A SEARCH FOR JET EXTINCTION IN THE INCLUSIVE JET p_{T} SPECTRUM IN PROTON-PROTON COLLISIONS WITH CENTER-OF-MASS ENERGY OF 7 TeV AT THE CMS EXPERIMENT

BY KEITH J. ROSE

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Amitabh Lath

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2013

ABSTRACT OF THE DISSERTATION

A Search for Jet Extinction in the Inclusive Jet p_{T} Spectrum in proton-proton Collisions with Center-of-Mass Energy of 7 TeV at the CMS Experiment

by Keith J. Rose

Dissertation Director: Amitabh Lath

The unification of the Standard Model of physics with the theory of general relativity is one of the most salient problems facing theoretical physicists today. A search is presented for strongly-coupled string theory, a possible resolution to this problem. At an accelerator experiment, evidence for this theory could manifest as a shutting off of all Standard Model processes. This phenomenon is referred to as *extinction*, beyond some energy threshold related to an extinction scale M. The search presented here utilizes 5.0 fb^{-1} of data taken in proton-proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. These data were taken by the Compact Muon Solenoid (CMS) experiment, situated at the Large Hadron Collider (LHC). The Standard Model predicts the formation of collinear streams of particles known as *jets*, with one of the highest production cross-sections of any process expected in proton-proton collisions. The transverse momentum distribution of all jets in data is constructed, and compared to that expected by the Standard Model, as well as that expected by extinction physics for various values of M. The standard model and extinction hypotheses are expressed in terms of an explicit calculation of perturbative quantum chromodynamics (pQCD) at next-to-leading order. No significant evidence of an extinction signature is observed, and an upper limit of 3.2 TeV is set on the scale M at 95% C.L.

Acknowledgements

I would like to give thanks to my advisor, Prof. Amitabh Lath, for support and guidance during a long and eventful graduate experience. I would also like to express gratitude to Profs. Eva Halkiadais and John Paul Chou for mentoring, and assistance in constructing and executing an extremely challenging analysis.

I would also like to acknowledge Drs. Vivian O'Dell and Konstantinos Kousouris, and Prof. Klaus Rabbertz for their invaluable help in teaching me the basic techniques used here, and for helping to formulate the framework for this study.

Lastly, I would like to thank all the other young scientists at CERN who made my four years there such an exciting and unique experience.

Dedication

For my family, and their endless support

Table of Contents

List of Tables

List of Figures

Definitions and Conventions

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC Apparatus
CDF	Collider Detector at Fermilab
CERN	European Center for Nuclear Research
CMS	Compact Muon Solenoid
D0	D0 Collaboration
EW	Electroweak
FNAL	Fermi National Accelerator Laboratory (Fermilab)
L1	Level 1 trigger
LHC	Large Hadron Collider
LHCb	LHC-beauty
LO	Leading Order
MSTW	Martin-Stirling-Thorne-Watt
NLO	Next to Leading Order
PDF	Parton Distribution Function
pQCD	perturbative QCD
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
SM	Standard model

Table 1: Acronyms and Abbreviations

GeV	one billion electron volts (1 eV = 1.602 $\times 10^{-19}$ joules)
TeV	one trillion electron volts
fb^{-1}	Inverse femtobarn. Cross-sections in particle physics are typically expressed in barns (1 barn = $10^{-28} m^{-2}$).
GeV/c^2	Common unit for expressing the mass of particles. For simplicity, it is conventional to assume units where $c = 1$, and thus GeV can be used interchnageably to express units of mass, momentum or energy.

Table 2: Units and Conventions

Chapter 1
Introduction and Motivation

1.1 Motivation

The study of physics is, at its heart, an endeavor to understand the world around us at a fundamental level. Likewise, the laws put forth to codify physics represent an effort to communicate this fundamental understanding in the broadest and most elegant manner possible.

Still, there are significant and deeply puzzling questions that remain. The Standard Model of physics explains in great detail the smallest known particles and their various interactions. It has substantial predictive power which has been validated by decades of experimental results. Nonetheless, it does not contain a candidate for *dark matter* or *dark energy*, both of which have been observed by inference from cosmological data, and combined, account for more than 90% of the mass of the known universe. Likewise, the theory of general relativity provides a proven framework for the structure of space-time on the universal scale, but as yet, has not been unified with any potential quantum theory of gravity.

This lack of unification between general relativity and the Standard Model remains one of the most salient open questions in theoretical physics today. There are several theories that put forth a framework by which this question may be answered. The search for evidence of one such theory will be the main focus of this thesis. This search will consider a particular variant of string theory, which is modeled as a direct modification of quantum chromodynamics (QCD). QCD governs all interactions in hadron collisions, such as the proton-proton collisions induced at the beam crossings of the Large Hadron Collider (LHC). In a region of string theory parameter space known as the strong-coupling limit, the various QCD production cross-sections will drop off exponentially beyond some energy threshold, directly related to a model input parameter M. This parameter will be referred to as the *extinction* paramter, as it leads to a complete suppression of all Standard Model processes, in favor of a high-entropy string state. This, as well as other related models, will be discussed in greater detail in Chapter 2.

1.2 Experimental Opportunities

The LHC is one of the largest and most impressive devices ever built for the primary purpose of scientific research. The collaborations which operate the Compact Muon Solenoid (CMS) and other experiments taking data at the LHC consist of thousands of dedicated scientists from dozens of member states. Any search on an LHC-based experiment will have access to the prior work of a number of man-hours unprecedented in particle physics.

This search in particular will measure the production of hard, collinear streams of hadrons known as *jets*, as predicted by the perturbative expansion of QCD. The primary observable will be the transverse momentum (p_{T}) of all jets in the dataset. Hadronic jet production is the physics channel with the highest production cross-section in hard interactions. Therefore, any search utilizing inclusive jet production will have a data sample with large statistics when compared to other searches looking for new physics on the same scale. The general philosophy will be to be as inclusive as possible with respect to data, when designing the analysis strategy. This will ensure the statistical power of the dataset is maximized.

Additionally, the inclusive jet p_{T} spectrum is traditionally a high-priority publication for any accelerator experiment in the first year of data-taking. Between the substantial manpower of the CMS experiment, and the staggering number of jets produced in LHC collisions, the inclusive jet p_{T} observable will be relatively well understood in short order, and therefore fertile ground for a potential search.

Finally, the jet p_{T} distribution may be calculated explicitly from a perturbative expansion of the theory. As a result, the observed spectrum may be compared directly to an absolute theoretical prediction, rather than an estimated prediction from simulation. This substantially increases the power of the search and ravoids any systematic uncertainties associated with simulation.

The experimental apparatus used to measure the energy of these hadronic jets, and the methods by which they are reconstructed are described in Chapters 3 and 4 respectively. The phase space of the analysis and the simulation of the model will be discussed in Chapter 5.

1.3 Experimental Challenges

Despite the advantages afforded the inclusive jet p_{T} observable in this search, there are a number of subtleties that are not inherently obvious. The inclusive jet p_{T} spectrum is exponentially falling, which traditionally carries a certain amount of baggage when performing the statistical analysis necessary for a search. Assuming the nature of extinction physics to rapidly supress

all Standard Model processes, this search will be conducted for an exponentially falling signal *within* an exponentially falling spectrum. Predictably, this is a challenging endeavor using conventional analysis techniques.

Even when assuming the Standard Model, given that the spectrum is exponentially falling, a small systematic uncertainty in the measured jet energy can propogate to large uncertainty in the measured production cross section. Furthermore, the dramatic difference in statistics between the low-p_{T} and high-p_{T} ends of the spectrum lead to difficulties when evaluating the compatability of the data and the model. Extinction physics will manifest first at the high-p_{T} end, where the statistical uncertainty alone will be dramatically higher than at the low-p_{T} end. If a likelihood is used to compare the observed p_{T} spectrum to that expected by the null and extinction hypotheses, it will be heavily constrained by the low-p_{T} data. In turn, if the correlations between systematic uncertainties across the full range of p_{T} are not properly taken into account, the systematic uncertainties will similarly be constrained, possibly to the point of overconstraint. A list of all sources of uncertainty and how the correlations are addressed can be found in Chapter 6, and the results of the evaluation of the likelihood can be found in Chapter 7.

Needless to say, given the predicted behavior of extinction physics to subsume all other processes, both known and unknown, the actual discovery of extinction would also paint a rather gloomy picture for the future of accelerator physics! The implications of this are discussed by Giddings & Thomas (2002).

Chapter 2

Theory and Related Searches

2.1 The Standard Model

The Standard Model is one of the most powerful theories available to physicists. It describes with a high degree of accuracy all known phenomena on the subatomic scale, and has substantial predictive power which has led to the discovery of much of what we currently believe to be the basic building blocks of matter. These fundamental particles are listed in table 2.1.

Fermions							
	Generation 1		Generation 2		Generation 3		Charge
Quarks	Up	u	Charm	c	Top	t	$+\frac{2}{3}$
	Down	d	Strange	s	Bottom	b	$-\frac{1}{3}$
Leptons	Electron	e	Muon	μ	Tau	τ	$+1$
	Electron Neutrino	ν_e	Muon Neutrino	ν_μ	Tau Neutrino	ν_τ	0

Bosons							
Electromagnetic force		Weak force			Strong force		Higgs Field	
Photon	γ	W, Z Boson	$W^{\pm}, Z^0$		gluon	g	Higgs Boson	H

Table 2.1: Table of Elementary Particles

The initial form of the Standard Model was composed in 1967 (Glashow (1961); Weinberg (1967); Salam (1968)) in an effort to unify the electromagnetic (EM) and weak interactions. The photon, which mediates the EM force, is massless. However, the $W^{\pm}$ and Z^0 bosons, which mediate the weak force, have a mass of 80.4 and 91.2 GeV/c^2 respectively (the proton, by comparison, has a mass of 938 MeV/c^2). The Higgs Field (Higgs (1964)) was then introduced as a means of explaining the difference in the masses of the vector bosons. The combined electroweak theory has been precisely tested over the subsequent decades. The fine-structure constant (α), which sets the relative strength of EM interactions, has been measured to 10 significant figures.

Electroweak theory is known to satisfy $SU(2) \times U(1)$ symmetry. An extension of the same

basic principles of symmetry can be used to enfold the strong nuclear force, which binds hadrons together, into the Standard Model. The existing theory of the strong force, quantum chromodynamics (QCD), satisfies $SU(3)$ symmetry. This is the simplest form that allows for the formation of both baryons (composed of three quarks) and mesons (composed of one quark and one anti-quark). As QCD is essential in predicting the formation of the hadronic jets studied in this thesis, it will be described in much greater detail in the next section.

The combined SM has shown extensive predictive power, further validated by the eventual discovery of the bottom and top quarks and tau neutrino. For many years, the Higgs Boson remained as the only particle predicted by the SM which had yet to be directly observed experimentally. The 2012 announcement of the discovery of a new boson with a mass of 125 GeV/c^2 by both general-purpose experiments at the LHC (CMS Collaboration Chatrchyan et al. (2012); ATLAS Collaboration Aad et al. (2012a)) is widely believed to have brought this search to a close.

2.2 Quantum Chromodynamics

QCD theory has been shown to explain the makeup and organization of the "hadron zoo", the wealth of unique particles observed in particle physics experiments. While the proton and neutron were originally considered among the basic building blocks of matter, as the zoo continued to grow it became apparent that even these phenomenally small particles must have some substructure. This next level down is now known to be the described by the quark model. Quarks have spin-$\frac{1}{2}$ and therefore behave as fermions, obeying Fermi statistics.

To allow quarks to exist within the bounds of the proton and other hadrons, they must be assigned an additional quantum number. This state is referred to as " color" and has three possible values, typically labelled as red, green and blue. These values are interchangeable by the $SU(3)$ symmetry group which describes QCD. For each color, there is a corresponding anticolor. A hadron must have zero total color charge, which be achieved with equal amounts of each color, or equal amounts of a color and its anticolor. A net zero color charge is considered white. These designations are purely conceptual and have no bearing on the physical qualities of the quarks themselves.

The color charge is carried by the gluon, the mediator of the strong force. There are eight gluons, each of which carries equal amounts of color and anticolor. At first glance, one would expect there to be six or nine gluons based solely on the combinatorics of the color charges. The existence of 8 gluons arises from the requirements of $SU(3)$ symmetry. If we require the

field to be invariant under an $SU(3)$ transformation:

$$\psi' = \mathbf{U}\psi \tag{2.1}$$

then the operator $\mathbf{U}$ can be written as:

$$\mathbf{U} = e^{i\mathbf{H}} \tag{2.2}$$

where $\mathbf{H}$ is a Hermitian matrix. Any 3×3 Hermitian matrix may be written as a linear combination of the identity matrix and eight orthogonal matrices. The eight non-identity matrices are referred to as Gell-Mann matrices, and in the most common representation of QCD, correspond to the combinations of color and anticolor carried by each type of gluon.

Unlike the photon or weak force carriers, the gluon is self-interacting. Due to this feature, the strong force potential increases with distance. This leads to a phenomenon known as *color confinement* which is unique to the strong force. Quarks cannot be observed individually outside of hadrons, as the increasing potential leads to the creation of quark-antiquark pairs and thus the production of additional hadrons, rather than unconfined quarks.

Given the topic of this thesis, we are most concerned with QCD as it applies to the internal structure of the proton. While the proton can always be assumed to contain three *valence* quarks (two u and one d), the actual content is continually shifting as virtual quark-antiquark pairs resolve and annihilate. A simple diagram is shown in figure 2.1.

The quarks and gluons within a proton are known in accelerator physics as partons, and the content of the proton is likewise described by Parton Distribution Functions (PDFs). These PDFs are used as input for QCD calculations and will be discussed in greater detail in Chapter 5.

2.2.1 Calculation of QCD

The general form for the calculation of cross-sections in accelerator collisions can be written as

$$\sigma = \int \frac{1}{\mathit{flux}} |\mathbf{M}^2| d\Phi \tag{2.3}$$

In this equation, *flux* is the incoming particle flow per unit area and $\mathbf{M}$ is the matrix element for the process being observed. This is can be calculated for each interaction from the corresponding Feynman diagram. The integral covers the entire phase space volume for the process, represented by Φ.

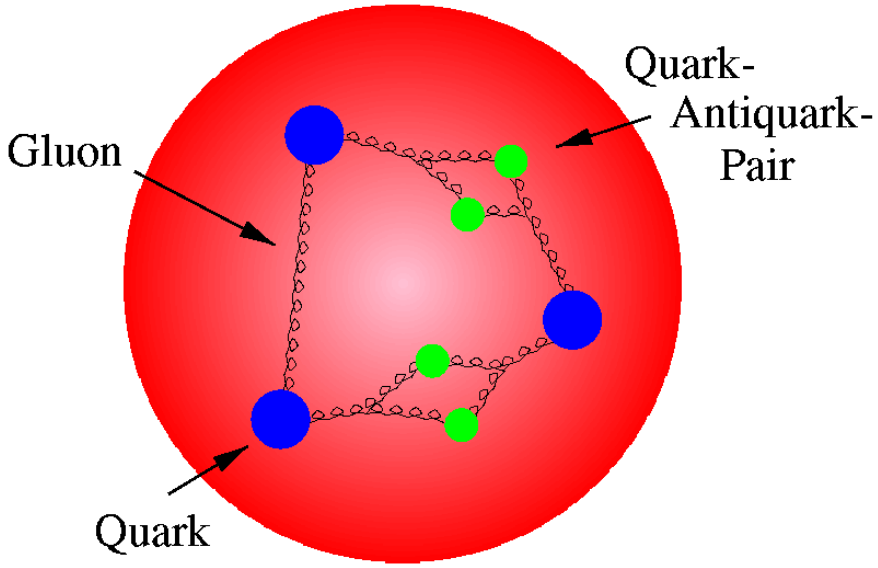


Figure 2.1: Simple diagram of valence and sea quarks within a hadron

The cross-section for QCD can be calculated using an expansion in perturbation theory, dependent on the running of the coupling constant for QCD, referred to as α_s. At leading-order (LO), the calculation considers only the simplestinteractions, in which there are two incoming interacting partons and two final-state partons. These interactions are known as $2 \rightarrow 2$ processes. The Feynman diagrams for these are shown in Figure 2.2 as shown by Voutilainen (2008). In these diagrams, straight lines represent quarks and looped lines represent gluons as per standard convention. This incorporates all possible combinations of four incoming and outgoing partons.

At higher orders, each additional vertex in the Feynman diagram adds a factor of α_s to $\mathcal{M}$. Unlike the QED coupling constant, where $\alpha \approx 1/137$, α_s has been shown experimentally to be greater than one. Normally, this would mean a series expansion in perturbation theory would diverge.

Additionally, the cross-section expansion at higher orders must account for the addition of loop diagrams, which occur when virtual particles appear and annihilate within the interaction. A simple loop diagram where a virtual particle pair is produced in a $q\bar{q} \rightarrow q\bar{q}$ interaction is shown in figure 2.3. The initial and final-state particles have energy restricted by the kinematics of the interaction, and at leading order the integral over the full phase space volume should converge. Virtual particles, however, exist only for a short time and have energy limited only by the uncertainty principle, $\Delta E \Delta t > \hbar$. This is an even greater concern for QCD, where

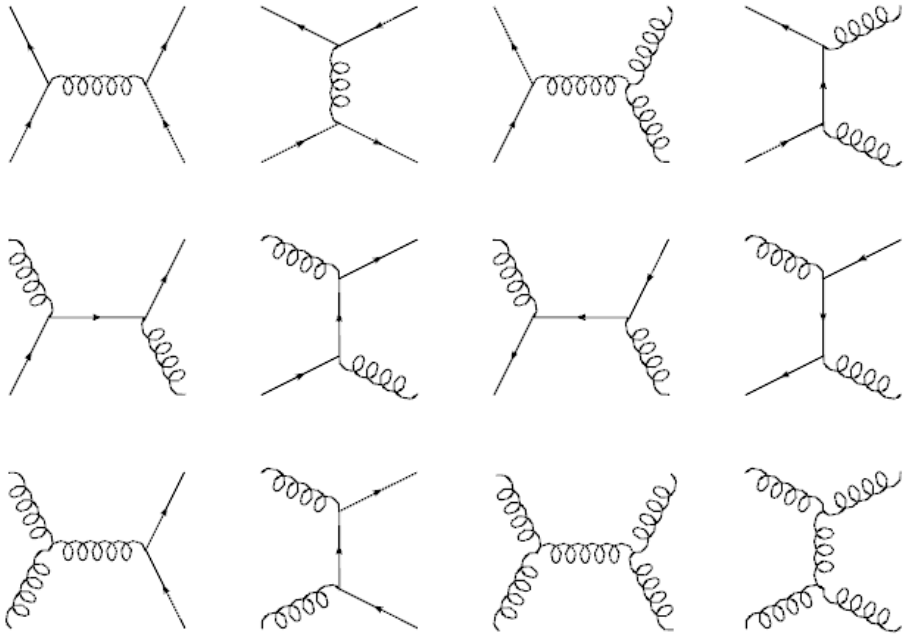


Figure 2.2: All leading order Feynman diagrams for QCD processes.

loops may also form from gluon self-interactions. As the total phase volume for these virtual particles covers an effectively infinite range of values for Energy, the higher-order terms in the cross-section integral will frequently diverge. In perturbation theory, this will lead to a sum containing several infinite terms.

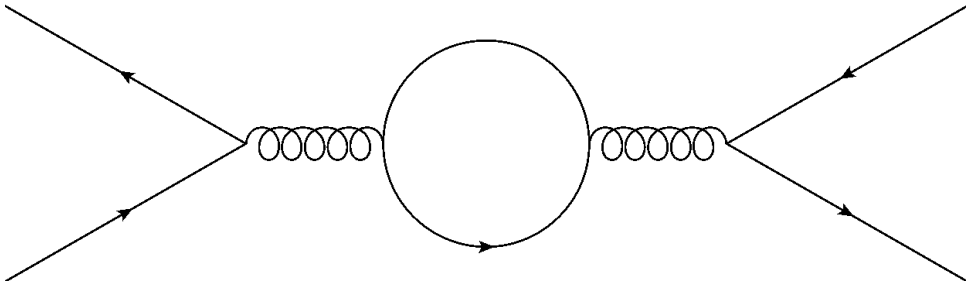


Figure 2.3: A one-loop diagram for $q\bar{q} \to q\bar{q}$ interactions

The solution to this problem is known as *renormalization*, and its application to solve these divergences in electroweak theory is considered one of the remarkable results particle physics in recent years ('t Hooft & Veltman (1972)).

The coupling constant is rewritten in terms of a *renormalization scale* μ_R. This choice of scale can be shown to have no impact on the final calculation, but consipres to yield a convergent result at all finite orders in perturbation theory. When solving for α_s in this case, the coupling constant now also has a dependence on Q^2, the momentum transfer of the interaction.

At leading order in renormalization (as opposed to perturbation theory), the coupling constant becomes:

$$\alpha_s(|Q^2|) = \frac{12\pi}{(11n_c - 2n_f)\ln(|Q^2|/\Lambda^2_{QCD})} \tag{2.4}$$

$$\ln(\Lambda^2_{QCD}) = \ln(\mu^2_R) - \frac{12\pi}{(11n_c - 2n_f)\alpha_s(\mu^2)} \tag{2.5}$$

Here n_c is the number of colors, and n_f the number of quark flavors. Extensive testing of QCD has thus led to significant restrictions on a possible fourth generation of fermions. This calculation of α_s can be shown to be independent of the choice of μ_R by substituting an arbitrary value of μ_R and solving for $\alpha_s(|Q^2|)$. Note that $\alpha_s(|Q^2|)$ grows rapidly at smaller Q^2. Given that small momentum transfer implies a large interaction distance by the uncertainty principle, the increasing potential leading to color confinement can be derived from the renormalized equation for α_s. This quality is known as *asymptotic freedom*. The scale Λ^2_{QCD} constrains the distance within which quarks are confined.

The running of the coupling constant with the energy scale allows for QCD calculations to be performed perturbatively. At the Z mass, $\alpha_s(M_Z) = 0.1184(7)$. Note that at low energy, the perturbative expansion will still be divergent. The point at which this is considered to be true is dependent on the kinematics of the interaction, and referred to as the *factorization scale*, μ_F. Given that the value of α_s is of order 10^{-1}, it is assumed before any calculations are performed that the next-to-leading order (NLO) contribution in perturbative QCD will be non-trivial for even relatively imprecise experimental measurements. As a result, all perturbative calculations used in this thesis are performed at NLO.

While the prediction of QCD must be identical at any scale, a partial calculation at any finite order will be heavily dependent on the choice of both μ_R and μ_F, though both are generally restricted by the energy scale of the interaction. The conventional choice by many large-scale experiments is to set both scales at or near Q. In this search, we will apply this convention, and vary each scale by a factor of 2 to determine the associated systematic uncertainty due to dependence of the pQCD prediction on the choice of scales. The theoretical uncertainty is discussed further in Chapter 6.

Given two initial partons with momentum p_1 and p_2, and final partons with momentum p_3 and p_4, we can write the contributions to the cross-section normalized to the coupling strength. For $2 \to 2$ interactions all containing 2 vertices, each term in $|\mathcal{M}|^2$ will carry a factor of α_s^4. In terms of the Lorentz-invariant Mandelstam variables ($s = (p_1 + p_2)^2, t = (p_1 - p_3)^2, u =$

$(p_1 - p_4)^2$), we have the following amplitudes for all $2 \to 2$ processes (Feynman et al. (1978)):

Process	$\frac{\lvert\mathcal{M}\rvert^2}{\alpha_s^4}$
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{u^2+t^2}{s^2}$
$qq' \to qq'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}$
$q\bar{q}' \to q\bar{q}'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}$
$\bar{q}\bar{q}' \to \bar{q}\bar{q}'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}$
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{u^2+t^2}{s^2} - \frac{2}{3}\frac{u^2}{st})$
$qq \to qq$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} - \frac{2}{3}\frac{s^2}{tu})$
$\bar{q}\bar{q} \to \bar{q}\bar{q}$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} - \frac{2}{3}\frac{s^2}{tu})$
$q\bar{q} \to gg$	$\frac{32}{27}\frac{t^2+u^2}{tu} - \frac{8}{3}\frac{t^2+u^2}{s^2}$
$gg \to q\bar{q}$	$\frac{1}{6}\frac{t^2+u^2}{tu} - \frac{8}{3}\frac{t^2+u^2}{s^2}$
$qg \to qg$	$-\frac{4}{9}\frac{s^2+u^2}{su} + \frac{s^2+u^2}{t^2}$
$\bar{q}g \to \bar{q}g$	$-\frac{4}{9}\frac{s^2+u^2}{su} + \frac{s^2+u^2}{t^2}$
$gg \to gg$	$\frac{2}{9}(3 - \frac{tu}{s^2} - \frac{su}{t^2} - \frac{st}{u^2})$

Table 2.2: Table of matrix element calculations for leading-order processes in perturbative QCD

2.2.2 Parton Distribution Functions

The cross-section in equation 2.3 may be rewritten as an integral over x_1 and x_2, the fractions of the momenta of the initial partons p_1 and p_2 carried by the final state particles:

$$\sigma = \sum_{ij} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \mathcal{M}_{ij}(\alpha_s(\mu_R^2), x_1 p_1, x_2 p_2, \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2}) \tag{2.6}$$

The indices i and j indicate a sum over all possible states of the initial partons. The functions f_i and f_j are the Parton Distribution Functions. At a given value of the factorization scale μ_F, the function $f_i(x_1, \mu_F^2)$ gives the probability of a particle i carrying a fraction of the initial hadron momentum between x_1 and $x_1 + \delta x_1$. There are several groups which offer models for the PDFs, based on empirical fits to data taken at other collider experiments. This analysis prefers the prediction of the MSTW PDF set (Martin, Stirling, Thorne and Watt), due to general compatibility between our statistical methods and those used by the MSTW group to

express uncertainty due to the empirical fits and variations in μ_R and μ_F. The PDFs at two values of Q^2 as given by the MSTW2008 set (Martin et al. (2009)) are shown in figure 2.4.

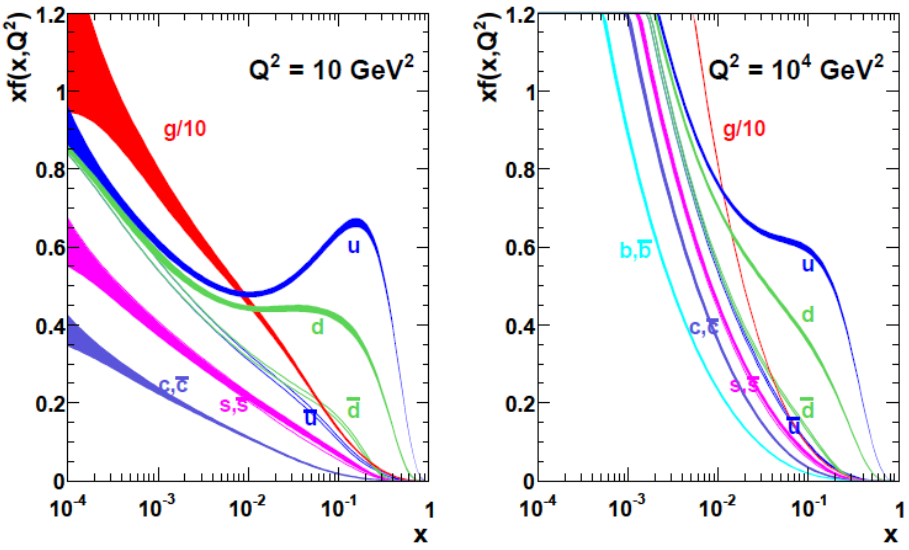


Figure 2.4: MSTW2008 Parton distribution functions for various partons at $Q^2 = 10\,\mathrm{GeV}^2$ and $10^4\,\mathrm{GeV}^2$

2.2.3 Showering

The measurement of the parton distribution functions and perturbative calculation of QCD matrix elements give us a fairly accurate picture of the relative amplitudes of $2 \to 2$ and $2 \to 3$ scattering processes in hadron collisions. However, the experimental picture is not quite as clean. A simple representation is seen in figure 2.5.

When struck in a collision, the initial partons will radiate gluons. Due to asymptotic freedom, at the energy scale of a typical collider these gluons will behave as free particles. This process is known as *showering*. These gluons emit further gluons and quarks, producing a cone of soft radiation around the initial parton. While showering may be described perturbatively, it is generally described by an approximation. An explicit calculation of pQCD including showering would require substantial computing resources.

2.2.4 Hadronization and Jet Formation

The showering process continuously lowers the energy scale of the emitted particles. Eventually, the constituent quarks and gluons are of sufficiently low energy that they may no longer be treated as asymptotically free. In a process known as *hadronization*, the individual particles

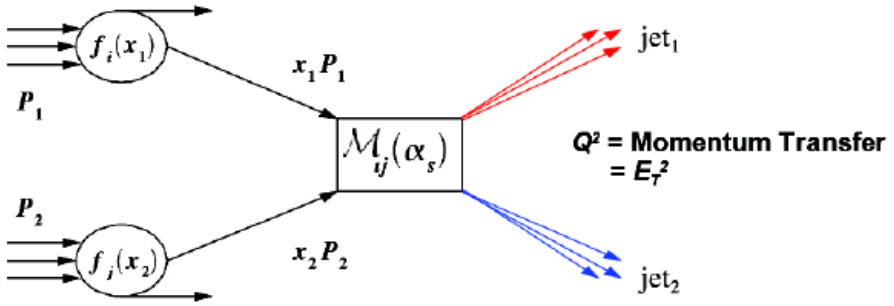


Figure 2.5: Simple diagram of dijet production in $2 \rightarrow 2$ hard scattering in hadron collisions

will combine into hadrons as the strong force becomes prevalent again. By the time these particles are detected, they will take the form of a collinear stream known as a *jet* with total energy roughly equal to that of the original final-state parton. As the scale between hadronizing partons is sufficient to create new hadrons, it is clearly in the regime where α_s is relatively large, and perturbation theory cannot be applied.

The analysis in this thesis will compare the measured jet transverse momentum distribution at the CMS detector to that expected by the MSTW2008 PDF set. It is preferable to have a direct theoretical prediction to compare the observed dataset to. The development of the model and the analysis strategy must be carefully formed to allow this comparison to remain meaningful. Hadronization must be accounted for in such a way that the measured jet observables can be traced back to the final-state partons of a given interaction. The method by which this search approaches this problem is described further in Chapter 5.

2.2.5 Recent Measurements of QCD

Electron-Proton Scattering at HERA

The most precise measurements of the PDFs for the proton come from data taken at the Hadron Electron Ring Accelerator (HERA), operated at the Deutches Electronen-Synchrotron (DESY) in Germany. The structure of the proton is probed in electron-proton collisions at $\sqrt{s} = 320\,\text{GeV}$. Four detectors (H1, ZEUS, HERMES, and Hera-B) collect data from charged- and neutral-current deep inelastic scattering (DIS). The combined measurement of the neutral-current DIS e/p cross-section is shown as a function of x and Q^2 in figure 2.6 (left), as published by Aaron et al. (2010).

This and other data taken by the detectors and fixed-target experiments are used as input to the HERAPDF1.0 PDF set, as shown for the up quark, down quark, sea quarks and gluons

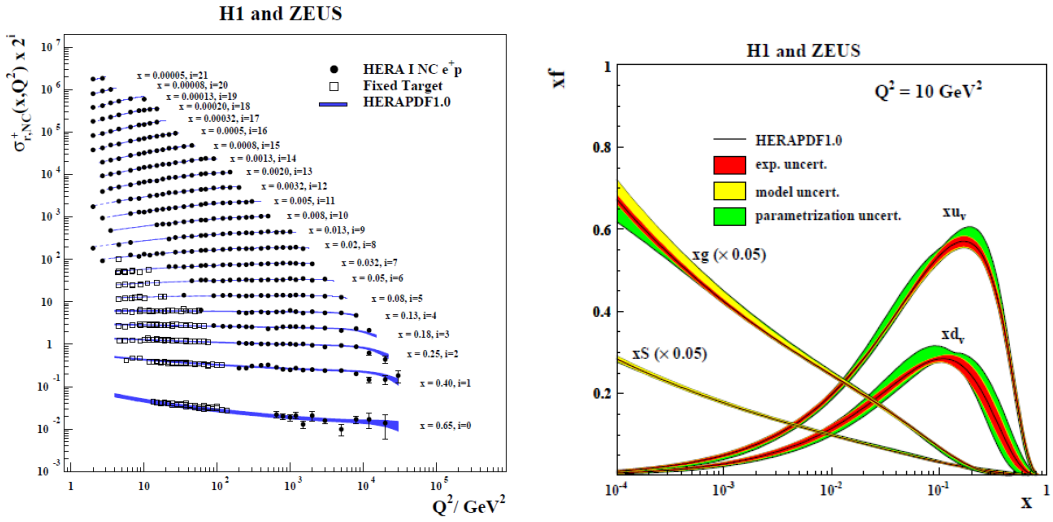


Figure 2.6: Combined H1 and ZEUS neutral-current deep inelastic scattering e/p cross-section (left) and subsequent PDFs (right)

in protons in figure 2.6 (right). Data from HERA is also used by other PDF groups to formulate their own predictions.

Selected QCD Measurements at the TeVatron

Prior to the startup of the Large Hadron Collider, the record for center-of-mass energy in a collider experiment had been set by the TeVatron at Fermilab. Collisions at the TeVatron can be split in two broad run periods; Run I at $\sqrt{s} = 1.8\,\text{TeV}$, and Run II at $\sqrt{s} = 1.96\,\text{TeV}$. Until 2010 the most energetic jets ever observed at a collider were seen at the two general-purpose experiments at the TeVatron, operated by the CDF and D0 collaborations. Both collaborations separately published (Aaltonen et al. (2008); Abazov et al. (2012)) the double differential measurement of the inclusive jet cross section in p_T and rapidity, shown in figure 2.7.

The differential in rapidity may be used to separate central and forward jets, which tend to have very different values for the overall systematic uncertainty and come from different scattering processes. The angular distribution of jets may also be used as evidence of new physics that produces no additional observable decay signatures, but creates a decorrelation between the momentum of the leading jets in a single event. The inclusive jet cross-section can be used as independent confirmation of the predictions of various PDF sets, and data from the TeVatron is already being incorporated in future iterations.

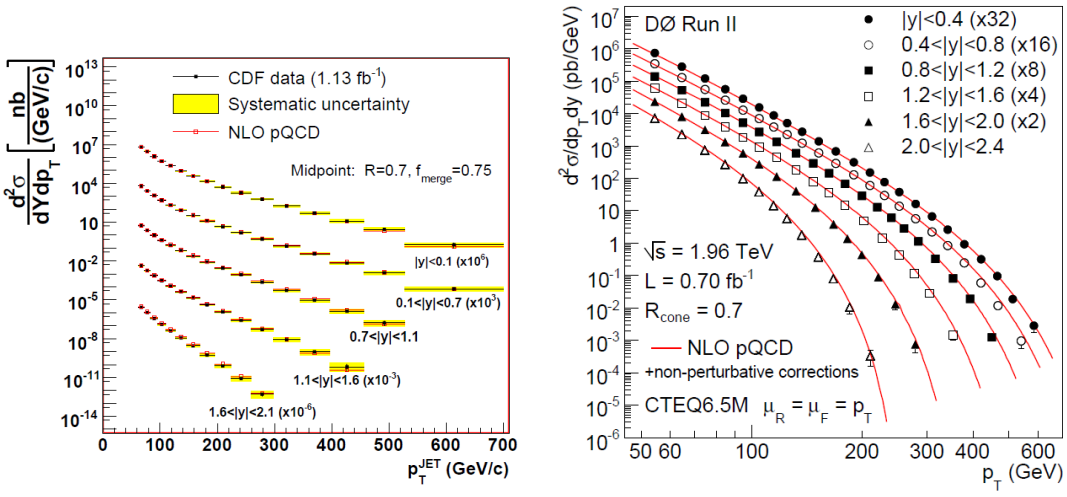


Figure 2.7: Double differential inclusive jet cross-section in p_{T} and rapidity, as measured by the CDF (left) and D0 (right) experiments

Selected QCD Measurements at the LHC

While these measurements are superficially very similar to measurements of the same observables at LHC experiments, the TeVatron is a proton-antiproton collider. This leads to different relative cross-sections of the various QCD scattering processes between TeVatron and LHC collisions, and therefore different production cross-sections for a variety of physics processes. An initial state of $q\bar{q}$, for example, requires at least sea quark at the LHC but can be produced with 2 valence quarks at the TeVatron.

As jets are produced in massive amounts in hadron collisions, the corresponding jet production cross-section measurement was a high-priority analysis for both general-purpose experiments at the LHC. Both the CMS and ATLAS collaborations published the double-differential jet production cross-sections on very small datasets taken during the initial run of the LHC in 2010 (CMS Collaboration Chatrchyan et al. (2011a); Aad et al. (2012)). Given the much larger center-of-mass energy of the LHC ($\sqrt{s} = 7\,\mathrm{TeV}$ in 2010 and 2011), one would expect the reach in p_{T} of the measurement to easily exceed that of the TeVatron experiments.

The published measurements are shown for both experiments in figure 2.8.

The integrated luminosity of the data sample used by each experiment is less than 3% of that used in the CDF measurement, which uses a slightly larger dataset than the D0 publication does. Nonetheless, the D0 measurement indicates a reach in p_{T} of 600 GeV/c, while the LHC measurements already show single jets with momentum on the TeV/c scale.

Other high priority jet measurements include the dijet invariant mass, taken from the mass of the two leading jets which can be assumed to come from an initial $2 \to 2$ interaction. The

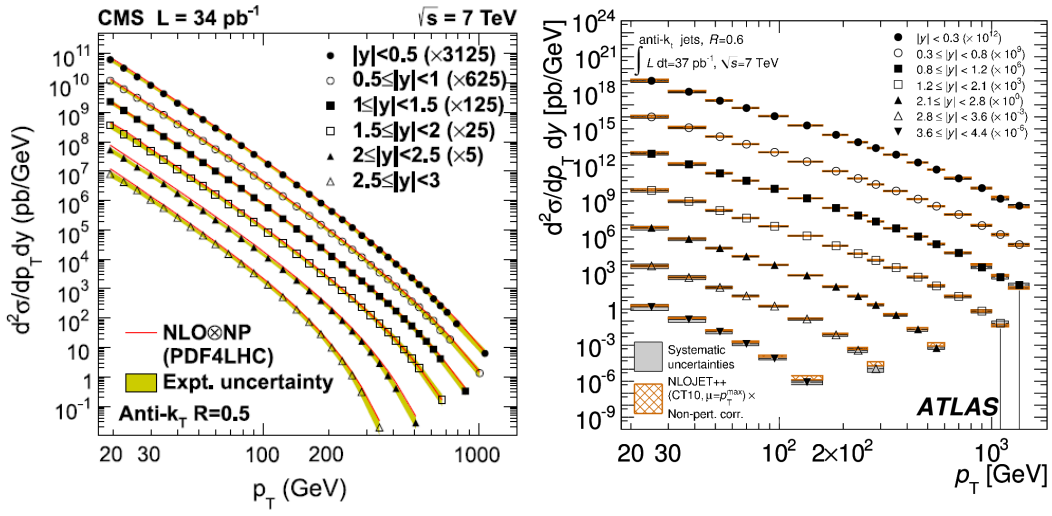


Figure 2.8: Double differential inclusive jet cross-section in p_{T} and rapidity, as measured by the CMS (left) and ATLAS (right) experiments in data taken in 2010.

dijet mass spectrum is predicted to be smooth. New physics at a specific mass point that decay to two jets will show up as a resonance peak within the mass spectrum, which is not expected by any SM processes. A "bump hunt" in the mass spectrum can be equated to a test of an extensive list of hypotheses for physics beyond the SM. The spectrum as observed the first fb^{-1} of data taken at the CMS detector (CMS Collaboration Chatrchyan et al. (2011b))is shown in figure 2.9. The corresponding measurement in a comparable amount of data at the ATLAS detector (ATLAS Collaboration Aad et al. (2012b)) is shown in figure 2.10.

A smooth fit to the observed spectrum is shown, along with the significance of deviations by the data from the fit. A significant deviation may be evidence of a resonance at that mass. Overlaid on the CMS version of the figure are the expected signal events assuming excited quark states and string resonances at various masses.

The search for string resonances in the dijet mass spectrum illustrates the type of analysis most physics experiments are optimized for. A pair of isolated objects likely to be jets are identified and cleaned, and their invariant mass reconstructed. If an excess is observed well above the predicted value, it may indicated evidence of new physics. Even the inclusive measurements presented above, as a whole, simply involve measuring the distribution and frequency of energy deposits that may be measured in a sufficiently advanced calorimeter.

The subsequent analysis presented in this thesis intends to explore a unique region of phase space where the signature is a deficit, rather than an excess of isolated objects. This search is intended to be complementary to existing searches for evidence of similar models, which will be discussed in the next section.

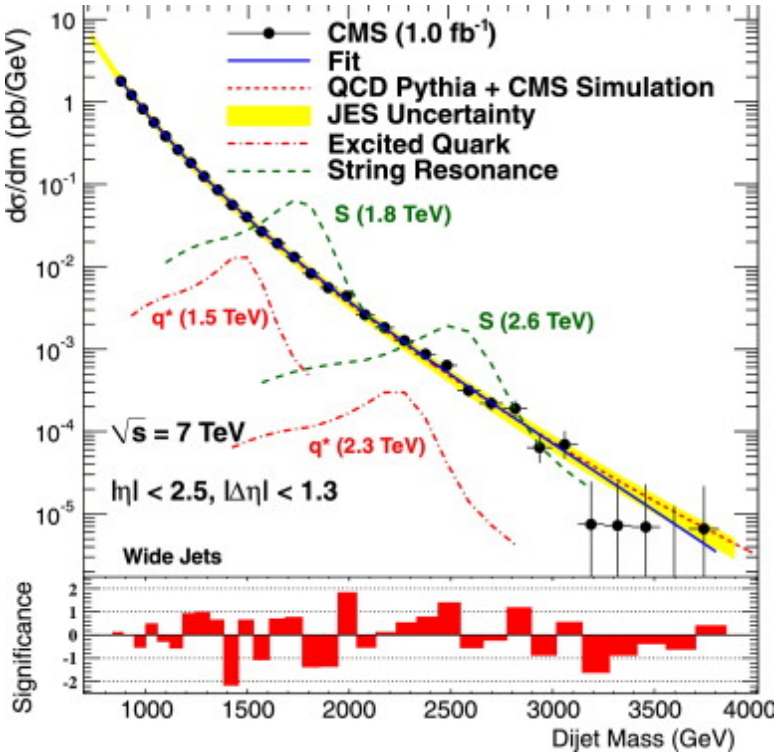


Figure 2.9: Differential dijet invariant mass spectrum as observed at CMS detector with significance of deviations from a smooth fit, overlaid with expected signal from string resonances and excited quark states.

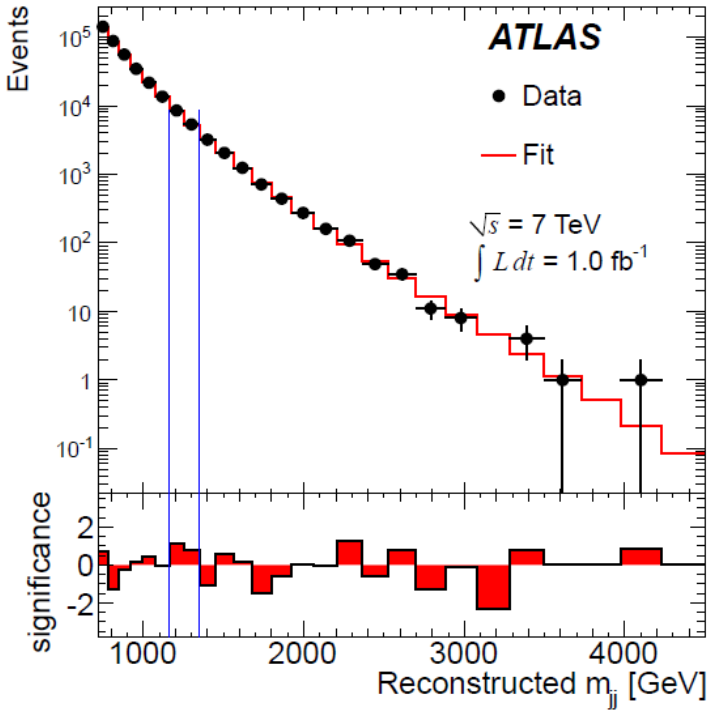


Figure 2.10: Dijet invariant mass distribution as observed at the ATLAS detector, with significance of deviations from a smooth fit.

2.3 Beyond the Standard Model

The difference in scale between gravity and the forces described in the Standard Model is a prime example of what is known as a *hierarchy problem* in particle physics. Such a problem is evident when the corresponding parameters for similar mechanisms vary by many orders of magnitude for no obvious reason. The large difference in mass between the photon and the weak force gauge bosons, for example, has been explained by the Higgs field.

The scale at which gravity becomes strong is known as the *Planck scale* $M_{Pl} \simeq 10^{19}\,\mathrm{GeV}$. It can be determined from the corresponding *Planck length*, the smallest distance which can be observed directly. Any photon with half-wavelength at or below that length scale will possess sufficient energy density that it will immediately collapse into a black hole. Clearly, any possible electromagnetic interactions will now be dominated by the gravitational singularity instead. Additionally, as information cannot escape a black hole, no observation may be performed on this state. As renormalization depends on the observation of the running of some constant (such as the electron charge, or α_s), this implies that any theory of quantum gravity beginning with general relativity is not renormalizable above the Planck scale.

Modern accelerator experiments are currently equipped to probe the electroweak scale ($M_{EW} \simeq 10^3\,\mathrm{GeV}$), many orders of magnitude below the Planck scale. Other than measurements of the cosmic background radiation, no similar experimental apparatus exists that tells us anything about the Planck scale. Furthermore, it is somewhat naive to assume that there are no additional modifications to the laws of physics in the 16 orders between M_{EW} and M_{Pl}. Given this large span, there is a correspondingly large phase space of hypotheses that suppose some evidence of strong gravity may appear at a scale much lower than M_{Pl}. For experimental physicists at the LHC, any consistent model which includes *terascale gravity*, or gravity which is strong close to M_{EW}, is of great interest.

An incidental consequence of the production of singularities above the Planck scale is a subsequent suppression of all low-multiplicity decay signatures. Black holes are expected to decay via Hawking radiation at energies well above the Planck scale. At this point, a semiclassical approximation of black hole behavior becomes valid. In this approximation, low-multiplicty scattering amplitudes such as the $2 \to 2$ processes of LO pQCD will be completely suppressed in favor of the isotropic energy distributions expected by decaying black holes.

This phenomenon is known as *jet extinction*. In the event of terascale gravity detectable at the LHC, it may be possible to search for jet extinction as an attentuation of the jet production cross-section as a function of increasing energy scale. The following sections describe the specific

model used here, as well as the existing phase space probed by other related searches at the LHC.

Giddings & Thomas (2002) refer to the extinction phenomenon as the end of short-distance physics. Beyond the Planck scale, particle accelerators effectively become black hole factories and searches based on the standard model are no longer possible. While there are many searches for model-specific signatures of black holes and other strong gravity effects, the extinction search is the only one ongoing at an accelerator experiment that is not dependent on a specific approximation for black hole behavior. The continued observation of very high-energy jets in scattering events sets an unavoidable limit on the phase space of strong gravity. The goal of the extinction search will then be to develop an analysis strategy with reach comparable to model-specific black hole searches.

2.3.1 Extra Dimensions

One of the earliest attempts to resolve the cosmological scale of gravity with the subatomic scale of electromagnetism resulted in Kaluza-Klein geometry, originally proposed by Kaluza (1921), in which the equations of general relativity are united with Maxwell's equations in a broader 5-dimensional picture of spacetime. The intervening decades have yielded a sequence of increasingly sophisiticated models, taking into account subsequent developments in theoretical physics. Many of these models of quantum gravity call for the existence of extra dimensions beyond the four currently observed.

In each of these models, the additional dimensions are "compactified", rolled up in some very small scale. Rather than a model where gravity is simply very weak, we instead consider one where gravity is strong, but propogates at a greater proportion than other forces in these extra dimensions. If the radius of compactification r_c is much smaller than the scale of a gravitational interaction, then given a number of additional dimensions D, the gravitational flux may be stronger than what we observe in 3+1 spacetime dimensions by a factor of r_c^D. The Planck scale would also be reduced by a corresponding factor.

In some models, the Planck scale can be removed as an independent parameter entirely by relating it directly to the electroweak scale. This has been shown to be possible through the addition of a single, small, "warped" dimension (the Randal-Sundrum model as proposed by Randall & Sundrum (1999)) or by two or more large extra dimensions with compactification scale as large as 1 millimeter (the ADD model, as proposed by Arkani-Hamed et al. (1999)). Many of these models are of great interest to LHC experimentalists, as they predict evidence

of terascale gravity through direct observation of new phenomena.

2.3.2 String Theory

String theory describes a broad class of models, all of which propose the existence of an object known as a string at scales at or above the Planck scale. The various particles of the Standard Model are subsequently assumed to be represented by excited states of these strings. While the number varies between models, these excited states are invariably expressed in extra dimensions through Kaluza-Klein geometry. As strings would exist beyond the Planck scale, there are very few experimental constraints on their parameters beyond those imposed by the Standard Model. String theory can therefore be rendered compatible with many other models, including those that predict terascale gravity. The carrier of the gravitational force, the graviton, is modeled as a string excitation with a large component in extra dimensions, and small or zero amplitude at energies below the Planck scale in the known 3+1 dimensions of spacetime.

Commonly in modern string theroy, these known dimensions are depicted as a surface called a brane. Extra dimensions in this case may be defined as parallel or perpendicular to our universe's brane. Strings may be considered open or closed, as shown in Fig. 2.11 (Lüst et al. (2009)). The extant mass of the known universe is attributed to and governed by strings which are bound to this brane by the boundary conditions of the theory.

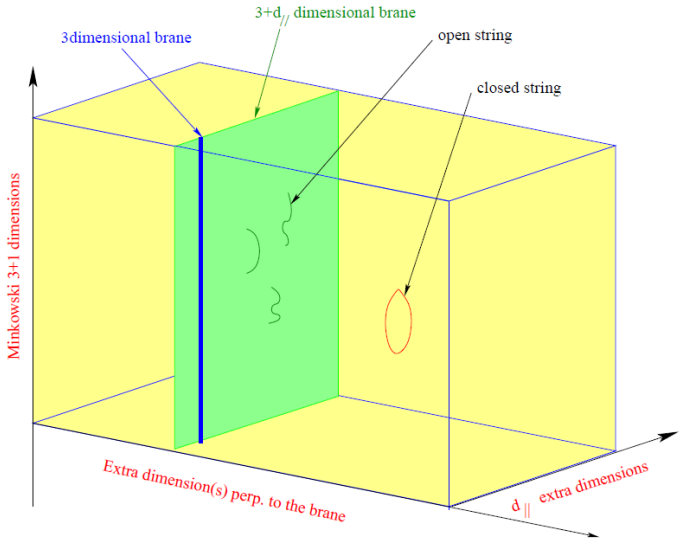


Figure 2.11: Depiction of a brane model as predicted by string theory

As we approach the string scale, we begin to see the presence of string excitations in the scattering processes of pQCD, beginning with LO amplitudes. A perturbative model of string

theory can be written in terms of the Veneziano form factor, as calculated by Veneziano (1968):

$$V(x,y) = \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)} \tag{2.7}$$

where x and y are replaced by Mandelstam variables s, t and u for different $2 \to 2$ subprocesses. The modified LO amplitudes for QCD processes are shown in Table 2.3.

Process	$\frac{\lvert\mathcal{M}\rvert^2}{\alpha_s^4}$
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{u^2+t^2}{s^2}\lvert V(s,t)\rvert^2$
$qq' \to qq'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}\lvert V(t,u)\rvert^2$
$q\bar{q}' \to q\bar{q}'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}\lvert V(s,t)\rvert^2$
$\bar{q}\bar{q}' \to \bar{q}\bar{q}'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}\lvert V(t,u)\rvert^2$
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{u^2+t^2}{s^2} - \frac{2}{3}\frac{u^2}{st})\lvert V(s,t)\rvert^2$
$qq \to qq$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} - \frac{2}{3}\frac{s^2}{tu})\lvert V(t,u)\rvert^2$
$\bar{q}\bar{q} \to \bar{q}\bar{q}$	$\frac{4}{9}(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} - \frac{2}{3}\frac{s^2}{tu})\lvert V(t,u)\rvert^2$
$q\bar{q} \to gg$	$\frac{32}{27}\frac{u^2+t^2}{s^2}(\frac{u}{t}\lvert V(s,t)\rvert^2 + \frac{t}{u}\lvert V(s,u)\rvert^2 - \frac{1}{4}Re(V(s,t)V*(s,u)))$
$gg \to q\bar{q}$	$\frac{1}{6}\frac{u^2+t^2}{s^2}(\frac{u}{t}\lvert V(s,t)\rvert^2 + \frac{t}{u}\lvert V(s,u)\rvert^2 - \frac{1}{4}Re(V(s,t)V*(s,u)))$
$qg \to qg$	$\frac{4}{9}\frac{s^2+u^2}{t^2}(-\frac{s}{u}\lvert V(t,u)\rvert^2 - \frac{u}{s}\lvert V(s,t)\rvert^2 + \frac{1}{4}Re(V(s,t)V*(t,u)))$
$\bar{q}g \to \bar{q}g$	$\frac{4}{9}\frac{s^2+u^2}{t^2}(-\frac{s}{u}\lvert V(t,u)\rvert^2 - \frac{u}{s}\lvert V(s,t)\rvert^2 + \frac{1}{4}Re(V(s,t)V*(t,u)))$
$gg \to gg$	$\frac{9}{4}(\frac{1}{s^2}+\frac{1}{t^2}+\frac{1}{u^2})(s^2\lvert V(t,u)\rvert^2 + t^2\lvert V(s,u)\rvert^2 + u^2\lvert V(s,t)\rvert^2)$ $-\frac{1}{3}(\frac{1}{s^2}+\frac{1}{t^2}+\frac{1}{u^2})(\lvert sV(t,u) + tV(s,u) + uV(s,t)\rvert^2)$

Table 2.3: Table of modified matrix element calculations for leading-order processes in perturbative QCD, including extinction physics

We may write a modified version of Veneziano form factor $\tilde{V}(x,y)$ (Kilic et al. (2012)) in terms of a string scale M and a coupling strength α:

$$\tilde{V}(x,y) = V(\frac{x(1+i\alpha)}{M^2}, \frac{y(1+i\alpha)}{M^2}) \tag{2.8}$$

For values of $M^2 >> Q^2$, the form factor approaches one, and the SM scattering amplitudes are restored. The behavior as we approach the string scale, however, is heavily dependent on α. Small values of α correspond to a region of phase space where gravity becomes strong well above

the string scale. This is known as the weak-coupling limit. The end result is a resonance near the string mass, which may be measured in jet production as a localized excess as projected on Figure 2.9. At a value of $\alpha = 1$, we are in the strong-coupling limit. In this limit, gravity becomes strong at the string scale. Quantum gravity effects subsume standard $2 \to 2$ scattering processes in favor of high-entropy final states. This is consistent with jet extinction, and may be observed as a rapid attenuation of the jet production cross-section as a function of the energy scale. The point at which the cross-section begins to attenuate will be dependent on the string scale M. In the context of a search for the signature of jet extinction, M will be referred to as the extinction scale.

The extinction search can be considered complimentary to searches that look for direct evidence of black hole production and other quantum gravity effects. The following sections will discuss other existing searches at the LHC, and the complementarity of their respective phase space to that of jet extinction.

2.3.3 Related Searches

The extinction search is already complementary to existing searches for string resonances in the weak-coupling limit. Both general-purpose experiments at the LHC (ATLAS Collaboration Aad et al. (2012b), CMS Collaboration Chatrchyan et al. (2011b)) have extensively probed the range of M available to the LHC at its current machine energy. The subsequent searches all look for direct evidence for strong terascale gravity, or decay products of processes that otherwise indicate the presence of strong gravity below the Planck scale.

Semi-classical Black Holes

One of the most striking signatures of terascale gravity can occur through the direct production of black holes in particle collisions. Black holes as astronomical entities are well-understood, and assumed to decay thermally by Hawking radiation. At the scale of accelerator physics, this is still assumed to be true if the mass of the black hole M_{BH} is much greater than the Planck scale M_{Planck}. Dimopoulos & Landsberg (2001) previously calculated the rate at which black holes may be produced at the LHC, as a function of the number of large extra dimensions and M_{BH}. Thermal decay of a black hole leads to a large, isotropic energy distribution that may be easily detected in an accelerator experiment. The total energy deposited may be much higher than that expected by electronic noise. Using standard reconstruction algorithms, the black hole signature will likely also appear as a high-multiplicity final state of very soft objects.

Chatrchyan et al. (2012a) have performed a full search for black holes in LHC collisions, as a function of the number of large extra dimensions n, and the extradimensional Planck scale M_D. Figure 2.13 shows the predicted production cross-section for black holes with $n = 2, 4$ and 6 large extra dimensions, as well as the regions of phase space excluded experimentally. The search includes the low-multiplicity states expected by the semiclassical approximation for black holes, as well as other final states that better describe black hole behavior close to MD. The measurement excludes black holes with a mass below 3.8 to 5.2 TeV, depending on the number of extra dimensions.

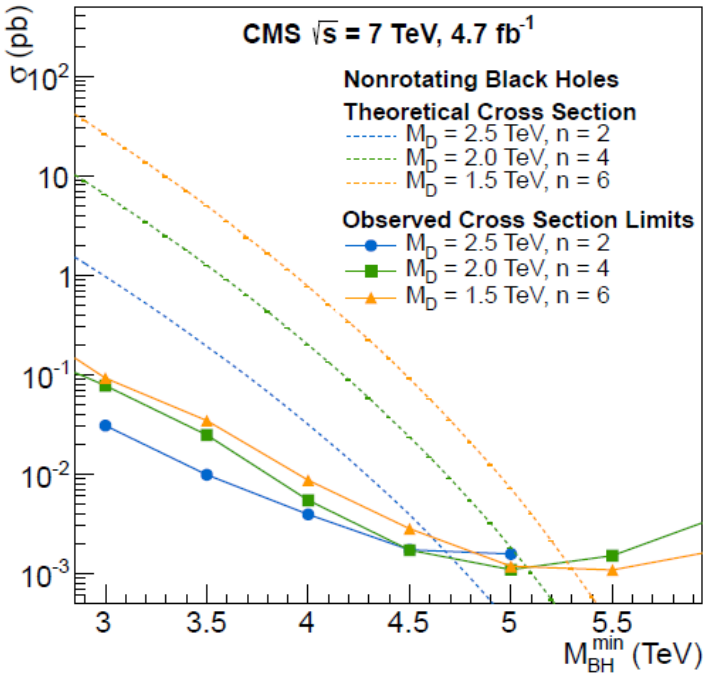


Figure 2.12: Expected theoretical cross-sections for black hole production in 2, 4 and 6 large extra dimensions, and experimental excluded levels of black hole production As

As with extinction physics, the minimum black hole mass is closely correlated to the Planck scale. The two models differ in that the black hole search requires observation of an energetic multiparticle final state not predicted by the Standard Model, while extinction physics requires only a deficit of well-understood SM processes. Model-independent limits in a black hole search cover only these highly isotropic final states. In this sense, the extinction and direct black hole searches are purely complementary in terms of a limit on M_D.

Gravitons

Additional searches for terascale gravity look for direct evidence of the graviton, the carrier particle for the gravitational force. Both the Randall-Sundrum (RS) and ADD models predict a spin-2 graviton which will then decay to 2 gauge bosons. Evidence of gravitons can therefore be seen as a deformation of the expected photon production rate.

The RS model is paramatrized in terms of the mass of the first graviton excitation, M_1, and a modified warp factor $\tilde{k} = k/M_{Planck}$ related to the warped extra dimension of the model. These excitations may be then observed as resonant peaks above the diphoton invariant mass spectrum.

The ADD model is parametrized in terms of an effective Planck scale M_S. Unlike the RS model, it predicts a continuum enhancement of the diphoton mass spectrum beyond a threshold related to M_S. Both models are the subject of published measurements of the diphoton mass spectrum at both CMS and ATLAS (Chatrchyan et al. (2012b); Aad et al. (2012)). Figure 2.13 shows the measured spectra from both experiments. Both figures include an overlay of the expected excesses for resonant and non-resonant models.

The highest excluded values of the effective Planck scale M_S range from 2.3 to 3.8 TeV, dependent on the number of extra dimensions and the parameters of the model. The limit on the mass of the gravition excitations in diphoton searches is currently the hardest limit set on M_1 at either experiment. As with the observation of direct black hole production, the limits on the Planck scale in this case are highly model-dependent and require observation of a specific final state. The extinction search, once again, is complimentary to the diphoton mass measurement.

2.3.4 Conclusion

The search for evidence of terascale gravity has received a substantial amount of attention at the LHC, even after only a single year of data-taking. While no evidence of the phenomenon has been observed, the experimental collaborations have nonetheless refined a wide variety of techniques for probing the limits of the Planck scale. The searches for diphoton and high-multiplicity final states are just two of many measurements that could indicate strong gravity on the scale of the LHC. Nonetheless, the overwhelming majority of these searches set a limit on the fundamental Planck scale through the observation of a model-specific final state not expected by the Standard Model. The goal of the jet extinction analysis will be to illustrate that a competitive search for the Planck scale may be performed without a restriction on a particular final state. It is effectively a black hole search that can be conducted without making

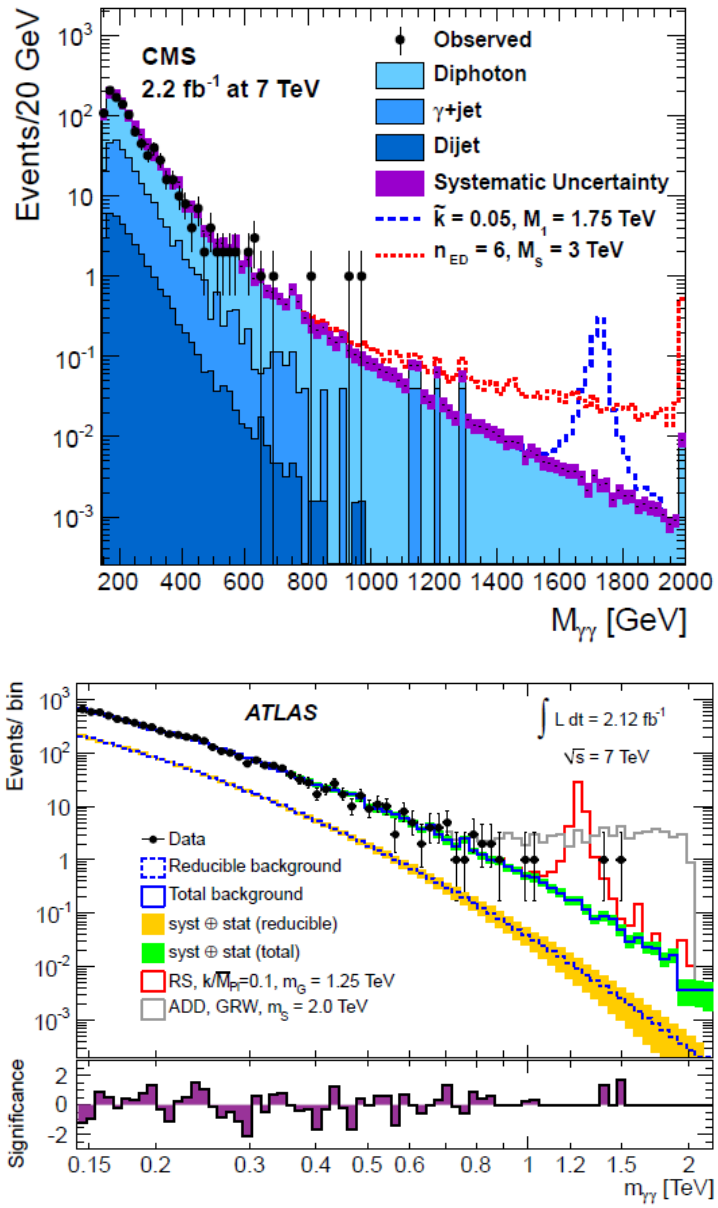


Figure 2.13: Diphoton production cross-sections as measured by CMS (above) and ATLAS (below) in 2011. A resonance from RS graviton production and continuum enhancement from ADD graviton production is overlaid.

any approximations regarding the behavior of black holes themselves.

Chapter 3
Experimental Apparatus

3.1 Large Hadron Collider

The Large Hadron Collider (Evans & Bryant (2008)) is a cyclic particle accelerator 27 km in diameter. It was designed and built by the European Center for Nuclear Research (CERN) and maintains a claim as the largest apparatus ever built for the primary purpose of scientific research. The collider is housed in a cavern an average of 100 meters below the surface, previously used to hold the Large Electron Positron collider (LEP). Most of the collider diameter resides below the French countryside, with a small fraction passing through Switzerland, near Geneva. The cavern is oriented 1.4 degrees off the normal to Earth's radius, mainly to avoid immovable underground structures.

The LHC consists of a two-ring system, with each ring circulating particle beams in opposite directions. The various experiments taking data at the LHC are located at interaction points between beams where collisions occur. The Compact Muon Solenoid (CMS) described by Adolphi et al. (2008) and the remainder of this chapter, and ATLAS, described by Aad et al. (2008) are general-purpose experiments optimized to look for evidence of the Standard Model Higgs Boson in the aftermath of proton-proton collisions. Other experiments at interaction points include LHCb, which is optimized for b-physics, and ALICE, which is optimized to take data on the collision of heavy ions such as lead nuclei. A general diagram of the LHC, including the relative position of the experiments, is shown in figure 3.1.

The LHC was originally designed to accelerate protons to a total energy of 7 trillion electronvolts (TeV), colliding them at a center-of-mass energy of 14 TeV. Protons are inserted into the LHC beam pipe at 450 GeV. A series of 1232 superconducting dipole magnets steer the beam, producing a field of up to 8.33 Telsa (T) to maintain a circular path for a 7 TeV beam. The particles in each beam are then accelerated by a sequence of 8 superconducting RF cavities with a voltage of 2 MV at 400 MHz. For protons, this translates to an addition of 0.5 MeV per proton per cavity. The RF cavities additionally group the protons into bunches spaced 25 ns apart, to be delivered at the experiments.

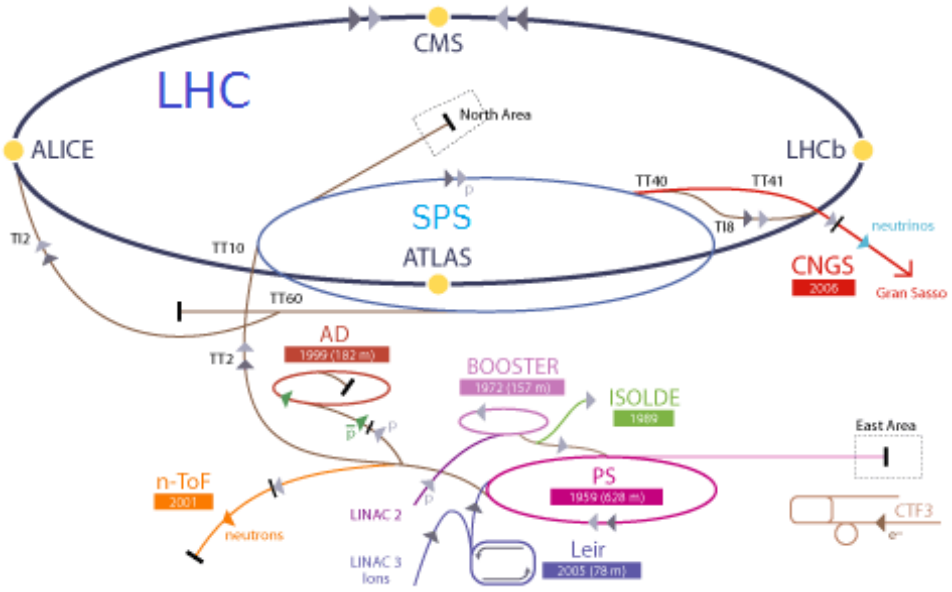


Figure 3.1: General diagram of the LHC, including relative position of experiments and multi-stage beam injection process.

The nominal design calls for the CMS experiment to receive 2808 bunches containing 1.15×10^{11} protons each, colliding at 25 ns with $\sqrt{s} = 14\,\text{TeV}$. This corresponds to an instantaneous luminosity of $\mathcal{L} = 10^{34}\ \text{cm}^{-2}\ \text{s}^{-1}$. However, during the initial attempt to accelerate the beams to 7 TeV, several magnets became non-superconducting. Due to deterioration in copper connecting busbars intended to siphon excess current in this case, parts of the collider overheated resulting in widespread damage. Following repair, subsequent runs have used more conservative beam energies and luminosities. Collisions began in 2010 at $\sqrt{s} = 900\,\text{GeV}$, and continue in 2012 at $\sqrt{s} = 8\,\text{TeV}$. The data used in this thesis represent the entirety of the data taken in proton-proton collisions delivered to the CMS experiment in 2011, all of which was done at $\sqrt{s} = 7\,\text{TeV}$. The instantaneous luminosity increased as the year went on, ending at $\mathcal{L} = 5 \times 10^{33}\ \text{cm}^{-2}\ \text{s}^{-1}$.

3.2 The CMS Detector

The CMS detector is one of two general-purpose experiments currently taking data at the Large Hadron Collider. While it is optimized for the detection of the decay products of the Higgs Boson, there are currently dozens of ongoing searches by members of the CMS collaboration for a variety of hypothesized particles and interactions beyond those predicted by the Standard Model.

The detector comprises several layers referred to as subdetectors, as shown in figure 3.2. The

innermost layer consists of several layers of silicon strips, which provide precise information on the location and direction of charged particles. Beyond this are the calorimeters, which measure the energy of the more stable decay products, including electrons, photons, and various hadrons. These subdetectors all lie within a solenoidal magnetic field, which applies the necessary curvature such that the charge and momentum of charged particles may be observed. Past the magnet in the outermost layers are a sequence of chambers which detect the presence and kinematics of muons. The full apparatus occupies a length of 29 m along the beam axis, with a cross-sectional diameter of 15 m. The coordinate system used in all CMS analyses is right-handed polar, with the z-axis corresponding to the beam line and the origin at the interaction point.

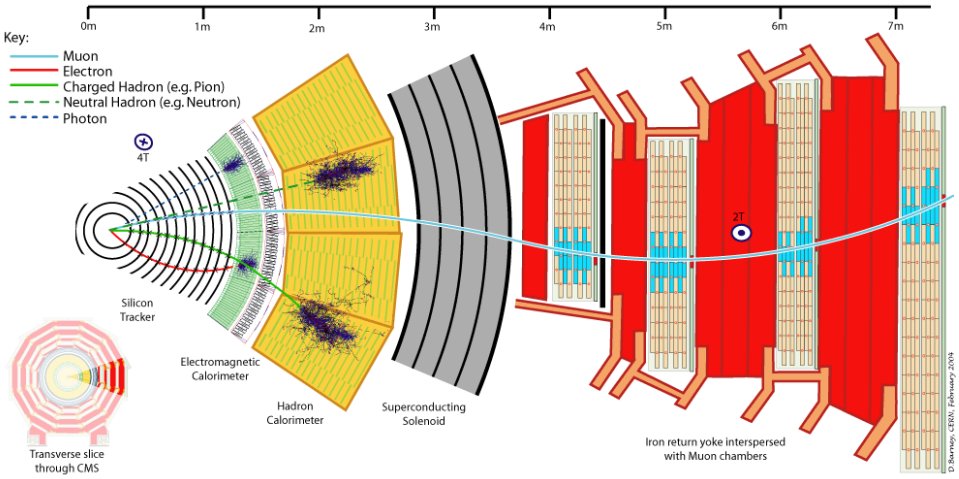


Figure 3.2: Cross-sectional slice of the Compact Muon Solenoid detector

The CMS detector has been designed with regard to the unique challenges of taking data on the products of the collisions of bunches of protons every 25 ns. As the instantaneous luminosity approaches the nominal value, the amount of noise in each event coming from incidental interactions between protons other than direct collisions (referred to as *pile-up*) increases exponentially. The detector firmware must be capable of selecting events likely to contain relevant physics signatures, while reconstructing particles with a high degree of efficiency and precision in the face of increasing pile-up and higher instantaneous luminosity.

As hadronic jets consist almost entirely of charged and neutral hadrons (mostly pions), the silicon tracker and calorimetry are required for their observation and reconstruction. However, the jets themselves will be reconstructed with the Particle Flow algorithm, as described in detail in CMS Collaboration (2010b) and discussed in Chapter 4, which attempts to identify every stable particle in the event. As this requires input from all subsystems of CMS, they will be

described in this section.

3.2.1 Solenoidal Magnet

The solenoidal magnet that envelops the calorimetry and tracking subdetectors is one of the defining characteristics of the CMS experiment. At 12 m long and 6 m in diameter, with a field strength of 3.8 T, it produces the largest closed-volume magnetic field in science today. The cold mass of the magnet consists of 4 layers of winding Niobium-Titanium superconducting cable, internally reinforced with aluminum. Internal reinforcing is preferable to support from external structures, as the width of the cold mass needs to be as thin as possible to minimize interference with particles coming from the interaction point. The magnetic flux is returned via three layers of iron yoke interspersed within the muon subdetectors. The position of the return yoke has the added benefit of shielding the muon system from particles other than muons and neutrinos.

3.2.2 Tracking system

The tracking system consists of two subsystems; an inner pixel detector and an outer strip tracker. Both systems are composed of silicon strips designed to withstand the massive radiation coming from the interaction point without significant loss of efficiency. The tracker is intended to record with extreme precision the paths of charged particles, which can be used to reconstruct their charge and momentum. These tracks are extrapolated back to the interaction point, where the primary vertex of the collision can be reconstructed. Tracking also features heavily in the identification of heavy-flavor interactions, which decay at a secondary vertex displaced from the primary one. In general, tracking is the primary contributor to the resolution of charged particles such as muons with momentum well below the 1 TeV/c scale, where the curvature becomes too small to measure accurately.

The Pixel detector is the innermost layer of the CMS apparatus, and closest to the interaction point. The detector contains 65 million pixels, nearly all of which are 100×150 μm^2. The total combined surface area is roughly 1 m^2. The general structure is shown in figure 3.3. The inner barrel pixel detector contains three cylindrical layers of pixels, situated at 4.4, 7.3 and 10.2 cm from the beam axis respectively. Outside the pixel barrel are 2 endcap pixel rings, at 32.5 and 46.5 cm from the interaction point in each direction. The pixel response is optimized for the measurement of tracks from particles with momentum of at least 1 GeV/c, and the position resolution of tracks is found to be 14 μm. The barrel and endcap provide coverage out

to $|\eta| < 2.5$, beyond which the CMS forward systems take over.

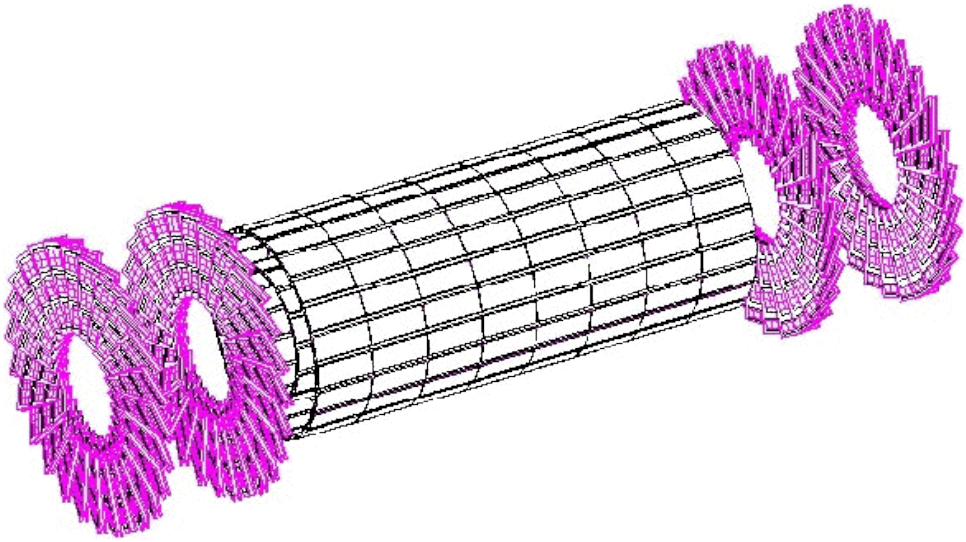


Figure 3.3: Schematic of the CMS Pixel detector.

Directly past the Pixel detector is the much larger Strip Tracker, containing 9.3 million silicon strips with a total active area of roughly 200 m^2. The strips are organized in 4 groups, as shown in figure 3.4. The Tracker Inner Barrel (TIB) contains 4 cylindrical layers, extending out to $|z| < 65$ cm with the outermost layer having a radius of 55 cm. Three Tracker Inner Disks (TID) lie on each side of the TIB, each containing a radial distribution of silicon strips that extend coverage in z to 110 cm. The outer 2 Strip Tracker layers are constructed similarly, with a 6-layer cylindrical Tracker Outer Barrel (TOB) covering out to $r = 116cm$ and $|z| < 118$ cm, and a 9-disc Tracker Endcap (TEC) on each end extending out to $z = 282$ cm. The momentum resolution of the endcap tracker is measured by observing J/ψ decaying to $\mu^+\mu^-$ pairs, and is found to be 2-3%. The residual position resolution of the tracker is 23 to 34 μm in the inner tracker layers and 35 to 53 μm in the outer layers.

3.2.3 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is designed first and foremost to detect the presence of photons decaying from the decay of a Higgs Boson, one of the particle's most fruitful potential channels for discovery. Additionally, electrons produced in collisions will deposit their energy here, and partial energy deposits from charged hadrons can be used to separate them from neutral hadrons when they appear in a jet.

The bulk of the ECAL is composed of lead tungstate ($PbWO_4$) scintillating crystals, which have radiation length $X_0 = 8.8$ mm and Moliere radius $R_M = 2.2$ cm. This allows for a relatively

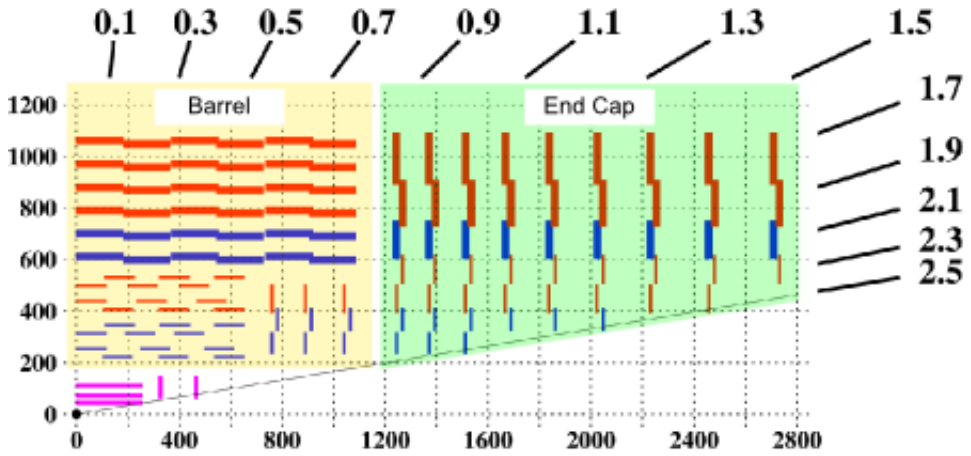


Figure 3.4: Schematic of the CMS Tracker, with coverage in pseudorapidity indicated.

compact design without sacrificing the resolution of the energy measured by the calorimeter. The crystals are arranged totalling 61200 in a barrel region and 7244 in each of two endcaps. The barrel crystals have a granularity of 360 crystals in ϕ and 170 in z, covering a range of $|\eta| < 1.48$. These crystals are further grouped into *supermodules* of 1700 crystals each. Each crystal has a square cross-sectional area 26 mm a side, and a length of 230 mm (26 times the radiation length) with the front face positioned 129 cm from the beam axis.

The granularity in the ECAL endcap is slightly worse, with each crystal having a square cross-sectional area of 28.6 mm a side and length 220 mm (25 times the radiation length). The crystals here are grouped into square *supercrystals* each containing 25 constituents. The endcap extends the coverage of the ECAL out to $|\eta| < 3.0$.

The endcap is preceded by a preshower detector, composed of two layers of lead followed by 2 silicon strip layers. The layers of lead have a combined thickness of 20 cm, three times the radiation length. Th preshower is intended to induce and subtract the decay of neutral pions (π_0), which are produced in large amounts in hadronic jets. As pions decay into photons, sufficiently energetic π_0 particles will produce photons too close together to be properly resolved by the lower granularity of the endcap, and can produce false positives for a Higgs search if allowed to decay inside the crystals. The silicon strip layers are intended to measure the energy lost in the event due to the preshower, and provide a correction.

The light from scintillation is collected using avalanche photodiodes (AVPs) in the barrel and vaccuum phototriodes (VPTs) in the endcap. As many of the components of the ECAL are extremely temperature sensitive, the system must be maintained at a steady temperature with a tolerance of 0.1 Kelvin (K). A schematic of the ECAL is shown in figure 3.5. Barrel supermodules are shown in yellow, endcap supercrystals in green, and the preshower in red.

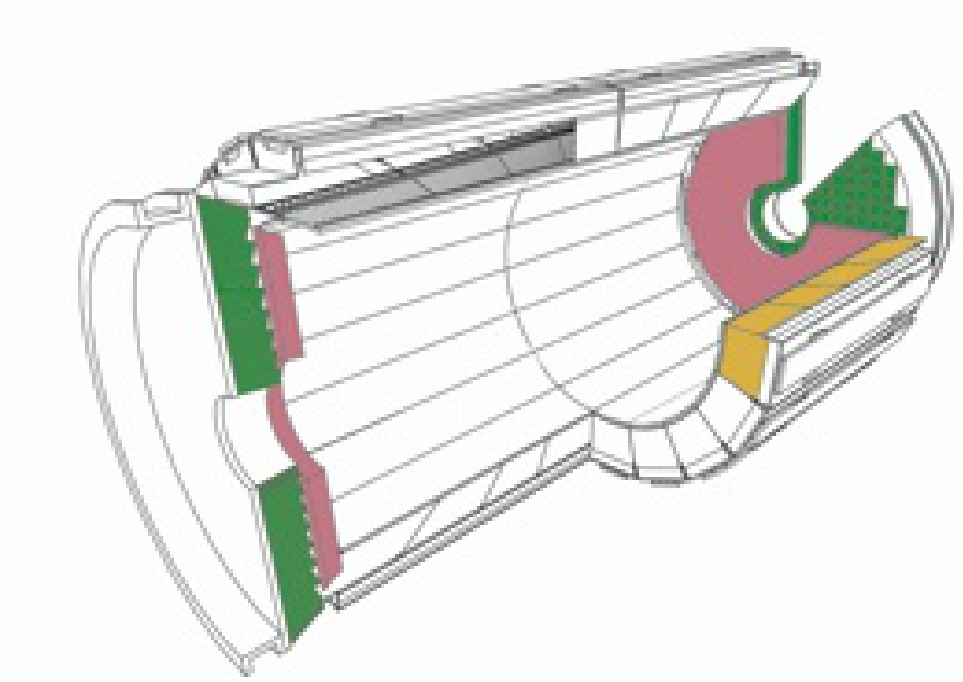


Figure 3.5: Cutout schematic of the Electromagnetic Calorimeter.

3.2.4 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) sits just outside the ECAL and fills the remainder of the space within the magnetic field. It is designed to absorb and measure the energy from hadrons, such as those produced in jets. As the majority of the energy in an event is deposited in the HCAL, it is also necessary for the calculation of missing energy, derived from the vector sum of energy deposits. Given the conservation of momentum, a non-trivial amount of missing energy indicates the presence of particles such as neutrinos that do not interact with any of the detector components. As calorimetry is of great importance to the measurement of the momentum of jets, the resolution will be discussed in detail in Chapters 4 and 6.

The HCAL barrel (HB) covers a region of $|\eta| < 1.4$ and contains 2304 towers. Each tower comprises layers of brass, interspersed with plastic scintillator. The brass is required to be non-magnetic to prevent the apparatus from moving withing the solenoid. Additional HCAL endcaps (HE) sit outside the magnetic field and extend the coverage out to $|\eta| < 3.0$. A steel/quartz forward calorimeter (HF) at $3.0 < |\eta| < 5.0$ combines with the HB and HE to provide near total coverage in $\eta - \phi$ space. As the size of the HCAL is limited by the space constraints of the magnet, an additional outer calorimeter (HO) sits outside the magnet to ensure the total hadronic content of the event is captured. A general diagram of the HCAL and its components is shown in figure 3.6, as found in a publication by Adolphi et al. (2008).

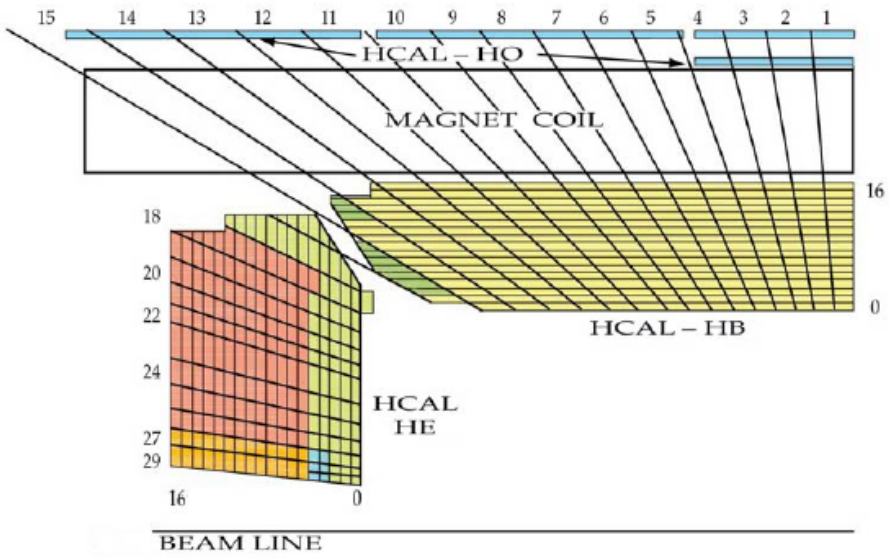


Figure 3.6: Composition and position of the Hadronic Calorimeter.

3.2.5 Muon Subsystems

The design of the muon subdetector is motivated, as with most of CMS, by the prevalence of certain Higgs decay channels. Above a mass of roughly 120 GeV/c^2, the production of the Higgs is dominated by $H \to ZZ \to 4\mu$ and $H \to WW \to 2\mu$ channels. As this mass is easily reachable in LHC collisions, the detector has been largely designed as a whole to leverage this fact. In inclusive jet production, on the other hand, the identification of muons is primarily used to subtract associated tracks, reducing the pool that can be used in reconstructing hadrons.

The muon chambers compose the remainder of the detector beyond the HO, along with the return yoke for magnetic flux. Three different types of gas chambers provide efficient measurement and identification of high-energy muons coming from the interaction point. The muon is longer-lived than the space between bunches and is not completely absorbed by the calorimeter. Thus, the chambers must be able identify muons rapidly to prevent out-of-time muons being associated with the wrong proton bunch crossing.

There are 250 Drift Tube (DT) chambers in 4 layers extending to $|\eta| < 1.5$. Each chamber is filled with 15% Carbon Dioxide and 85% Argon gas. Active wires within the chambers detect charge left behind when passing muons ionize the gas. Outside the DTs, in the region $0.9 < |\eta| < 2.4$, are a sequence of Cathode Strip Chambers (CSCs) containing 13 alternating cathode and anode wire planes. Complementing the two sets are Resistive Plate Chambers (RPCs), in which the ionizing gas lies between two charged plates. The plate structure provides lower resolution than the wire chambers, but faster response time that allows the muon to

be identified within the 25 ns bunch window. The difference in structure between the various systems is also required by the nonuniformity and strength of the magnetic field at higher values of η.

3.2.6 Trigger System and Data Acquisition

Given the bunch spacing of 25 ns, the CMS experiment will collide batches of protons at a rate as high as 40 MHz. At design luminosity, this results in about 20 proton-proton collisions per bunch crossings, or nearly 10^9 collisions per second. It is nearly impossible to record the outcome of every single one of these collisions, and a great many of them contain little of interest to any current physics analysis. Given this situation, the experiment uses a two-stage trigger system to decide which events will be worth future consideration.

The level-one trigger system (L1) selects events based only on information from the muon systems and calorimeter. If a muon, jet, or other object is found in the event that passes a certain energy threshold, the event is passed to the next level. Events containing lower-energy objects may be selected on a prescale, meaning only a certain percentage satisfying the requirements are selected. The trigger system is located 90m from the detector, creating a latency of about 3 μs from the collision. The trigger system can store up to 128 events in a buffer, which is roughly the maximum that can be taken during this latency period. The L1 trigger outputs events at a nominal rate of 100 kHz.

The High Level Trigger (HLT) system then applies a similar selection on fully reconstructed objects using information from the entire dectector. Each HLT trigger is seeded by an L1 trigger with a similar but less restrictive threshold. The HLT event output rate is 100 Hz, with a corresponding data storage rate of 150 MB/s.

Subsets of the triggers are grouped into Primary and Secondary Datasets (PDs and SDs) and recorded to magnetic tape in a raw data format. These trigger subsets are generally selected to be generally useful to physics analysis groups. The dataset used in the extinction search, for example, contains all data recorded on triggers which require reconstructed jets. More advanced offline reconstruction is then performed on trigger-selected data, and made available on tape for general consumption by CMS researchers.

Chapter 4
Event Simulation and Jet Reconstruction

Monte Carlo generators are currently extensively used in accelerator experiments today as a means of studying the response of the detector to a variety of stimuli. These studies are our primary means of understanding the true technical capacity of the apparatus, as well as its viability in terms of differentiating the signatures of new physics from Standard Model processes.

The following sections of this chapter contain the details of the Monte Carlo generators used to produce simulated collision events and model the signatures they leave in the detector. Additional Monte Carlo software separate from the event generators are also used to calculate the parton-level production cross-sections as predicted by perturbative QCD.

Lastly, the algorithms by which energy deposited in the detector (both in simulation and in observation) are reconstructed into jets will be discussed. The end goal will be to connect the output of perturbative QCD as discussed in Chapter 2 to the jet momentua observed and expected at the CMS detector. This will permit the comparison between pQCD and the observed jet p_{T} spectrum to be performed.

4.1 Physics Simulation

4.1.1 PYTHIA

PYTHIA (Sjöstrand et al. (2006)) is a leading-order (LO) event generator in wide use at accelerator experiments. It can be used to generate initial-state distributions for a wide variety of physics signatures, both known and hypothesized. The properties and lifetimes of known particles generally are taken from the standard observed values, but can be set internally to allow for the development of new models. The initial parton distributions may also be specified explicitly on an event-by-event basis through use of a Les Houches (LHE) file described by Alwall et al. (2007). The strongly-coupled string model used in this analysis is not one of the standard PYTHIA routines, and must be generated through these LHE files. This was done by calculating the modifications to the QCD scattering amplitudes by hand and throwing initial

parton momenta accordingly.

More complex states as expected by higher-order interactions may be produced via a process internal to PYTHIA, similar to Brehmsstrahlung. Hadronization is approximated bye the Lund String model described by Andersson et al. (1983), and depicted in figure 4.1.

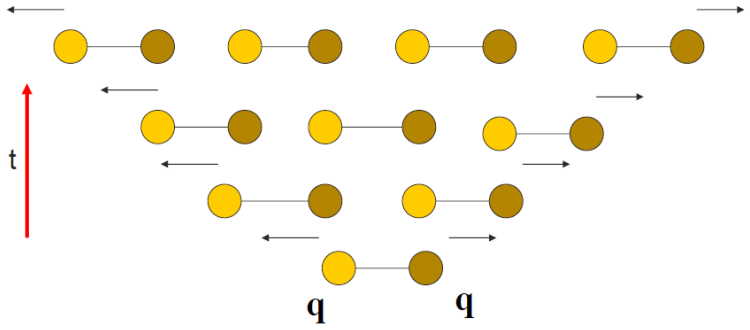


Figure 4.1: The Lund String model as it applies to hadronization within jets

In this model, hadrons are assumed to be connected by a relativistic massless string, with potential that increases linearly with the distance between quarks. At sufficient energy, the strings snap, forming additional hadrons. This continues until the remaining particles are sufficiently stable and at low enough energy to prevent further hadronization. In jets, these particles are almost universially pions ($\pi^0, \pi^{\pm}$).

An additional effect not predicted by perturbation theory is the contribution to a jet by the underlying event. As collisions at the LHC involve large bunches of protons, there are a variety of soft interactions in every bunch crossing which produce particles that obviously are not related to any hard-scattering process. This is accounted for in PYTHIA by a "tune". A given tune is developed by measuring the underlying event in *zero-bias* crossings with no hard interaction and then matching the underlying event generated by PYTHIA accordingly. A tune is generally unique to a specific accelerator.

The tunes most widely used by the CMS experiment are known as "Z2" (Chatrchyan et al. (2011)) and "D6T"(Field (2010)) , though Z2 is generally preferred due to greater agreement with observation. Any Monte Carlo events generated for this analysis will use a Z2-tuned underlying event.

4.1.2 Detector Simulation

Propogation of generated particles through a simulation of the CMS detector is the most computationally expensive part of event generation. As the primary generated particles pass through

the material of the detector, they produce showers that may contain thousands of secondary particles. This includes the products of propogation through both active and dead materials. Agostinelli et al. (2003) describe the software package commonly used to model the detector substructure in simulation.

As this aspect of the simulation is extremely time-consuming, it is often useful to develop an ansatz that describes the detector resolution for general use. As both signal and background MC samples for the extinction search contain only SM QCD events, we will assume the detector response is the same in both cases. The spectrum of generated primary particles will be convolved with a function which models the approximate detector response in jet p_{T}. The uncertainty on the parametrization of this ansatz will be taken as a source of systematic uncertainty and will be discussed in greater detail in Chapter 6.

4.1.3 pQCD Calculations

The MC-generated extinction events and the inclusive jet production as observed in data will eventually be compared to an absolute theory prediction. The theoretical jet production cross-section is typically calculated explicitly using MC integration. The current standard program for doing so is NLOJet++. A second related program, fastNLO, is then used to produce final cross-section distributions (Nagy (2002, 2003); Kluge et al. (2006)).

For a given PDF set, variations on the PDF parameters will be treated as a source of systematic uncertainty when performing a serach. Performing the full cross-section calculation for each variation is extremely time-consuming. Instead, the integration is performed once in NLOJet++ for a grid of values in x and Q^2. These values are then used as weights when performing a cross-section calculation for an arbitrary PDF in fastNLO. The PDF variations will be discussed in greater detail in Chapter 6.

4.2 Particle Flow Jet Reconstruction

The Particle Flow (PF) algorithm (CMS Collaboration (2010b)) is an event reconstruction algorithm designed for high-resolution measurements of the decay products of particles from accelerator collisions. As the name alludes to, Particle Flow aspires to cleanly reconstruct every particle stable enough to pass through or deposit energy in the detector. Previous measurements of inclusive jet observables as recent as those published by the TeVatron (Aaltonen et al. (2008); Abazov et al. (2012)) use jets reconstructed primarily from energy deposits in the detector calorimeters.

In PF reconstruction, this information is enhanced using input from all subdetectors (primarily tracking, for hadronic jets) to improve jet resolution and efficiency, and provide a reduced uncertainty on the jet energy scale. The PF algorithm can also be shown to improve understanding of missing energy in an event, as more comprehensive reconstruction leads to a more complete vector sum of event energy. Additionally, high-multiplicity decays of exotic particles at the TeV scale can have stable final-state particles with energy as low as a GeV. Precise reconstruction of these constituents help to isolate these signal events against the dominating background of Standard Model processes.

4.2.1 Track Reconstruction

It is generally a requirement of any reasonable reconstruction algorithm to maximize efficiency while minimizing the number of false positives, also known as a *fake rate*. It is particularly important, however, for PF tracks to have a very high efficiency with a negligible fake rate. Charged hadrons from jets which are not reconstructed in the tracker will then only be seen in the calorimeter, in which case tracking will be critical to differentiate them from neutral hadrons. Additionally, fake tracks with an arbitrary momentum rapidly pollute the sample and make it extremely difficult to obtain a precise energy resolution.

In order to keep fake rates low and efficiency high, the PF algorithm employs an iterative tracking method (CMS Collaboration (2010e)). Initially, tracks are reconstructed and recorded using very tight identification critieria. The hits in the tracker associated with these tracks are then removed from consideration. With each iteration, the identification criteria loosens slightly. As hits are removed after each iteration, the number of fakes due to combinatorics is greatly reduced. Ultimately, the algorithm correctly identifies 99.5% of all tracks associated with muons, and more than 90% of all tracks associated with charged hadrons from jets.

4.2.2 Calorimeter Clustering

In addition to measuring the raw energy of the constituent particles of jets, calorimetry is also used in the Particle Flow algorithm to differentiate charged hadrons from neutral hadrons. A collinear stream of particles, such as a jet, will leave deposits in both the ECAL and HCAL with very little separation. As the HCAL has lower granularity than the ECAL, additional information is helpful to separate the hadrons belonging to a jet from pile-up or other background particles. Additionally, the information from the calorimeter helps to bound the measurement of the energy of high-p_{T} charged particle. Due to lack of curvature in the magnetic field, these

particles cannot always be accurately measured in the tracker

PF Clusters are seeded about any calorimeter cell with an energy measurement at least two standard deviations above the mean electronic noise for that cell. For central jets, this threshold is about 0.08 GeV in the ECAL and 0.8 GeV in the HCAL.

4.2.3 Linking Clusters to Tracks

From a given collection of clusters and tracks in an event, it is now possible to begin reconstructing individual particles. Clusters and tracks will be linked into blocks, which may then be passed to various high-level algorithms such as those used to reconstruct full jets.

Each track in the collection is extrapolated to the preshower, and to one radiation length in the ECAL and HCAL. A track is linked to a calorimeter cluster if the endpoint of the extrapolated track lies within the cluster boundaries, taking into account cracks and spaces in the calorimeter. Similarly, a cluster in the preshower or ECAL is linked to an HCAL cluster if the cluster if the prior cluster falls within the bounds of the latter. In both cases, a tolerance factor of about one cell is applied to account for cracks and spaces between sections of the subdetector. Additional linking is performed to associate tracks with muons and electrons with potential radiated photons, but these are largely unrelated to jet reconstruction. These linked objects are then passed to a series of particle idenfication algorithms.

4.2.4 Particle Identification

The first stage of Particle Flow identification removes all muons and associated tracks from the collection. A small amount of energy is subtracted from the calorimeter, as measured from observation of cosmic muons. Electrons and radiated photons are subsequently removed from the remaining candidates, and associated tracks are removed from the collection.

The remaining reconstructed deposits will be associated with photons, or with the charged and neutral hadrons. Charged hadrons are separated from neutral hadrons by comparing the energy deposited in the ECAL and HCAL with the momentum of any linked tracks. Only the closest calorimeter cluster to a track's trajectory is used for the separation of hadrons. Clusters with a total energy more than three standard deviations below the momentum of the track are assumed to belong to muons which were not picked up during the initial step of identification. Each remaining track is then associated with a charged hadron.

Any remaining excesses in the ECAL are attributed to photons, and leftover excesses in the HCAL are attributed to neutral hadrons. The reconstructed momentum of the various particles

are derived from calibrations using both cluster energy and track momentum.

4.2.5 Jet Clustering

Once the individual particles in an event are identified, they can now be clusetered into jets. Ideally the properties of this jet will closely match the kinematics of one of the final state partons in the original scattering process. Therefore, the clustering algorithm by which we group observed energy deposits into jets must be relatively insensitive to non-perturbative effects such as hadronization. In particular, the observed momentum of a jet should be invariant if the initial parton radiates a gluon which is soft (infrared) or parallel (collinear) to the jet. These interactions cannot be precisely predicted by perturbative QCD and must be accounted for in jet reconstruction. Clustering algorithms which satisfy these criteria are considered **infrared and collinear safe** (IRC-safe).

The most common algorithm currently used in jet reconstruction is known as the anti-k_{T} algorithm, as defined by Cacciari et al. (2008). The extinction search will use jets reconstructed with this algorithm, to allow synchronization of the inclusive jet observable with other analysis groups looking at the same data.

The anti-k_{T} algorithm is an iterative clustering algorithm. The clustering loop is given a list of all observed particles in an event. For each pair of objects i and j, a distance parameter d_{ij} is defined as:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{4.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the distance between the two objects in rapidity-phi space, k_{ti} is the transverse momentum of particle i, R is a distance parameter, and p is a power which defines the algorithm. A distance d_{iB} between the particle i and the beam B is also defined to be k_{ti}^{2p}. If $d_{ij} < d_{iB}$, the particle is added to the jet. Otherwise, the combined object is labelled a jet, and all constituents are removed from consideration for the next iteration. This process continues until all observed particles are clustered.

The distance parameter R determines the overall size of the reconstructed jet area, and is commonly referred to as the *cone size*. The power p sets the sensitivity of the algorithm to soft radiation relative to the distance Δ_{ij}^2. In the anti-k_{T} algorithm, $p = -1$. Nearly all analyses conducted at the CMS experiment set $R = 0.5$ or 0.7. The extinction analysis sets $R = 0.7$ in reconstruction. The larger cone size is preferred as it increases the statistics of the sample and

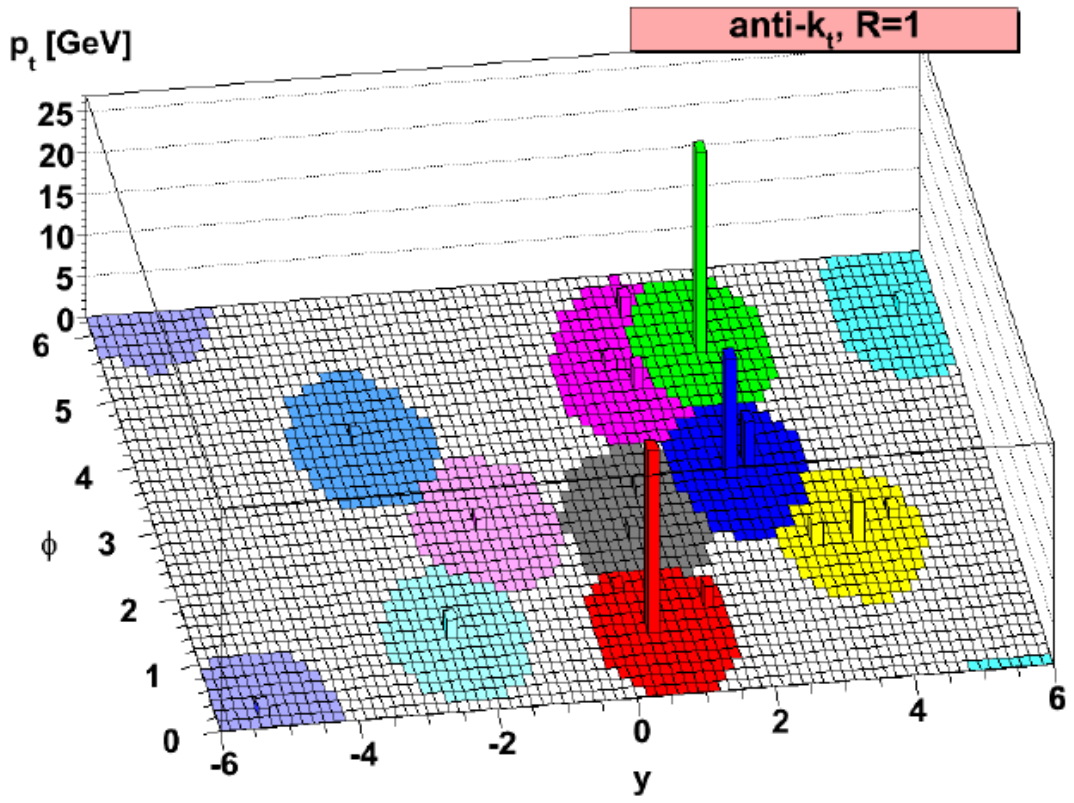


Figure 4.2: Map of energy deposits in an $anti-k_{\mathrm{T}}$-reconstructed event with $R = 1$ for a simple scattering event. Shaded regions represent particles associated with a single jet.

reduces the loss of energy due to possible out-of-cone effects.

An isolated jet consisting of a hard parton and clusters coming from soft radiation, with no other objects of comparable momentum within a distance $2R$, will be perfectly conical. Two hard particles with $\Delta_{ij} < R$ will be clusetered into a single jet. The intermediate case where two hard particles are observed with $R < \Delta_{ij} < 2R$ will yield 2 reconstructed jets, with nearby soft particles within a radius R of each hard particle split between the two jets. The final reconstructed jets will be *soft-resilient*, where the shape of each jet is not dependent on the amount or location of soft radiation.

Figure 4.2 shows a map of the energy deposits for a simple scattering event in simulation. Jets are reconstructed using the anti-k_{T} algorithm with $R = 1$. The jets are shown by shaded region. Each individual jet is roughly circular, illustrating the insensitivity of the algorithm to soft radiation.

A single event in data observed at the CMS detector, fully reconstructed by the PF algorithm with anti-k_{T} clustering, is shown in Figure 4.3. Charged particles are indicated by solid lines,

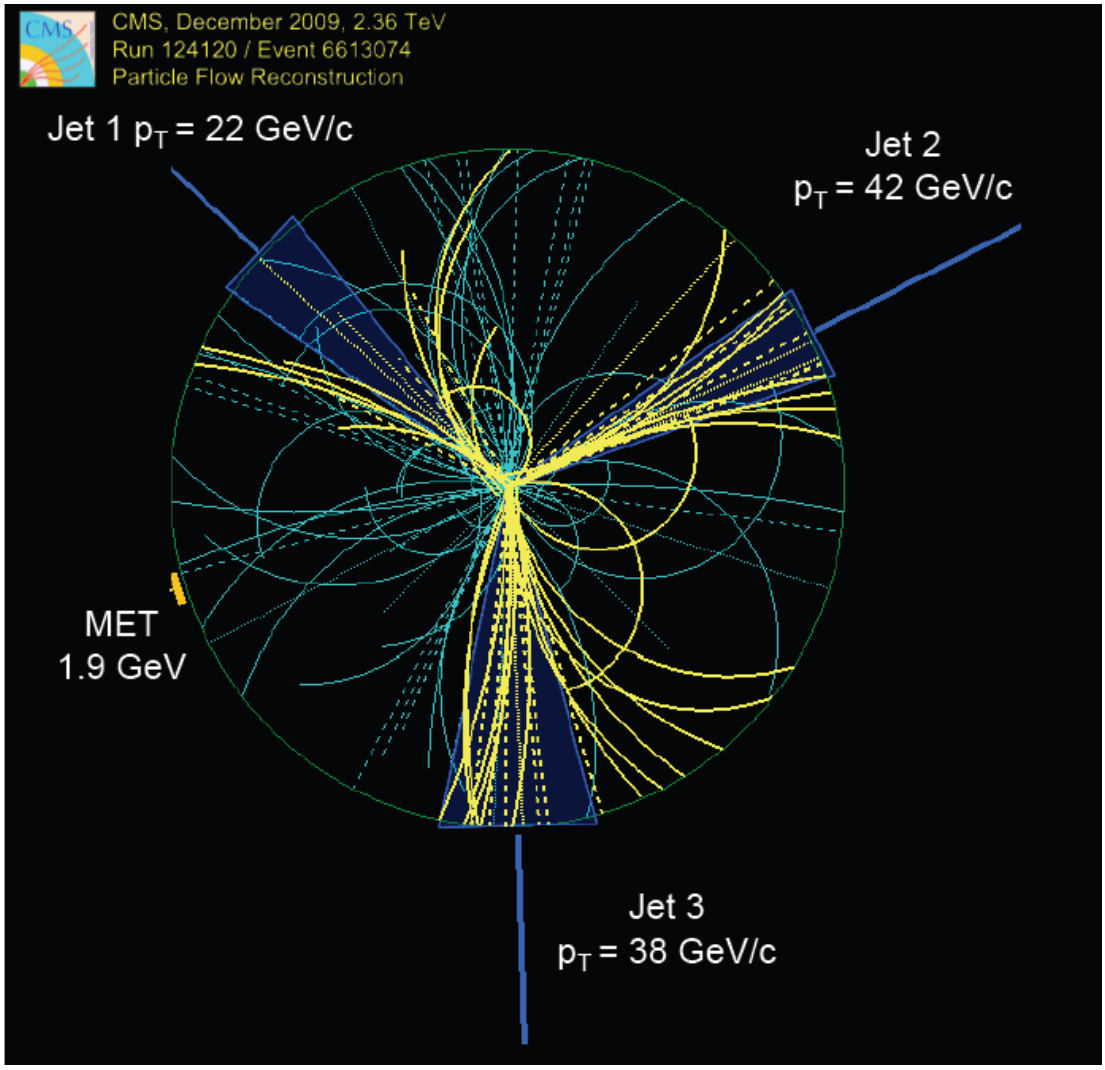


Figure 4.3: Particle Flow-reconstructed event, with charged particles shown by solid lines and neutral particles by dashed lines. Particles clustered into jets are shown in yellow, with the jet cone indicated by the shaded blue region.

and neutral particles by dashed lines. The cone of a reconstructed jet is shown by the shaded region, with included particles in blue. Isolated particles not associated with jets are shown in yellow.

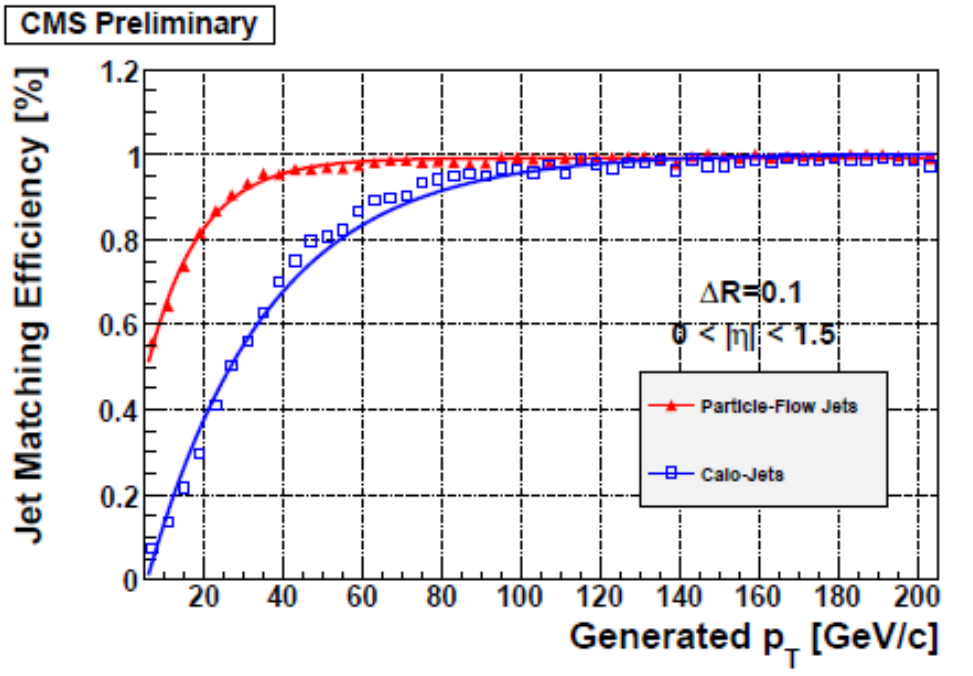
The pQCD calculations performed by NLOJet++ and fastNLO will also assume anti-k_{T} clustering with $R = 0.7$. As the initial calculations only specify the final distributions of partons before hadronization, this assumption only causes partons with $\Delta_{ij} < R$ to be grouped into jets. Non-perturbative effects will be applied as a series of corrections to the parton-level jet distribution. These corrections will be determined in simulation and discussed in Chapter 5.

4.2.6 Particle Flow Performance

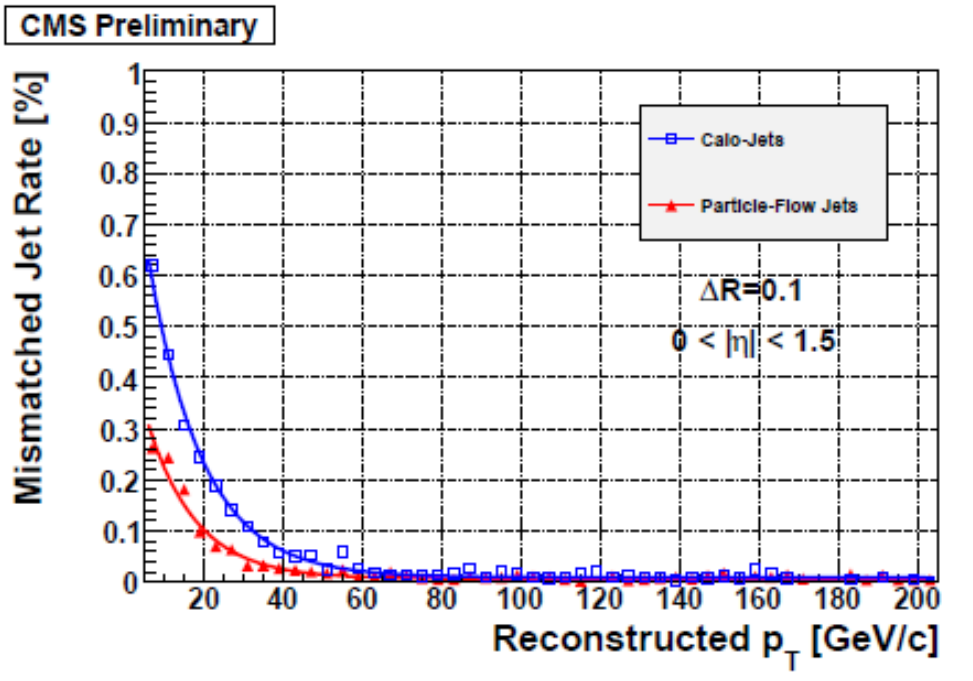
The Particle Flow response to jets has been studied at great length in simulation. Figure 4.5 shows the matching and mis-matching rate between generated and reconstructed jets within 0.1 in $\eta - \phi$ space for both PF and pure calorimeter jets. For central values of η and values of p_{T} above 100 GeV, the matching efficiency is effectively unity. The corresponding fake rate is negligible. As a result, the jet efficiency and fake rate in reconstruction can be assumed not to contribute any experimental uncertainty to the extinction search.

Assuming all jets are correctly identified, the bulk of the remaining work for a given analysis lies in correctly measuring the jet response. The difference between the generated and reconstructed jet momentum in simulation is shown as a fraction of the generated momentum in Figure 4.5. In data, a series of corrections are applied to the reconstructed jet momentum to bring the jet response to unity.

Over all phase space, the jet response is much closer to unity for Particle Flow reconstruction than for pure calorimeter jets. PF jets therefore require a much smaller residual correction. The corresponding uncertainty on the jet energy scale will therefore be smaller. The jet energy scale is known to be the dominant source of detector-related experimental uncertainty in an inclusive jet p_{T} measurement, which suggests that PF jets are preferable to calorimeter jets for use in the extinction search.

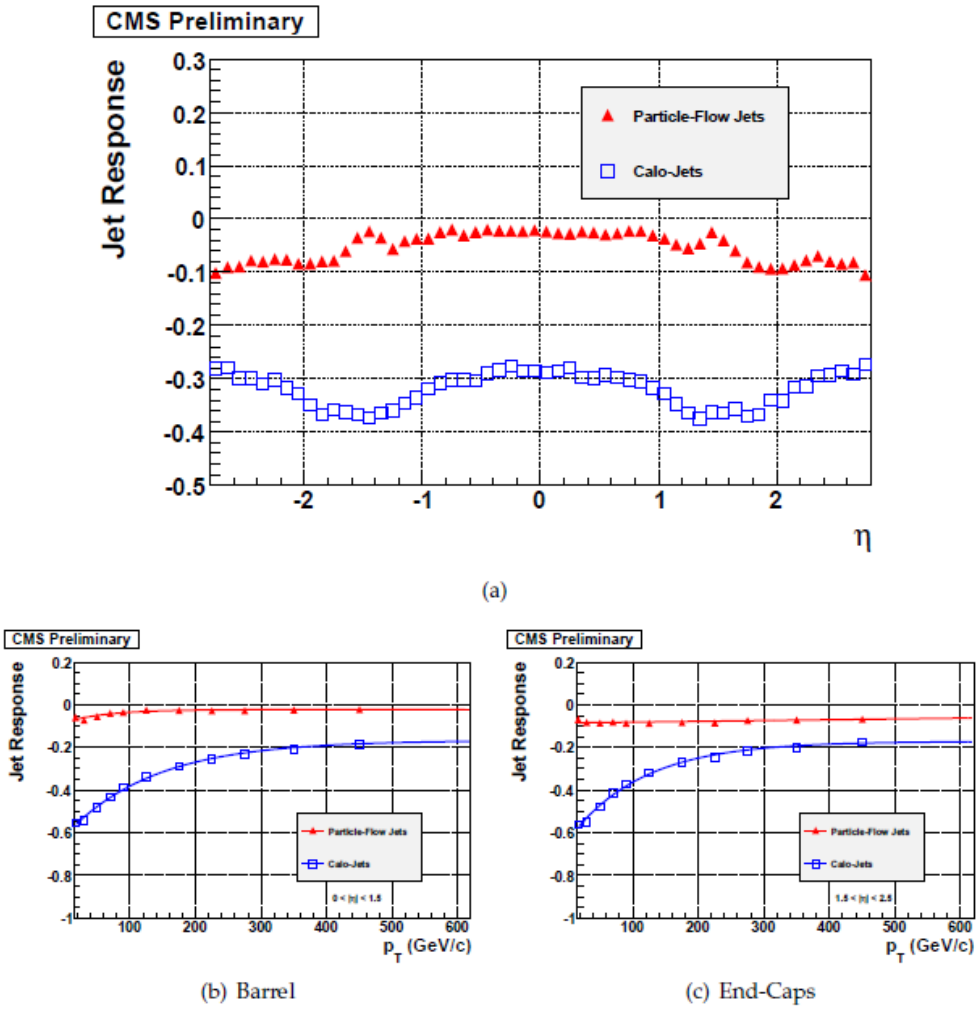


(a) $-1.5 < \eta < 1.5$



(c) $-1.5 < \eta < 1.5$

Figure 4.4: Jet matching efficiency (top) and mis-matching rate (bottom) with a distance parameter of = 0.1, as a function of jet p_{T} for both PF and calorimter jets.



(a)

(b) Barrel

(c) End-Caps

Figure 4.5: Jet Response as a function of η (top) and p_{T} (bottom) for PF and calorimeter jets.

Chapter 5
Analysis Methodology

5.1 Introduction

5.1.1 Observable

The choice of the inclusive jet p_{T} spectrum as an observable for an extinction search is motivated by the tremendous amount of statistics expected in a year's worth of data taken at the CMS experiment. The observable is defined as:

$$\frac{d\sigma_{jets}}{dp_{\mathrm{T}}} = \frac{1}{\mathcal{L} \cdot \epsilon} \frac{N_{jets}}{\Delta p_{\mathrm{T}}} \tag{5.1}$$

where:

* $\mathcal{L}$ is the total recorded luminosity
* ϵ is the overall efficiency of jet identification
* N_{jets} is the number of jets recorded in a given p_{T} range
* Δp_{T} is the width of that p_{T} range.

Due to the large number of events collected, the analysis is restricted to subsets of the data where the efficiency ϵ is unity. This avoids potential additional sources of systematic uncertainty which would otherwise decrease the sensitivity to an extinction effect. Within this subset, selection of events is as inclusive as possible to maximize the statistical power of the dataset.

5.1.2 Summary

The null hypothesis of SM jet production will be modeled by an explicit calculation of pQCD at NLO, and the alternative hypothesis of extinction will be expressed in terms of the SM. The effect of extinction manifests as an exponential attenuation of jet production from LO

scattering. This effect will be modeled as a sigmoid function, alternately referred to as an error function. The extinction p_{T} spectrum will be expressed as the product of the SM p_{T} spectrum and this sigmoid. This is fairly unusual for an analysis at an accelerator experiment. Typically, measurements of an single-variate observable expect new physics to manifest as an excess of events. The alternative hypothesis can then be modeled as signal + background. Extinction physics manifests as a deficit of SM processes, and is modeled as signal multiplied by background.

The pQCD calculation is performed by NLOJet++ and fastNLO assuming the anti-k_{T} clustering algorithm. Non-perturbative effects are accounted for by corrections to the parton-level distribution and determined in MC simulation. The theoretical spectra are then convolved with a function representing the detector resolution. Finally, the data will be compared to both hypotheses taking into account all non-trivial sources of systematic uncertainty. This analysis is systematics-limited, so the application of various uncertainties in this comparison must be performed very carefully. Details of these applications will be found in Chapter 6.

5.2 Observed Dataset

The extinction search is performed in 5.0 fb^{-1} of data recorded at the CMS detector throughout 2011. The events are selected from the Jet secondary dataset (SD), a skim of the data stream in which all events where at least one jet-related trigger signals acceptance. The data streams are continuously monitored for stability of both hardware performance and physics response. The total luminosity used in the extinction analysis does not include runs rejected during this monitoring process.

Dataset Name	total lumi
/Jet/Run2011A-May10ReReco	205 pb^{-1}
/Jet/Run2011A-PromptReco-v4	930 pb^{-1}
/Jet/Run2011A-Aug05ReReco	371 pb^{-1}
/Jet/Run2011A-PromptReco-v6	661 pb^{-1}
/Jet/Run2011B-PromptReco-v1	2.8 fb^{-1}
Total	5.0 fb^{-1}

Table 5.1: Datasets and corresponding integrated luminosity

5.2.1 Event Selection

From this 5.0 fb^{-1} sample, a very basic event selection is applied. Each event is required to pass the HLT_Jet370 trigger, which requires the presence of at least one jet in the event with p_{T}

Selection Level	Event Count
Total Sample	22872114
Trigger	3104180
Vertex Selection	3075201

Table 5.2: Event counts following each selection criterion

above 370 GeV. Many of the triggers in the Jet SD have a prescale factor, where only a fraction of the events which would pass the trigger are recorded. The HLT_Jet370 is the inclusive jet trigger with the lowest p_{T} threshold which does not have a prescale. This is consistent with the stated goal of maintaining a selection efficiency ϵ of unity. The CMS Collaboration (2010c) shows that the dataset on this trigger is relatively insensitive to pile-up from additional vertices and soft interactions in the event.

Simple selection criteria are applied to the reconstructed vertex in each event. These vertices are associated with the initial scattering processes of the event. The location of the vertices are determined from an extrapolation of the particle trajectories back to the beam line. Events with poorly reconstructed vertices may have been reconstructed out-of-time, in the wrong bunch crossing. To summarize, including trigger selection, the event selection criteria are:

1. HLT_Jet370 trigger selected
2. Vertex selection:
   (a) At least one good vertex
   (b) Position of the vertex on the z-axis < 24 cm from the beam spot
   (c) Degrees of freedom > 4 for the vertex fit.

Table 5.2 shows the total number of events in the sample, and the number which pass the trigger and vertex selection criteria. After event selection, there are more than three million events remaining in the data available for the extinction search.

5.2.2 Jet Selection and Spectrum Construction

The measurement of the inclusive p_{T} observable has already been performed by CMS Collaboration (2010c). The extinction search inherits a large portion of the analysis strategy from this previous iteration of the measurement, including the binning in p_{T} and jet selection criteria.

Jets observed in data are first corrected by a factor equal the difference in jet response between the energy measured by the detector and the "true" momentum in simulation. The

derivation of, and uncertainty on these corrections is a significant source of systematic uncertainty in this search and will be discussed at length in Chapter 6.

The corrections are applied in several stages. First, an offset correction is applied to remove the expected contribution from particles from the underlying event and electronic noise. A relative correction is then done as a function of jet p_{T}. Finally, an absolute correction to the p_{T} scale is performed dependent on $|\eta|$. As these corrections are largely derived from simulation, an additional residual correction is required. This residual correction is derived from simple closure between simulation and data.

In summary, the following corrections are applied to the observed jet sample:

1. Offset Correction
2. Relative Correction in η
3. Absolute Correction in p_{T}
4. Additional Residual corrections.

Corrected jets are first restricted by their kinematics, corresponding as usual to the region where the acceptance of the sample is unity. The search is performed in a region in detector rapidity of $|\eta| < 1.5$. Jet extinction has no angular dependence as shown in Figure 5.1, but the reach of the dataset in p_{T} decreases as jets become less transverse. The jet response also becomes less accurate with increasing η, leading to increased systematic uncertainty. The specific choice of 1.5 as a threshold is required by the pQCD calculations, which are binned in rapidity increments of 0.5.

A p_{T} cut is applied, based on the efficiency of the HLT_Jet370 trigger. While the trigger accepts all events containing at least one jet with p_{T} above 370 GeV, this measurement of the momentum is not identical to that used by most analyses. At trigger level, the jets are uncorrected. As a result, the final corrected spectrum will not be completely efficient exactly at the threshold of 370 GeV.

The exact p_{T} cut on the dataset is again determined by the pQCD calculations. The inclusive jet p_{T} measurement performed by CMS Collaboration (2010c) uses 50 p_{T} bins, spanning the full range of possible jet momenta up to 3.5 TeV. The width of each bin corresponds roughly to the detector resolution for jets with p_{T} at the center of that bin. The pQCD calculation has been performed assuming the same binning, in order to allow a direct comparison with the observed spectrum. The extinction search uses the same binning, in order to enable the same comparison with existing QCD theoretical predictions.

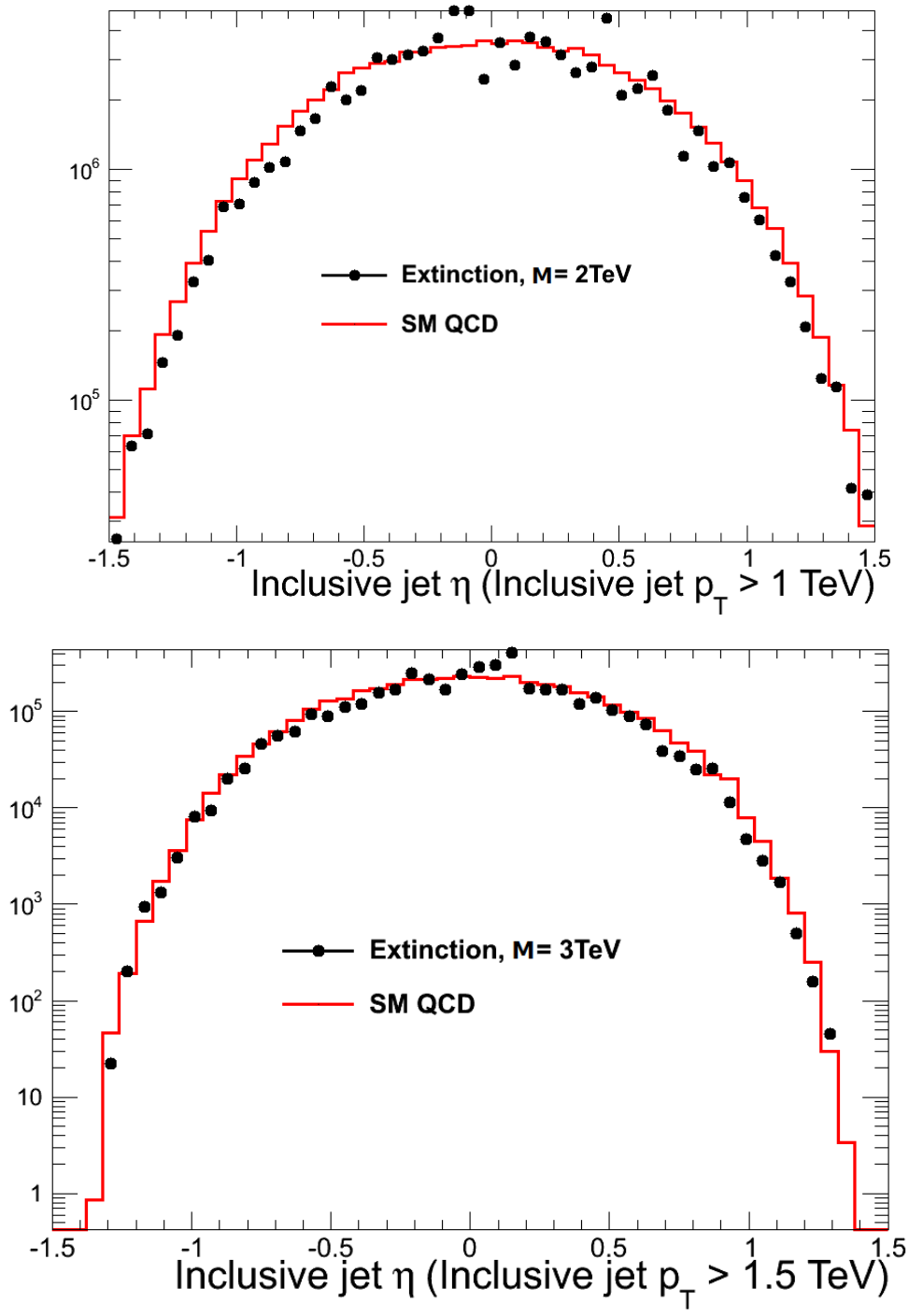


Figure 5.1: Pseudorapidity distribution for Extinction MC (points) and SM QCD MC (red line) for all jets with $p_{\mathrm{T}} > M/2$.

The extinction search is conducted in 19 bins in p_{T}, all of which have been shown by the CMS Collaboration (2010c) to be completely efficient when the HLT_Jet370 trigger is applied. The bin edges are:

$$p_{\mathrm{T}} = \{592.,638.,686.,737.,790.,846.,905.,967.,1032.,1101.,1172.,$$
$$1248.,1327.,1410.,1497.,1588.,1684.,1784.,1890.,2000.\} \text{ GeV}.$$

Note that the previous measurement is a double differential cross-section, binned in p_{T} and rapidity (y) rather than $|\eta|$. The jet counts observed in the region $|\eta| < 1.5$ have been compared directly with the jet counts in the three central rapidity bins in the double differential measurement. The difference between the two counts are negligible. A comparison between the jets selected in the extinction search and the combined pQCD calculations for the three central rapidity bins will therefore still be valid.

The remaining selection criteria are simple jet identification cuts. They are used predominantly to reject electronic noise in the HCAL which may have been misidentified as a jet. All jets containing only a single particle or containing only HCAL deposits are rejected. The remaining cuts are applied on the relative fraction of jet energy coming from neutral and charged hadrons, as well as from photons and electrons. An electron energy fraction near unity is evidence of an electron misidentified as a jet. The hadronic and photonic cuts are based on the average makeup of a typical jet and the likelihood that a jet composed heavily of neutral particles may be either noise or a misidentified photon.

To summarize, the full sequence of jet selection criteria are:

1. $p_{\mathrm{T}} > 592\ GeV$
2. $|\eta| < 1.5$
3. Jet ID for Particle Flow:
   (a) $N_{particles} > 1$
   (b) Neutral Hadronic Energy fraction (NHF)< 0.90
   (c) Photonic Energy fraction (PHF) < 0.90
   (d) $N_{chargedhadrons} > 0$
   (e) Charged Hadronic Energy fraction (CHF) > 0
   (f) Electron Energy fraction (ELF) < 0.99.

Figure 5.2 shows the distribution of the fractional jet identification variables as seen in data and in simulation. In each plot, all other selection criteria are applied besides the one being shown. In each case, the selection is shown to remove only jets which are not expected by simulation, while leaving the overwhelming majority of "true" jets. This is again consistent with maintaining a selection efficiency of unity.

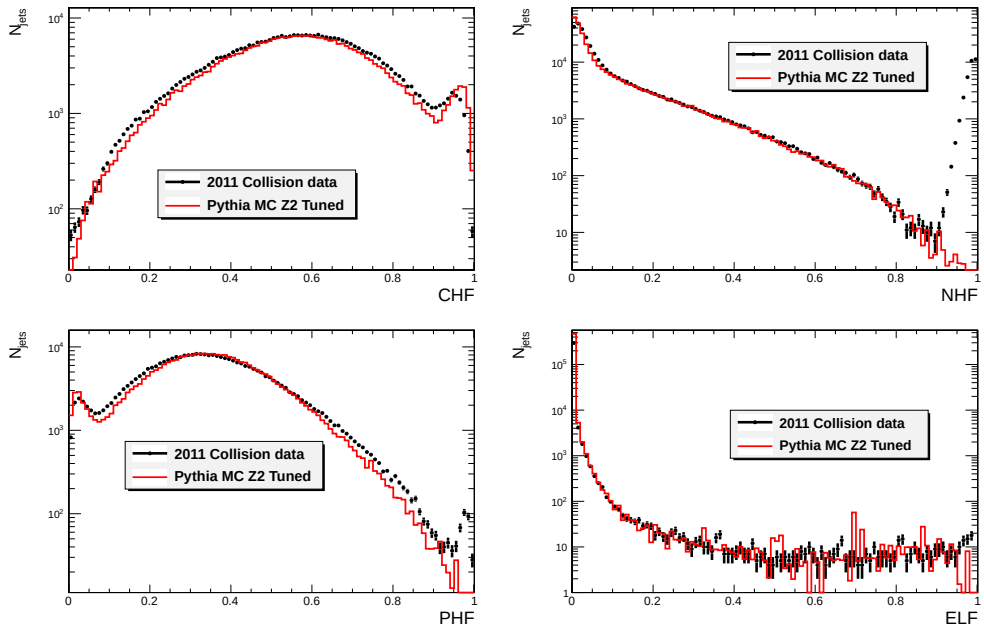


Figure 5.2: Jet ID variable distributions for data vs. Pythia Z2-tuned MC. In each figure, all other tight Jet ID criteria are applied except the one displayed.

5.3 Monte Carlo Datasets

5.3.1 Standard Model Prediction

The extinction search is designed to allow a direct comparison between the jet p_{T} distribution observed in data, and the distribution predicted by pQCD at NLO. To construct the SM prediction, we begin as discussed in Chapter 4 with a calculation of the QCD scattering amplitudes using NLOJet++. A sample of final parton distributions is then generated from these amplitudes using fastNLO, assuming the MSTW PDF set developed by Martin et al. (2009) as input. The *parton-level* jet p_{T} spectrum is constructed by assuming anti-k_{T} clustering with distance parameter $R = 0.7$. The clustering at parton level does little other than grouping sufficiently close partons into single jets, as hadronization has not yet been simulated.

The parton-level distribution is corrected for non-perturbative effects, including hadronization and contributions from multi-parton interactions (MPI). These cannot be accounted for reasonably using pQCD and modern computing techniques, and are instead derived from simulation. SM QCD events are generated in PYTHIA with hadronization turned on and off. The p_{T} spectra at are then compared at generator-level, before the detector response is modeled. A bin-to-bin NP correction is produced by normalizing the p_{T} spectrum with NP effects to the pQCD-only spectrum.

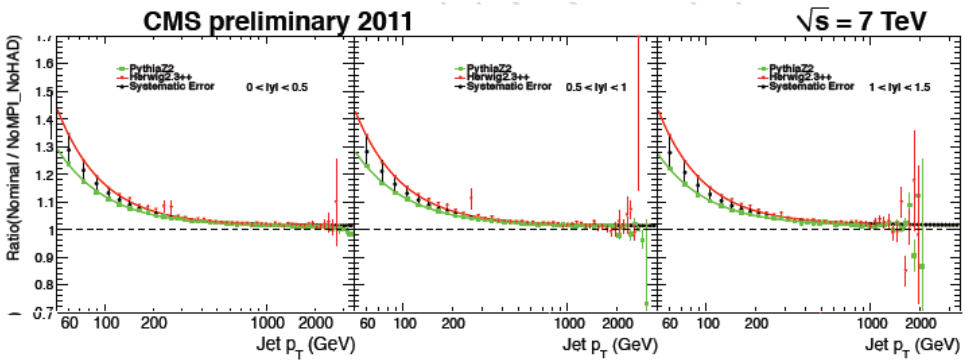


Figure 5.3: Relative uncertainty of all sources of theoretical uncertainty at 1 σ as a function of jet p_{T}.

Figure 5.3 shows the corrections as derived by the CMS Collaboration (2010c) for the three central rapidity bins. The corrections are derived using two separate generators. The difference between the two sets of corrections is taken as a source of systematic uncertainty. For all p_{T} bins above 592 GeV, this uncertainty is negligible. The total NP corrections are found to fall from 3% to 1% across the 19 p_{T} bins considered in the extinction search.

The particle-level spectrum is then convolved with a function representing the detector resolution. This technique is known as *forward smearing*. A fit to this function is shown in Figure 6.9. The width of the resolution is assumed to have a 10% uncertainty. This is included as a source of systematic uncertainty in the search as discussed in Chapter 6.

5.3.2 Extinction Samples

The modification of the QCD scattering amplitudes has been calculated explicitly as described by Kilic et al. (2012). These amplitudes are used as input to a series of Les Houches (LHE) files in accordance with Alwall et al. (2007), which then generate corresponding final parton distributions. These parton distributions are used as input to PYTHIA with NP effects turned off and clustered using the anti-k_{T} algorithm with $R = 0.7$. Each extinction sample comprises

13 smaller samples binned in $\hat{p_{\mathrm{T}}}$, where $\hat{p_{\mathrm{T}}}$ is the transverse momentum of the hard-scattered partons. The bin boundaries is as follows:

$$\hat{p_{\mathrm{T}}} = \{500,800,1100,1400,1700,2000,2300,2600,2900,3200,3500,3800,4100,\mathrm{inf}\} \text{ GeV}.$$

At least 20,000 events are generated in each $\hat{p_{\mathrm{T}}}$ bin. The samples are weighted by their respective production cross-sections, and the p_{T} distribution is constructed. Each sample is then advanced to particle-level and smeared with the detector resolution in the same fashion as the fastNLO pQCD calculation. As the extinction sample contains only QCD events, the same NP corrections and forward smearing is applied to both the extinction and SM samples.

Five Extinction MC samples are produced, corresponding to values of the extinction scale $M = 2, 2.25, 2.75, 3$ and $4\,\mathrm{TeV}$. For reference, a sample of SM QCD is produced in the same fashion. However, as the direct SM pQCD calculation is preferred in the comparison with data, this additional PYTHIA sample is only used for systematic studies.

5.4 Likelihood Comparison

The comparison between the data and the hypotheses is performed using a likelihood function. The jet p_{T} observable is assumed to be Poisson-distributed in each bin. Given a binned dataset and a set of systematic uncertainties, the likelihood of the data given is defined as:

$$L(\vec{x}|\vec{\theta}) = \prod_i \frac{\mu_i(\vec{\theta})^{n_i} \cdot e^{-\mu_i(\vec{\theta})}}{n_i!} \tag{5.2}$$

where:

* $L(\vec{x}|\vec{\theta})$ is the likelihood of $\vec{x}$ given $\vec{\theta}$
* $\vec{x}$ is a vector of data. In this case, the binned jet p_{T} spectrum.
* $\vec{\theta}$ is a vector of nuisance parameters. Each source of systematic uncertainty is treated as a separate nuisance parameter.
* $\mu_i(\vec{\theta})$ is the expected content for a given hypothesis in bin i, assuming the given value of the nuisance parameters.
* n_i is the observed content of bin i.

The nuisance parameters can be divided into four categories:

$$\vec{\theta} = (\theta_{PDF}, \theta_{JES}, \theta_{JER}, \mathcal{L}) \tag{5.3}$$

* θ_{JES} is the uncertainty in the residual corrections applied to jets after clustering. This is typically referred to as the Jet Energy Scale (JES).

* θ_{JER} is the 10% uncertainty in the resolution function used to smear the particle-level spectra.

* θ_{PDF} is the associated uncertainty in the Parton Distribution Functions (PDF). This is the only theoretical source of uncertainty relevant to the extinction search.

* $\mathcal{L}$ is the luminosity.

Jet extinction is modeled as an exponential attenuation of SM jet production. For a given value of the extinction scale M, the expected content of each p_{T} bin can be expressed in term of the differential SM jet production cross-section σ_{SM} and an attenuating function $F(p_{\mathrm{T}}, M)$:

$$\mu_i(\vec{\theta}) = \sigma_{sm}(\theta_{PDF}, \theta_{JES}, \theta_{JER}) \cdot \mathcal{L} \cdot F(p_{\mathrm{T}}, M, \theta_{JES}, \theta_{JER}) \cdot dp_{Ti}. \tag{5.4}$$

Each term is written including dependencies on the various nuisance paramters. Both the SM cross-section and the extinction function $F(p_{\mathrm{T}}, M)$ are dependent on the jet energy scale and resolution. As the extinction function is a simple multiplicative modification of the SM cross-section, it is not dependent on the PDF uncertainties.

5.4.1 Extinction Function

Extinction effects, if evident at the LHC, will first be observed in the asymptotic turn-on region where the scale M can only barely be reached given the machine energy. In this regime, Kilic et al. (2012) state that only LO scattering amplitudes will be noticeably affected. The extinction function $F(p_{\mathrm{T}}, M)$ must then have some minimum value dependent only on the contributions to the jet production cross-section from higher-order scattering processes. In the case of the inclusive jet p_{T} spectrum, only contributions up to NLO are significant.

The minimum value term $C_{\mathrm{NLO}}(p_{\mathrm{T}})$ is then defined in terms of the differential cross-section at LO ($(d\sigma/dp_{\mathrm{T}})_{\mathrm{LO}}$) and at NLO ($(d\sigma/dp_{\mathrm{T}})_{\mathrm{NLO}}$):

$$C_{\mathrm{NLO}}(p_{\mathrm{T}}) = 1 - \frac{(d\sigma/dp_{\mathrm{T}})_{\mathrm{LO}}}{(d\sigma/dp_{\mathrm{T}})_{\mathrm{NLO}}}. \tag{5.5}$$

The contribution due solely to NLO scattering processes is shown in Figure 5.4. The full differential cross-section due to both LO and NLO scattering is included as reference.

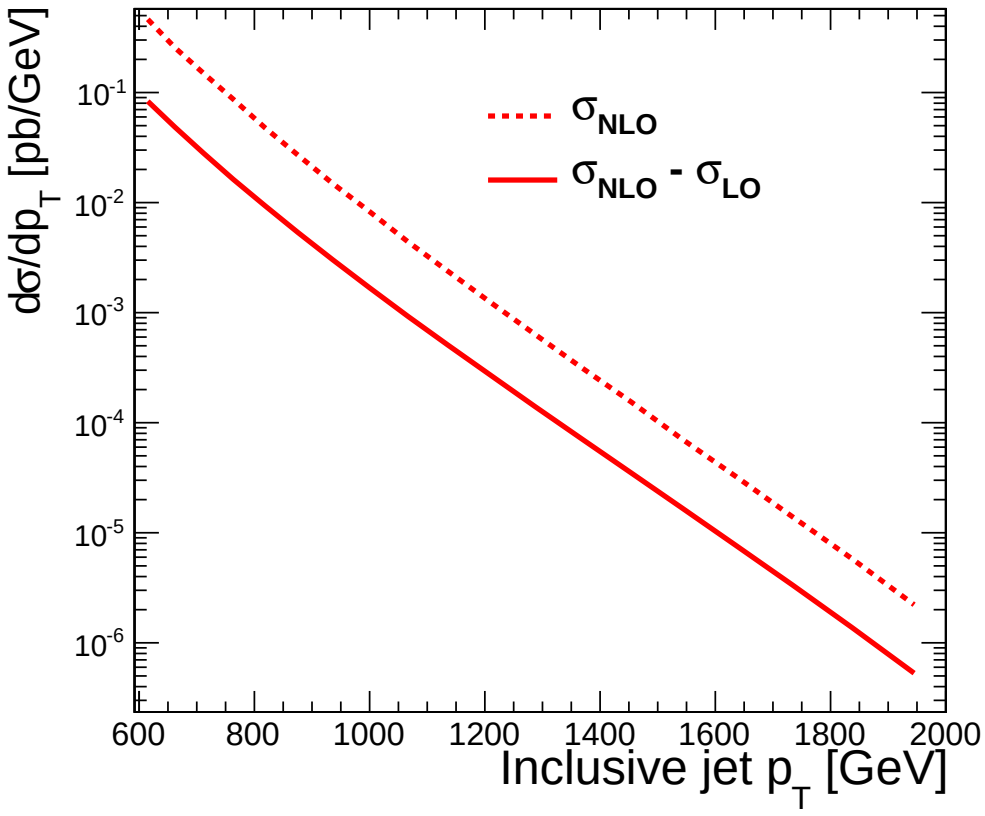


Figure 5.4: The absolute inclusive jet cross-section as a function of jet p_{T} at NLO, and the contribution due solely to NLO scattering processes.

Kilic et al. (2012) have shown that the extinction function takes the form of a sigmoid, also known as an error function. Taking into account the minimum value $C_{\mathrm{NLO}}(p_{\mathrm{T}})$, the function takes the form:

$$F(p_{\mathrm{T}}, M) = \frac{1 - C_{\mathrm{NLO}}(p_{\mathrm{T}})}{1. + \exp\frac{p_{\mathrm{T}} - p_{\mathrm{T},1/2}}{p_{\mathrm{T},0}}} + C_{\mathrm{NLO}}(p_{\mathrm{T}}). \tag{5.6}$$

Here, $p_{\mathrm{T},1/2} = f_{1/2}(M)$ describes the p_{T} threshold at which the extinction manifests, while $p_{\mathrm{T},0} = f_0(M)$ indicates how quickly beyond $p_{\mathrm{T},1/2}$ the LO cross-section becomes negligible. The values of $f_{1/2}(M)$ and $f_0(M)$ are extracted from fits of the likelihood to the generated p_{T} spectrum using the minimization program MINUIT developed by James & Roos (1975). This fit is repeated for each the five available values of M. Each fit is performed relative to the MC uncertainty of each bin. In 5.0 fb^{-1} of data, the first three bins will have a lower Poisson uncertainty than the MC uncertainty of the same bins in simulation. To ensure the fit will be accurate in this region, these three bins are weighted by a factor of 5 above the other bins in evaluating the likelihood.

The five fits of the extinction function are shown in Figure 5.5. MC statistical uncertainty is shown in black, and expected Poisson uncertainty at 5.0 fb^{-1} is shown in green. There is no evident bias in any of the fits.

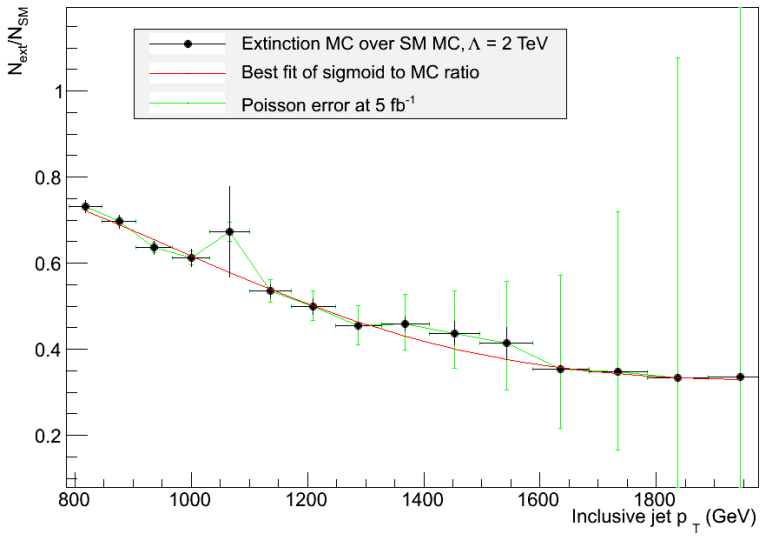
$f_0(M)$ is expected to be roughly independent of the extinction scale, while $f_{1/2}(M)$ should increase linearly with M. A linear fit is performed to the extracted values for the five values of M, as shown in Figure 5.6. This linear parametrization is used to express the extinction function completely in terms of p_{T} and M.

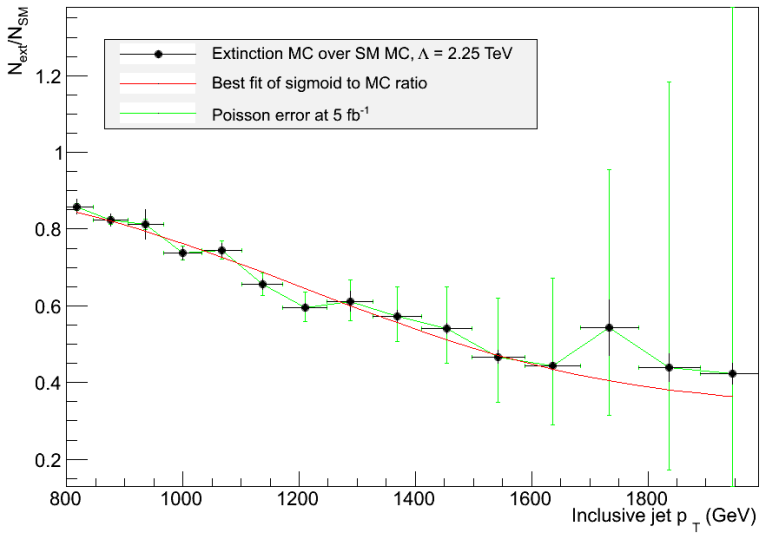
The impact of the nuisance parameters must also be accounted for in the extinction function. As the jet energy scale and jet energy resolution systematic uncertainties are assumed to be uncorrelated, the extinction function is factorized as:

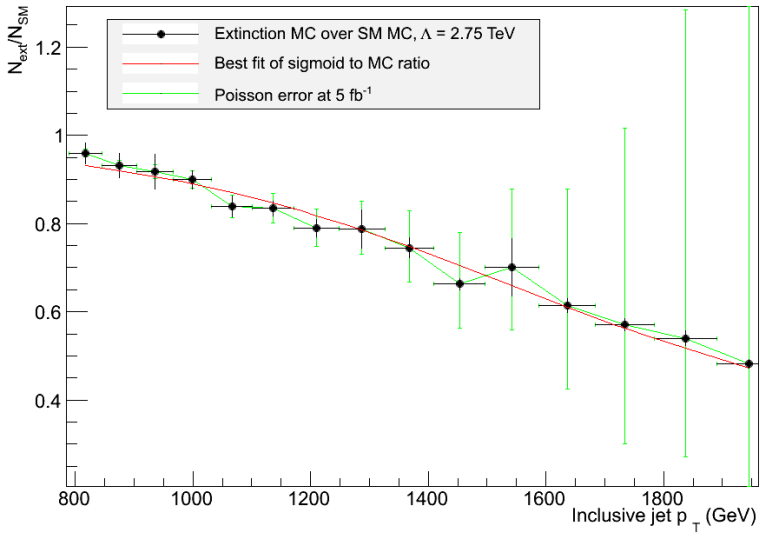
$$F(p_{\mathrm{T}}, M, \theta_{JES}, \theta_{JER}) = F(p_{\mathrm{T}}, M) D_1(\theta_{JES}) D_2(\theta_{JER}). \tag{5.7}$$

Here D_1 and D_2 are the deformation functions that represent the effect the nuisance parameters have on our ability to observe the extinction function. These deformation functions are calculated analytically in each p_{T} bin as:

$$D_n = 1 + \frac{\Delta\sigma_{\mathrm{QCD}}}{\sigma_{\mathrm{QCD}}} - \frac{\Delta\sigma_{\mathrm{EXT}}}{\sigma_{\mathrm{EXT}}}. \tag{5.8}$$


Extinction MC over SM MC, Λ = 2 TeV
Best fit of sigmoid to MC ratio
Poisson error at 5 fb-1
N_ext/N_SM
Inclusive jet p_T (GeV)


Extinction MC over SM MC, Λ = 2.25 TeV
Best fit of sigmoid to MC ratio
Poisson error at 5 fb-1
N_ext/N_SM
Inclusive jet p_T (GeV)


Extinction MC over SM MC, Λ = 2.75 TeV
Best fit of sigmoid to MC ratio
Poisson error at 5 fb-1
N_ext/N_SM
Inclusive jet p_T (GeV)

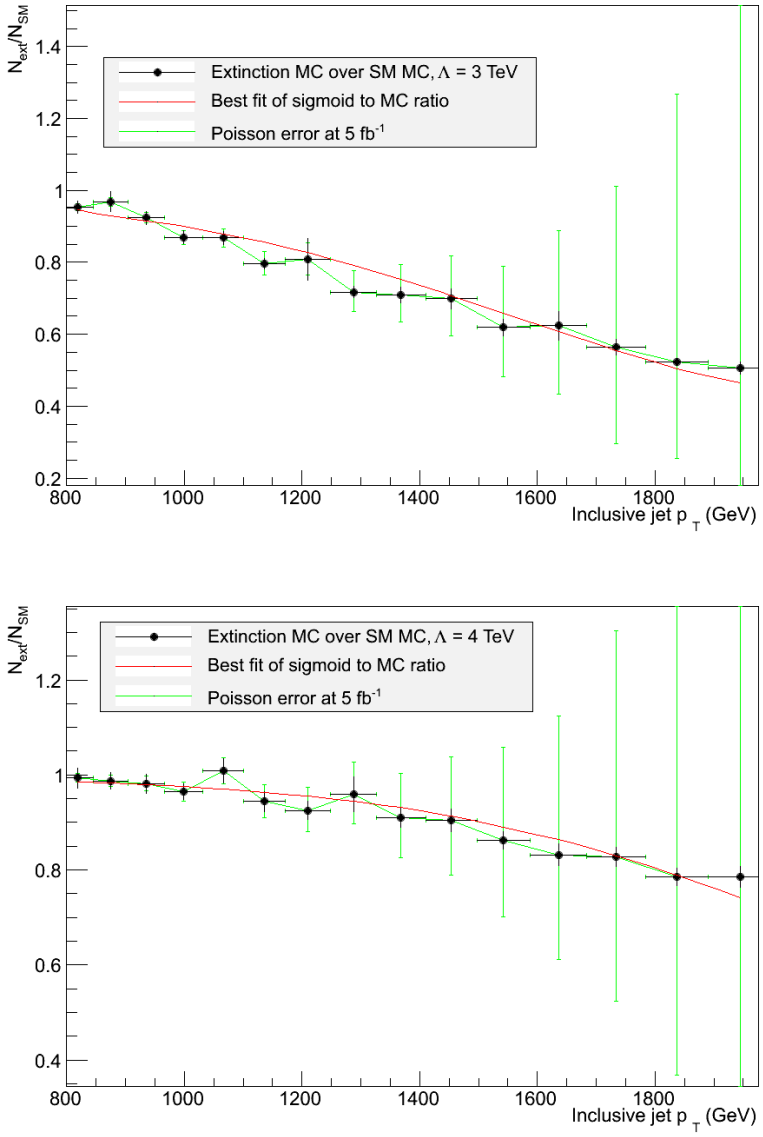


Figure 5.5: Fits of the extinction function to simulated data samples for 5 values of M, including both MC statistical error (black) and expected Poisson error at 5.0 fb^{-1} (green).

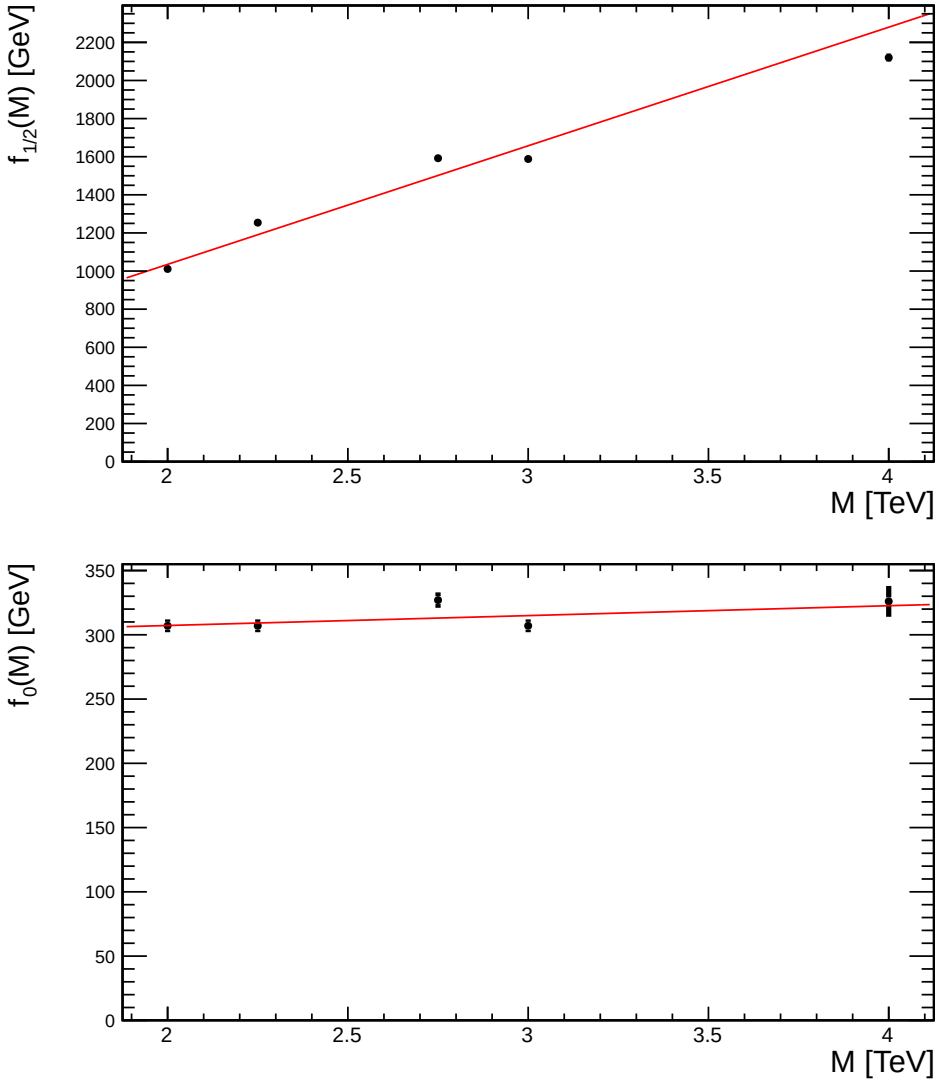


Figure 5.6: Values of $f_0(M)$ and $f_{1/2}(M)$ extracted from fits of the likelihood to the signal p_{T} spectrum, with linear fit overlaid.

σ_{QCD} is the expected occupancy of the bin given the SM, while σ_{EXT} is the occupancy assuming jet extinction. $\Delta\sigma_{\mathrm{QCD}}$ and $\Delta\sigma_{\mathrm{EXT}}$ are the respective changes in the occupancy of each bin for a given value of the nuisance parameter.

Chapter 6
Systematic Uncertainty

The measurement of the inclusive jet p_{T} observable is systematics-limited. For a given p_{T} bin, the occupancy N_{jets} will have a substantially higher associated systematic uncertainty compared to the statistical Poisson error. This will be especially obvious in the first few p_{T} bins, given the exponentially falling nature of the p_{T} distribution.

The extinction search is therefore primarily a measurement of whether or not any deviations between the predicted and observed spectra may be accounted for by a systematic effect. The correlations between these various sources of systematic uncertainty must be taken into account in a principled manner. The comparison between the observed data and the model will be heavily constrained by the low-p_{T} bins, where the statistical power of the sample is the highest. If the correlations between p_{T} bins are assumed to be too high, the systematic uncertainty may be overconstrained by the low-p_{T} data. On the other hand, if the correlations are too low, much of the information from the systematics will be lost.

The dominant sources of uncertainty in the extinction search are attributed to measurement of the jet energy scale (JES) and theoretical uncertainty in the PDFs. It is common to express these uncertainties as singular sources, completely correlated in p_{T} across the full range of the observable. The CMS Collaboration (2010c) presents the most recent measurement of the double-differential jet production cross-section with the uncertainties treated in such a manner.

The JES and PDF uncertainties contain a tremendous amount of information from a variety of component sources. In the extinction search, these component sources are each treated as a separate source of systematic uncertainty. This treatment ensures that the low-p_{T} data can be covered without decreasing our sensitivity at the high-p_{T} end due to overconstraint. Extinction is an effect that manifests predominantly at high-p_{T}. This treatment of the systematics is therefore required to allow a search to occur without needing to ignore the information provided by the low-p_{T} bins.

All sources of uncertainty in the extinction search are assumed to be Gaussian, with log-normal constraints in the likelihood function.

6.1 Detector Uncertainty

6.1.1 Jet Energy Scale

Comparisons between observed data and pQCD predictions rely on the assumption that most of the information at parton-level can be measured in instrumentation and associated with the correct partons. The reconstruction and clustering algorithms described in Chapter 5 have been chosen to ensure that this comparison will be insensitive to non-perturbative effects. For observables such as inclusive jet p_{T}, the energy of each jet is then corrected to account for differences between the detector response and the true jet momentum.

The study of these corrections is known as the jet energy scale (JES), and the full set of corrections is referred to as jet energy corrections (JEC). The uncertainty associated with this corrections is typically small. The CMS Collaboration (2010d) shows that for jets in the kinematic range of the extinction search, the true jet response is known to the percent level.

However, due to the exponentially falling nature of the p_{T} spectrum, a small uncertainty in JES translates to a large uncertainty in the occupancy of each p_{T} bin. Figure 6.1 shows a cartoon represenation of events migrating bin-to-bin due to JES systematic effects. The uncertainty on the occupancy of a p_{T} bin is usually five to ten times larger than the JES uncertainty for jets within that p_{T} range.

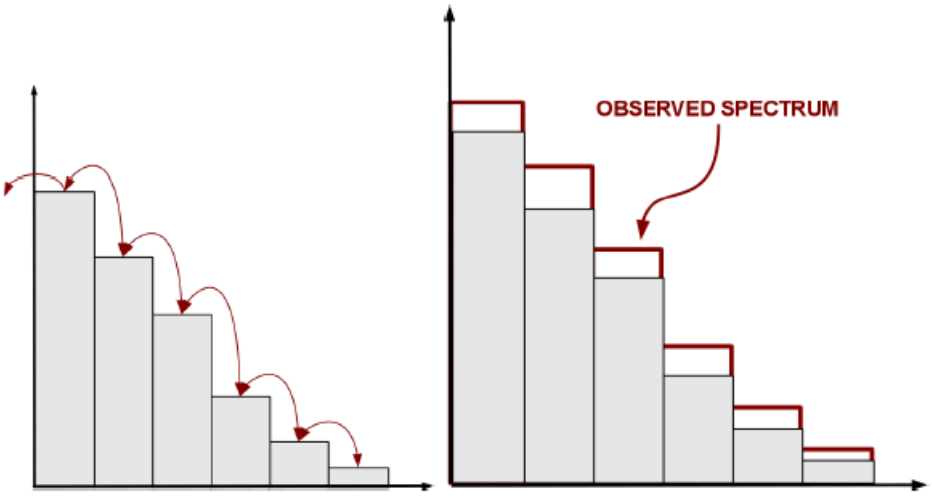


Figure 6.1:

In Chapter 5, the reconstruction sequence includes four jet energy corrections. The corrections are applied using a factorized approach in which each of the four corrections are multiplied

by the observed jet momentum in order. The final jet momenta after all corrections are assumed to be directly comparable to the predictions of pQCD at detector level.

The corrections include an offset correction, a relative correction in η, an absolute correction in p_{T}, and a residual correction to achieve closure with simulation. Each of these corrections has several sources of associated uncertainty, and each of these sources is assumed to be orthogonal. The corrected momentum of each jet is related to the "raw" measured momentum by:

$$p_{\mathrm{T}}^{\mathrm{cor}} = \mathcal{C} \cdot p_{\mathrm{T}}^{\mathrm{raw}} \tag{6.1}$$

Where $\mathcal{C}$ is the product of all four component corrections. The correction product is not commutative, with the following dependencies:

$$\mathcal{C} = C_{\mathrm{offset}}(p_{\mathrm{T}}^{\mathrm{raw}}) \cdot C_{\mathrm{MC}}(p_{\mathrm{T}}^{\mathrm{off}}, \eta) \cdot C_{\mathrm{rel}}(\eta) \cdot C_{\mathrm{abs}}(p_{\mathrm{T}}^{\mathrm{rel}}) \tag{6.2}$$

where $C_{\mathrm{offset}}, C_{\mathrm{MC}}, C_{\mathrm{rel}}$, and $C_{,\mathrm{abs}}$ are the offset, residual, relative and absolute corrections respectively, $p_{\mathrm{T}}^{\mathrm{off}}$ is the jet p_{T} after applying the offset correction, and $p_{\mathrm{T}}^{\mathrm{rel}}$ is the jet p_{T} after applying all corrections other than the absolute scale.

Offset Correction

The offset correction is intended to remove energy from the jet not associated with hard scattering. This includes energy deposits from soft interactions, as well as electronic noise. For a given jet j, an energy density ρ is defined as the jet transverse momentum $p_{\mathrm{T}j}$ divided by A_j, the total area covered by the clustered jet energy in $\eta - \phi$ space. The energy density is derived using the anti-k_{T} clustering algorithm as discussed in Chapter 5, which is known to be insensitive to variations in soft radiation. An average underlying event energy density $\langle \rho_{UE} \rangle$ is calculated from events with no pileup events, where only one vertex is reconstructed. The offset correction is defined as:

$$C_{\mathrm{offset}}(p_{\mathrm{T}}^{\mathrm{raw}}, A_j, \rho) = 1 - \frac{(\rho - \langle \rho_{UE} \rangle) \cdot A_j}{p_{\mathrm{T}}^{\mathrm{raw}}}. \tag{6.3}$$

The uncertainty related to this pileup correction comes from several orthogonal sources. The correction has a p_{T} dependence, which is determined by varying ρ and $\langle \rho_{UE} \rangle$. Further uncertainty comes from an observed bias in the corrections between zero-bias and QCD events

in simulation. A third source appears when the jet production rate $\langle N_{jets} \rangle$ as a function of the number of reconstructed pile-up vertices is averaged over all jet triggers in the datastream. Each of these sources contributes at least a 1% uncertainty to the occupancy of one or more p_{T} bins in the inclusive jet p_{T} spectrum.

Two additional sources of uncertainty are considered as input to the JES systematic, but are trivial in the phase space of the extinction search and can be ignored. These sources are related to the contribution to the pileup correction from out-of-time bunch crossings, and the difference in pileup response between data and zero-bias simulation.

Residual Corrections

The residual corrections are calculated from a comparision of jets before and after detector response in Monte Carlo simulation. Reconstructed jets are matched to generated jets spatially by requiring a separation ΔR of less than 0.25 in $\eta - \phi$ space. For a given generated p_{T} bin with center $p_{\mathrm{T\,gen}}$, the reconstructed transverse momentum $p_{\mathrm{T\,reco}}$ and response factor $\mathcal{R} = p_{\mathrm{T\,reco}}/p_{\mathrm{T\,gen}}$ are recorded. The residual correction is defined as the inverse of the average response, $C_{\mathrm{MC}} = \langle \mathcal{R} \rangle^{-1}$.

The QCD simulation used to construct this response matrix has a jet composition with a higher proportion of low-p_{T} gluon jets than would be normally expected in data. In general, gluons and heavy-flavor quarks produce higher particle multiplicities and softer particle momenta in jets compared to those produced from light-flavor quarks. This difference in flavor mixture between various Monte Carlo generators and data is taken as a source of systematic uncertainty, and contributes to the uncertainty of the inclusive jet p_{T} spectrum.

Relative Corrections

A di-jet balancing technique is used to derive the relative jet energy scale in η. The balancing technique requires one jet to be reconstructed in the barrel, with $|\eta| < 1.3$. The second jet can be at arbitrary η and is labeled the *probe* jet. The barrel is preferred as reference point, as barrel jets have the smallest range in response and highest reach in p_{T}. Central jets can also be calibrated much more easily to the absolute p_{T} scale reference object, as discussed in the next section.

A balance quantity is defined in terms of the p_{T} of the 2 jets and their average:

$$\mathcal{B} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} - p_{\mathrm{T}}^{\mathrm{barrel}}}{p_{\mathrm{T}}^{\mathrm{avg}}} \tag{6.4}$$

The variable $\mathcal{B}$ is binned in terms of $p_{\mathrm{T}}^{\mathrm{avg}}$ and η^{probe}. The response used in the relative correction is defined in terms of the average value of the balance in each bin as:

$$\mathrm{R} = \frac{2 + \langle \mathcal{B} \rangle}{2 - \langle \mathcal{B} \rangle} \tag{6.5}$$

This nonstandard formula for the response is used to minimize the impact of the detector resolution. A direct comparison defining $\mathrm{R} = p_{\mathrm{T}}^{\mathrm{probe}}/p_{\mathrm{T}}^{\mathrm{barrel}}$ will be biased towards the object with the worse resolution. Redefinition of the balance in terms of $p_{\mathrm{T}}^{\mathrm{avg}}$ ensures that the resolution bias cancels out.

Uncertainty on the relative jet energy scale comprises six orthogonal systematic effects. Five of these sources relate to the dependency of the correction on the rapidity of the probe jet, while the sixth arises due to the η-dependence on the correction due to final-state radiation. All of these sources are defined relative to the barrel, where the vast majority of the jets in the extinction search dataset will be found. As a result, the relative correction uncertainty is not a significant contributor to the systematic uncertainty of the inclusive jet p_{T} observable in this analysis.

Absolute Correction

The absolute jet p_{T} scale is defined in terms of well-understood reference objects, such as Z^0 bosons or photons. These objects generally have a much lower resolution than jets and are extremely useful for defining a reference scale at low p_{T}.

In γ+jet events, the photon is required to be in the barrel, while the balancing jet must be separated by $\Delta\phi > 2.7$. Photons are required to be isolated within a cone of > 0.4 in the calorimeter and tracking to prevent pollution of the sample by final state radiation. The detector response is calculated using the same methods discussed previously. This exercise is performed both in data and in MC.

The data-driven corrections are limited by the statistics of the sample available. Above a p_{T} value of a few hundred GeV, the corrections must be determined entirely from an extrapolation in simulation. The uncertainty on the data-driven corrections and on the extrapolation in the high-p_{T} region are considered orthogonal sources of systematic uncertainty on the jet energy scale. Additional uncertainty arises in simulation due to differences in the single particle response between data and MC. All three of these sources are significant contributors to the

systematic uncertainty on the inclusive jet p_{T} measurement.

Summary

In addition to the listed sources of systematic uncertainty, there is a time dependence on the corrections due to out-of-time energy deposits in the very forward tracker. As the extinctions search is limited to $|\eta| < 1.5$, this is not a relevant consideration. A summary of all 16 orthogonal sources of uncertainty on the jet energy scale, and their reference name in JEC softward, is shown below. Sources that contribute at least a 1% uncertainty to any p_{T} bin in the inclusive jet p_{T} observable for $p_{\mathrm{T}} > 592\,\mathrm{GeV}$ and $|\eta| < 1.5$ are marked in **bold**.

* **Absolute scale uncertainty**; an experimental uncertainty derived from jet balancing studies that define the reference scale for jets.

* **High p_{T} extrapolation**; Extrapolation of corrections beyond the data-driven reference scale. Uncertainty comes from the comparison between the corrections derived from different generators.

* **Single Pion**; a correction due to the difference in single particle response as measured in Particle Flow jets.

* **Flavor**; corrections for the difference in jet flavor mixing (quark jets, gluon jets, b-jets, charm jets) between different generators and proper QCD mixture.

* Time; Time dependence on JEC. Only relevant for very forward jets.

* Relative JER; η dependence of uncertainty due to jet resolution. Trivial in magnitude for jets in the barrel. Comprises 3 of the 16 sources:

    - EC1; correlated uncertainty for endcap calorimeter within tracking.
    - EC2; correlated uncertainty for endcap calorimeter outside tracking.
    - HF; correlated uncertainty for forward calorimetry.

* Relative FSR; η-dependence of uncertainty due to final-state radiation

* Relative Stat; statistical uncertainty on determination of η-dependence of corrections. Only relevant in the very forward regions. Contributes 2 of the 16 sources:

    - EC2; correlated uncertainty for endcap calorimeter outside tracking.
    - HF; correlated uncertainty for forward calorimetry.

* **PileUp**; various uncertainties related to pileup corrections. Comprises 5 of the 16 sources:
  - DataMC; difference in pileup response between data and zero-bias MC
  - OOT; uncertainty due to out-of-time pileup. Only contributes for MC scaled to data on prescaled triggers.
  - **pT**; p_{T} dependence on the pileup correction
  - **Bias**; Difference in pileup correction between zero-bias MC and QCD MC
  - **JetRate**; Observed variation in N_{jets} vs. N_{vertex} per collision over single jet triggers

Systematic effects due to the jet energy scale have the effect of moving events from bin to bin, as shown in Figure 6.1. 1σ in a given JES uncertainty is defined by varying all jets in an MC-simulated sample coherently by the respective uncertainty and comparing the shifted spectrum to the central one. The systematic effect of the uncertainty is measured in the range between $+2\sigma$ and -2σ in steps of 0.5σ.

The effect of each source of uncertainty is expressed as a transfer matrix, as shown in Figure 6.2 to 6.8. These matrices measure the fraction the events in each p_{T} bin that migrate to an adjacent bin given a systematic shift. At positive values of σ, events migrate towards higher p_{T} bins, while the opposite is true for negative values of σ. Only the seven transfer matrices that contribute more than a 1% systematic effect are shown. Most of the remaining sources are only applicable outside of the phase space of the extinction search.

By definition, this construction of the uncertainty assumes each individual source is 100% correlated across all p_{T} bins. While this is not necessarily true, it is a reasonable approximation which allows deviations between the observed and expected p_{T} spectra to be attributed to specific systematic effects.

6.1.2 Jet Energy Resolution

The extinction and SM jet p_{T} spectra are convolved with a function representing the detector resolution in the final stage of model development, as discussed in Chapter 5. This function is taken from MC simulation of the detector response. The detector response $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ is constructed as a function of inclusive jet p_{T} as shown in Figure 6.9. An ansatz is fit to the distribution of the form:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = (\frac{N^2}{p_{\mathrm{T}}^2} + S^2 p_{\mathrm{T}} + C)^{1/2} \tag{6.6}$$

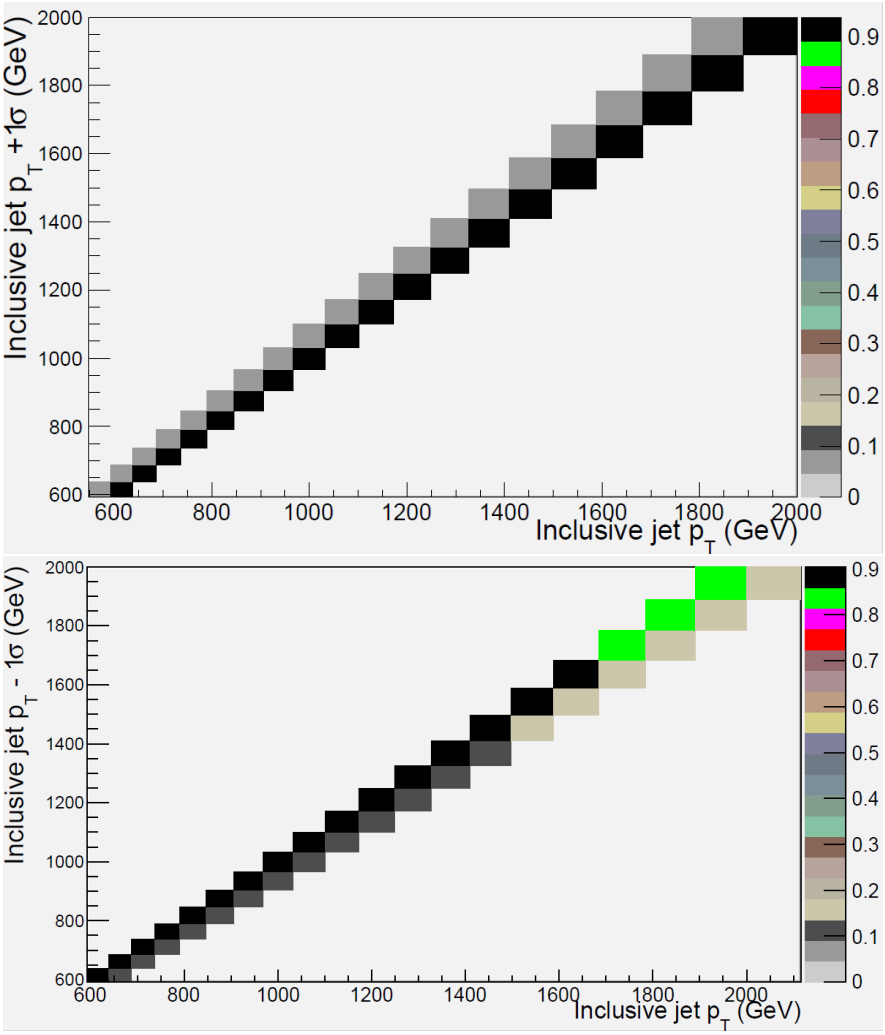


Figure 6.2: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in absolute p_T JES correction uncertainty

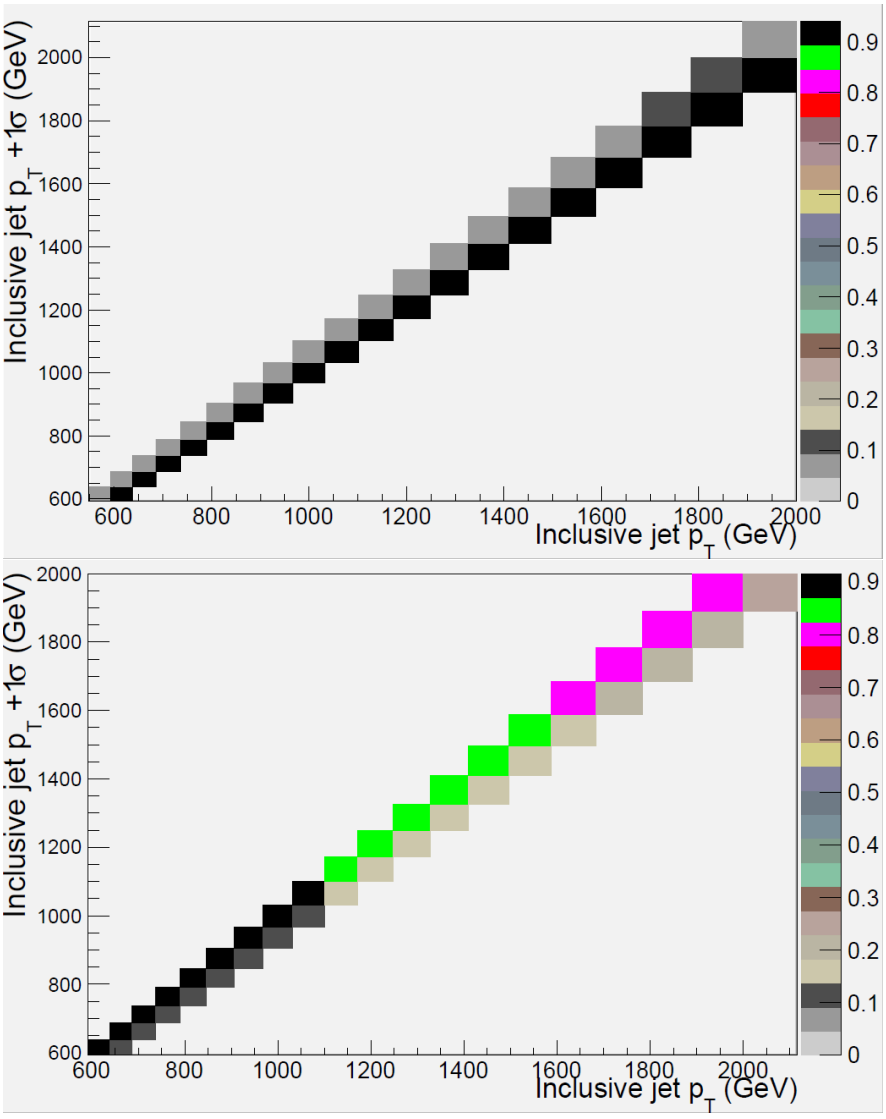


Figure 6.3: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in high p_T extrapolation uncertainty

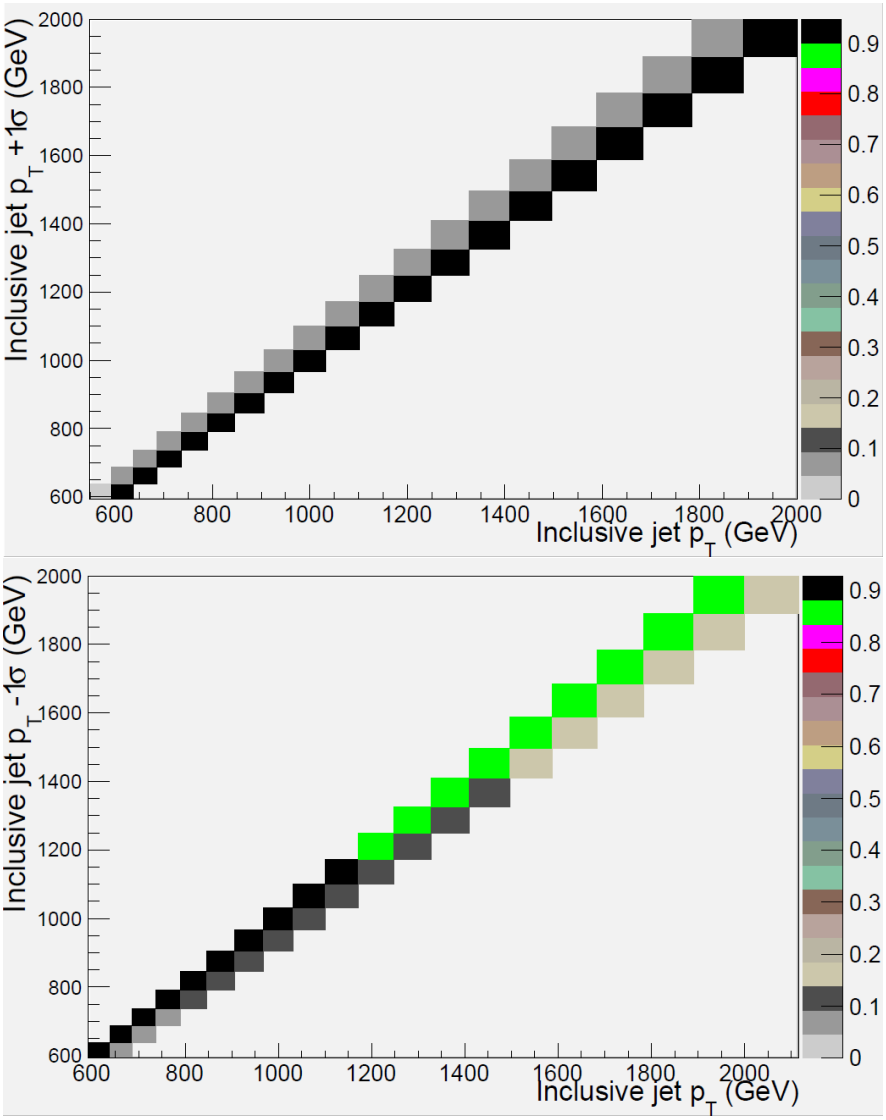


Figure 6.4: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in single-pion response uncertainty

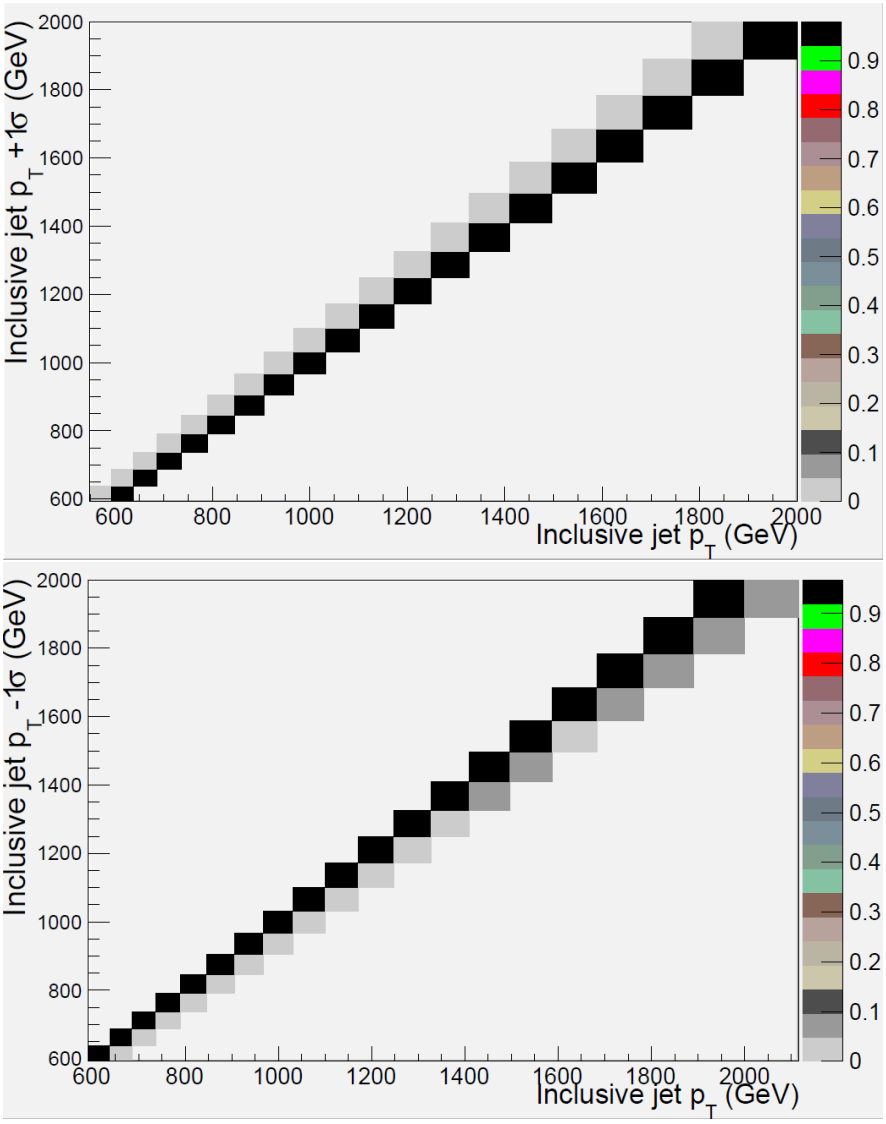


Figure 6.5: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in jet flavor correction uncertainty

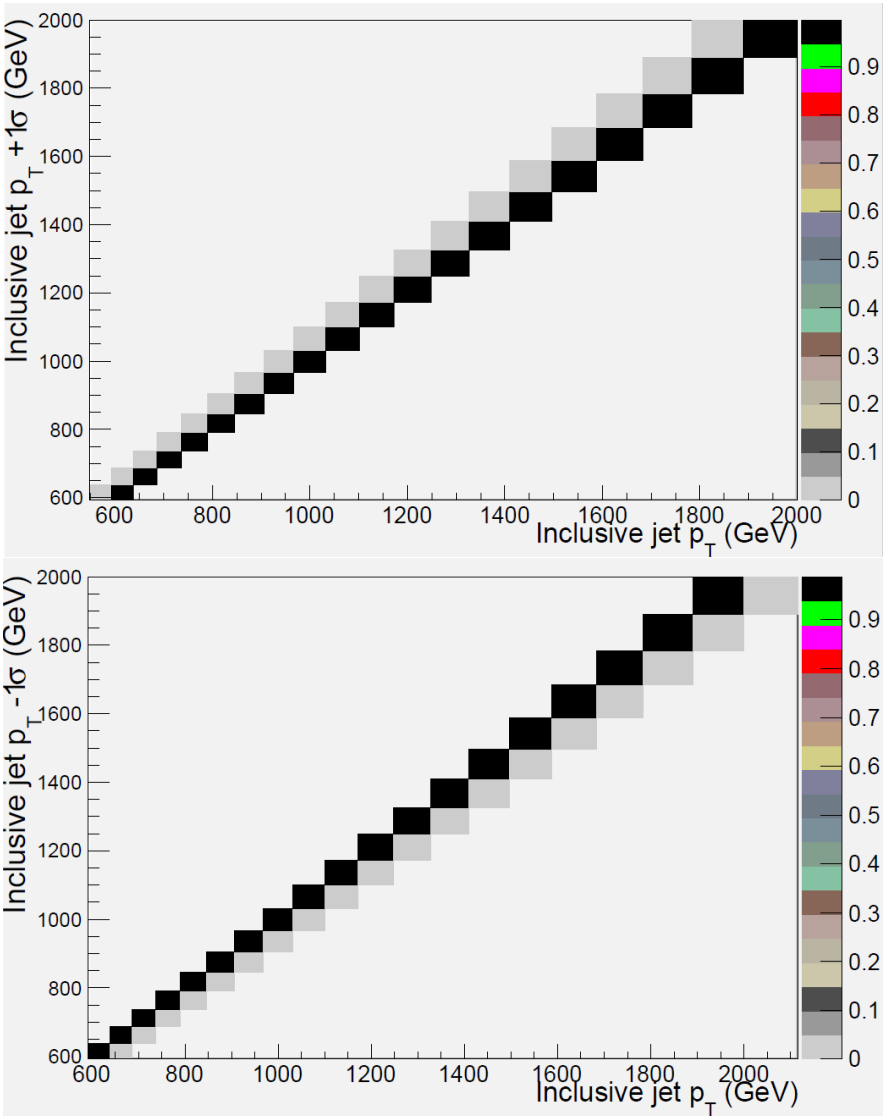


Figure 6.6: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in pileup p_T uncertainty

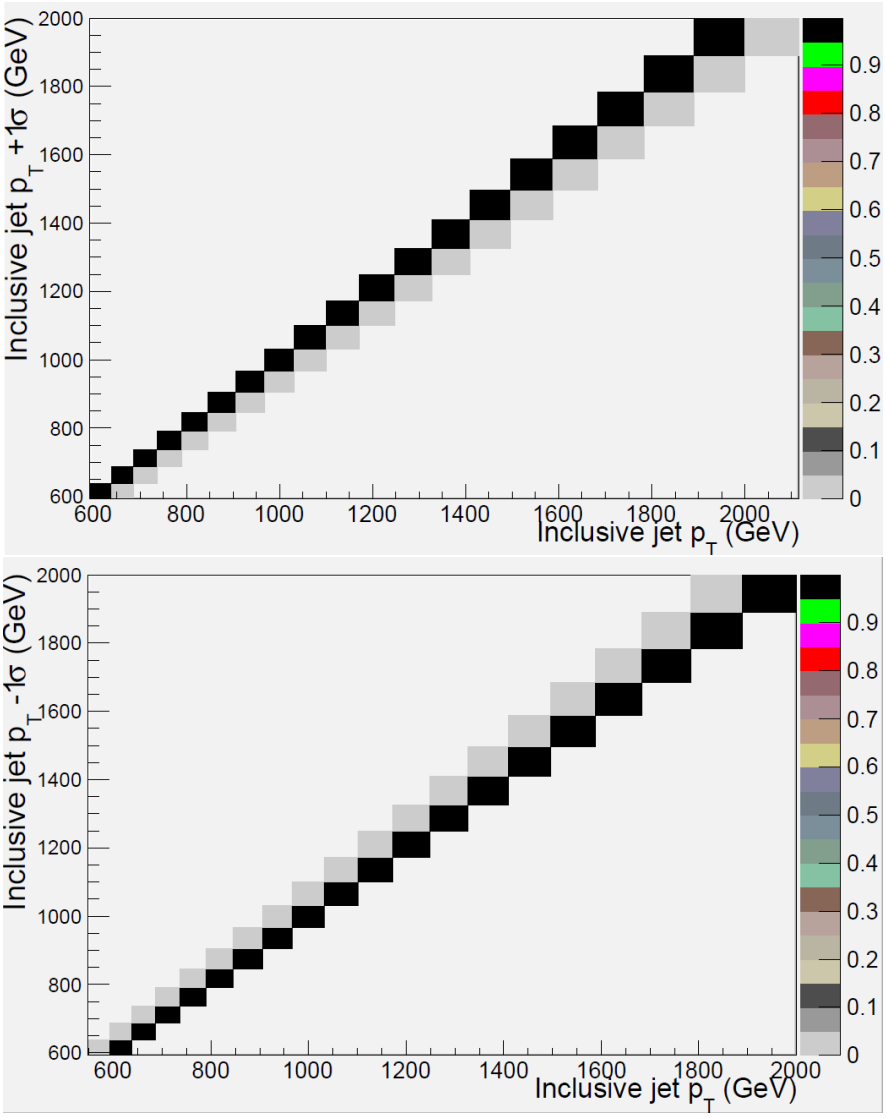


Figure 6.7: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in pileup bias

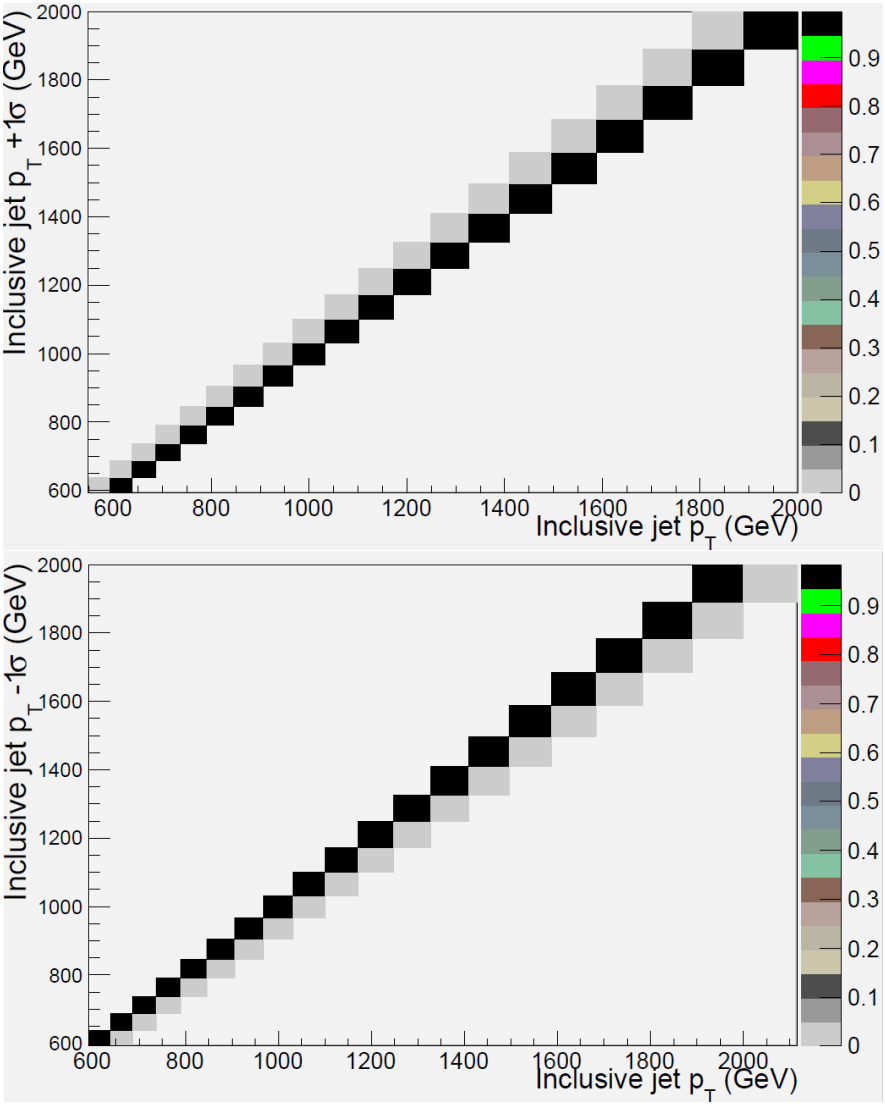


Figure 6.8: Bin-to-bin correlation matrix for nominal jet p_{T} versus jet p_{T} shifted up by 1σ and down by 1σ in pileup rate uncertainty

where N is an electronic noise term, S is a stochastic term for the calorimeter, and C is a constant. The resolution function creates a smearing effect, similar to the transfer matrix which models the JES systematic. The resolution, however, can cause jets to migrate bin-to-bin in both directions in p_{T} at any value of the uncertainty.

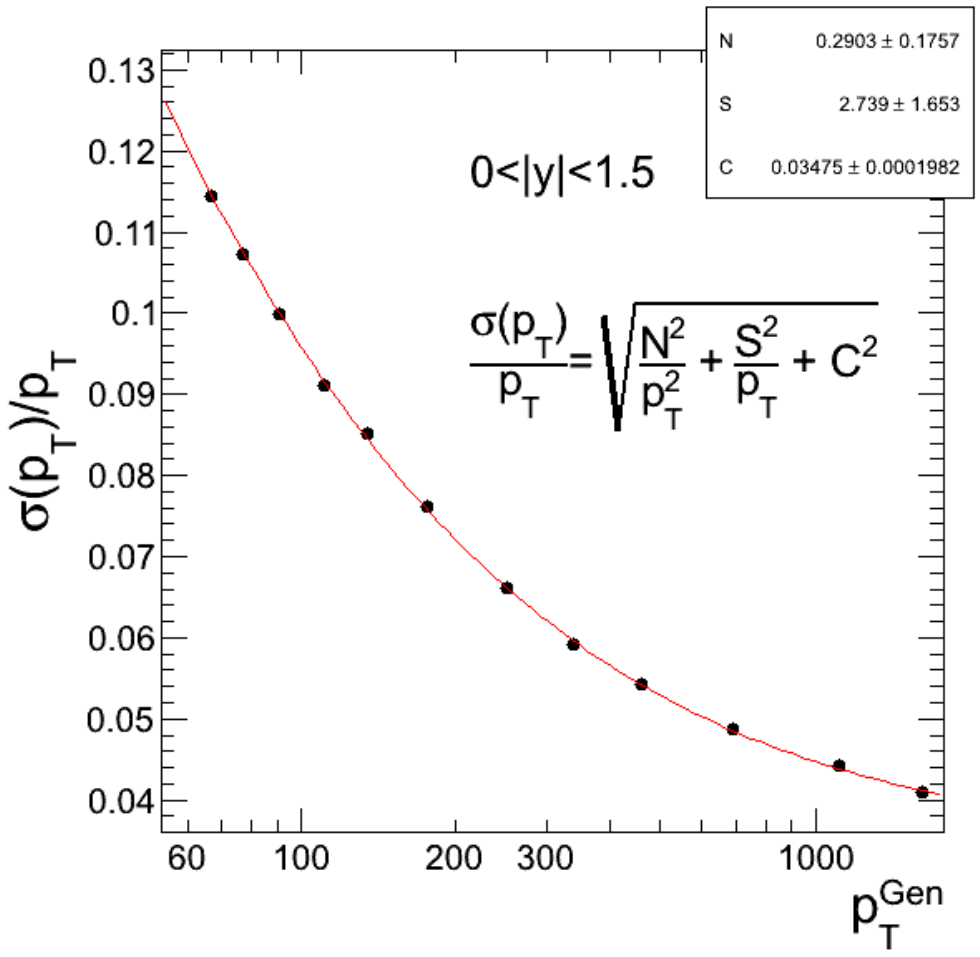


Figure 6.9: Detector response with respect to jet p_{T} and functional fit

The uncertainty due to misunderstanding of the jet energy resolution in smearing is modeled as a 10% uncertainty on the fit parameters N, S and C. The effect of the uncertainty due to JER at 1σ is modeled by repeating the smearing process assuming all fit parameters are increased by 10%. The systematically shifted spectrum is then compared to the central value.

As with the JES uncertainty, the effect of the JER uncertainty is measured between $+2\sigma$ and -2σ in steps of 0.5 σ. The systematic effect is then applied as a transfer matrix in the same fashion. The transfer matrix for the JER systematic at $+1\sigma$ and -1σ is shown in Figure 6.10.

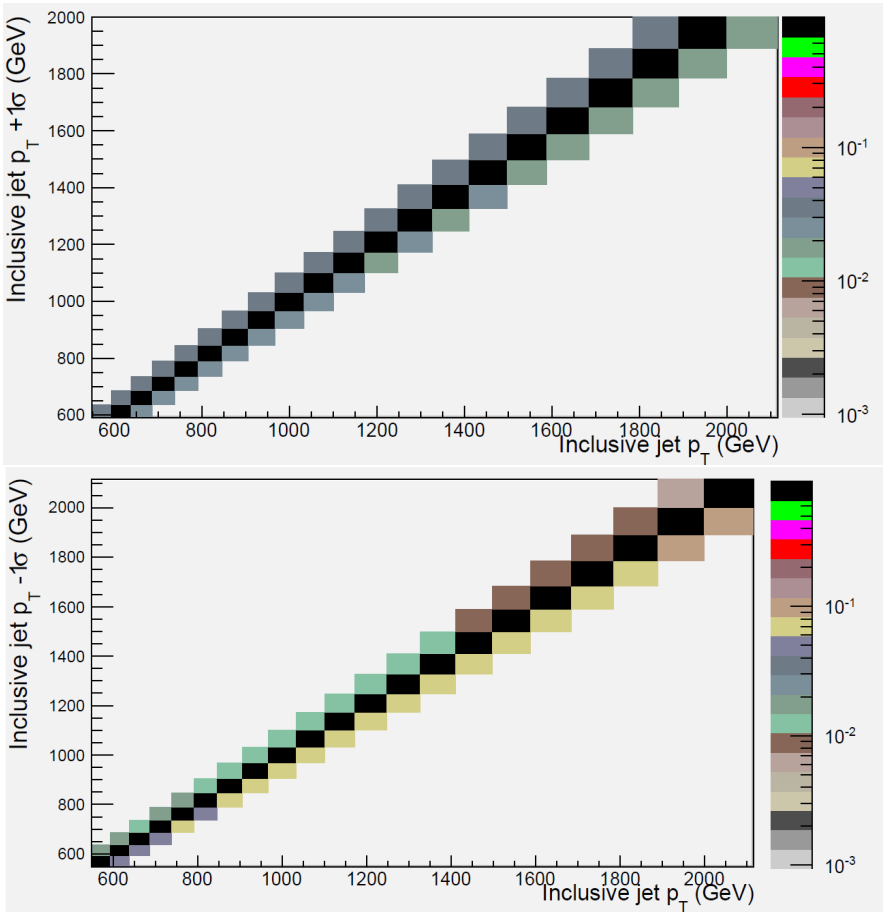


Figure 6.10: Bin-to-bin correlation matrix for jet p_{T} smeared by nominal resolution versus jet p_{T} smeared by the resolution shifted in by 1σ and out by 1σ

6.1.3 Luminosity

The uncertainty on the total luminosity of the dataset is derived from a count of the number of clusters appearing in the pixel detector (CMS Collaboration (2010a)). As the pixel tracker has an extremely high resolution, it is highly unlikely that two tracks from the same beam crossing will pass through the same pixel. The number of observed clusters should therefore be a directly linear function of the luminosity.

The number of clusters are measured using a technique known as a Van der Meer scan. In this technique, the width of the colliding beams is measured by moving them across one another in the x and y planes. Events are selected on a zero-bias trigger, which only requires that a bunch crossing of protons takes place. The number of pixel clusters observed over the course of the scan takes a roughly gaussian form in both x and y. The instantaneous luminosity, or the luminosity delivered per unit time ($d\mathcal{L}/dt$), can be inferred from the measured shape of the beam in this scan. The expected cross section for pixel clusters σ_{pixel} is then given as:

$$\sigma_{pixel} = \langle N_{clusters} \rangle f (d\mathcal{L}/dt)^{-1} \tag{6.7}$$

where $\langle N_{clusters} \rangle$ is the average number of clusters per zero-bias trigger at the peak of the beam scan, and f is the frequency of the LHC.

The uncertainty on this measurement of the luminosity is taken to be 2.2%. It is derived both from uncertainty on the number of pixel events observed in zero-bias events, as well as on the results of the Van der Meer scan used to derive the instantaneous luminosity.

We apply this uncertainty as an absolute scale factor on the expected p_{T} spectrum.

6.2 Summary of Detector Effects

The relative uncertainty in each p_{T} bin as a function of all non-trivial detector-related systematic effects is shown in Figure 6.11. In terms of the raw magnitude of the effect on the p_{T} spectrum, uncertainty to to the high-p_{T} extrapolation, difference in single pion response, and data-driven absolute p_{T} scale are the dominant systematic effects, with a relative 8 to 10% uncertainty at high p_{T}.. The jet energy resolution uncertainty is relatively small by comparison, maximally about 3% at high p_{T}.

In the likelihood function, the expected spectrum for a given value of the detector systematic uncertainties is generated by applying each transfer matrix to the central expected spectrum.

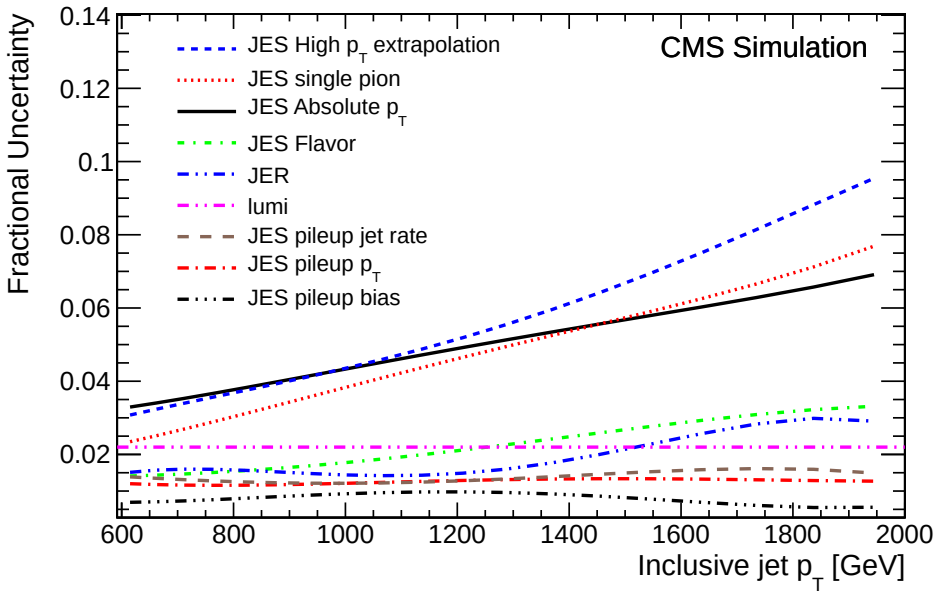


Figure 6.11: Relative uncertainty of all sources of detector uncertainty at 1 σ as a function of jet p_{T}.

The uncertainty on the luminosity is then applied as a simple scale factor.

6.3 Theoretical Uncertainty

The theoretical uncertainty is entirely attributed to associated uncertainty in the parton distribution functions used as a basis for the SM pQCD calculation. As the extinction p_{T} spectrum is expressed in terms of the SM, the theoretical uncertainty will only affect the SM expectation.

The MSTW2008 PDF set as formuated by is used as an input to derive the expected SM inclusive p_{T} spectrum. Monte Carlo integration of the QCD matrix elements is performed, and then weighted in accordance with the central prediction of MSTW2008.

The PDF $f(x, Q^2)$ and coupling constant $\alpha_s(Q^2)$ are fixed at a reference scale $Q_0^2 = 1\,\mathrm{GeV}^2$. The running of the function and constant are then expressed from this reference scale by a series of equations developed by Martin et al. (2009). This parametrization is done independently for u and d quarks, gluons, and sea quarks, and is derived from fits to the data as shown from HERA and the TeVatron in Chapter 2. By definition, this means that any associated uncertainty in the parametrization of the PDFs will be 100% correlated in energy, and therefore jet p_{T}.

The parametrization has 20 free parameters, each of which has an associated uncertainty. The MSTW2008 prediction comprises 41 separate PDF sets. The first is the central value, while the remainder correspond to each of the 20 free parameters at 1σ in positive and negative

uncertainty.

The theoretical uncertainty on the inclusive jet p_{T} spectrum is derived separately for each of the 20 free parameters. The p_{T} spectrum is constructed using each of the 41 PDF sets, and the relative uncertainty for a given PDF parameter is determined by comparing the shifted spectra to the central value. 1σ in a given theoretical source of uncertainty is defined as the relative shift in the p_{T} spectrum with a parameter shifted by its positive uncertainty. The uncertainty is applied in the likelihood as a scale factor, linear in the uncertainty σ.

There is no correlation between the sign of this scale factor and the sign of σ. In many cases, the p_{T} spectrum decreases or increases for both systematic shifts of a given PDF parameter. As with the detector uncertainties, a PDF uncertainty is only considered relevant if it produces at least a 1% relative uncertainty in any p_{T} bin. Of the 20 eigenvector pairs, seven satisfy this criterion. The relative uncertainty due to a systematic shift of 1σ in each of these seven PDF parameters is shown in Figure 6.12.

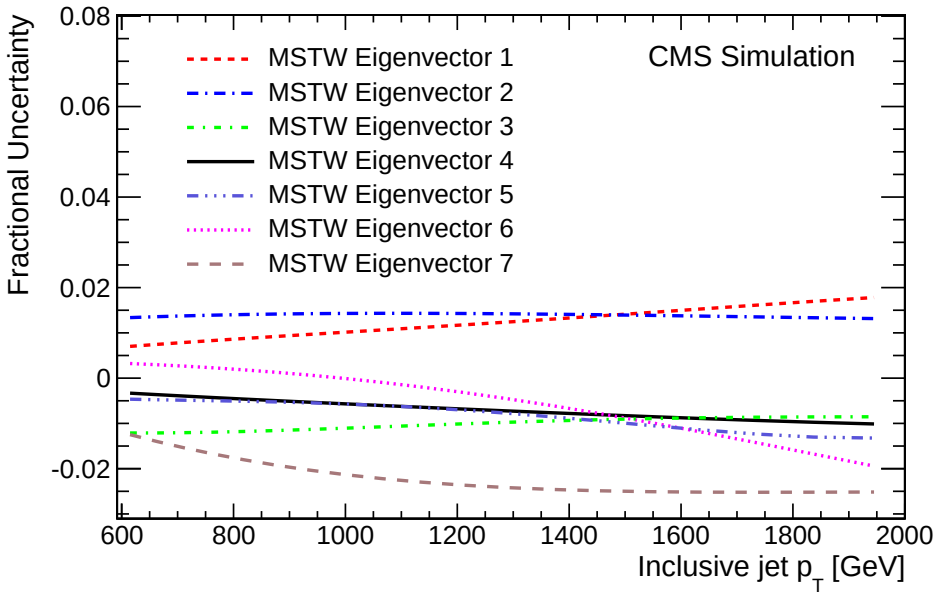


Figure 6.12: Relative uncertainty of all sources of theoretical uncertainty at 1 σ as a function of jet p_{T}.

The relative component uncertainty from all 40 non-central MSTW PDF sets can be seen in Appendix B.

Chapter 7
Presentation of Results

7.1 Inclusive jet p_{T} spectrum

The results of the inclusive jet p_{T} measurement for all jets with $p_{\mathrm{T}} > 592\,\mathrm{GeV}$ and $|\eta| < 1.5$ in 5.0 fb^{-1} of data taken at the CMS detector is shown in Table 7.1. The observed count is shown for each p_{T} bin, as well as the expected count as predicted by the central MSTW2008 PDF set. The associated statistical uncertainty is included, along with the sum of all sources of systematic uncertainty added in quadrature.

In all p_{T} bins below 1.4 TeV, the systematic uncertainty is much larger than the statistical uncertainty. In any given likelihood comparison, the differences between the predicted and measured spectra will nearly always be covered by systematics rather than attributed to the data.

The differential quantity $dN_{jets}/dp_{\mathrm{T}}$ as a function of inclusive jet p_{T} is shown in Figure 7.1. The prediction of MSTW2008 in this figure is normalized to the number of jets observed in the data. In this logarithmic scale, there is no obvious visible deviation between data and expectation within uncertainty.

The measurement in Figure 7.1 is again shown in Figure 7.2 normalized to the MSTW2008 central prediction. The latter figure includes additional spectra from simulation assuming extinction with $M = 2, 3$, and $4\,\mathrm{TeV}$. The data continue to show good agreement with the predictions of pQCD, though there is a small deficit of jets in many of the last few p_{T} bins.

7.2 Extinction

The data is compared to the extinction model using a likelihood function. The likelihood is discussed in Chapter 5, and the details of the comparison are discussed in Appendix C. As part of this comparison, a maximum likelihood fit is performed of the model to the data. The likelihood is expressed in terms of a modified version of the extinction scale, $\beta \equiv M^{-2}$. The maximum likelihood value of β is considered the measured value of the extinction scale in this

p_{T} range (GeV)	$\mathrm{N}_{jets}(observed)$	$\mathrm{N}_{jets}(expected)$	Stat. Unc.	Sys. Unc.
$592 < p_{\mathrm{T}} < 638$	99977	105377	0.3%	9.1%
$638 < p_{\mathrm{T}} < 686$	60421	60828	0.4%	9.2%
$686 < p_{\mathrm{T}} < 737$	37047	37843	0.5%	9.6%
$737 < p_{\mathrm{T}} < 790$	21980	22761	0.7%	9.9%
$790 < p_{\mathrm{T}} < 846$	13159	13323	0.9%	9.9%
$846 < p_{\mathrm{T}} < 905$	7890	7897	1.1%	10.2%
$905 < p_{\mathrm{T}} < 967$	4530	4572	1.5%	10.4%
$967 < p_{\mathrm{T}} < 1032$	2586	2674	2.0%	11.0%
$1032 < p_{\mathrm{T}} < 1101$	1502	1504	2.6%	11.2%
$1101 < p_{\mathrm{T}} < 1172$	906	832	3.3%	11.9%
$1172 < p_{\mathrm{T}} < 1248$	495	465	4.5%	12.3%
$1248 < p_{\mathrm{T}} < 1327$	217	247	6.8%	12.8%
$1327 < p_{\mathrm{T}} < 1410$	129	129	8.8%	13.4%
$1410 < p_{\mathrm{T}} < 1497$	60	67	12.9%	13.6%
$1497 < p_{\mathrm{T}} < 1588$	30	32	18.3%	15.2%
$1588 < p_{\mathrm{T}} < 1684$	24	15	20.4%	16.5%
$1684 < p_{\mathrm{T}} < 1784$	5	7	44.7%	16.3%
$1784 < p_{\mathrm{T}} < 1890$	3	3	57.7%	17.9%
$1890 < p_{\mathrm{T}} < 2000$	3	1	57.7%	19.9%

Table 7.1: Expected and observed jet counts by p_{T} bin, including systematic and statistical uncertainty.

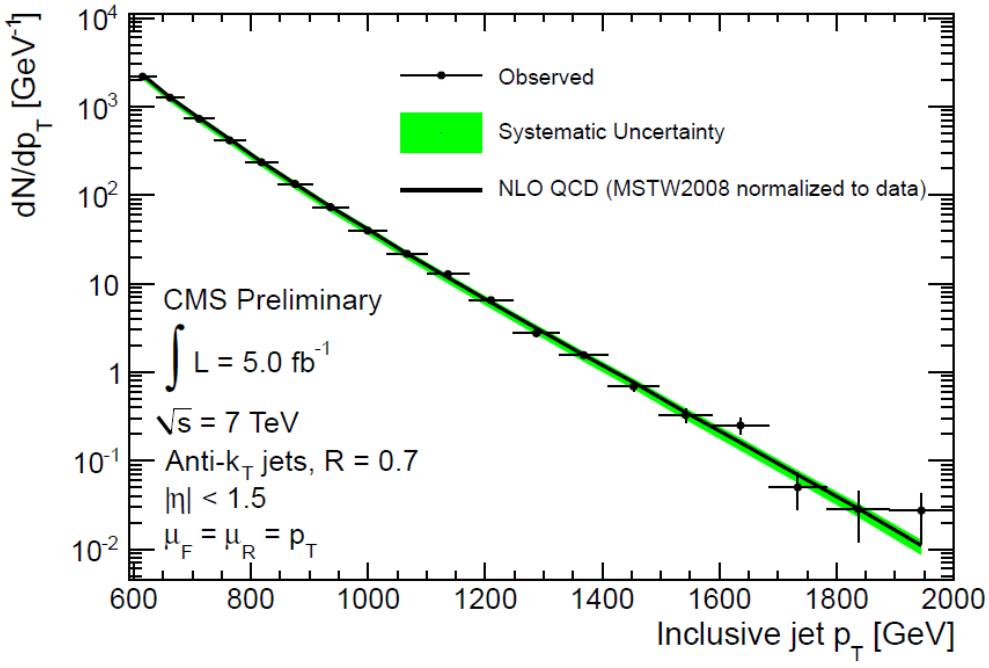


Figure 7.1: Inclusive jet p_{T} spectrum as observed in 5.0 fb^{-1} of data taken at the CMS detector. The prediction at NLO of the MSTW2008 PDF set is included, as well as the overall magnitude of the systematic uncertainty.

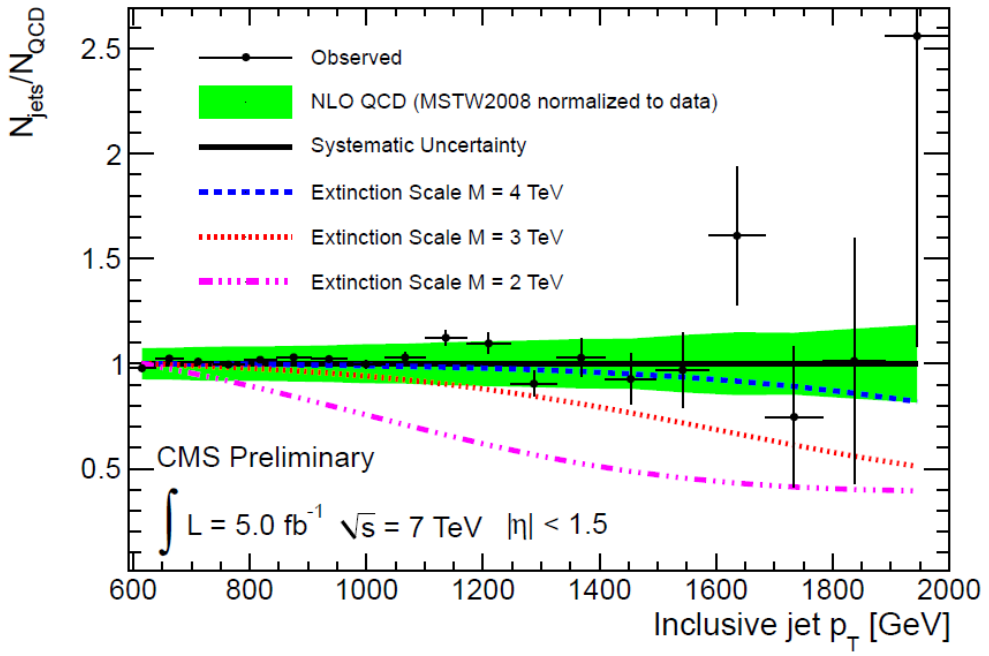


Figure 7.2: Observed inclusive jet p_{T} spectrum normalized to the NLO prediction of the MSTW2008 PDF set. Similarly normalized spectra as expected given extinction masses of $M = 2, 3$ and 4 TeV are also shown.

analysis. The error on this value of β then corresponds to the error on the observed extinction scale. A observation of $\beta = 0$ correspons to the standard model.

The fit is first performed taking into account only statistical error, with all sources of systematic uncertainty fixed at the central values. The residuals of this fit are shown in Figure 7.3. This is not terribly meaningful as the analysis is statistics limited, but it is a useful measure of the general properties of the dataset and a good sanity check of the comparison. The fit returns a maximum likelihood value of $\beta = 0.091 \pm 0.002$. Assuming a perfect detector and perfect understanding of the jet energy scale and luminosity, the small deviation between the expected and observed spectra would indicate the presence of extinction at a very high degree of significance.

When the fit is performed with all systematics varied between $\pm 2\sigma$, the fit returns a maximum likelihood value of $\beta = 0.045 \pm 0.055$, with residuals shown in Figure 7.4. The small deviations in the observed spectrum are reflected in the slight positive value. The large uncertainty relative to the best fit value indicates good agreement with the standard model. The maximum likelihood value of the JES uncertainty due to the high-p_{T} extrapolation is close to 2σ, with most other systematic sources at or near the central value. The differences between the pQCD calculation and the observed jet occupancy can then be attributed to a systematic mismeasurement of the jet energy scale in the very high-p_{T} region. Given that the phase space of the analysis is restricted solely to central high-p_{T} jets, this is a reasonable conclusion.

Figure 7.5 shows the result of a $\mathrm{CL_s}$ scan on the modified extinction scale β using the method described in Appendix C. The range of β for which $\mathrm{CL_s} < 0.05$ is considered to be excluded at 95% confidence level (C.L.) For the dataset considered in this search, the region of $\beta > 0.094\,\mathrm{TeV}^{-2}$ is excluded at this level. This is consistent with the observed maximum likelihood value of β and its associated uncertainty. The exclusion region corresponds to an observed lower limit of 3.2 TeV on the extinction scale M. The scale in this case is directly comparable to the Planck scale M_{Planck}. This limit is quite competitive with the model-specific limits on M_{Planck} set in searches for extra dimensions and black hole signatures discussed in Chapter 2, which range from 2 to 5 TeV depending on the model.

7.3 Conclusion

A full search for evidence of strong gravity manifesting as an extinction effect has been performed. A comparison has been made between the observed inclusive jet p_{T} spectrum and the predictions of perturbative QCD at NLO, and a lower limit has been set on the extinction

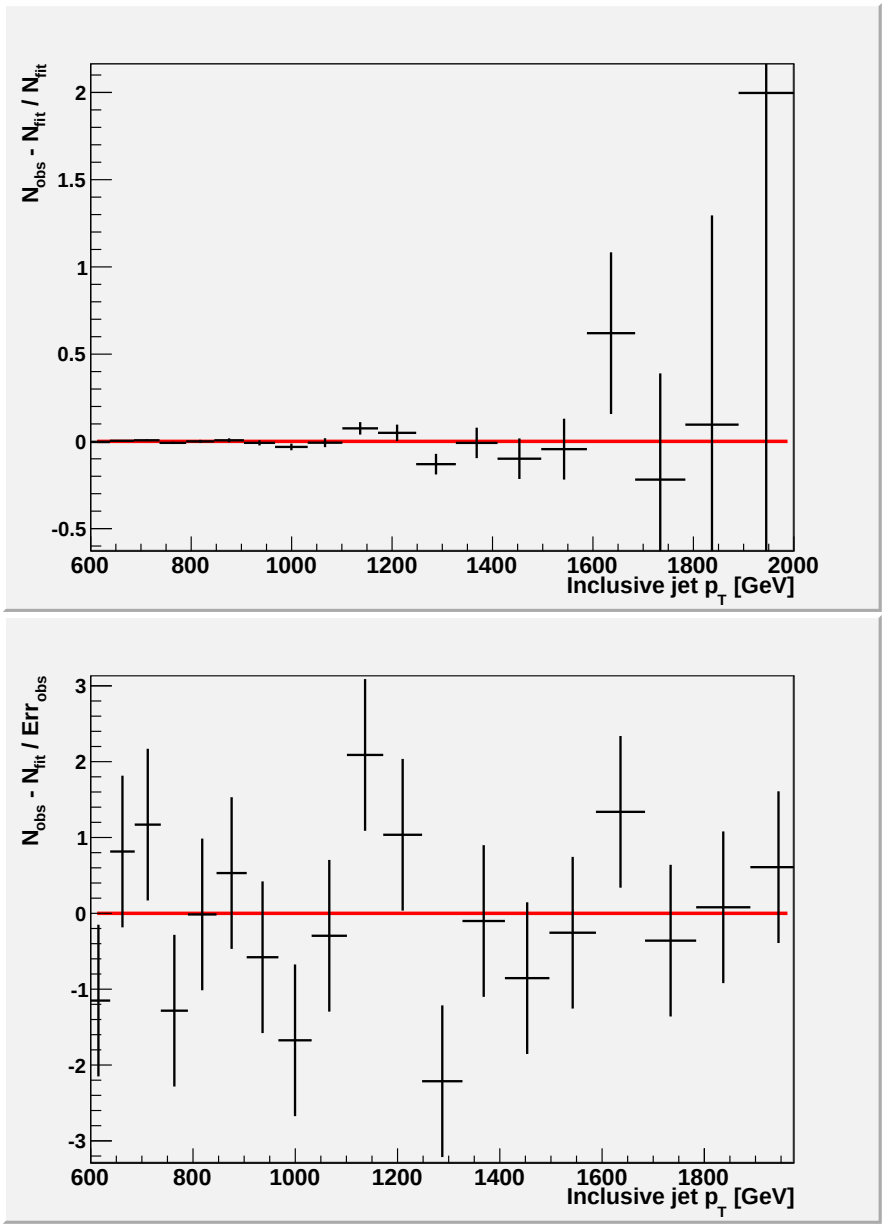


Figure 7.3: Residuals of the best fit for our Extinction model given the data, statistical errors only.

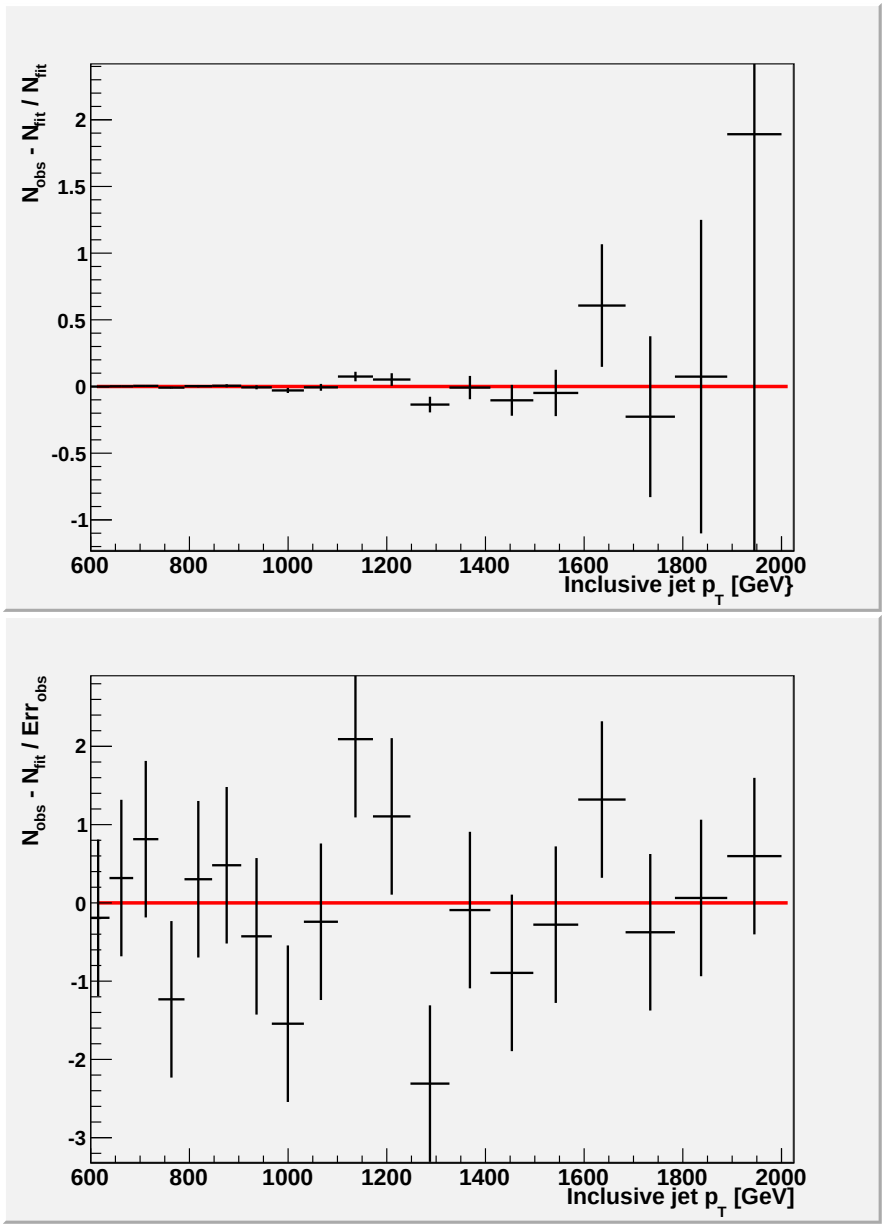


Figure 7.4: Residuals of the best fit for our Extinction model given the data and all systematics.

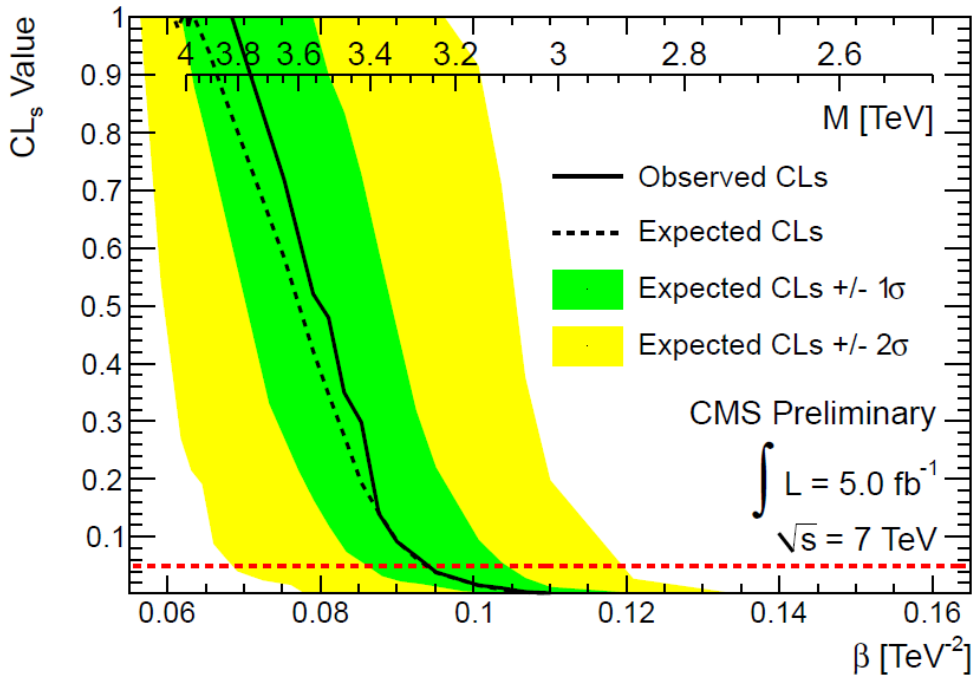


Figure 7.5: Frequentist criterion CL_s as a function of the extinction parameter M and the modified parameter $\beta \equiv M^-2$

scale M of 3.2 TeV. This limit is competitive with other similar searches designed to look for model-specific signatures of strong gravity.

It is a testament to the strength of the tremendous collaborative effort at CMS and other accelerator experiments that a precision measurement of inclusive jet production may be performed, with the jet energy scale understood in the high-p_T regime to the percent level. The extreme statistical power of the jet dataset may now be leveraged to perform a smooth shape analysis out to the edge of our reach in jet momentum. The extinction search may be performed in fairly short order as data continues to accrue at the LHC and beyond, and will continue to serve as a complimentary limit to black hole searches and other similar analyses in the future.

Bibliography

Aad, G., et al. 2008, Journal of Instrumentation, 3, S08003

Aad, G., Abbott, B., Abdallah, J., et al. 2012, , 86, 014022

Aad, G., et al. 2012, Phys. Lett. B, 710, 538

Aaltonen, T., Adelman, J., Akimoto, T., et al. 2008, Phys. Rev. D, 78, 052006

Aaron, F. D., Abramowicz, H., Abt, I., et al. 2010, Journal of High Energy Physics, 1, 109

Abazov, V. M., Abbott, B., Acharya, B. S., et al. 2012, Phys. Rev. D, 85, 052006

Adolphi, R., et al. 2008, JINST, 0803, S08004

Agostinelli, S., Allison, J., Amako, K., et al. 2003, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506, 250

Alwall, J., Ballestrero, A., Bartalini, P., et al. 2007, Computer Physics Communications, 176, 300

Andersson, B., Gustafson, G., Ingelman, G., & Sjöstrand, T. 1983, , 97, 31

Arkani-Hamed, N., Dimopoulos, S., & Dvali, G. 1999, Phys. Rev. D, 59, 086004

ATLAS Collaboration Aad, G., Abajyan, T., Abbott, B., et al. 2012a, Physics Letters B, 716, 1

ATLAS Collaboration Aad, G., Abbott, B., Abdallah, J., et al. 2012b, Physics Letters B, 708, 37

Cacciari, M., Salam, G. P., & Soyez, G. 2008, Journal of High Energy Physics, 2008, 063

Chatrchyan, S., Khachatryan, V., Sirunyan, A. M., et al. 2011, Journal of High Energy Physics, 9, 109

Chatrchyan, S., et al. 2012a, JHEP, 1204, 061

—. 2012b, Phys. Rev. Lett., 108, 111801

CMS Collaboration. 2010a, Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update, CMS Physics Analysis Summary CMS-PAS-SMP-12-008

—. 2010b, Commissioning of the Particle-flow Event Reconstruction with the First LHC Collisions Recorded in the CMS Detector, CMS Physics Analysis Summary CMS-PAS-PFT-10-001

—. 2010c, Jet cross sections and PDF constraints, CMS Physics Analysis Summary CMS-PAS-QCD-11-004

—. 2010d, Jet Energy Scale performance in 2011, CMS Detector Performance Summary CMS-DP-2012-006

—. 2010e, Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET, CMS Physics Analysis Summary CMS-PAS-PFT-09-001

CMS Collaboration Chatrchyan, S., Khachatryan, V., Sirunyan, A. M., et al. 2011a, Phys. Rev. Lett., 107, 132001

—. 2011b, Physics Letters B, 704, 123

—. 2012, Physics Letters B, 716, 30

Cowan, G., Cranmer, K., Gross, E., & Vitellis, O. 2011, European Physical Journal C, 71, 1554

Dimopoulos, S., & Landsberg, G. 2001, Physical Review Letters, 87, 161602

Evans, L., & Bryant, P. 2008, Journal of Instrumentation, 3, S08001

Feynman, R. P., Field, R. D., & Fox, G. C. 1978, Phys. Rev. D, 18, 3320

Field, R. 2010, ArXiv e-prints

Giddings, S. B., & Thomas, S. D. 2002, Phys. Rev., D65, 056010

Glashow, S. 1961, , 22, 579

Higgs, P. W. 1964, Physical Review Letters, 13, 508

James, F., & Roos, M. 1975, Comput. Phys. Commun., 10, 343

Junk, T. 1999, Nucl.Instrum.Meth., A434, 435

Kaluza, T. 1921, Zum Unitätsproblem der Physik, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin

Kilic, C., Lath, A., Rose, K., & Thomas, S. 2012, ArXiv e-prints

Kluge, T., Rabbertz, K., & Wobisch, M. 2006, in 14th International Workshop on Deep Inelastic Scattering (DIS 2006), 20-24 Apr 2006, Tsukuba, Japan, 483

Lüst, D., Stieberger, S., & Taylor, T. R. 2009, Nuclear Physics B, 808, 1

Martin, A. D., Stirling, W. J., Thorne, R. S., & Watt, G. 2009, European Physical Journal C, 63, 189

Nagy, Z. 2002, Phys. Rev. Lett., 88, 122003

—. 2003, Phys. Rev. D, 68, 094002

Randall, L., & Sundrum, R. 1999, Phys. Rev. Lett., 83, 3370

Read, A. L. 2002, J. Phys. G, 28, 2693

Salam, A. 1968, in Elementary Particle Theory, ed. W. Svartholm (Stockholm: Almquist and Wiksell), 367

Sjöstrand, T., Mrenna, S., & Skands, P. Z. 2006, JHEP, 05, 026

't Hooft, G., & Veltman, M. 1972, Nuclear Physics B, 44, 189

Veneziano, G. 1968, Nuovo Cimento, 57, 190

Voutilainen, M. A. 2008

Weinberg, S. 1967, Physical Review Letters, 19, 1264

Appendix A

Table of Extinction Cross-sections

A.1 Tables of Cross-sections used in Extinction simulation

The extinction physics in this thesis are not currently in the standard library of physics models included with PYTHIA. To simulate the effects of extinction on QCD, the leading-order cross-sections of all $2 \to 2$ scattering processes are calculated by hand using the matrix elements shown in Chapter 2. These cross-sections are then added by hand as input into PYTHIA. Events are simulated accordingly as discussed in Chapter 4.

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.007×10^{3}	1.011×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.395×10^{2}	1.411×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	3.044×10^{1}	3.134×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	7.690×10^{0}	8.573×10^{0}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	2.082×10^{0}	2.585×10^{0}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	6.101×10^{-1}	8.178×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	1.837×10^{-1}	2.658×10^{-1}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	5.577×10^{-2}	8.581×10^{-2}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.632×10^{-2}	2.666×10^{-2}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	4.664×10^{-3}	8.093×10^{-3}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	1.276×10^{-3}	2.300×10^{-3}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	3.233×10^{-4}	6.120×10^{-4}
$4100 < \hat{p_{\mathrm{T}}}$	9.561×10^{-5}	1.920×10^{-4}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.008×10^{3}	1.008×10^{3}	1.008×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.422×10^{2}	1.408×10^{2}	1.409×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	3.139×10^{1}	3.188×10^{1}	3.129×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	8.806×10^{0}	8.822×10^{0}	8.947×10^{0}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	2.761×10^{0}	2.758×10^{0}	2.855×10^{0}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	9.040×10^{-1}	8.799×10^{-1}	9.515×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	3.007×10^{-1}	2.873×10^{-1}	3.329×10^{-1}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	1.008×10^{-1}	9.336×10^{-2}	1.160×10^{-1}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	3.335×10^{-2}	3.002×10^{-2}	3.978×10^{-2}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	1.056×10^{-2}	9.521×10^{-3}	1.330×10^{-2}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	3.160×10^{-3}	2.882×10^{-3}	4.247×10^{-3}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	9.044×10^{-4}	8.308×10^{-4}	1.292×10^{-3}
$4100 < \hat{p_{\mathrm{T}}}$	3.096×10^{-4}	2.886×10^{-4}	4.831×10^{-5}

Table A.1: Table of cross-sections for $qq \to qq$ production for all simulated samples

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.525×10^{1}	1.621×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.770×10^{0}	1.867×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	2.736×10^{-1}	3.141×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	4.982×10^{-2}	5.602×10^{-2}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	5.941×10^{-3}	9.050×10^{-3}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	6.266×10^{-4}	1.183×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	1.001×10^{-4}	1.855×10^{-4}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	2.039×10^{-5}	3.525×10^{-5}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	2.358×10^{-6}	5.580×10^{-6}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	2.413×10^{-7}	1.614×10^{-6}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	7.659×10^{-8}	6.229×10^{-8}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	1.159×10^{-9}	3.977×10^{-8}
$4100 < \hat{p_{\mathrm{T}}}$	9.215×10^{-10}	7.660×10^{-9}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.613×10^{1}	1.622×10^{1}	1.697×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.924×10^{0}	1.777×10^{0}	1.983×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	3.655×10^{-1}	3.221×10^{-1}	3.435×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	6.571×10^{-2}	6.568×10^{-2}	6.346×10^{-2}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	1.339×10^{-2}	1.705×10^{-2}	1.768×10^{-2}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	3.132×10^{-3}	3.130×10^{-3}	3.751×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	4.837×10^{-4}	6.384×10^{-4}	1.179×10^{-3}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	7.609×10^{-5}	1.105×10^{-4}	2.571×10^{-4}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.015×10^{-5}	1.552×10^{-5}	6.640×10^{-5}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	1.117×10^{-6}	4.551×10^{-6}	1.399×10^{-5}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	2.350×10^{-7}	5.561×10^{-7}	2.338×10^{-6}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	7.057×10^{-8}	4.544×10^{-8}	4.556×10^{-7}
$4100 < \hat{p_{\mathrm{T}}}$	4.151×10^{-9}	2.148×10^{-8}	5.049×10^{-8}

Table A.2: Table of cross-sections for $q\bar{q} \to q\bar{q}$ production for all simulated samples

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.608×10^{1}	1.599×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.746×10^{0}	1.844×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	3.179×10^{-1}	3.216×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	5.551×10^{-2}	5.724×10^{-2}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	1.014×10^{-2}	1.272×10^{-2}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	1.872×10^{-3}	2.371×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	3.631×10^{-4}	3.978×10^{-4}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	7.381×10^{-5}	1.158×10^{-4}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.042×10^{-5}	2.079×10^{-5}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	2.837×10^{-6}	4.941×10^{-6}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	3.512×10^{-7}	5.472×10^{-7}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	1.002×10^{-7}	1.298×10^{-7}
$4100 < \hat{p_{\mathrm{T}}}$	2.048×10^{-8}	3.044×10^{-8}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	1.494×10^{1}	1.479×10^{1}	1.685×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.819×10^{0}	1.777×10^{0}	1.878×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	3.357×10^{-1}	3.113×10^{-1}	3.221×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	7.199×10^{-2}	6.852×10^{-2}	7.492×10^{-2}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	1.609×10^{-2}	1.732×10^{-2}	1.717×10^{-2}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	3.615×10^{-3}	3.865×10^{-3}	5.090×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	6.886×10^{-4}	9.671×10^{-4}	1.350×10^{-3}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	1.858×10^{-4}	2.499×10^{-4}	3.550×10^{-4}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	4.514×10^{-5}	5.000×10^{-5}	7.311×10^{-5}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	8.195×10^{-6}	1.211×10^{-5}	1.893×10^{-5}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	2.288×10^{-6}	1.791×10^{-6}	5.380×10^{-6}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	2.447×10^{-7}	4.972×10^{-7}	1.041×10^{-6}
$4100 < \hat{p_{\mathrm{T}}}$	1.039×10^{-7}	9.485×10^{-8}	1.610×10^{-7}

Table A.3: Table of cross-sections for $q\bar{q} \to gg$ production for all simulated samples

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	3.352×10^{3}	3.356×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	3.245×10^{2}	3.256×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	5.237×10^{1}	5.322×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	1.039×10^{1}	1.097×10^{1}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	2.273×10^{0}	2.525×10^{0}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	5.406×10^{-1}	6.103×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	1.333×10^{-1}	1.583×10^{-1}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	3.382×10^{-2}	4.124×10^{-2}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	8.357×10^{-3}	1.078×10^{-2}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	2.051×10^{-3}	2.728×10^{-3}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	4.884×10^{-4}	6.848×10^{-4}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	1.075×10^{-4}	1.607×10^{-4}
$4100 < \hat{p_{\mathrm{T}}}$	2.742×10^{-5}	4.328×10^{-5}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	3.360×10^{3}	3.392×10^{3}	3.362×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	3.266×10^{2}	3.298×10^{2}	3.283×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	5.401×10^{1}	5.400×10^{1}	5.417×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	1.140×10^{1}	1.157×10^{1}	1.164×10^{1}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	2.778×10^{0}	2.858×10^{0}	2.928×10^{0}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	7.237×10^{-1}	7.603×10^{-1}	8.064×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	1.943×10^{-1}	2.063×10^{-1}	2.352×10^{-1}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	5.328×10^{-2}	5.808×10^{-2}	6.955×10^{-2}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.465×10^{-2}	1.634×10^{-2}	2.031×10^{-2}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	4.005×10^{-3}	4.457×10^{-3}	5.860×10^{-3}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	1.047×10^{-3}	1.214×10^{-3}	1.663×10^{-3}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	2.642×10^{-4}	3.101×10^{-4}	4.461×10^{-4}
$4100 < \hat{p_{\mathrm{T}}}$	7.744×10^{-5}	9.386×10^{-5}	1.442×10^{-4}

Table A.4: Table of cross-sections for $qg \to qg$ production for all simulated samples

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	8.685×10^{1}	8.537×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	5.474×10^{0}	5.422×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	6.089×10^{-1}	6.310×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	8.258×10^{-2}	9.430×10^{-2}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	1.184×10^{-2}	1.578×10^{-2}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	1.930×10^{-3}	2.912×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	3.632×10^{-4}	4.421×10^{-4}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	5.725×10^{-5}	7.577×10^{-5}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	9.108×10^{-6}	1.743×10^{-5}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	1.623×10^{-6}	3.648×10^{-6}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	3.107×10^{-7}	5.472×10^{-7}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	5.890×10^{-8}	1.117×10^{-7}
$4100 < \hat{p_{\mathrm{T}}}$	1.133×10^{-8}	3.222×10^{-8}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	8.452×10^{1}	8.594×10^{1}	8.679×10^{1}
$800 < \hat{p_{\mathrm{T}}} < 1100$	5.544×10^{0}	5.480×10^{0}	5.395×10^{0}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	6.894×10^{-1}	6.227×10^{-1}	6.475×10^{-1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	1.053×10^{-1}	1.011×10^{-1}	1.005×10^{-1}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	1.755×10^{-2}	2.056×10^{-2}	2.240×10^{-2}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	3.542×10^{-3}	4.747×10^{-3}	4.528×10^{-3}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	7.736×10^{-4}	8.403×10^{-4}	1.193×10^{-3}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	1.486×10^{-4}	1.947×10^{-4}	2.826×10^{-4}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	3.275×10^{-5}	4.818×10^{-5}	8.598×10^{-5}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	7.044×10^{-6}	7.791×10^{-6}	1.749×10^{-5}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	2.107×10^{-6}	2.275×10^{-6}	4.449×10^{-6}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	3.208×10^{-7}	4.079×10^{-7}	9.006×10^{-7}
$4100 < \hat{p_{\mathrm{T}}}$	6.290×10^{-8}	8.644×10^{-8}	2.763×10^{-7}

Table A.5: Table of cross-sections for $gg \to q\bar{q}$ production for all simulated samples

	Production cross-section (pb)	
p_{T} bin (GeV)	$M = 2\,\mathrm{TeV}$	$M = 2.25\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	2.414×10^{3}	2.400×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.524×10^{2}	1.511×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	1.778×10^{1}	1.806×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	2.800×10^{0}	2.850×10^{0}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	5.042×10^{-1}	5.194×10^{-1}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	1.024×10^{-1}	1.049×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	2.213×10^{-2}	2.352×10^{-2}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	4.968×10^{-3}	5.308×10^{-3}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.142×10^{-3}	1.218×10^{-3}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	2.561×10^{-4}	2.783×10^{-4}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	5.610×10^{-5}	6.487×10^{-5}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	1.186×10^{-5}	1.404×10^{-5}
$4100 < \hat{p_{\mathrm{T}}}$	2.859×10^{-6}	3.464×10^{-6}

p_{T} bin (GeV)	$M = 2.75\,\mathrm{TeV}$	$M = 3\,\mathrm{TeV}$	$M = 4\,\mathrm{TeV}$
$500 < \hat{p_{\mathrm{T}}} < 800$	2.410×10^{3}	2.404×10^{3}	2.395×10^{3}
$800 < \hat{p_{\mathrm{T}}} < 1100$	1.547×10^{2}	1.535×10^{2}	1.543×10^{2}
$1100 < \hat{p_{\mathrm{T}}} < 1400$	1.838×10^{1}	1.841×10^{1}	1.837×10^{1}
$1400 < \hat{p_{\mathrm{T}}} < 1700$	2.977×10^{0}	3.035×10^{0}	3.052×10^{0}
$1700 < \hat{p_{\mathrm{T}}} < 2000$	5.924×10^{-1}	6.143×10^{-1}	6.242×10^{-1}
$2000 < \hat{p_{\mathrm{T}}} < 2300$	1.289×10^{-1}	1.365×10^{-1}	1.439×10^{-1}
$2300 < \hat{p_{\mathrm{T}}} < 2600$	2.897×10^{-2}	3.249×10^{-2}	3.567×10^{-2}
$2600 < \hat{p_{\mathrm{T}}} < 2900$	6.832×10^{-3}	7.937×10^{-3}	9.253×10^{-3}
$2900 < \hat{p_{\mathrm{T}}} < 3200$	1.663×10^{-3}	2.064×10^{-3}	2.410×10^{-3}
$3200 < \hat{p_{\mathrm{T}}} < 3500$	4.155×10^{-4}	5.118×10^{-4}	6.105×10^{-4}
$3500 < \hat{p_{\mathrm{T}}} < 3800$	1.009×10^{-4}	1.277×10^{-4}	1.594×10^{-4}
$3800 < \hat{p_{\mathrm{T}}} < 4100$	2.250×10^{-5}	3.041×10^{-5}	4.043×10^{-5}
$4100 < \hat{p_{\mathrm{T}}}$	6.247×10^{-6}	8.604×10^{-6}	1.167×10^{-5}

Table A.6: Table of cross-sections for $gg \to gg$ production for all simulated samples

Appendix B

Uncertainty from MSTW eigenvectors

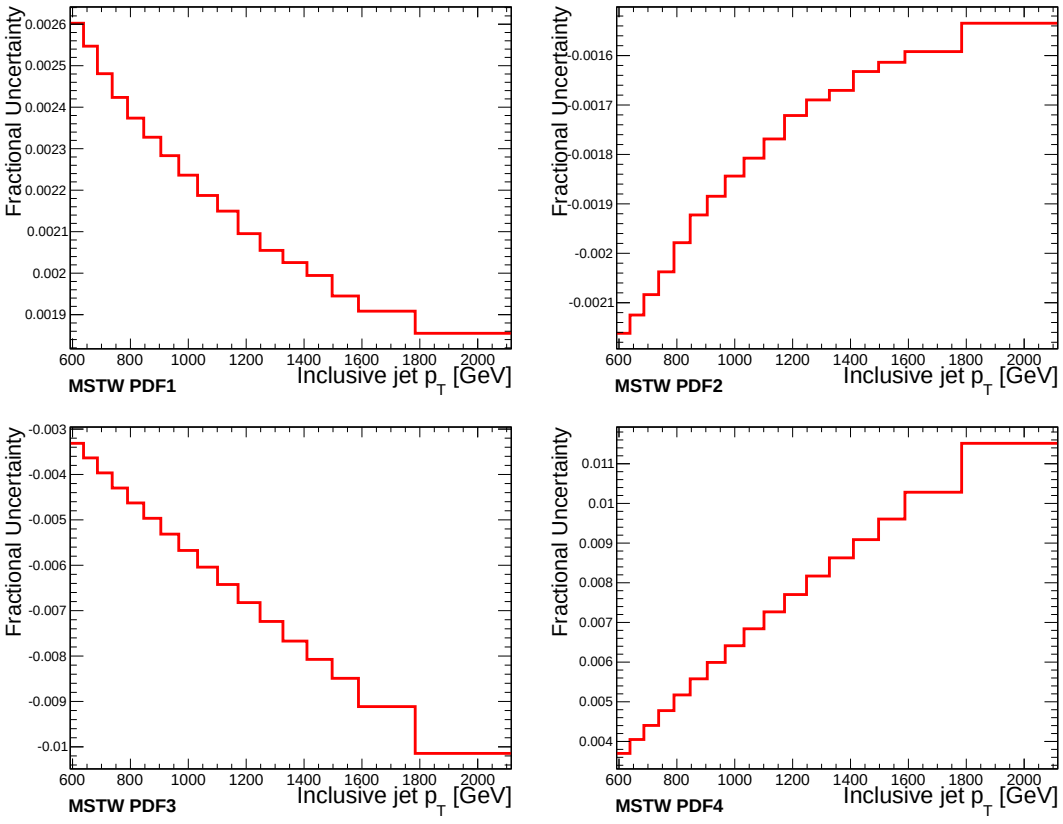


Figure B.1: Difference between MSTW PDF Set 1-4 and the central PDF set, normalized to the central set.

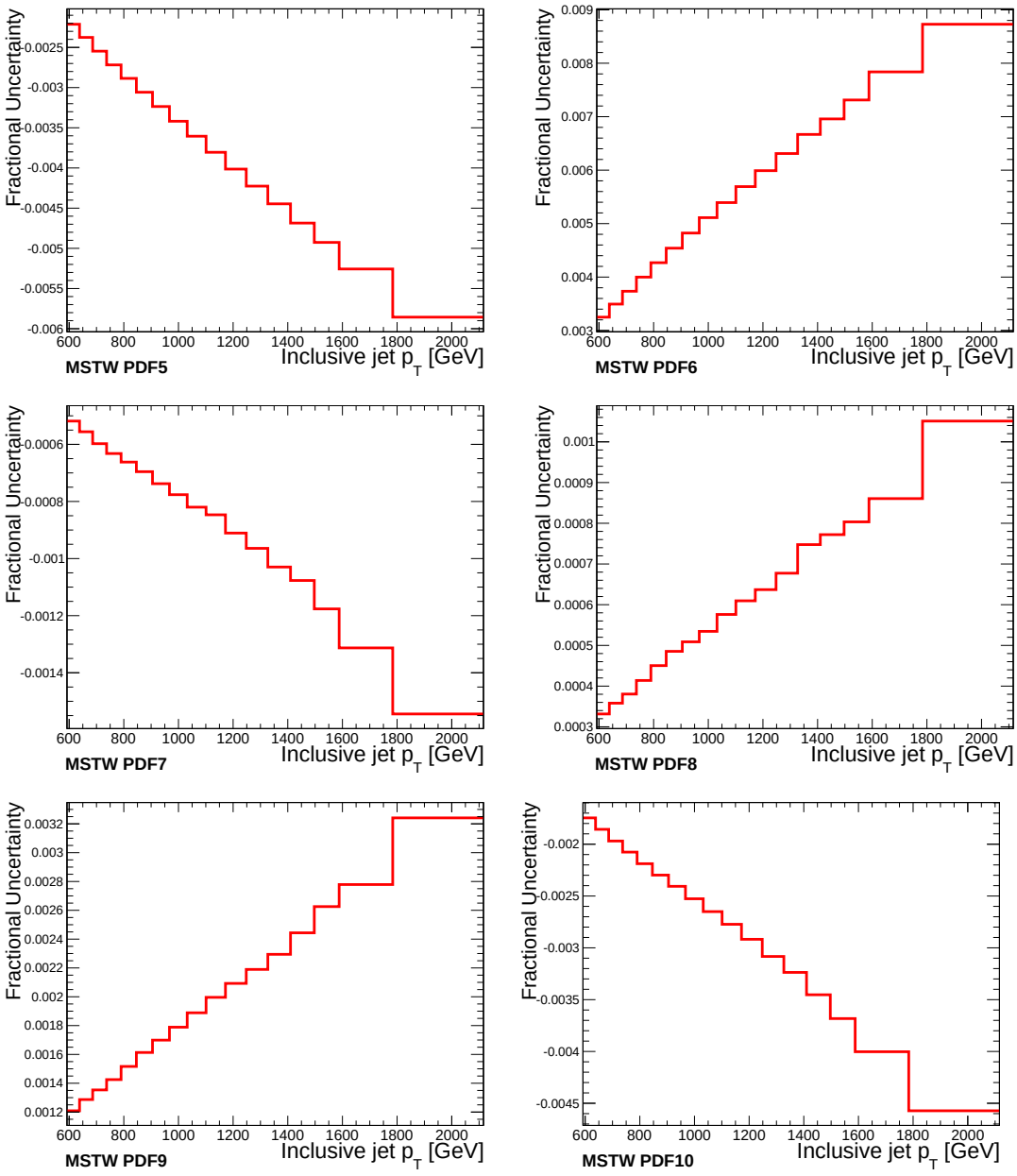


Figure B.2: Difference between MSTW PDF Set 5-10 and the central PDF set, normalized to the central set.

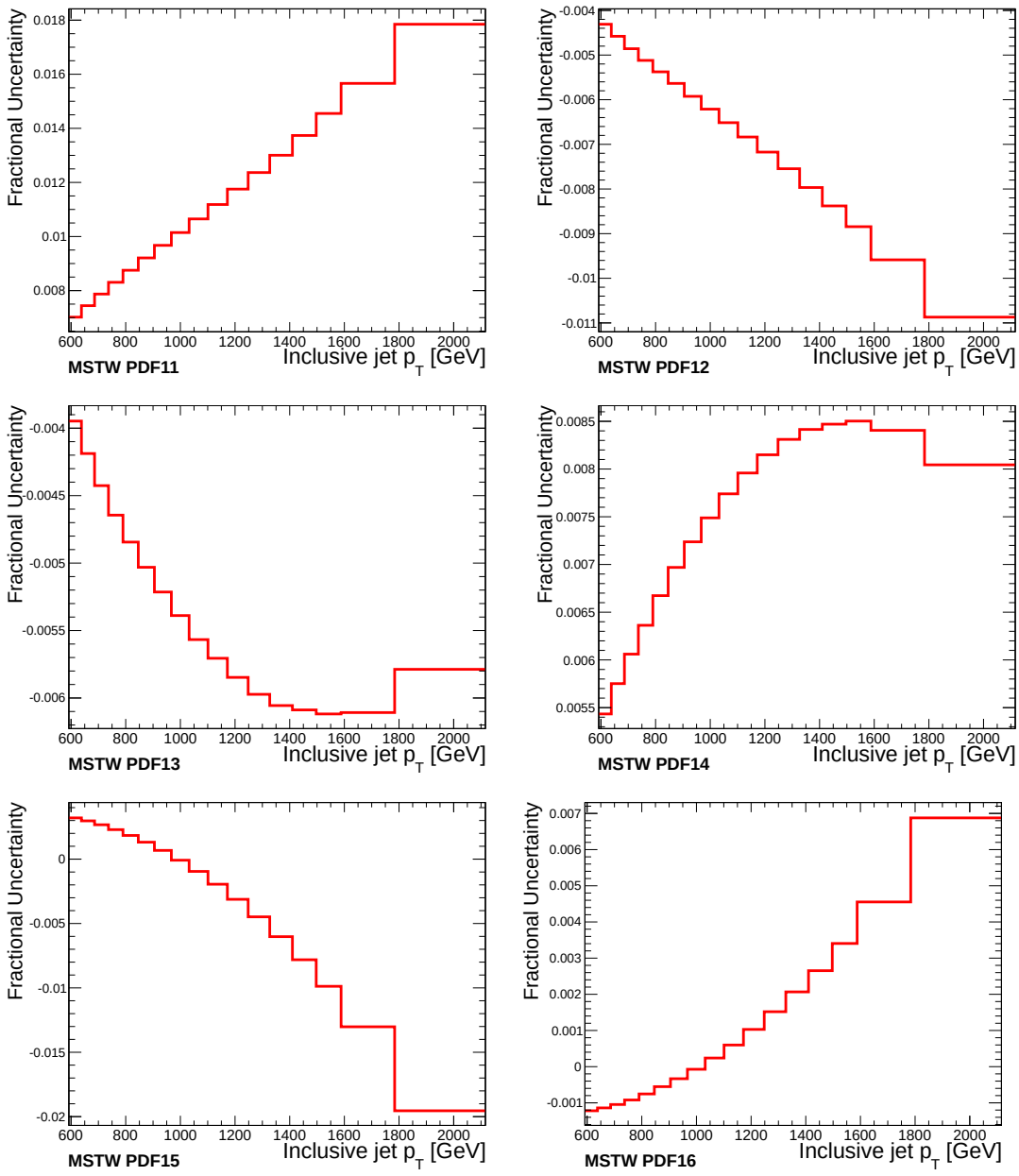


Figure B.3: Difference between MSTW PDF Set 11-16 and the central PDF set, normalized to the central set.

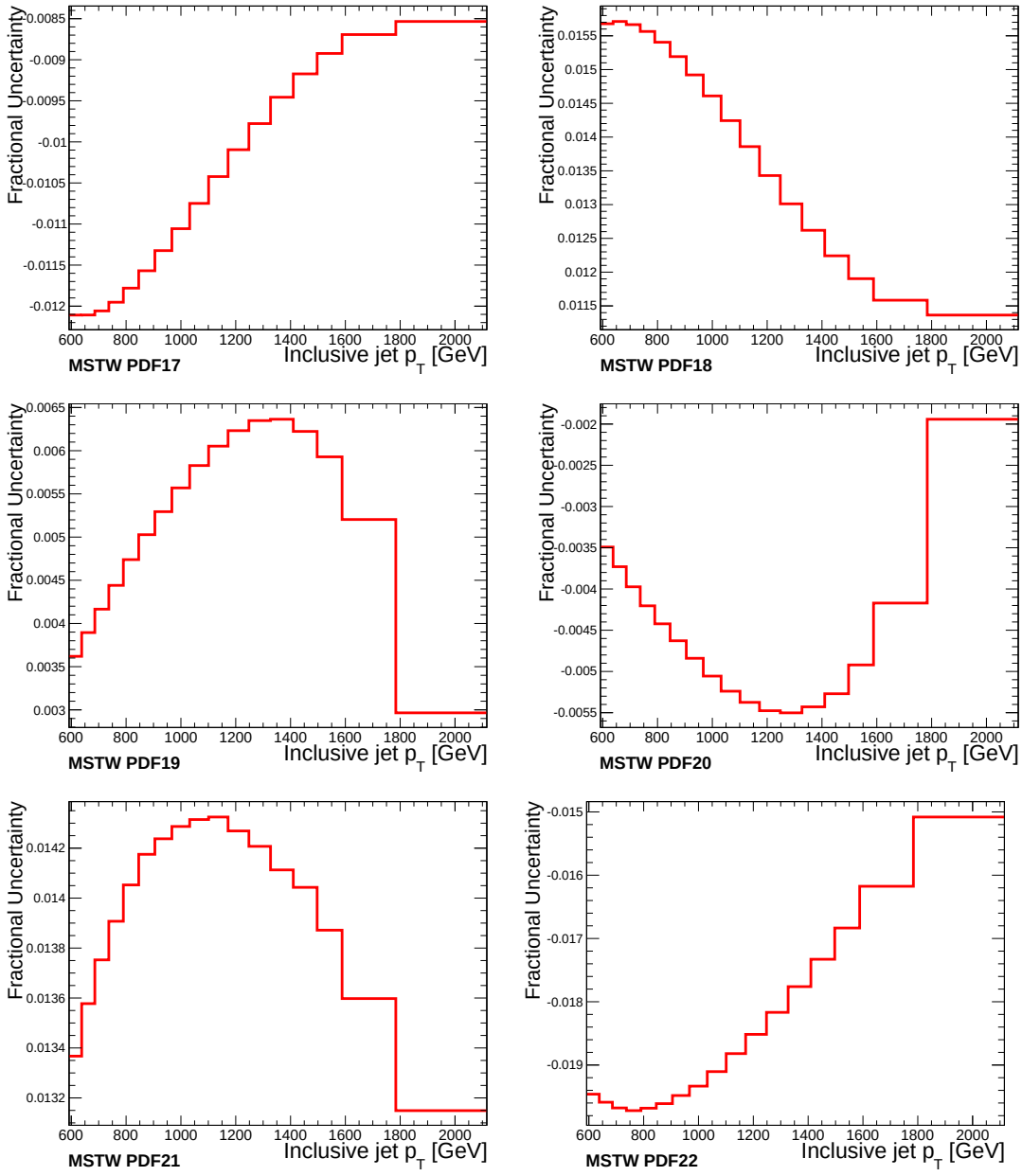


Figure B.4: Difference between MSTW PDF Set 17-22 and the central PDF set, normalized to the central set.

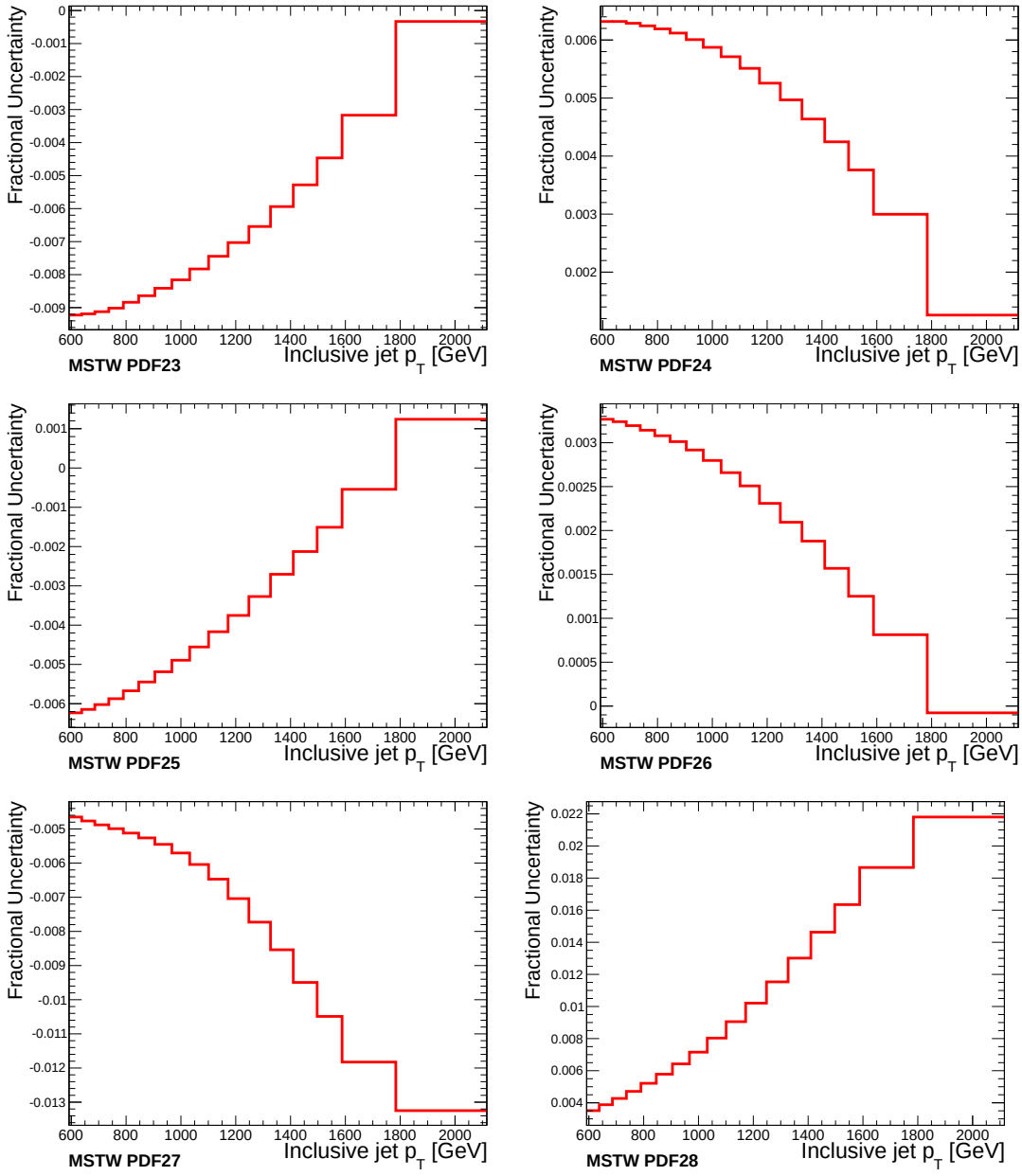


Figure B.5: Difference between MSTW PDF Set 23-28 and the central PDF set, normalized to the central set.

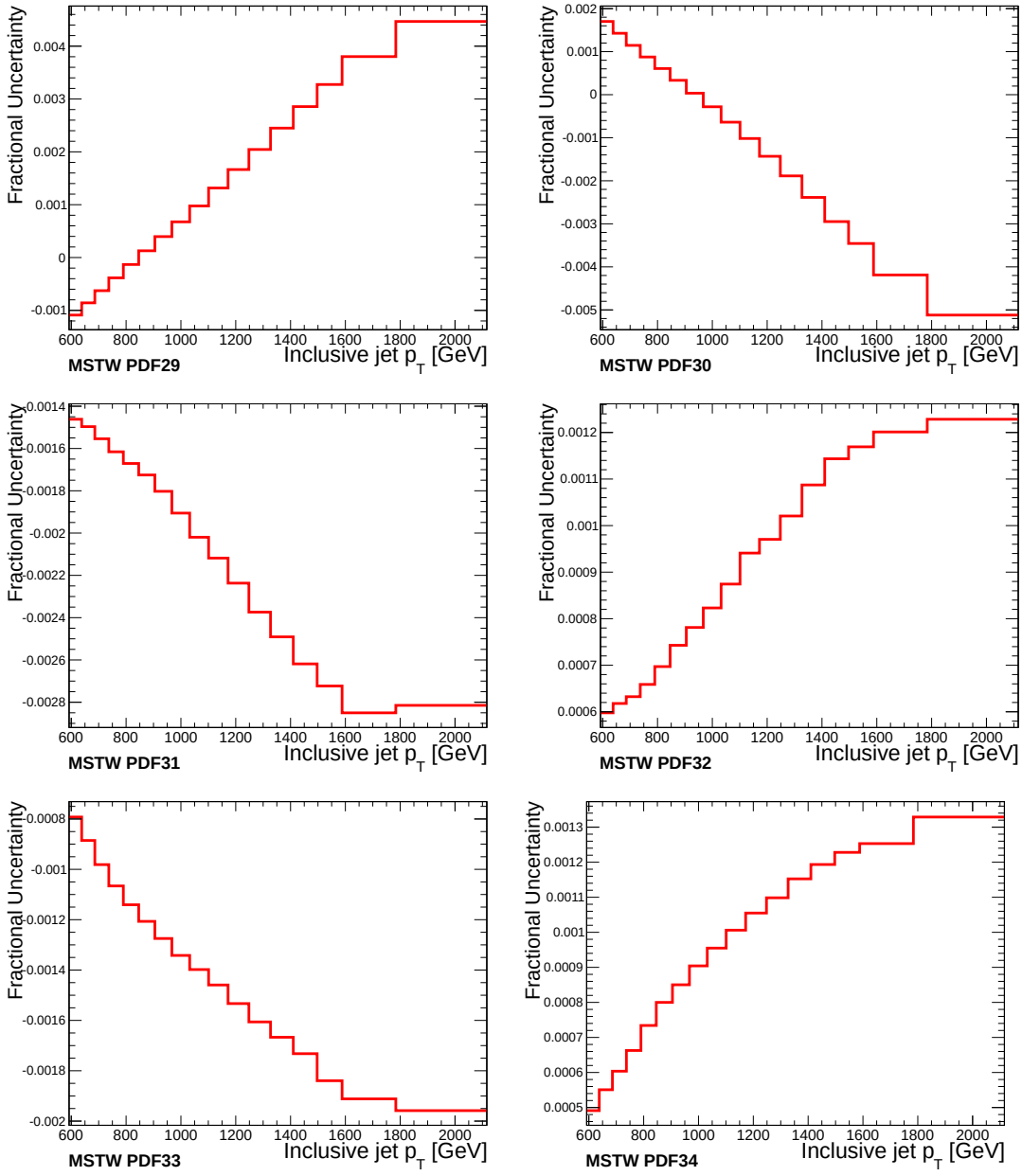


Figure B.6: Difference between MSTW PDF Set 29-34 and the central PDF set, normalized to the central set.

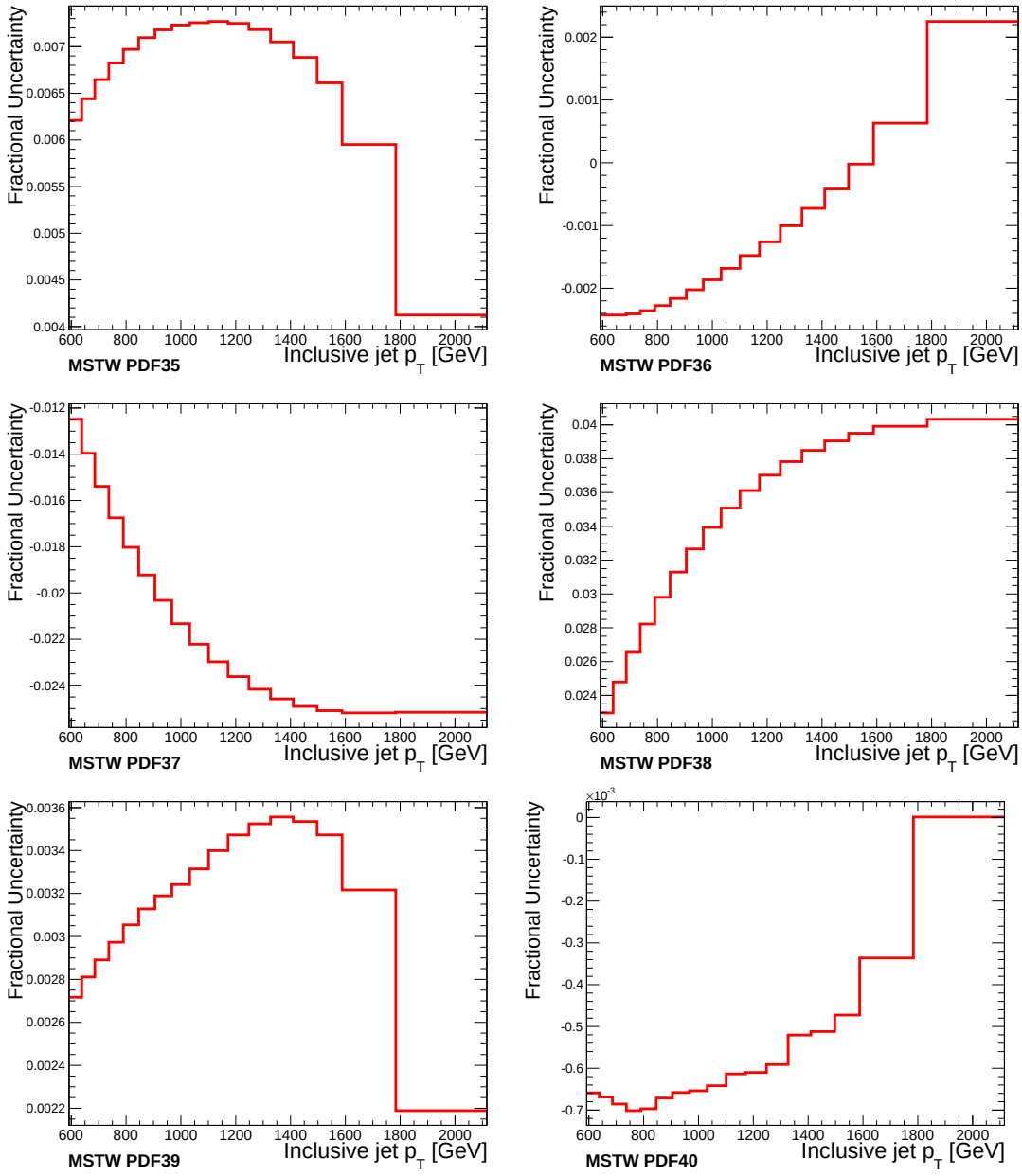


Figure B.7: Difference between MSTW PDF Set 35-40 and the central PDF set, normalized to the central set.

Appendix C
Likelihood Function

C.1 Definition of Terms and Conventions

In this chapter, the method used to compare the null and extinction hypotheses to data will be discussed. Using a likelihood function, we will define a region in the extinction parameter M which is considered to be excluded at 95% confidence level (C.L.) To determine this, we use the modified frequentist criterion CL_s Junk (1999); Read (2002). Any value of M for which $CL_s < 0.05$ will be considered excluded at this C.L. There are a variety of conventions which must be chosen by the user when applying CL_s. Unless otherwise noted, we use the conventions defined by the CMS and ATLAS collaborations (Cowan et al. (2011)) for use by the Higgs Boson combined analysis.

* $L(a|b)$ = likelihood of b given a
* μ = general variable representing an arbitrary signal strength
* $\vec{x}$ = a vector of data. In this analysis, this is the binned inclusive jet p_{T} spectrum.
* $\vec{\theta}$ = The vector of all nuisance parameters.
* λ = general variable representing a likelihood ratio
* $L(\mu|\hat{\hat{\vec{\theta}}})$ = The likelihood, given a fixed value of the signal strength, for the values of the nuisance parameters that maximize likelihood.
* $L(\hat{\mu}|\hat{\hat{\vec{\theta}}})$ = the likelihood, given the value of the signal strength and nuisance parameters that maximize the likelihood for a global fit.
* M = variable representing a specific value of the extinction scale in TeV.
* β = variable representing the extinction scale, where $\beta = M^{-2}$
* Q = common reference for the test statistic in the CL_s criterion.

C.2 Definition of Test Statistic

The CL_s criterion requires definition of a test statistic Q. Using the conventions defined above, we can calculate the following likelihood ratio $\lambda(\mu)$:

$$\lambda(\mu) = \frac{L(\mu|\hat{\hat{\vec{\theta}}})}{L(\hat{\mu}|\hat{\vec{\theta}})} \tag{C.1}$$

The likelihood used in the extinction search is defined in Chapter 5. The numerator of this ratio is more commonly known as a *profile likelihood*, referring to the practice of finding the maximum-likelihood values of the nuisance parameters for a given μ. The test statistic Q will then be given by:

$$Q = -2\ln(\lambda(\mu)) \tag{C.2}$$

By definition, the denominator of $\lambda(\mu)$ is the maximum likelihood for a given dataset. As the likelihood L is positive-definite, this means the ratio will be bound between zero and one. Likewise, the test statistic Q will be positive-definite. Values of Q closer to zero indicate good agreement between the dataset and the model prediction given the signal strength μ.

For a given value of the modified extinction parameter $\beta \equiv M^{-2}$, we will throw 2 ensembles of pseudoexperiments. The first ensemble will assume the background hypothesis, where $\beta \equiv 0$. The second ensemble assumes a fixed value of the extinction parameter $\beta = \mu$ for all pseudoexperiments. For each pseudoexperiment in both ensembles, all nuisance parameters will be randomly thrown using a log-normal probability distribution. A binned inclusive jet p_{T} spectrum will then be produced given the predictions of the model for those values of β and $\vec{\theta}$, which we will call $\vec{x_e}$. The content of each p_{T} bin in the final spectrum for the pseudoexperiment will then be randomly thrown from a Poisson distribution, with mean equal to the content of that bin in $\vec{x_e}$.

Each ensemble will contain 6000 pseudoexperiments. The test statistic Q will be calculated for the data, and compared to the set of all values of Q computed for each pseudoexperiment.

We then further define:

* p_μ = the fraction of the pseudoexperiments assuming signal strength $\beta = \mu$ for which Q is greater than that observed for the data.

* p_0 = the fraction of the pseudoexperiments given purely background, $\beta = 0$, for which Q is greater than that observed for the data.

Our value of CL_s for a given value of the signal strength is defined as:

$$CL_s = \frac{p_\mu}{p_0} \tag{C.3}$$

We exclude all values of the signal strength β for which $CL_s < 0.05$, corresponding to exclusion at 95% C.L.