigma_i^b > \sigma_0$,

  then $\delta\sigma_i^+ = max(\sigma_i^a, \sigma_i^b) - \sigma_0$ and $\delta\sigma_i^- = 0$;

• if $\sigma_i^a < \sigma_0$ AND $\sigma_i^b < \sigma_0$,

then $\delta\sigma_i^+ = 0$ and $\delta\sigma_i^- = min(\sigma_i^a, \sigma_i^b) - \sigma_0$;

• else (i.e. $\sigma_i^a > \sigma_0$ and $\sigma_i^b < \sigma_0$, or *vice versa*);

then $\delta\sigma_i^+ = max(\sigma_i^a, \sigma_i^b) - \sigma_0$ and $\delta\sigma_i^- = min(\sigma_i^a, \sigma_i^b) - \sigma_0$,

where $\sigma_i^{a,b}$ labels the cross-sections computed with each pair of PDFs associated with eigenvector i, σ_0 is the cross-section calculated with the central value PDF and $\delta\sigma_i^{\pm}$ is the upper/lower cross-section uncertainty for that eigenvector. It should be stressed that this procedure is carried out on a bin-by-bin basis.

Finally, the upper and lower total error bounds for each bin may be calculated as follows:

$$\sigma^+ = \sigma_0 + \left(\sum_{i=1}^{20}(\delta\sigma_i^+)^2\right)^{1/2}; \qquad \sigma^- = \sigma_0 - \left(\sum_{i=1}^{20}(\delta\sigma_i^-)^2\right)^{1/2}. \tag{7.7}$$

7.1.6 Computing cross-section uncertainties due to PDF uncertainties

The process by which the uncertainties in the PDF may be used to calculate the resultant uncertainties in the reconstructed jet cross-sections is not trivial. An initial attempt was made using events generated with each of the 40 uncertainty PDFs in the CTEQ6.1M set, reconstructing and analysing them using Atlfast and then plotting the jet quantities thereby derived with the quantities calculated using the central value PDF as a benchmark. However this approach met with limited success, as the statistical uncertainty outweighed the uncertainty due to the PDFs even with a large data set ($\sim 10^5$ events per J-region per PDF). This is due to the necessity of adding cross-section uncertainties in quadrature, such that both positive and negative statistical fluctuations in cross-section tend to accumulate, whereas they would naturally cancel to at least some degree if they were added linearly.

This necessitates a different approach to evaluating these uncertainties; namely, reweighting. In this technique, only a single set of events is generated, using the central value PDF, and the generator-level event information is read out and stored. (In Athena this is achieved by accessing the object class `McEventCollection` using the `GEN_AOD` Storegate key.) The kinematic quantities of the primary vertex are located by selecting the $2 \rightarrow 2$ vertex with the highest CoM energy for

the incoming partons; the flavours and four-momenta of the incoming partons are read out and the parton density for each parton, $f_i(x, Q^2)$, may then be calculated using LHAPDF [23], for $x = p_z/7$ TeV and flavour i. The value of Q^2 is calculated according to the default prescription used in `PYTHIA` for QCD events:

$$Q^2 = \frac{1}{2}(m_3^2 + m_4^2) + \frac{\hat{t}\hat{u} - m_3^2 m_4^2}{\hat{s}}, \tag{7.8}$$

in which the masses of c and b quarks are taken into account while u, d and s quarks, as well as gluons, are treated as massless. Then the *weight* for each event may be calculated as follows:

$$w_j = \left(\frac{f_j^{i_1}(x_1, Q^2)}{f_0^{i_1}(x_1, Q^2)}\right) \cdot \left(\frac{f_j^{i_2}(x_2, Q^2)}{f_0^{i_2}(x_2, Q^2)}\right), \tag{7.9}$$

in which j runs from 1 to 40, for the 40 PDFs of CTEQ6.1M, f_0 denotes the central PDF and $i_{1,2}$ and $x_{1,2}$ are the flavour and fractional momentum of partons 1 and 2 respectively. Thus a table of 40 weights is generated for each event, and the jet quantities may be used to create 40 sets of histograms which may then be used to estimate the uncertainty. In this approach, statistical fluctuations are not a problem when it comes to adding uncertainties in quadrature, since they are the same for all 40 data sets and therefore cancel out rather than accumulating.

The result of these calculations is shown in figure 7.9, the dotted uncertainty bands either side of the Standard Model QCD prediction showing the quadratically added uncertainty contributions from the 20 pairs of PDFS.

7.2 Theoretical energy scales

7.2.1 Renormalisation scale

In the historical development of the first fully realised quantum field theory, namely quantum electrodynamics, it soon became apparent that corrections[1] to scattering amplitudes calculated naïvely using the formalism of Feynman diagrams contained terms which involved divergent integrals; in other words, the predicted probability for the interaction process blows up to infinity.

[1] Specifically, virtual particles forming loops inside propagators.

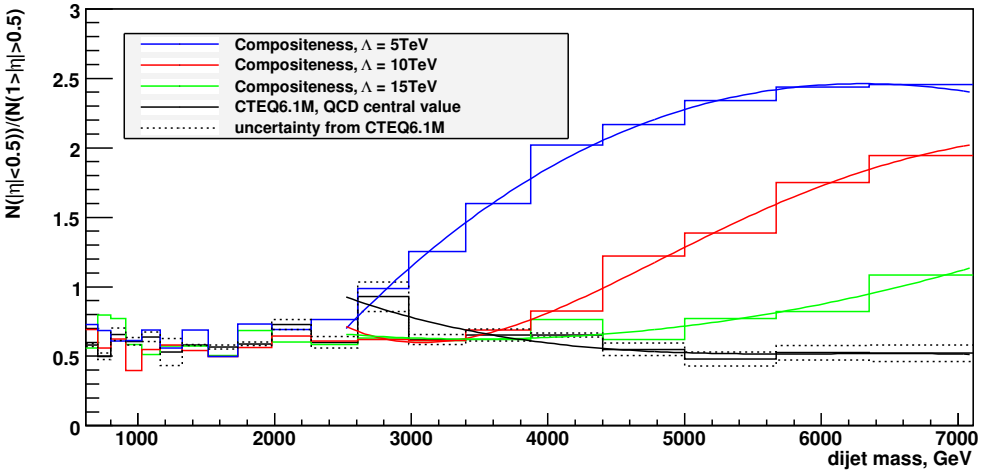

Figure 7.9: Dijet ratio for the three compositeness models, compared to the Standard Model QCD prediction with uncertainty due to the CTEQ6.1M PDF family.

This problem was solved by realising that certain fundamental properties of particles as measured in the laboratory (such as the electronic mass and charge, in the case of QED) do *not* correspond to the 'bare' quantities of the particle considered alone, but incorporate quantum field effects due to virtual particles in the vacuum; for electric charge, this involves the phenomenon of vacuum polarization, which leads to a running coupling constant as discussed in section 2.3.

However, in order for meaningful calculations to be made, particle properties must be used in the equations of quantum field theories with values that correspond to measurements made at a certain characteristic energy scale, called the renormalisation scale. This scale is arbitrary, but in practice, when using Monte Carlo generators, it is either chosen to be equal to a fixed quantity appropriate to the energy scale of the collider, or some energy scale used in the calculations, such as the CoM energy of the incoming partons or the mass of the heaviest quark flavour considered; it is conventional to use a squared renormalisation energy scale, denoted μ_R^2. In `PYTHIA` 6.4 the default is a fixed value 10^4GeV^2. The effect of an uncertainty in this energy scale can then be investigated by comparing data generated with the scale set to some multiple or fraction of the default value to that generated using the default.

7.2.2 Factorisation scale

When hadrons are involved in a collision, partons given off by those involved in the hard scatter may be treated in two ways: they may be considered as final state radiation, in which case they give rise to jets that can be recorded by the detector just like the jets from the hard subprocess, or they can be absorbed into the hadron's PDF. The parton is treated in the former way if its $E_\perp$ is greater than some energy scale, and in the latter if not. This energy scale, again usually treated as a squared energy, is the factorisation scale, μ_F^2. It appears in the formula used to calculate hadron-level (i.e. observable) cross-sections from partonic-level cross-sections and PDFs:

$$\sigma(\mu^2) = \sum_{i,j=1}^{n_f} \int_{x_1=0}^{1} \int_{x_2=0}^{1} \hat{\sigma}_{1,2\to X}(p_1, p_2, \alpha_s(\mu^2), \mu^2/\mu_F^2) f_1^i(x_1, \mu^2) f_2^j(x_2, \mu^2) dx_1 dx_2, \qquad (7.10)$$

in which the quantities are as defined in section 3.2. This is a very general formula known as the *factorisation theorem*.

The quadratically combined effects of factor-of-two uncertainty in both the renormalisation and factorisation scales on the Standard Model QCD dijet ratio is shown in figure 7.10, along with the three compositeness models for comparison. This figure is shown as a set of plotted functions rather than histograms as it was adapted from calculations made using Frixione's NLO jet program, rather than Atlfast/Athena.

7.3 Total experimental and theoretical uncertainty

Figure 7.11 shows the dijet ratio predictions for four physics scenarios, namely Standard Model QCD behaviour and the three models of quark compositeness. The uncertainty band on the SM QCD ratio is calculated by adding in quadrature the two kinds of experimental uncertainty and three kinds of theoretical uncertainty considered in this analysis. To recap:

$$\Delta\sigma_{tot} = \Delta\sigma_{abs.cal.} \oplus \Delta\sigma_{rel.cal.} \oplus \Delta\sigma_{CTEQ} \oplus \Delta\sigma_{renorm} \oplus \Delta\sigma_{fact} \qquad (7.11)$$

in which:

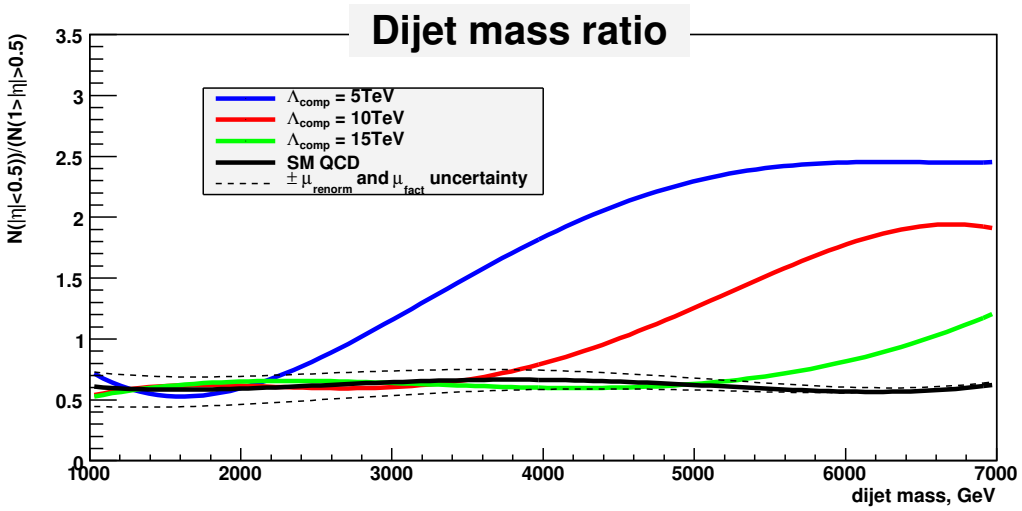

Figure 7.10: The dijet ratio for Standard Model QCD and the three compositeness models, with the QCD prediction showing uncertainty bands due to varying the renormalisation and factorisation scales to 50% and 200% of their default values. See text for an explanation of the smooth functional form of the lines.

- $\Delta\sigma_{abs.cal.}$ is the dijet ratio uncertainty due an uncertainty of 3% in the absolute energy response of the hadronic calorimeter;
- $\Delta\sigma_{rel.cal.}$ is the dijet ratio uncertainty due to a difference of 1% in the relative energy response of the calorimeter regions corresponding to pseudorapidities of $\eta < 0.5$ and $1.0 < \eta < 0.5$;
- $\Delta\sigma_{CTEQ}$ is the uncertainty due to the CTEQ6.1M PDF set;
- $\Delta\sigma_{renorm}$ is the uncertainty due to a factor-of-two uncertainty in the renormalisation scale; and
- $\Delta\sigma_{fact}$ is the uncertainty due to a factor-of-two uncertainty in the factorisation scale.

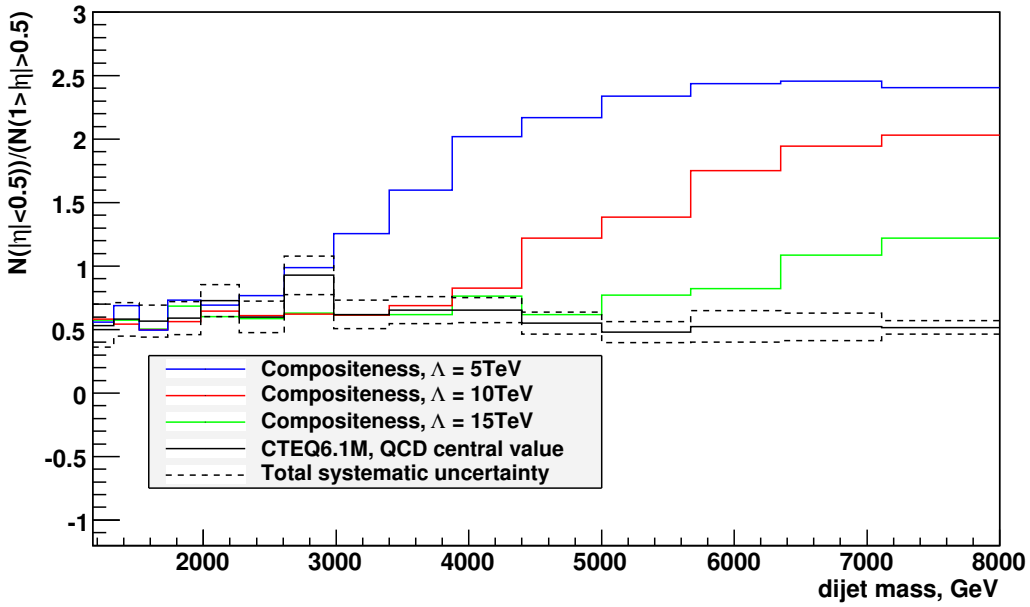

Figure 7.11: The dijet ratio for the the four physics models under consideration, with combined systematic (experimental $\oplus$ theoretical) uncertainties shown on the QCD ratio.

Chapter 8

Conclusion and summary

The sensitivity of ATLAS to quark compositeness may be quantified by a figure called the *significance*, which may be thought of as the difference between a positive signal (evidence for compositeness) and the 'null' signal (Standard Model QCD behaviour), relative to the overall resolution of the analysis.

A χ^2 value is calculated for each of the three compositeness models under consideration, according to the following formula:

$$\chi^2(L_{int}) = [(R_{\mathrm{QCD}} - R_\Lambda)/\sigma_{\Lambda_{comp}}(L_{int})]^2 \tag{8.1}$$

where R_{QCD} and $R_{\Lambda_{comp}}$ are the values for the dijet mass pseudorapidity ratio (equation 6.4) for Standard Model QCD and quark compositeness (for the three given values of Λ_{comp}) respectively, and σ_Λ is the combined systematic and statistical uncertainty on R_Λ as a function of integrated luminosity. The significance, $S_\Lambda(L_{int})$, is then simply

$$S_\Lambda(L_{int}) = \sqrt{\chi^2(L_{int})}, \tag{8.2}$$

In order to maximise the significance of the discrimination, only data in the high-m_{jj} part of the spectrum (specifically, dijets with an invariant mass greater than around 1 TeV) have been used, as it is this region where phenomenological divergences between the various models start to become significant compared to systematic uncertainties.

As may be seen from figure 8.1, statistical uncertainty dominates systematic uncertainties for low integrated luminosities, to the point where the curves cannot be distinguished since they coincide, but as accumulated total approaches 100fb^{-1}, the systematic uncertainties start to become significant. This is, of course, a general feature of any experiment that collects progressively larger amounts of total data over time.

The horizontal black lines on the plots in figure 8.1 show significances of 2 and 5 standard deviations. The 2-σ line marks a potential early indication of quark compositeness, while the 5-σ line represents (as is conventional) potential for actual discovery. An integrated luminosity of 100pb^{-1} shows no potential even for an indication of quark compositeness for any value of Λ_{comp} within the range considered here; at 1fb^{-1}, there is potential for an early indication of compositeness at values of Λ_{comp} at around 5TeV, the smallest value considered in this analysis, but no potential for a more definite discovery. At 10fb^{-1}, indication of possible compositeness begins to look feasible for Λ_{comp} below about 8 TeV, with discovery potential below 5.5 TeV, while for 100fb^{-1}, discovery appears feasible for all values below around 6 TeV.

Figure 8.2 shows the dependence on integrated luminosity of the upper bound on Λ_{comp} to which ATLAS is sensitive in this analysis, for significances of 2σ and 5σ. Note that the limit on Λ_{comp} is only significantly affected by theoretical and experimental uncertainties at integrated luminosities larger than around 10fb^{-1}[1]; below that level the limit on signal sensitivity is dependent entirely on statistical uncertainty. As more data is accrued, however, limits on ATLAS's sensitivity to quark compositeness will become increasingly dominated by systematic uncertainties, and these will need to be reduced in order to further improve the experiment's sensitivity to this physics channel. This may be achieved for experimental uncertainties by (for instance) calibrating the calorimetry, and for theoretical uncertainties through using PDFs based on improved models of hadron structure and novel computational techniques in QCD.

[1]This amount of integrated luminosity corresponds to approximately four months' continuous data-taking at a luminosity of $10^{32}\text{cm}^{-2}\text{s}^{-2}$.

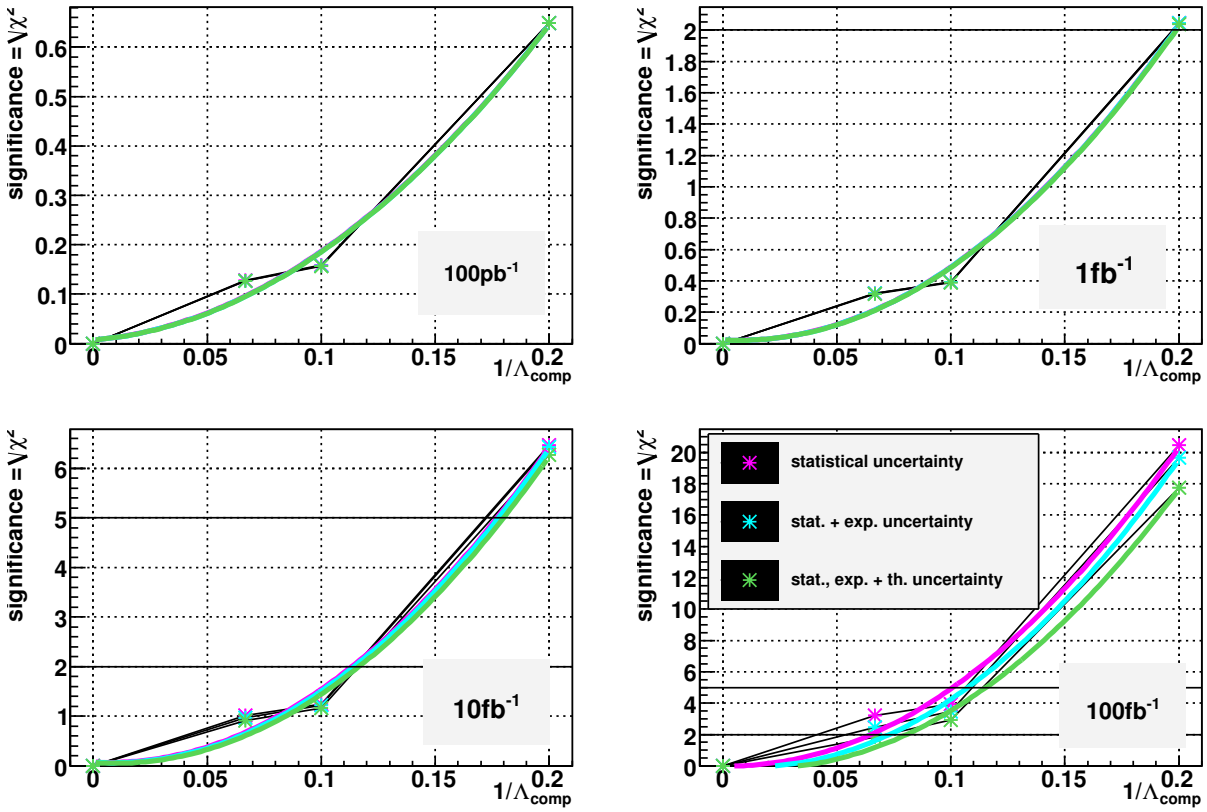

Figure 8.1: The significance of the discovery signal as a function of $1/\Lambda_{comp}$, for four integrated luminosities; values for statistical uncertainty only, statistical plus experimental and combined statistical, experimental and theoretical uncertainties are shown separately. The curves are fitted as a quadratic function of $1/\Lambda_{\rm comp}$ with the constant term set equal to zero, which ensures the significance is identically zero when $1/\Lambda_{\rm comp} = 0$, corresponding to Standard Model QCD with elementary quarks.

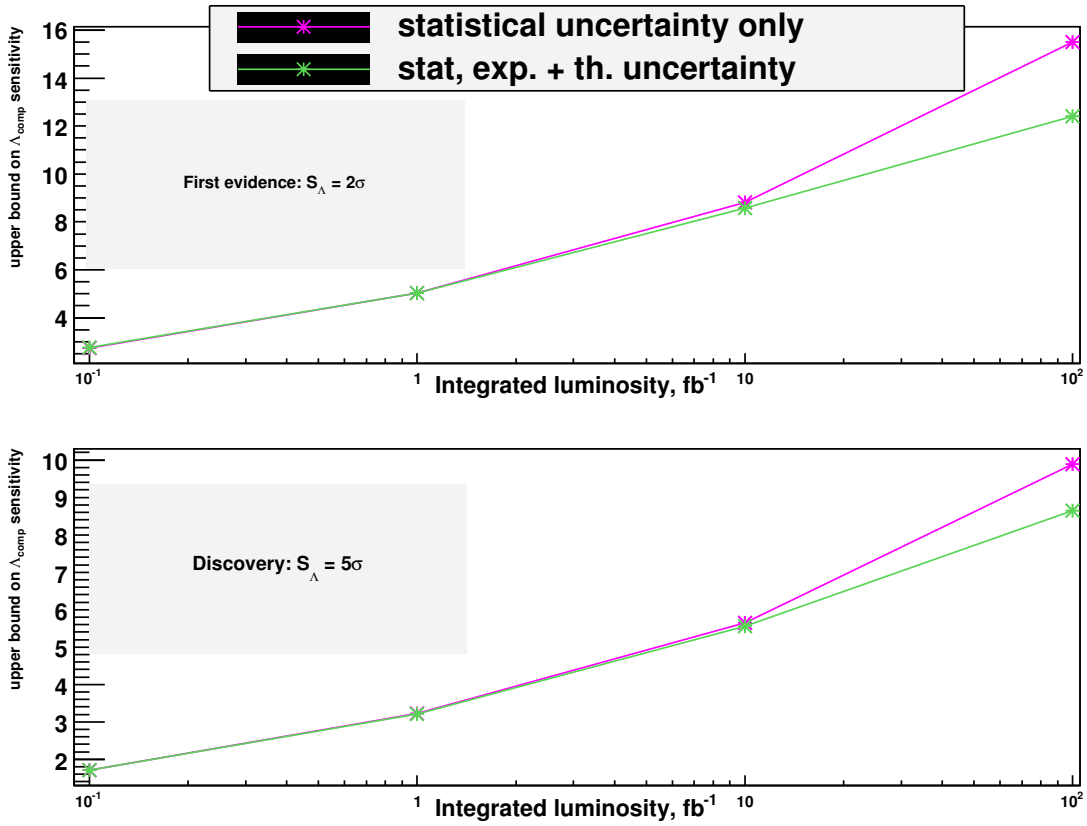

Figure 8.2: The upper bound on the value of Λ_{comp} to which ATLAS is sensitive in this analysis, as a function of integrated luminosity. The upper plot shows the value for a 'first evidence' significance of 2σ, the lower plot a 'discovery' signal of 5σ. This Λ_{comp} limit is shown for both statistical significance only and for significance with combined statistical, experimental and theoretical uncertainties.

Chapter 9

Acknowledgements

I would like to acknowledge the help of Amanda Cooper-Sarkar, Claire Gwenlan and Mark Sutton in this analysis, both in technical aspects and for general help and guidance.

Chapter 10

Appendix A: structure functions and parton densities

It may be useful at this point to give a rigorous definition of the variables used in the physics of hadron interactions. The following discussion is based on [22]. First, consider the simplest possible inelastic interaction between a hadron and another particle: the scattering of an electron off a proton by the exchange of a photon, and the subsequent fragmentation of the proton. This is called an inclusive process, since all possible final states are summed over, or included. Let the incoming electron have a four-momentum k, and the outgoing electron k'. (Indices have been suppressed for the sake of simplicity.) The virtual photon exchanged by the electron and proton has four-momentum q. Since it is meaningless to ask which particle has emitted the photon and which has absorbed it, it is conventional to use q^2 instead. Because of the photon's virtuality, this quantity is negative, so it is more common to define a positive quantity $Q^2 \equiv -q^2$. Now let the incoming proton have a four-momentum p, and the outgoing proton fragments have an aggregate invariant mass W. At this stage, we are intentionally ignoring the proton's internal structure, so the process is simply 'inelastic scattering' rather than 'deep inelastic scattering'. Now we define

three new variables, one with dimensions of energy (in natural units) and two without dimension:

$$\nu \equiv \frac{p \cdot q}{M}; \qquad x \equiv \frac{-q^2}{2p \cdot q} = \frac{Q^2}{2M\nu}; \qquad y \equiv \frac{p \cdot q}{p \cdot k}, \tag{10.1}$$

in which M is the proton's invariant mass. x is the same Björken x encountered earlier. The squared invariant mass of the inclusive hadronic final state (i.e. all outgoing particles except the electron) is now

$$W^2 = (p+q)^2 = M^2 + 2M\nu + q^2 \tag{10.2}$$

Now we start to consider the concept of pointlike objects inside the proton. At this stage, we are only interested in quarks - we are using a photon as a probe, and gluons, being electrically neutral, do not couple directly to photons. Historically, evidence for a model of the hadron as a collection of pointlike objects came from the phenomenon of Björken scaling, namely the fact that, at medium to high x ($x > 0.05$, say) the electrostatic structure function (see below) is very nearly independent of Q^2. This is a consequence of asymptotic freedom. Conversely, the strong potential gets larger as the distance between two quarks increases, until (at a distance of just a few fm [1]) enough energy is latent in the gluonic field to allow the production of real quark-antiquark pairs. This phenomenon is called confinement, and the production of $q\overline{q}$ pairs to form new hadrons is called hadronization. This is why free gluons and quarks are never observed, and therefore why QCD is the hardest fundamental interaction (with the arguable exception of gravity!) to study.

This reversal of potential dependence of distance (from the intuitive $V \propto 1/r$ behaviour of the Coulomb potential) means that perturbative QCD is suitable for the phenomenological description of high-energy hadronic interactions, while this approach does not work for low-energy ('soft') interactions. Appropriately, this is the reverse of the electromagnetic interaction, in which perturbative methods only work for *small* potentials.

In pre-quark-model theory, the interaction of a hadron and an electron was calculated in terms of the scalar product of two tensors. The proton tensor was written as the sum of several tensors, each with a dimensionless coefficient (structure function) whose value was not specified as the proton's structure was a unknown at this time. When the quark model is adopted, the proton

[1] 1fm = 10^{-15}m = 1 femtometre, or fermi.

structure functions become:

$$W_1^{point} = \frac{Q^2}{m_q^2}\delta\left(\nu - \frac{Q^2}{2m_q}\right); \qquad W_2^{point} = \delta\left(\nu - \frac{Q^2}{2m_q}\right) \tag{10.3}$$

Here, "point" is used to indicate that we are now dealing with pointlike objects inside the proton, while m_q is the mass of those objects (quarks). It becomes apparent that, in the limit of large Q^2, these structure functions are not functions of Q^2 and ν independently, but only of the ratio $Q^2/2m_q\nu$ (this is not the case for elastic ep scattering). After some work, it may be seen that

$$MW_1(\nu, Q^2) = \lim_{large\ Q^2} F_1(\omega); \qquad \nu W_2(\nu, Q^2) = \lim_{large\ Q^2} F_2(\omega) \tag{10.4}$$

where

$$\omega \equiv \frac{2M\nu}{Q^2} = \frac{1}{x}. \tag{10.5}$$

F_1 and F_2 respectively are the structure functions for magnetic and electrostatic (Coulomb) interaction. The link between these quantities and quark PDFs comes about in the following way: let $q_i(x)$ be the $f_i(x)$ for quarks and $\overline{q}_i(x)$ be the $f_i(x)$ for antiquarks, let e_i be the electric charge of each quark flavour and e be the electronic charge:

$$F_2 \propto x\sum_i \left(\frac{e_i}{e}\right)^2 (q_i(x) - \overline{q}_i(x))^2 \tag{10.6}$$

Now we have two dimensionful structure functions which, in the limit of high photon virtuality, are dependent on a single kinematic variable; the ratio of energy lost by the electron to photon virtuality, since in the proton's rest frame,

$$\nu = E - E', \tag{10.7}$$

in which E and E' are the initial and final electron energies, respectively.

It is possible to trace the historical development of the theory of hadron structure with diagrams representing the supposed internal structure of the proton and the appropriate PDF one might expect if this was actually the case. This progression is graphically illustrated in figure 10.1.

From the particle's discovery in the early 20th century until the 1960s, the proton was thought to be an elementary particle - in this case, the PDF is simply a delta function at $x = 1$ (top box).

If the proton is instead considered to consist of three quarks of equal mass, but they are considered to be held rigidly in place, the PDF looks like a delta function at $x = \frac{1}{3}$ (second box). When the quarks are considered to be held together by a finite classical potential of some kind, the PDF still peaks at $x = \frac{1}{3}$, but now has finite height and width - it is smeared out somewhat by the elasticity of the potential binding the quarks together (third box). Finally, when full QCD is 'switched on' and virtual quark-antiquark pairs are 'allowed' to condense out of the gluonic field, there is an additional feature: the PDF now has an important contribution from the sea at low x (bottom box). At high Q^2, the sea contributions is dominant, and the PDF no longer peaks even locally at $x = \frac{1}{3}$, but the curve still has its point of inflection here, due to the three valence quarks.

An official ZEUS PDF is shown in figure 10.2. Partons are labelled d_v (valence, down), u_v (valence, up), S (sea, all flavours) and g (gluons). The latter two lines are scaled down by a factor of 20, giving an impression of the dominance of these contributions at low x. The PDF sets H1-O and MRST 2001 are plotted for comparison.

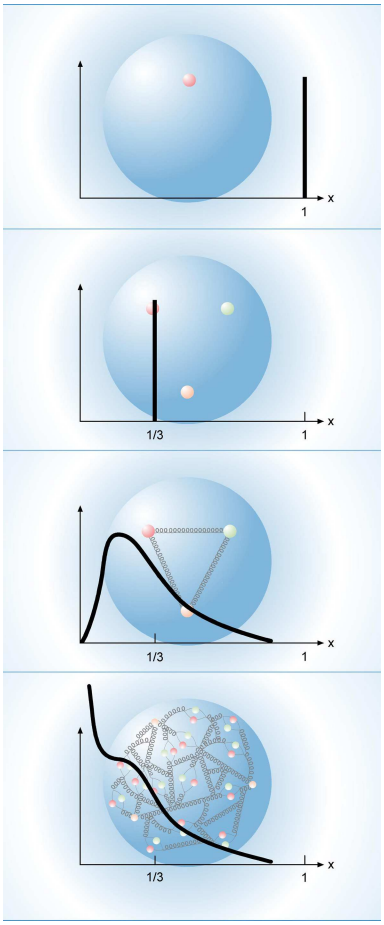

Figure 10.1: The evolution of our picture of hadron structure: see text for explanation.

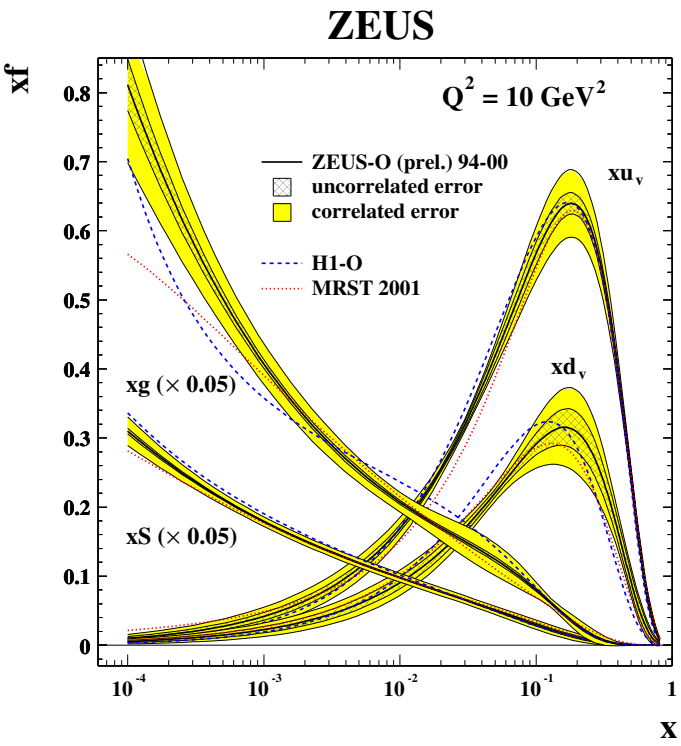

Figure 10.2: A PDF based on data taken with ZEUS, at $Q^2 = 10$ GeV$^2/c^4$.

Chapter 11

Appendix B: the eigenvectors of ZEUS-JETS

The ZEUS-JETS family of PDFs is based on an 11-dimension Hessian uncertainty matrix, whose eigenvectors represent orthogonal combinations of the PDF parameters. Each of the four parton types considered is described by four parameters, giving a total of 16 possible parameters as shown:

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3}(1+p_5 x), \tag{11.1}$$

where $xf(x)$ is the PDF-weighted normalised momentum of the parton of flavour f; $f = u_v$ (u valence quark), d_v (d valence quark), S (the total $q\overline{q}$ sea) or g (the gluon). Thus p_1 determines the overall normalisation, p_2 is sensitive to low x, p_3 is sensitive to high x and p_5 is a fine tuning parameter, sensitive mainly to medium and high x. An additional parameter, p_4, is used in a variant of the ZEUS parameterisation, not considered here:

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3}(1+p_4\sqrt{x}+p_5 x) \text{ [3]}. \tag{11.2}$$

It is worth mentioning that this form of the PDF parameterisation has been empirically found to give a good fit and is not theoretically motivated; PDF sets by other authors use different parameterisations. The approximate correspondence between eigenvectors and partons/x-regimes is given below:

eigenvector:	corresponding part of PDF:
1	low-x S
2	low-x u
3	low-x g
4	total S
5	high-x u
6	high-x d
7	mid-x u; high-x S; high-x g
8	high-x g
9	mid-x u; mid-x d; high-x S
10	mid-x d
11	mid-x g

Chapter 12

Appendix C: pseudorapidity distribution

Apart from the ratio of dijet mass distributions in different η regions, another potential discriminant between Standard Model QCD behaviour and quark compositeness is the distribution of jets in η. This sub-analysis has not been subjected to the same statistical treatment as the distribution of dijet pseudorapidity as a function of invariant mass, and is therefore left as an appendix.

An excess (compared to the SM prediction) of jets in the central η region should be directly measurable from the cross-section vs. η distribution; a large $p_\perp$ cut will need to be made on the data, since deviation from the Standard Model is observed (in the models considered here) in jets with a $p_\perp$ of several TeV. Such events have a very low cross-section, so systematic and theoretical uncertainties will be critical to the measurement due to large statistical uncertainties.

The η distributions for the four physics models under consideration are shown in figures 12.1 and 12.2, with three $p_\perp$ cuts. The uncertainty is due to a simulated $\pm 1\%$ or $\pm 3\%$ uncertainty in measured $p_\perp$. Figure 12.3 shows the combined effects of a $\pm 3\%$ uncertainty in calorimeter response and statistical uncertainty on 30fb^{-1} of data, for jets with a minimum $p_\perp$ of 3,200 GeV.

An analysis of theoretical uncertainties on dijet distributions in η is not performed in this analysis, though it may be noted from the discussion of the proton PDF in relation to jet pseudorapidity

spectra (chapter 7) that by far the largest contribution to uncertainty in dijet cross-sections in the central region, which is where deviation from Standard Model behaviour is most apparent in contact interactions, comes from the mid-to-high-x gluon. It is therefore clearly in this part of the PDF that future parameterisations will have to make the most improvement in order to significantly impact on the LHC's potential for discovering new physics channels with signatures in this kinematic region, *viz*. multi-TeV dijets with $|\eta| < 1.0$.

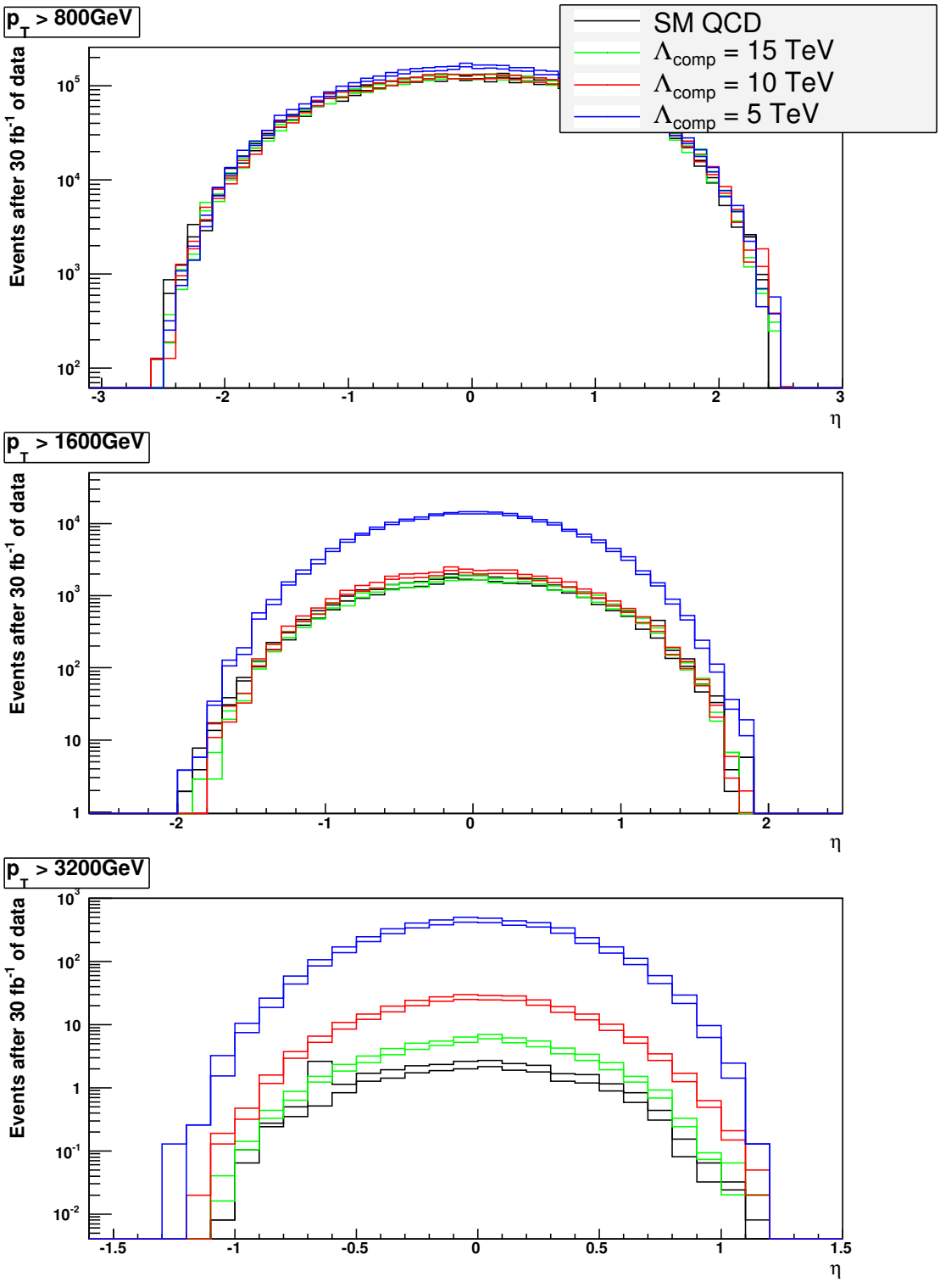

Figure 12.1: Pseudorapidity jet distributions for the SM and three compositeness models after three different $p_\perp$ cuts, showing the effects of a $\pm 1\%$ calorimeter uncertainty.

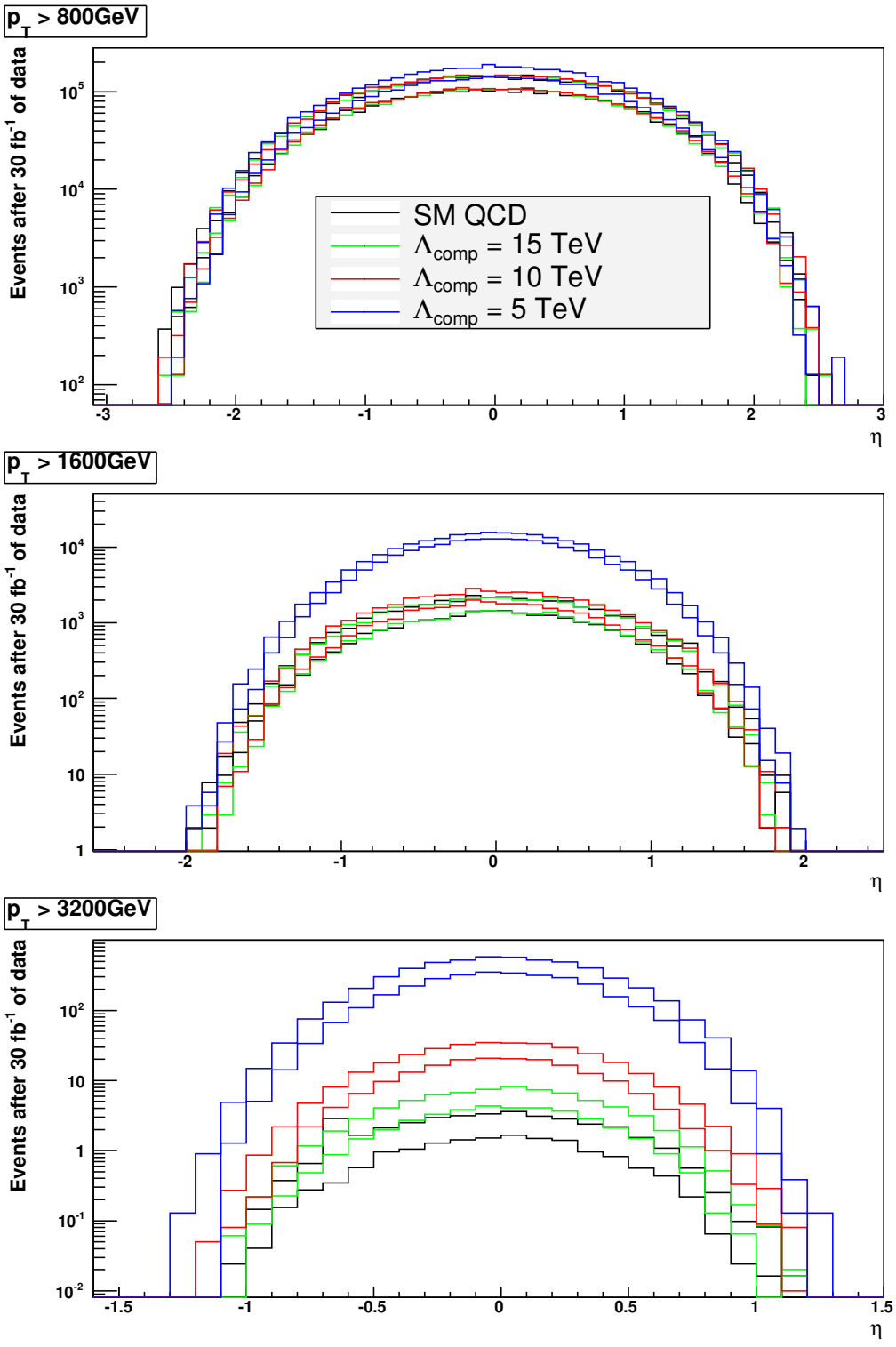

Figure 12.2: Pseudorapidity jet distributions for the SM and three compositeness models after three different $p_\perp$ cuts, showing the effects of a $\pm 3\%$ calorimeter uncertainty.

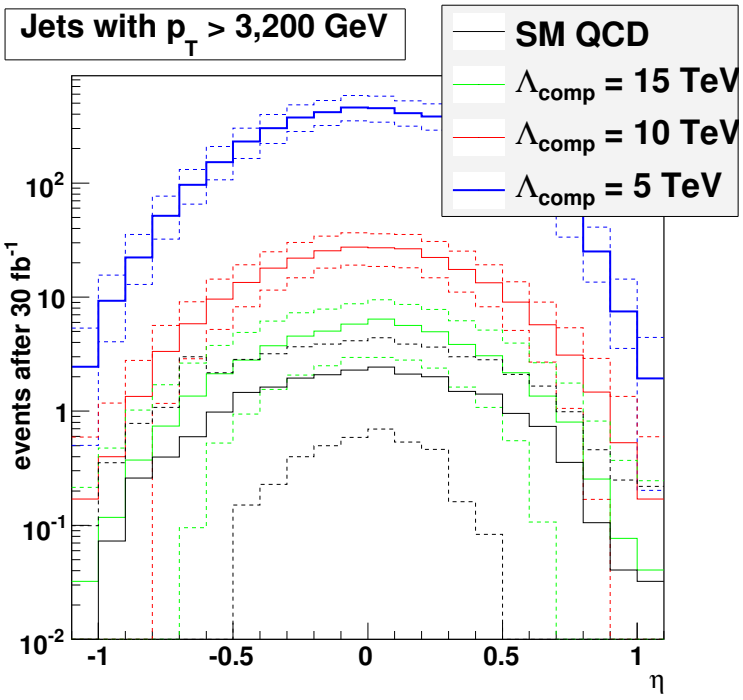

Figure 12.3: Jet pseudorapidity distributions for the four models, showing effects of a ±3% calorimeter uncertainty combined with statistical uncertainty for 30fb^{-1} of data.

Bibliography

[1] Francis Halzen, Alan D. Martin, 'Quarks & Leptons: An Introductory Course in Modern Particle Physics', published by John Wiley & Sons, 1984.

[2] David J. Gross, 'Twenty Five Years of Asymptotic Freedom': hep-th/9809060.

[3] ZEUS Collaboration, 'An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment', DESY-**050-50**; 2005.

[4] The Coordinated Theoretical/Experimental QCD collaboration: http://www.phys.psu.edu/∼cteq/

[5] ATLAS Collaboration, 'ATLAS Inner Detector Technical Design Report', CERN, 1997, CERN-LHCC-97-16.

[6] ATLAS Collaboration (Calorimetry and DAQ), 'Hadron Energy Reconstruction for the ATLAS Calorimetry in the Framework of the Non-parametrical Method', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **Volume 480, Issue 2-3**, 508-523.

[7] ATLAS Collaboration, 'ATLAS Muon Spectrometer: Technical Design Report', CERN, 1997, CERN-LHCC-97-22.

[8] ATLAS Collaboration, 'ATLAS Level 1 Trigger: Technical Design Report', CERN, 1998, ATLAS-TDR-12.

[9] ATLAS Collaboration, 'ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report', CERN, 2003, CERN-LHCC-2003-022.

[10] 'GEANT4 - a simulation toolkit', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **Volume 506, Issue 3**, 250-303.

[11] Elzbieta Richter-Was, 'Atlfast 2.0 a fast simulation package for ATLAS': ATL-PHYS-98-131.

[12] Edward E. Smith, *Skylark Three*, serialised in *Amazing Stories*, 1930, later published in book form by Fantasy Press (1948).

[13] A. Salam, J.C. Pati, *Physics Review*, **D 10**, 275, 1974.

[14] 'Combination of CDF and D0 Results on the Mass of the Top Quark': hep-ex/0608032.

[15] Particle Data Group 2008.

[16] Haim Harari, 'A SCHEMATIC MODEL OF QUARKS AND LEPTONS', SLAC-PUB-2310, 1979.

[17] 'Quantum Gravity and the Standard Model', Sundance O. Bilson-Thompson, Fotini Markopoulou, Lee Smolin, hep-th/0603022, 2006.

[18] Quoted in New Zealand's *Sunday Star Times*, 1911.

[19] S. Jain, A.K. Gupta and N.K. Mondal, 'Search for quark-lepton compositeness at Fermilab Tevatron and CERN LHC', Physical Review D, Vol **62**, 095003; 2000.

[20] A.A. Pankov, N. Paver, 'MODEL-INDEPENDENT LIMITS ON FOUR-FERMION CONTACT INTERACTIONS AT LC WITH POLARIZATION', IC/98/40, 1998.

[21] D. Stump, J. Huston, J. Pumplin and W.Tung, 'Inclusive Jet Production, Parton Distributions and the Search for New Physics', hep-ph/0303013v1, 2003.

[22] R.K. Ellis and J.C. Sexton, 'QCD RADIATIVE CORRECTIONS TO PARTON-PARTON SCATTERING', *Nuclear Physics*, **B269** 445;1986.

[23] The Les Houches Accord PDF Interface, http://hepforge.cedar.ac.uk/lhapdf/

[24] M. Bötje, 'QCDNUM version 16.12', preprint ZEUS-**97-066**; 1997.

[25] QCDNUM, a QCD evolution program by M. Bötje: http://www.nikhef.nl/∼h24/qcdnum/

[26] Yu. L. Dokshitzer, *Sov. Phys. JETP* **46**, 641; 1977.

[27] V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys. 15*, **438**, 675; 1972.

[28] G. Altarelli and G. Parisi, Nucl. Phys. **126**, 297; 1977.

[29] R.S. Thorne and R.G. Roberts, 'A variable flavour number scheme for charged current heavy flavour structure functions', *European Journal of Physics*, **C19**, 339-349; 2001.

[30] C. Gwenlan, 'Impact of future HERA data on PDF uncertainties [and] predictions for high-$E_\perp$ jet cross sections at the LHC', presented to the ATLAS Standard Model working group, July 14th 2006.

[31] S. Catani, M.L. Mangano, P. Nason, and L. Trentadue, 'The Top cross section in hadronic collisions' CERN-TH/96-21, hep-ph/9602208; 1996.

[32] E. Eichten, I. Hinchcliffe, K. Lane and C. Quigg, 'Supercollider physics', Rev. Mod. Phys., **56**, No.4, 1984.

[33] P. Chiappetta and M. Perrottet, (in) 'Large Hadron Collider Workshop', eds. G. Jarlskog and D. Rein, CERN 90-10, Vol. II, 806, 1990.