

Chapter 1

Introduction

Particle physics deals with the study of the basic constituents of matter and the forces governing the interactions among them. The Standard Model (SM) is the most accepted theory describing the nature and properties of the fundamental particles and their interactions. The elementary particles leptons and quarks, known as fermions, interact through the exchange of the gauge bosons. The gauge bosons acquire masses in the process of electroweak symmetry breaking whereas the masses of the fermions are generated through Yukawa interactions with the field associated to the scalar Higgs boson. The gauge bosons are the mediators of the four fundamental forces of interaction existing in nature : the electromagnetic force, the strong force, the weak force and the gravitational force. Quantum Chromodynamics (QCD) is the theory of the strong interactions between the quarks mediated by the massless gluons. The quarks and gluons, together known as partons, have a peculiar property of “color” charge. Due to confinement property of QCD, the quarks cannot exist freely in nature but bind themselves into colorless particles called hadrons such as protons and neutrons together known as nucleons, pions etc. The structure and the properties of sub-atomic particles can be explored by first accelerating them using particle accelerators and then colliding at very high energies. The end products of these collisions are recorded in the ~~real~~ particle detectors constituting the real data.

*slang
measurements?*

These data sets are analyzed in detail to reveal the structure and characteristic properties of the fundamental particles.

To search for the very rare particles, to investigate the physics beyond SM, and to explore the regime of undiscovered physical laws, the particle accelerators have become bigger and complex over the past few decades. The Large Hadron Collider (LHC) is one of the biggest and the most powerful particle collider in which the protons are accelerated and collided at extremely high center-of-mass energies to probe their internal structure and the parton distribution functions (PDFs). The PDFs give the probability to find a parton at an energy scale Q carrying a fractional momentum x of the proton. Since the proton is not elementary and is made up of partons, the proton-proton (pp) collisions are viewed as interactions between their constituent partons. The final products of the scattering are observed by Compact Muon Solenoid (CMS), one of the four detectors of the LHC, located around the interaction points of the collisions. The scattering cross-section can be expressed as a sum in terms of increasing powers of the strong coupling constant α_S convoluted with PDFs. The lowest-order α_S^2 term represents the production of two partons in final states whereas terms of higher-order α_S^3 , α_S^4 etc. signify the existence of multi-partons in final states. The highly energetic final state partons emit quarks and gluons with lower energies and give rise to a parton shower (PS). The colored products of parton shower hadronize to a spray of colorless hadrons known as jets. The jets are the final structures observed in the detector. So they carry the significant information of the energy and direction of the initial partons and hence are important to study. The final partons also have the probability to radiate more gluons and quarks which also hadronize and result in multijets in the final state. At LHC, such events are produced in large number and are an important source for testing the predictions given by QCD. They also serve as an important background in the searches for new particles and physics beyond SM.

The inclusive multijet event cross-section σ_{i-jet} , given by the process

Vorlesung
 $\text{pp} \rightarrow i\text{jets} + X$, is proportional to α_s^i . The study of inclusive jet cross-sections in terms of jet transverse momentum p_T and rapidity y is very important because it provides the essential information about the PDFs and the precise measurement of α_s . Also the ratio of cross-sections given by Eq. 1.1 is proportional to the QCD coupling constant α_s and hence can be used to determine the value of α_s .

$$R_{mn} = \frac{\sigma_{m-jet}}{\sigma_{n-jet}} \propto \alpha_s^{m-n} \quad (1.1)$$

Instead of studying inclusive cross-sections, the cross-section ratio is more useful because of the partial or complete cancellation of many theoretical and experimental uncertainties in the ratio. The CMS Collaboration has previously measured the ratio of the inclusive 3-jet cross-section to that of the inclusive 2-jet as a function of the average transverse momentum, $\langle p_{T1,2} \rangle$, of the two leading jets in the event at 7 TeV [1]. This study leads to an extraction of $\alpha_s(M_Z) = 0.1148 \pm 0.0055$, where the dominant uncertainty stems from the estimation of higher-order corrections to the next-to-leading order (NLO) prediction. In this thesis, a measurement of inclusive 2-jet and 3-jet event cross-sections as well as ratio of 3-jet event cross-section over 2-jet R_{32} , is performed using an event sample collected during 2012 by the CMS experiment at the LHC and corresponding to an integrated luminosity of 19.7 fb^{-1} of pp collisions at a center-of-mass energy of 8 TeV. The event scale is chosen to be the average transverse momentum of the two leading jets, referred to as $H_{T,2}/2$

*Unclear:
 It is predicted
 if known
 at one
 scale, e.g. M_Z*
 in this thesis. The strength of the strong force, α_s at a given energy scale Q is not predicted and has to be extracted from the experiment. Hence, the value of the strong coupling constant at the scale of the Z boson mass $\alpha_s(M_Z)$ is extracted from the measurements performed in this thesis. The value of α_s depends on the energy scale Q and it decreases with the increase of Q scale. The running of α_s with scale Q is also studied and compared with other CMS measurements as well as results from different experiments. This checks the consistency with QCD via the

renormalization group equation (RGE)¹, which precisely describes the evolution of α_S at the renormalization scale of QCD.

The organization of this thesis² is as follows :

Chapter 2 gives a brief overview of the Standard Model of particle physics and the theory of strong interactions QCD, theory of hadron collisions as well as formation of jets and jet algorithms.

Chapter 3 deals with experimental apparatus which covers the details of the geometry of the CMS detector and its various sub-detectors.

Chapter 4 describes the methods of event generation used in different Monte-Carlo event generators, detector geometry simulation and reconstruction of the particles in the detector. This chapter also gives the details of the different approaches of jet reconstruction at CMS and applied jet-energy corrections along with the description of the software framework used in the analysis presented in the current thesis.

Chapter 5 presents the measurement of differential inclusive multijet event cross-sections and the cross-section ratio. The measurements are corrected for detector effects by unfolding procedure which is discussed in detail in this chapter. The sources of the experimental uncertainties are studied in detail.

Chapter 6 contains a detailed description of the NLO perturbative QCD theory predictions obtained using different PDF sets. The NLO predictions are corrected with the non-perturbative and electroweak corrections. The theoretical uncertainties are calculated from various sources. At the end of this chapter, the unfolded measurements are compared with the predictions at NLO in pQCD as well as with the predictions from several Monte Carlo event generators.

¹According to the RGE, the strong force becomes weaker at short distances corresponding to large momentum transfers. This is referred to a property of QCD called asymptotic freedom.

²The common unit convention based on International System of Units (SI) as followed in particle physics will be used throughout the thesis. In addition, the units electron volt (eV) and barn (b) are used for energy and interaction cross-section, respectively. The reduced Planck constant (\hbar) and speed of light (c) are set to unity, i.e. $\hbar = c = 1$.

Chapter 7 describes the method to extract the strong coupling constant at the scale of mass of Z boson $\alpha_s(M_Z)$ from the measurements of differential inclusive multijet cross-sections and the cross-section ratio R_{32} . Also, the running of α_S with energy scale Q is presented along with the previous measurements from different experiments.

Chapter 8 summarizes the results and conclusions of the work done in this thesis.

Chapter 9 mentions the participation in other hardware and software activities.

Chapter 2

Theoretical Background

Since 1930s, many theories and discoveries in particle physics have revealed the fundamental structure of matter. The matter is made up of fundamental particles and their interactions are mediated by four fundamental forces [2]. The theoretical models describe all the phenomena of particle physics as well as predict the properties of particles. These models must be either confirmed experimentally or totally excluded giving hints of new physics. This interplay between experimental discoveries and the corresponding theoretical predictions leads to a theoretical model called Standard Model, which describes the fundamental particles and their interactions. The world's most powerful particle accelerators and detectors are used by physicists to test the predictions and limits of the Standard Model where it has successfully explained almost all experimental results. This chapter describes the Standard Model with main focus on the theory of strong interactions called Quantum Chromodynamics and its features which serve as the theoretical base of this thesis.

maybe? strive to describe, masses are not predicted

2.1 Standard Model

The Standard Model (SM) of particle physics [3–5] was developed in 1970s. It is a mathematical framework which describes the nature and properties of the funda-

fermions need not be elementary. Nuclei are either fermions or bosons for example.

Misleading

mental particles and the three of the four known forces of interactions between them, as summarized in Fig. 2.1. According to the SM, the basic constituents of matter are the elementary particles i.e. without any internal structure, known as fermions and bosons. The fermions have half integral spin and obey Fermi-Dirac statistics. They follow the Pauli exclusion principle according to which two or more identical fermions cannot occupy the same quantum state. Each fermion has an associated anti-particle having the same properties but opposite-sign quantum numbers.

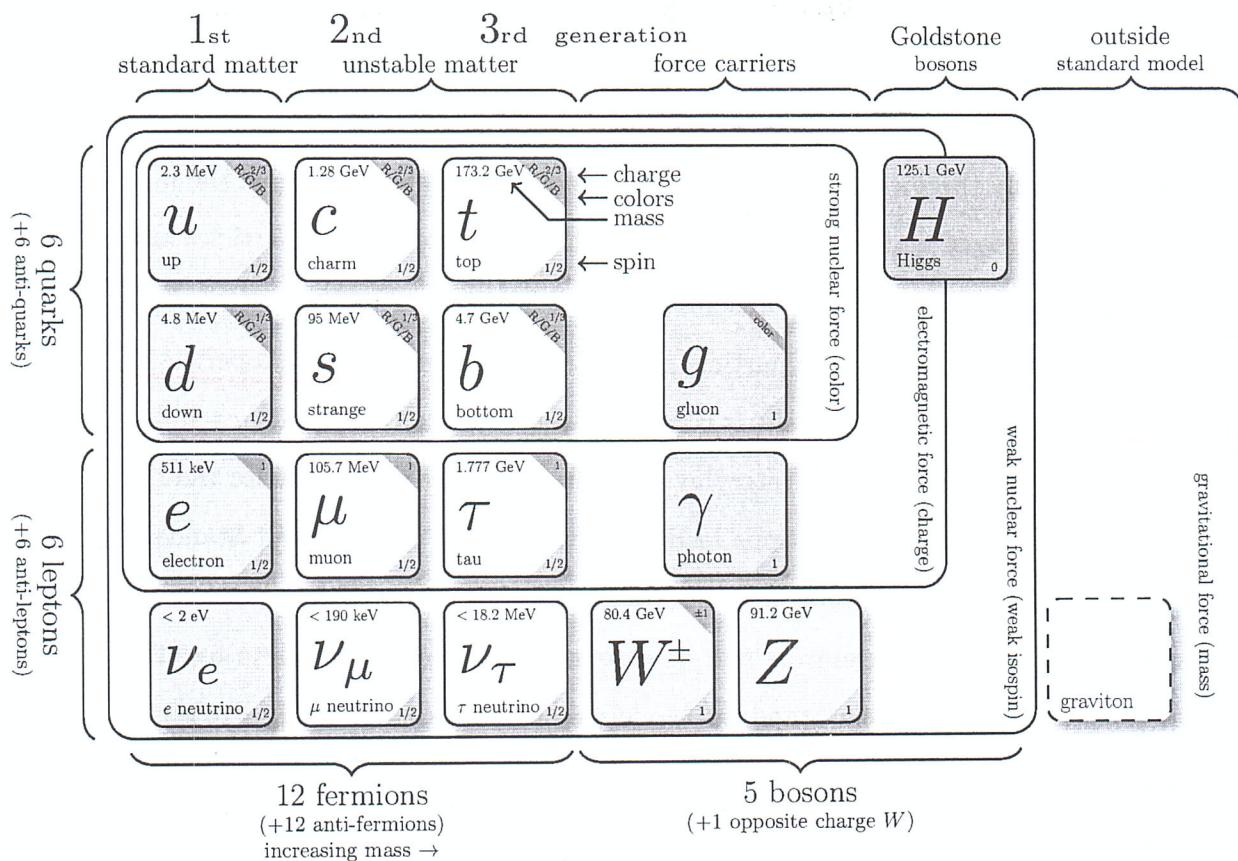


Figure 2.1: The Standard Model³ summarizing the properties of elementary particles known as fermions (leptons and quarks) grouped into three generations, gauge bosons as mediators for the interactions, the scalar Higgs boson and not incorporated graviton for the gravitational force.

incorporated?

³Source : <http://www.texample.net/tikz/examples/model-physics>

There are also 8 gluons...
if W^\pm are counted,

Depending on how the fermions interact, these are classified into two categories - leptons (ℓ) and quarks (q). The leptons are of six types : electron (e), muon (μ) and tau (τ) with electric charge $Q = \pm 1$ ⁴ and the corresponding neutrinos : electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ) having electric charge $Q = 0$. The quarks exist in six “flavors” : up (u), down (d), strange (s), charm (c), bottom (b) and top (t). u , c and t carry electric charge $Q = \pm \frac{2}{3}$ whereas d , s and b carry $Q = \pm \frac{1}{3}$. The quarks and leptons are categorized into three generations. The first generation has the lightest and the most stable particles whereas the heavier and less stable particles belong to the second and third generations.

The elementary bosons have integral spin and obey the Bose-Einstein statistics. These are further of two types : gauge bosons having non-zero integral spin and a scalar boson with zero spin. The gauge bosons are the force carriers which mediate the electromagnetic, strong, weak and gravitational forces. Every interaction involves the exchange of a gauge boson : the massless photon (γ) for the electromagnetic force, massless gluons (g) for the strong force, massive W^\pm and Z for the weak force and the graviton (not yet found) for the gravitational force. However, the gravitational force has not been incorporated into SM yet. Along with this, the existence of dark matter or dark energy and the matter-antimatter asymmetry are still missing pieces in the SM. The interaction between fundamental particles acts because of some peculiar property of the particles - charge for the electromagnetic force, color for the strong force and flavor for the weak force. *Usually Beyond SM.*

In the SM, the forces of interaction except gravity are unified into one quantum field theory [6], known as Grand Unified Theory (GUT) [7–9]. The SM framework based on quantum field theories is described by $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry where C stands for the color charge, L for weak isospin and Y for hypercharge. Here $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$ terms give rise to strong, weak and

*Undefined
 $Y = ?$
Where is Q^3 ?*

⁴ all charges are expressed in units of elementary charge e

Define together with $t=c=1$?

electromagnetic forces, respectively. $U(n)$ are the unitary and $SU(n)$ are the special unitary groups of degree n . The $SU(3)_C$ term defines the strong interaction between quarks and gluons mediated by gluons, with the three degrees of freedom of the color charge (C). The electromagnetic interaction of particles is explained by a ~~well es-~~ ^{the most precise} ~~known today~~ established modern theory called Quantum Electrodynamics (QED). In SM, the weak and electromagnetic interactions are combined by an electroweak symmetry theory, described by $SU(2)_L \otimes U(1)_Y$ gauge group. But this electroweak unification could not explain the occurrence of massive weak gauge bosons. This problem was solved by Brout-Englert-Higgs mechanism [10, 11]. The Higgs boson, named after Peter Higgs, is the field quantum of the Higgs field responsible for electroweak symmetry breaking. In SM, the Higgs field is a $SU(2)$ doublet which is a scalar under Lorentz transformations. The coupling of the bosons to the scalar Higgs field causes the spontaneous symmetry breaking which triggers the Higgs mechanism. After symmetry breaking, three of the four degrees of freedom in the Higgs field interact with the three weak gauge bosons (W^\pm and Z) and allows them to be massive, while the remaining one degree of freedom becomes the Higgs boson. Its existence was confirmed by the CMS [12] and ATLAS [13] collaborations in 2012, with ~~K~~ properties consistent with the SM. In contrast to the electroweak symmetry, the $SU(3)_C$ of the strong interaction is an exact symmetry and hence the gluons are massless. The strong interaction between quarks and gluons is described by ~~theory called~~ quantum chromodynamics (QCD), explained in detail in the next section.

2.2 Quantum Chromodynamics

The strong interactions between the quarks and gluons are described by a non-abelian gauge theory called quantum chromodynamics (QCD) [14, 15]. The gauge group of QCD is the special unitary group $SU(3)_C$ with color charges C as the generators of the gauge group. Color charge is the peculiar property of QCD and

Similar

has a same role as the electric charge in electromagnetic interactions. However, the mediator of electromagnetic interactions i.e. the photon itself does not carry any electric charge whereas the gluon itself carry color charge. This allows the self coupling of gluons and hence makes the theory non-abelian. Both the quarks and gluons carry three types of color charges : red (r), green (g) and blue (b), and three types of anti-color charges : anti-red (\bar{r}), anti-green (\bar{g}) and anti-blue (\bar{b}). The quarks carry a single color charge whereas gluons carry a combination of color charges. There are nine eigen states of gluons but one of them $\frac{1}{\sqrt{3}}(r\bar{r} + g\bar{g} + b\bar{b})$ is a totally symmetric color singlet which has no net color charge and does not take part in interaction.

The remaining eight eigen states of the gluons are :

$$r\bar{b}, r\bar{g}, g\bar{r}, g\bar{b}, b\bar{g}, b\bar{r}, \frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}), \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g}) \quad (2.1)$$

The dynamics of the quarks and gluons are controlled by the gauge invariant QCD Lagrangian \mathcal{L}_{QCD} which is composed of four terms as :

$$\mathcal{L}_{QCD} = \underbrace{-\frac{1}{4}F_{\mu\nu}^A F_A^{\mu\nu}}_{\mathcal{L}_{gluons}} + \underbrace{\sum_{flavors} \bar{q}_a (i\gamma^\mu (D_\mu)_{ab} - m_q) q_b}_{\mathcal{L}_{quarks}} + \mathcal{L}_{gauge} + \mathcal{L}_{ghost} \quad (2.2)$$

where \mathcal{L}_{gluons} describes the kinetic term of the gluon fields A_μ^A ; \mathcal{L}_{quarks} defines the interaction between spin- $\frac{1}{2}$ quark fields q_a of mass m_q and spin-1 gluon fields A_μ^A summing over all presently known six flavors of quarks; \mathcal{L}_{gauge} describes the chosen gauge and \mathcal{L}_{ghost} is the so-called ghost term required to treat the degeneracy of equivalent gauge field configurations in non-abelian gauge theories. In Eq. 2.2, the Greek letters $\mu, \nu, \dots \in \{0,1,2,3\}$ are the space-time indices; $a,b,c \in \{1,2,3\}$ and $A,B,C \in \{1,\dots,8\}$ are the indices of the triplet and octet representations, respectively,

color

of the gauge symmetry group $SU(3)_C$. The field tensor $F_{\mu\nu}^A$ is defined as

$$F_{\mu\nu}^A = \partial_\mu A_\nu^A - \partial_\nu A_\mu^A - g_s f_{ABC} A_\mu^B A_\nu^C$$

Summation implied
over identical indices
(2.3)
should be removed.

where g_s is the coupling constant determining the strength of the interaction between colored partons and f_{ABC} are the structure constants of the $SU(3)_C$ group. The last term in Eq. 2.3 is a non-abelian term which distinguishes QCD from QED and gives rise to a three- and a four-gluon vertex. In the term \mathcal{L}_{quarks} , $(D_\mu)_{ab}$ is the covariant derivative given by Eq. 2.4 and γ_μ are the Dirac γ -matrices.

$$(D_\mu)_{ab} = \partial_\mu \delta_{ab} + ig_s T_{ab}^A A_\mu^A \quad (2.4)$$

A_μ^A are the gluon fields with factors T_{ab}^A factors corresponding to the generators of the $SU(3)_C$ gauge group. The generators are represented via $T^A = \lambda^A/2$ by the Hermitian and traceless Gell-Mann matrices λ^A [16]. The generator matrices T^A follow the commutation relations :

$$[T^A, T^B] = if_{ABC}^{\uparrow} T^C \quad (2.5)$$

Summation C

In \mathcal{L}_{QCD} , the classical contribution comes from \mathcal{L}_{gluons} and \mathcal{L}_{quarks} terms which give rise to the free quark- and gluon-field terms, and the quark-gluon interaction terms presented in Fig. 2.2. The cubic and quartic gluon self-interaction vertices proportional to g_s and g_s^2 , respectively, come into play due to the non-abelian property of QCD.

It is impossible to use perturbation theory on a gauge invariant Lagrangian without choosing a specific gauge in which to calculate. The usual gauge-fixing term

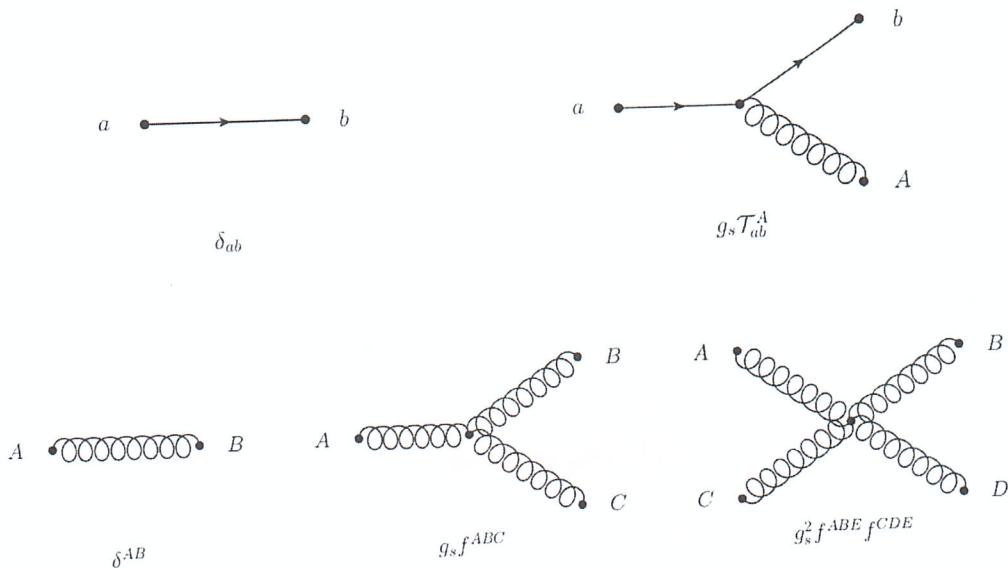


Figure 2.2: The fundamental Feynman rules of a free quark-field term (top left), cubic quark-gluon interaction term (top right), free gluon-field term (bottom left), quartic gluon self-interaction term (bottom middle) and quartic gluon self-interaction term (bottom right). Taken from [17].

is given by

$$\mathcal{L}_{gauge} = -\frac{1}{2\xi} (\partial^\mu \mathcal{A}_\mu^A)^2 \quad (2.6)$$

where ξ may be any finite constant. This choice fixes the class of covariant gauges with ξ as the gauge parameter. As QCD is non-abelian, the gauge fixing term must be supplemented by a ghost Lagrangian as

$$\mathcal{L}_{ghost} = \partial_\alpha \eta^{A\dagger} (D_{AB}^\mu \eta^B) \quad (2.7)$$

where η^A is a complex scalar field which obeys Fermi-Dirac statistics. The ghost fields cancel unphysical degrees of freedom arising due to use of covariant gauges. This completes the QCD Lagrangian shown in Eq. 2.2.

The strength of an interaction is given by a fundamental parameter called the coupling constant α . In QED, the coupling constant $\alpha_e = e^2/4\pi = 1/137$ is constant. In contrast to this, in QCD, the coupling constant $\alpha_S(Q) = g_s^2/4\pi$ is

No, also α_e is running because of renormalisation!

e.g.: $\alpha_e(M_Z) \approx 1/128$

✓ Same for QED.

not constant and depends on the separation between the interacting particles. It increases with the increase in the distance or decrease in the energy scale Q . At large distances or low energies, the quarks can never be found as free particles but exit in color neutral bound states known as hadrons. Hadrons are of two types : baryons and mesons. According to the quark model [2] every (anti-)baryon is made up of three (anti-)quarks and every meson is made up of a quark-antiquark pair. When the colored partons within a hadron are pulled farther and farther apart, there is an increase in the strength of force between them. This results in creation of new quark-antiquark pairs making it impossible to liberate a free quark or gluon. This property of QCD is known as confinement according to which at low energy, the partons are forever bound into hadrons such as protons (uud), neutrons (udd) etc. Although the gluons are massless but the confinement leads to the finite range of the strong interactions. On the other hand, at small distances, the strength of coupling decreases. The quarks and gluons interact very weakly and behave as free particles. This property is known as asymptotic freedom. This indicates that perturbative theory is only applicable at high energies or small distances.

2.2.1 Perturbative Quantum Chromodynamics

At high energies, the property of asymptotic freedom allows a perturbative treatment in QCD calculations. In perturbative quantum chromodynamics (pQCD), any physical observable X such as cross-section of a scattering process, can be expanded as a perturbative series in terms of coupling constant α_s as :

$$X = \sum_{i=0}^N \alpha_s^n c_i = c_0 + \alpha_s^1 c_1 + \alpha_s^2 c_2 + \dots \quad (2.8)$$

where c_i are the perturbative coefficients. In a process, the pQCD calculation of X is determined by summing over the amplitudes of all Feynman diagrams contributing

to that process. For a given Feynman diagram, the power of α_s is determined by the number of vertices associated with quark-gluon or gluon-gluon interactions. A leading order (LO) prediction sums over only the lowest-order contribution whereas next-to-leading order (NLO) includes terms with ~~the additional powers~~^{one} of α_s . The next-to-next-to-leading order (NNLO) includes emission of another gluon or a virtual gluon loop. The different order of the QCD processes are shown in Fig. 2.3. The calculations become complex with the loop diagrams where the momenta of

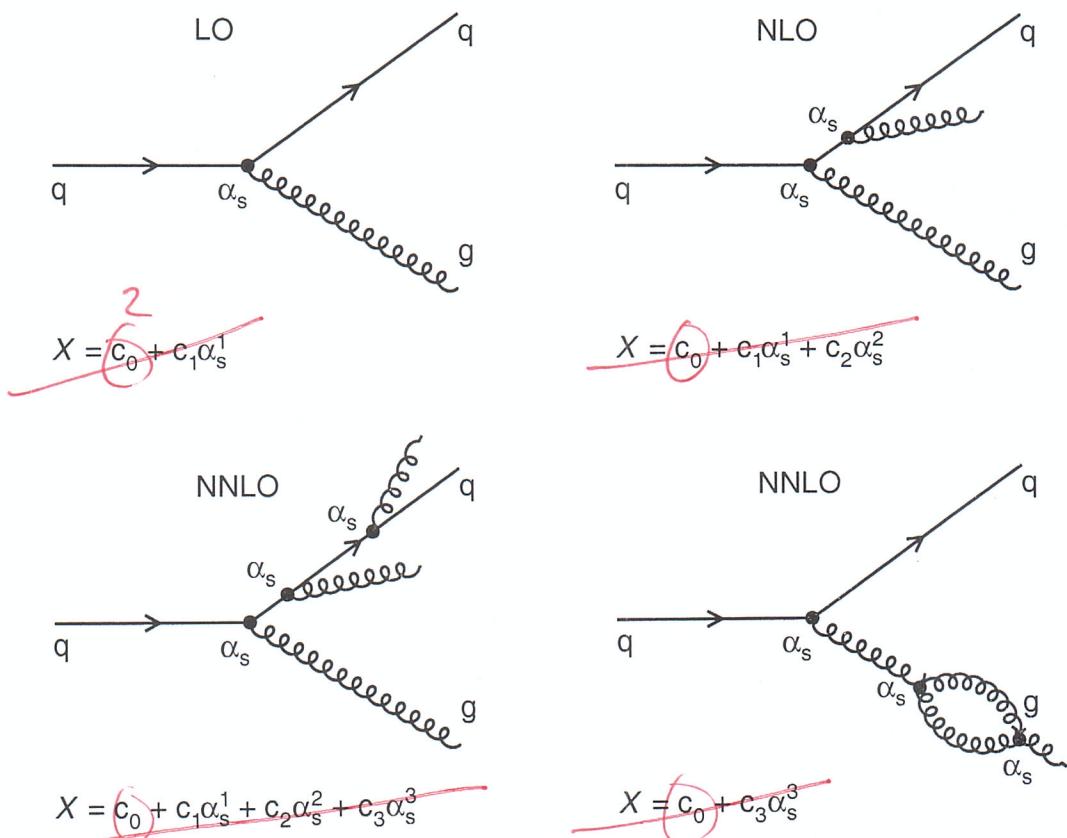


Figure 2.3: ~~The~~⁵ Feynman diagrams⁵ of leading-order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) processes in quantum chromodynamics along with the perturbative expansion of any observable X in powers of the strong coupling constant α_s . At each successive step in perturbation series, the emission of an additional gluon take place.

Power counting not clear. Better show concrete diagrams for your case. See after page 31!

the virtual particles in a loop are not fully constrained by four-momentum conservation and the associated integrals are divergent. Such ultraviolet (UV) divergences

⁵Drawn using ROOT

enter the calculations beyond LO either due to loop or vertex corrections. These are overcome by a procedure known as renormalization, described in next section. Apart from the UV divergences, the QCD also suffers from infrared and collinear divergences (IRC) due to the presence of massless gluons and neglected quark masses. These need to be handled in pQCD calculations. The observable to be studied must be IRC safe.

2.2.2 Renormalization and Running of the Strong Coupling

The renormalization is a mathematical procedure which allows the finite calculation of momenta integrals of virtual loop by removing UV divergences. It introduces a regulator for the infinities, the renormalization scale μ_r . At first, the divergences are regularized temporarily by introducing a cut-off to the loop momenta at μ_r scale. Then the free parameters of the Lagrangian, i.e. the coupling constant are redefined or renormalized to absorb the UV divergences. Due to this, both $\alpha_s(Q)$ and observable X become a function of μ_r . The exact dependence of $\alpha_s(\mu_r^2)$ on μ_r is described by the renormalization group equation (RGE) [18] which determines the running of $\alpha_s(\mu_r^2)$. According to RGE, the dependence of X on μ_r must cancel. Mathematically this can be expressed as :

$$\mu_r^2 \frac{d}{d\mu_r^2} X \left(\frac{Q^2}{\mu_r^2}, \alpha_s(\mu_r^2) \right) = \left(\mu_r^2 \frac{\partial}{\partial \mu_r^2} + \mu_r^2 \frac{\partial \alpha_s(\mu_r^2)}{\partial \mu_r^2} \frac{\partial}{\partial \alpha_s(\mu_r^2)} \right) X = 0 \quad (2.9)$$

Using beta function $\beta(\alpha_s) = \mu_r^2 \frac{\partial \alpha_s(\mu_r^2)}{\partial \mu_r^2}$, Eq. 2.9 can be re-written as

$$\left(\mu_r^2 \frac{\partial}{\partial \mu_r^2} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s(\mu_r^2)} \right) X = 0 \quad (2.10)$$

By setting the renormalization scale equal to the physical scale i.e. $\mu^2 = Q^2$, $X(1, \alpha_s(Q))$ is a solution to above equation. Q -dependence of ~~the~~ X is only from the renormalization of the theory which is present in the ~~classical~~? theory. Hence measuring the Q -dependence of X will directly probe the quantum structure of the theory. The β function in QCD has a perturbative expansion as :

$$\beta(\alpha_s) = -\alpha_s^2 \left(b_0 + b_1 \alpha_s + b_2 \alpha_s^2 + \mathcal{O}(\alpha_s^3) \right) \quad (2.11)$$

where b_n is the $n+1$ -loop β -function coefficients giving the dependence of the coupling on the energy scale Q . In the modified minimal subtraction ($\overline{\text{MS}}$) scheme [19, 20], the beta coefficient functions have following values :

$$b_0 = \frac{33 - 2n_f}{12\pi}, \quad b_1 = \frac{153 - 19n_f}{24\pi^2}, \quad b_2 = \frac{77139 - 15099n_f + 325n_f^2}{3456\pi^3} \quad (2.12)$$

where n_f is the number of quark flavours with masses $m_q < \mu_r$. On integration of Eq. 2.11, the energy dependence of α_s is yielded. Neglecting the higher orders, the first order solution of RGE is :

$$\alpha_s(\mu_r^2) = \frac{1}{b_0 \ln(\mu_r^2/\Lambda_{QCD}^2)} \quad (2.13)$$

where Λ_{QCD} is the constant of integration. The perturbative coupling becomes large at the scale Λ_{QCD} and the perturbative series diverge. With $b_0 > 0$, the coupling becomes weaker at higher scales Q , i.e. the effective color charge is small at small distances or large energies. This allows the quarks to behave as free particles within the hadron, leading to the property called asymptotic freedom. It is always convenient to express α_s at some fixed scale. Since some of the best measurements come from Z decays, it is common practise to determine the strong coupling at the $\overline{c^2}$.

scale of the Z boson mass $\alpha_s(M_Z)$. So, Eq. 2.13 can be expressed as :

$$\alpha_s(\mu_r, \alpha_s(M_Z)) = \frac{\alpha_s(M_Z)}{1 + \alpha_s(M_Z) b_0 \ln(\mu_r^2/M_z^2)} \quad (2.14)$$

Since α_s is ~~a~~ the free parameter of QCD theory, it is ~~always~~ extracted from ~~the~~ experimental measurements and evolved to the scale of the Z boson. According to Particle Data Group (PDG) [21], the current world average value of the strong coupling constant at the scale of mass of Z boson is $\alpha_s(M_Z) = 0.1181 \pm 0.0011$.

$$\alpha_s(M_Z) = 0.1181 \pm 0.0011 \quad (2.15)$$

This value is derived using data from deep inelastic scattering process, electron-positron annihilation processes, hadronic τ lepton decays, lattice QCD calculations and electroweak precision fits. The different experimental determinations of the strong coupling constant evolved at the scale Q are shown as a function of Q in Fig. 2.4 which describe the running of the α_s up to the 1 TeV scale.

2.3 Hadronic Collisions

At ~~a~~ large momentum transfer, the collision between two hadrons can be visualized as an interaction between their constituents - quarks and gluons. In this thesis, we are studying the proton-proton collisions taking place at the Large Hadron Collider (LHC). A proton is a complex composite particle consisting of three valence quarks (uud), gluons for the exchange of the strong force and the sea quarks. The sea quarks consist of quark-antiquark pairs coming into and out of existence rapidly and continuously due to gluon colour field splitting. In any collision, ~~one of the~~ ~~most important~~ quantities to evaluate is the cross-section (σ) of a certain process which gives the probability that the two hadrons interact and give rise to that

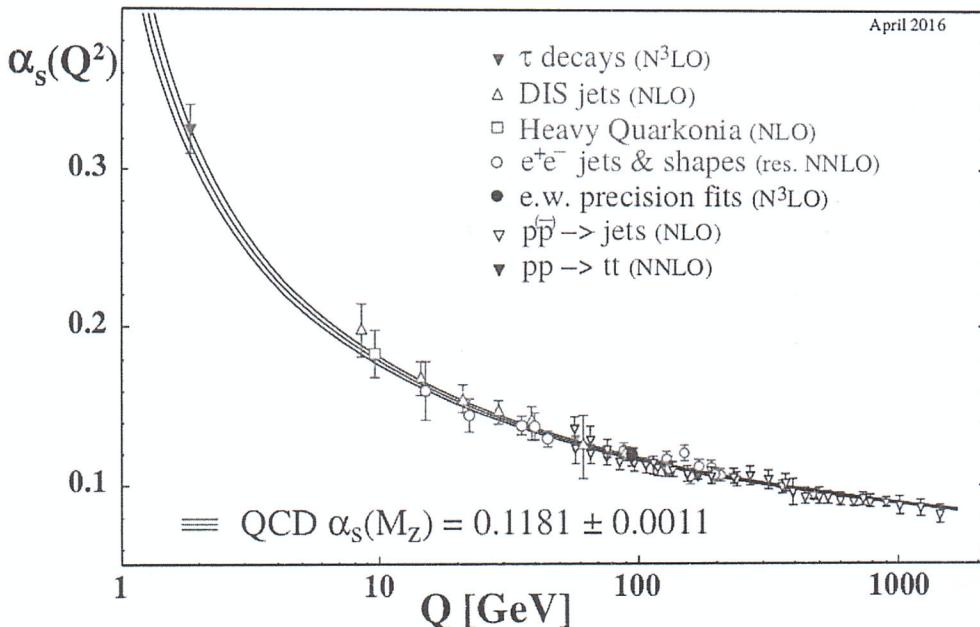


Figure 2.4: The different experimental measurements of the strong coupling constant α_S evolved at the energy scale Q are shown as a function of Q . These describe the running of the α_S up to the 1 TeV scale. Taken from [21].

final state?

specific process. In a hadronic collision, the perturbation theory is only valid at the parton-level but due to property of confinement at low energies, free partons cannot exist in nature. Only hadrons with a complex internal structure are available for the high energy collisions. Here, a factorization theorem of QCD [22] comes into play which allows the calculation of σ by separating into two parts : a short-distance partonic cross-section calculable with pQCD, and a non-perturbative long-distance part described by parton distribution functions $f_i(x, \mu_f)$ (PDFs). The PDFs describe the partonic content of the colliding hadrons and give the probability to find a parton i with momentum fraction x within a hadron. μ_f is a factorization scale which corresponds to the resolution with which the hadron is being probed. The particles which are emitted with transverse momenta $p_T > \mu_f$ are considered in the calculation of hard scattering perturbative coefficients. The particles emitted with $p_T < \mu_f$ are accounted for within the PDFs. Applying the factorization theorem

to

in a proton-proton collision, the cross-section of a hard scattering process can be written as :

$$\sigma_{P_1 P_2 \rightarrow X} = \sum_{i,j} \int dx_1 dx_2 f_{i,P_1}(x_1, \mu_f) f_{j,P_2}(x_2, \mu_f) \times \hat{\sigma}_{ij \rightarrow X} \left(x_1 p_1, x_2 p_2, \alpha(\mu_r^2), \frac{Q^2}{\mu_f^2} \right) \quad (2.16)$$

where f_i and f_j are the proton PDFs which depend on momentum fractions x_1 and x_2 of parent protons P_1 and P_2 respectively as well as on the factorization scale μ_f . The sum extends over all contributing initial-state parton flavours i, j . The cross-section for the production of final state X at parton level ($\hat{\sigma}_{ij}$) depends on the final state phase, the factorization scale μ_f and the renormalization scale μ_r . Figure. 2.5 illustrates the factorization into the PDFs and the hard scattering cross-section in a proton-proton collision.

The PDFs of the proton are a necessary input to almost all theory predictions of a proton-proton collision. The QCD does not predict the parton content of the proton. So the shapes of PDFs are determined in fits to experimental measurements of different experiments. The dependence of PDFs on μ_f is given by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [23–25] equations which use α_s and the RGE as inputs. The knowledge of proton PDFs mainly comes from the Deep-Inelastic Scattering (DIS) HERA, fixed-target and Tevatron data. The LHC data has a potential to improve constraints of the PDFs further as done in one of the recent CMS measurements [26]. There are several groups which determine the PDFs using various strategies the different minimization methods, phenomenological approaches, and the methods to estimate the uncertainties. The global PDFs are the CTEQ [27], MMHT [28], NNPDF [29], ABM [30] and HERAPDF [31] at LO, NLO and NNLO.

*not global
since only HERA data!*

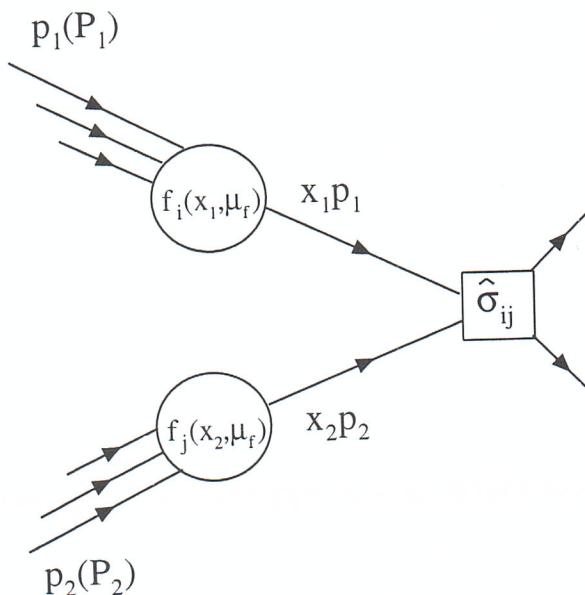


Figure 2.5: Schematic illustration⁶ of the factorization theorem in a collision of two protons P_1 and P_2 having momenta p_1 and p_2 , respectively. In a hard-scattering process at a scale Q^2 , the two partons x_1 and x_2 participate with momenta $x_1 p_1$ and $x_2 p_2$. The total cross-section is factorized into the hard scattering cross-section $\hat{\sigma}_{ij}$ calculable using perturbative quantum chromodynamics and the PDFs $f_i(x_1, \mu_f)$ and $f_j(x_2, \mu_f)$ with factorization scale μ_f .

p QCD

2.3.1 Parton Shower and Hadronization

The partons involved in a hard scattering process get accelerated due to large momentum transfers. These accelerated partons emit QCD radiation in the form of gluons with successively lower energy. Unlike the uncharged photons in QED, the gluons themselves carry color charge and hence also emit further gluons. The emitted gluons in turn split into $q\bar{q}$ pairs. This successive emission of partons leads to a parton shower. In a parton shower, the main contribution is by the collinear parton splitting and the soft gluon emissions. The parton showers mimic the effect of higher-order corrections to the hard process. These cannot be calculated exactly and are taken into account using the parton shower approximation. The two incom-

*Unity
color
vs.
colour
usage.*

⁶Drawn using ROOT

ing partons which are constituents of two colliding hadrons and taking part in hard scattering process can also develop parton showers, commonly known as Initial-State Radiation (ISR). The initial partons produce showers till they collide to initiate the hard scattering process. The final outgoing partons produced from a hard scattering process can also undergo parton showering giving rise to Final-State Radiation (FSR). A parton shower terminates when the scale Q is below the hadronization scale ~ 1 GeV for QCD.

At the end of the shower, there is a decrease in the energy of partons due to successive emission of gluons. Due to this, the coupling constant of QCD α_s grows and QCD is not applicable any more. This leads to the confinement of colored quarks and gluons into the color-neutral composite particles called hadrons and this process is known as hadronization. The hadronization takes place at low momentum transfer and hence is non-perturbative in nature. Although no exact theory for hadronization is known, the different phenomenological models have been developed to simulate the hadronization process. The two main models implemented in Monte Carlo event generators to simulate the hadronization process are :

Lund String Model - In the Lund string model of hadronization [32], the highly energetic gluons are treated as field lines. Due to the gluon self-interaction, the gluons are attracted to each other forming a narrow tube or string of strong color field between a $q\bar{q}$ pair. This model is based on an observation that at distances greater than about a femtometre (fm)⁷, the potential energy $V(r)$ of colored quarks grows linearly with the increase in distance between them (r) as :

$$V(r) = \kappa r \quad (2.17)$$

where $\kappa \sim 1$ GeV/fm² is the tension of the string connecting the quarks. When the q and \bar{q} are pulled apart from each other move apart, the gluonic string stretches.

⁷ 1 femtometre = 1×10^{-15} metres

A consequence

Due to this, the potential energy of the string grows at the expense of the kinetic energy of the quarks. As the potential energy becomes of the order of hadron masses, the string breaks at some point along its length, creating a new $q\bar{q}$ pair. The newly formed two string segments again stretch and break producing further $q\bar{q}$ pairs. This process of stretching and breaking continues until all the potential energy gets converted to $q\bar{q}$ pairs. This whole process is illustrated in Fig 2.6. The $q\bar{q}$ pairs then undergo hadronization due to confinement property. PYTHIA Monte Carlo generator uses the Lund string model.

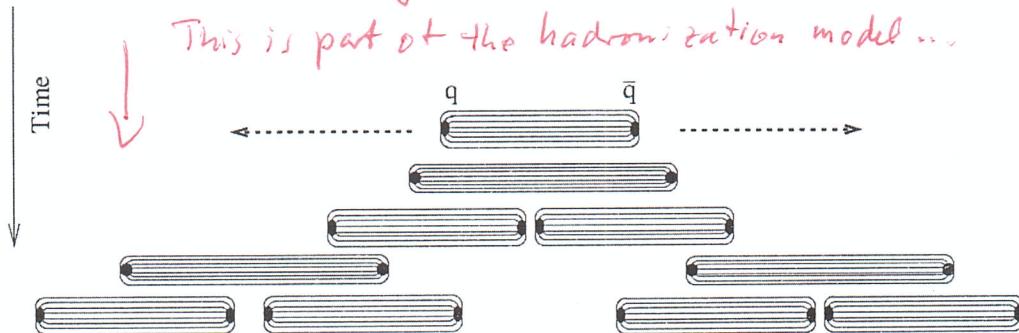


Figure 2.6: Illustration of the hadronization process in Lund string model⁸. When the quark q and anti-quark \bar{q} are pulled apart from each other, the potential energy of the gluonic string connecting the quarks increases. As it becomes of the order of hadron masses, the string breaks and a new $q\bar{q}$ pair is created. The breaking of string and creation of $q\bar{q}$ continues till all the potential energy gets converted to $q\bar{q}$ pairs which then get hadronized.

Cluster Model - The cluster model of hadronization [33, 34] is based on preconfinement property of QCD [35]. According to this property, at evolution scales Q_0 much less than the hard process scale Q , the partons produced in a shower are clustered in colourless groups with an invariant mass distribution, depending on nature of hard process and Q_0 , not on Q . This model contains two steps : firstly all gluons split into $q\bar{q}$ pairs at the end of the parton shower and in the second step, a new set of low-mass color-singlet clusters are obtained which decay into either secondary clusters or directly into hadrons. The generator HERWIG is based on the cluster

⁸Source : <http://inspirehep.net/record/806744>

fragmentation model. However, this model has problems in dealing with the decay of very massive clusters. *You don't discuss problems of the string model either...*

2.3.2 Underlying Event

Due to the composite nature of the protons, their collisions are not clean events.

The event structure is significantly more complex than that of the lepton collisions.

The final states of the collisions involve the multi-particle calculations. In a high energy proton-proton collisions, the underlying event (UE) includes the effects which are not coming from the primary hard scattering process. The UE includes the contributions from relatively small momentum transfer processes : initial and final-state radiations (ISR, FSR), leftover partons in the collisions called beam remnants and multiple parton interactions (MPI). Due to composite nature of proton, the remaining two partons which do not participate in a hard collision may also interact giving rise to multiple parton interactions. The UE induces an additional energy in an event which is not related to the main hard interaction. This acts as an unavoidable background which needs to be removed. Hence, it is very crucial to study and understand the UE. The UE activity increases with Q and the center-of-mass energy \sqrt{s} . Figure 2.7 shows the complex variety of processes taking place in a single proton-proton collision.

The bunch of hadrons, produced from hadronization of quarks and gluons, gets collimated in the form of "jets" with the direction towards the direction of the initial parton that originated them. The jets are the final structures observed experimentally in the detectors. These act as a bridge between the elementary quarks and gluons of QCD and the final hadrons produced in high energy collisions. Therefore, at large momentum transfer of the interacting partons, the jets and their observables are the best tools to test the predictions of perturbative QCD. Also, the jet production is sensitive to the strong coupling constant α_s . The precise knowledge of the jet production cross-section can help to extract the value of α_s .

If you define acronyms, you have to use them...

Slang

that can be

✓ Yes.

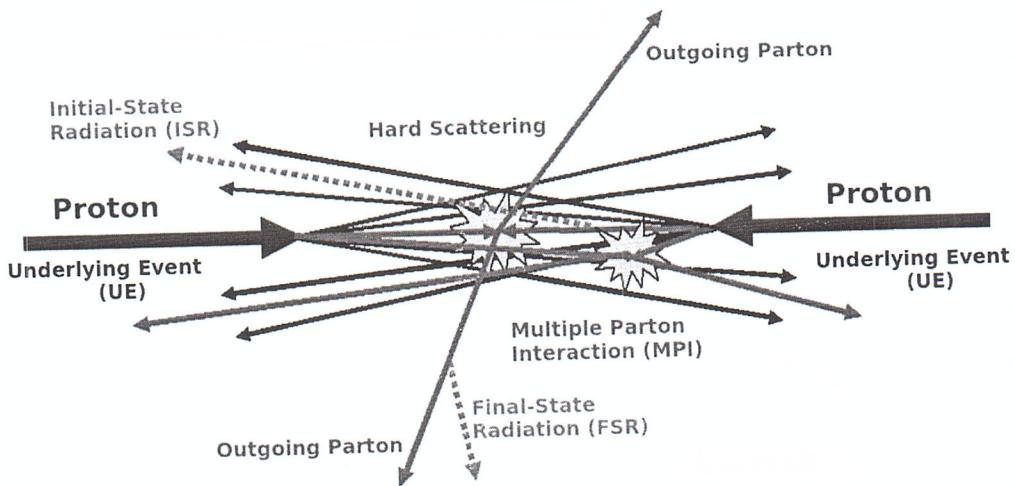


Figure 2.7: A proton-proton collision⁹ involving the main hard scattering process along with the low momentum transfer underlying event (UE) contributions coming from initial- and final-state radiations (ISR and FSR) complemented with multiple parton interactions (MPI) and collisions from leftover partons called beam remnants.

At the

and also to reduce the uncertainties of the PDFs of proton. In LHC, the simplest jet production process is $2 \rightarrow 2$ scattering process at leading-order giving dijet events. But the partons originating from ISR, FSR or MPI can also hadronize to produce jets greater than 2 in a single proton-proton collision. This results in the production of multijet events. The investigation of inclusive multijet event cross-sections permits more elaborate tests of QCD. Also, a precise study of jet variables helps to understand the signal and background modelling for the new physics searches in hadronic final states. In this thesis, the inclusive multijet event cross-sections as well as the ratio of cross-sections are exploited to extract the value of strong coupling constant α_s . In the next section, we focus on the definition of a jet.

In your analysis the UE multijet events are cut out using $P_T > 100 \text{ GeV}$.

2.4 Jets

Jets [36] are the conical structures which group the hadrons into a single physics entity. Figure 2.8 shows the the outgoing partons of the hard scattering process

⁹Source : The Energy Dependence of Min-Bias and the Underlying Event at CDF

This is not
the
point!
Only at
low P_T !!
Then jets
are back
ground.
The real
multijet
is from
 $2 \rightarrow 3$
 $2 \rightarrow 4$
...
parton
reaction!
in QCD!

in a proton-proton collision, undergoing fragmentation and hadronization processes and forming a conical jet with radius R. The jet structure was observed for the first time in hadron production of e^+e^- annihilation process at SLAC in 1975 [37]. The partons can not be measured directly by the experiments because they can not exist freely in nature. The information about the dynamics of the partons can be obtained indirectly from jets. The configurations of high-energy quarks and gluons at short distances are truly reflected in the energy and angular distributions of the jets. Hence the jets are important to study. To perform the clustering of particles, a certain set of rules *is defined* are followed in the form of jet algorithms.

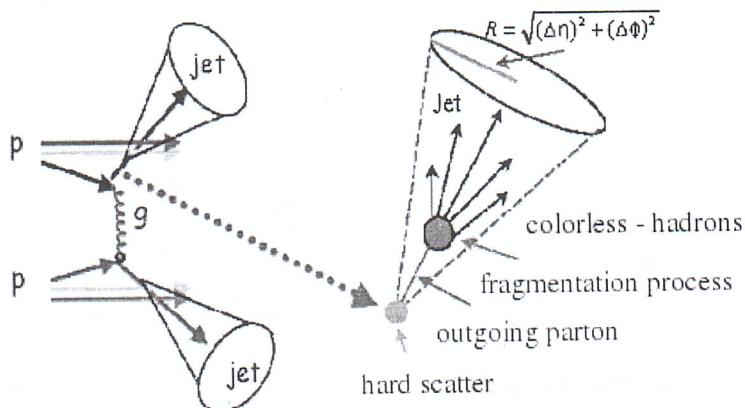


Figure 2.8: In a proton-proton collision, the outgoing partons of the hard scattering process undergo fragmentation and hadronization processes producing a shower of partons which get collimated into a conical jet with radius R.

2.4.1 Jet Algorithms

Jet algorithms [38] provide a set of rules which determine how the particles can be clustered into a jet. In a jet algorithm, usually one or more parameters are involved that indicate how close two particles must be for them to belong to the same jet. These parameters can either measure closeness in coordinate space (cone algorithms) or in momentum space (sequential algorithms). The jet algorithms are applicable on parton, particle and calorimeter levels. A recombination scheme is always associated with the detector.

ated with a jet algorithm which calculates the momentum assigned to the combined particles. A jet algorithm along with its parameters and a recombination scheme forms a “jet definition”. A jet definition [39] must be simple to implement in an experimental analysis as well as in the theoretical calculation. It should be defined at any order of perturbation theory and must yield finite cross-sections that is relatively intensive to hadronization. In addition to these requirements, a jet algorithm must be infrared and collinear (IRC) safe. Infrared safety is the property by which the addition of a soft emission i.e. addition of a soft gluon should not change or modify the number of hard jets found in an event. In an infrared unsafe algorithm, a soft gluon emission in the middle of two cone jets can lead to overlap of the two initial cones, as shown in Fig. 2.9 (top). This produces a single jet instead of initial

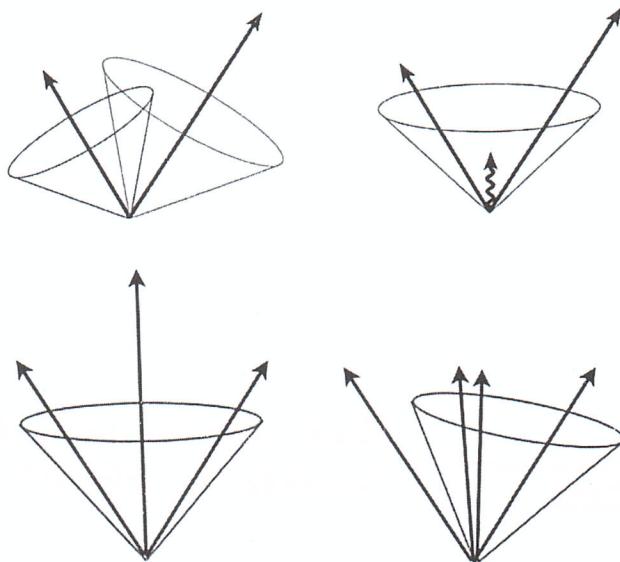


Figure 2.9: Top : Infrared unsafe behaviour of jet algorithm is illustrated where the presence of soft radiation between two jets may cause a merging of the jets that would not occur in the absence of the soft radiation. Bottom : Collinear unsafe behavior of jet algorithm is shown in which the number of jets change due to a collinear splitting¹⁰.

two jets resulting in the change of number of jets. Collinear safety is the property by virtue of which the collinear splitting i.e. replacement of one parton by two at

¹⁰Source : <http://inspirehep.net/record/1251416>

the same place should not modify the number of jets formed in an event. This implies that the output of the jet algorithm should remain the same if the energy of a particle is distributed among two collinear particles. According to the collinear safety property, the two cases shown in Fig. 2.9 (bottom) should always produce a single jet. If an algorithm produces zero or two jets after collinear splitting, then it is not collinear safe. The jet algorithms can be classified mainly into two types :

Cone algorithms - In the iterative cone (IC) algorithm [40], the jet is defined as a cone with fixed radius R in $\eta\phi$ space drawn around the highest energy seed. The relative distance (d) of all the particles is iteratively calculated and compared with R . If the calculated $d < R$, the considered particles are clustered together in a jet and the directions of the clustered particles give the direction of the jet. On the other side i.e. if $d > R$, the considered particles initiate two different jets. The algorithm iterates until the cone is stable which means that the direction of sum of momentum of all the particles is same as that of the center of cone. But IC algorithm is not IRC safe. There is another cone algorithm, Seedless Infrared-Safe cone (SIS-Cone) [41], which is an exact seedless i.e. does not rely on seed threshold and is IRC safe. This is a complex approach which tests the stability of all subsets of particles and has a complexity of $\mathcal{O}(N2^N)$ for N particles. But this algorithm is much slower and hence not preferred.

recombination

*But it's the only cone
usable in pQCD!*

Sequential algorithms - The sequential algorithms [42] cluster the particles by defining a distance between pairs of particles and recombine the pair of closest particles successively. This is collinear and infrared safe algorithm. It is possible for ~~the jet cones~~ ^{jets} to overlap such that one particle is contained in more than one jet but the sequential algorithm never assigns a particle to more than one jet. The sequential algorithm is based on transverse momentum p_T of the particles and follows the below procedure :

1. First the distance d_{ij} between two particles i and j and distance d_{iB} of the

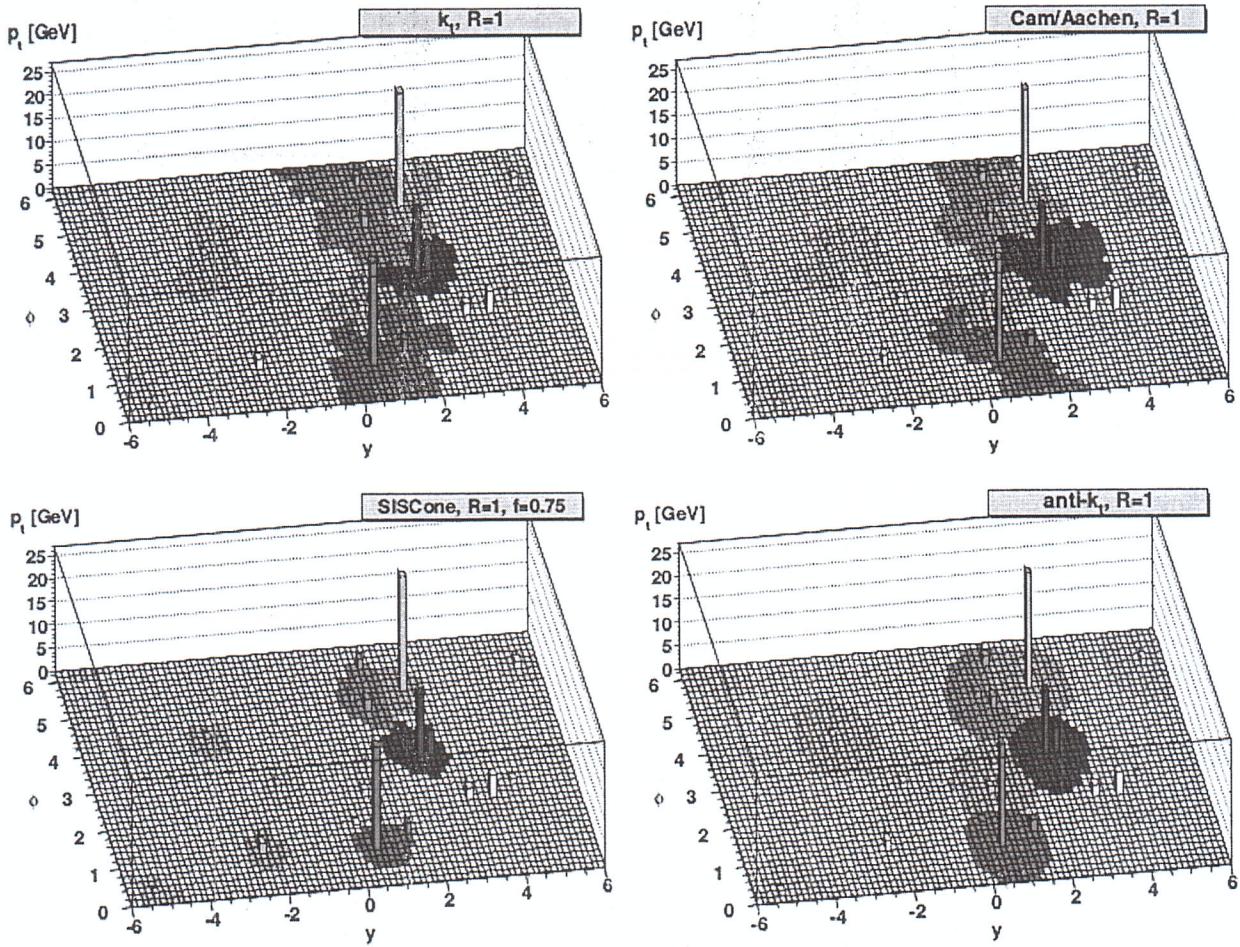


Figure 2.10: The clustering of particles, in $y\phi$ space at the parton level, into jets clustered with the k_t (top left), Cambridge/Aachen (top right), SIScone (bottom left) and anti- k_t (bottom right) algorithms with $R = 1$. The towers represent the jet p_T . The anti- k_t algorithm gives circular jets while the jets produced with other three algorithms have irregular shapes. Taken from [38].

corresponds to vector addition of four-momenta where the four-momenta of the jet is obtained by simply adding the four-momenta vector of merging particles.

The sequential clustering algorithms have ~~always~~^{traditionally} been favoured by theorists but not by experimentalists because of slow computational performance. However, the introduction of the FASTJET program [47] enhanced the speed of clustering algorithms and hence are preferred by experimentalists as well. This thesis studies the particles produced in proton-proton collisions by clustering them in to jets using

particle to the beam are calculated.

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = p_{Ti}^{2p} \quad (2.18)$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$

y_i y_j ! jct object

2. If $d_{ij} < d_{iB}$, then the particles i and j are merged into a new single particle k , summing four-momenta of two initial particles by recombination scheme and step 1 is repeated.
3. If $d_{iB} < d_{ij}$, particle i is declared as a final-state jet and the particle gets removed from the list.

This procedure continues until all particles get clustered into jets. The value of the parameter p defines the three different sequential algorithms having distinct properties. For $p = 1$, we have k_t algorithm [43, 44], $p = 0$ gives the Cambridge/Aachen (C/A) algorithm [45] whereas $p = -1$ defines the anti- k_T algorithm [46]. The k_t algorithm involves clustering of soft particles first resulting in an area that fluctuates considerably. This algorithm is susceptible to the underlying and pileup events. The C/A algorithm involves energy independent clusterings. Both k_t and C/A produce jets of irregular shapes. Instead of jet analysis, these are widely considered for studying the jet substructure. The anti- k_T algorithm tends to cluster hard particles first and produces jets with more circular regular shapes. It is less sensitive to underlying and pileup events. It is the most preferred algorithm for jet studies at the LHC. Figure 2.10 shows the clustering of same particles but using the different jet algorithms.

A jet algorithm must specify how to combine the momenta of different partons or particles going to be clustered into a jet. This is given by the recombination scheme. The most widely used recombination scheme is the E -scheme [40] which

anti- k_t algorithm with distance parameter $R = 0.7$. These jets are observed in the Compact Muon Solenoid detector of the Large Hadron Collider, the details of which are discussed in the following chapter.

To Fig. 2.3

The bottom right diagram (loop) contributes to the NLO via interference!

The listed diagrams do not represent e.g. a scattering process:

~~parton + parton \rightarrow 2 + partons~~ (e.g. $pp \rightarrow 2\text{jets} +$)
 $g_s^2 = 4\pi\alpha_s$

$$\left| \begin{array}{c} \diagup \text{eeee} \diagdown \\ g_s \quad g_s \end{array} \right|^2 \sim g_s^4 \sim \alpha_s^2 \quad \text{LO}$$

$$\left| \begin{array}{c} \diagup \text{eeeeee} \diagdown \\ \text{eeeeeee}, j_2 \end{array} \right|^2 \sim \dots \alpha_s^2 \quad \text{LO}$$

$$\left| \begin{array}{c} \diagup \text{eeee} \diagdown \\ g_s \quad g_s \quad g_s \quad \text{eee} \end{array} \right|^2 \sim g_s^6 \sim \alpha_s^3 \quad \text{NLO}$$

interference term! (identical initial & final state)

$$\left(\begin{array}{c} \diagup \text{eeee} \diagdown \\ g_s \quad g_s \end{array} \right) \otimes \left(\begin{array}{c} \diagup \text{eeeeeee} \diagdown \\ g_s \quad g_s \quad g_s \quad g_s \quad \text{eee} \end{array} \right) \sim g_s^6 \sim \alpha_s^3 \quad \text{NLO}$$

$$\left| \begin{array}{c} \diagup \text{eeeeeeee} \diagdown \\ \text{eeeeeeee} \end{array} \right|^2 \sim g_s^8 \sim \alpha_s^4 \quad \text{NNLO}$$

Chapter 5

Measurement of the Differential Inclusive Multijet Cross-sections and their Ratio

In a proton-proton collision, the inclusive jet cross-section studied as a function of jet properties, provides essential information about the parton distribution functions of the proton and the strong coupling constant. This chapter describes the measurement of differential inclusive multijet event cross-sections and the cross-section ratio. The event and jet selections, trigger studies, spectrum construction, corrections applied and calculation of the experimental uncertainties are discussed in detail.

The differential inclusive multijet event cross-sections, given by Eq. 5.1, are studied as a function of the average transverse momentum, $H_{\mathrm{T},2}/2 = \frac{1}{2}(p_{\mathrm{T},1} + p_{\mathrm{T},2})$, where $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ denote the transverse momenta of the two leading jets.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}(H_{\mathrm{T},2}/2)} = \frac{1}{\epsilon \mathcal{L}_{\mathrm{int,eff}}} \frac{N_{\mathrm{event}}}{\Delta(H_{\mathrm{T},2}/2)} \quad (5.1)$$

where N_{event} is the number of inclusive n -jet events counted in an $H_{\mathrm{T},2}/2$ bin, ϵ

is the product of the trigger and jet selection efficiencies, which are greater than 99%, $\mathcal{L}_{\text{int,eff}}$ is the effective integrated luminosity, and $\Delta(H_{\text{T},2}/2)$ are the bin widths which increase with $H_{\text{T},2}/2$ and are proportional to the $H_{\text{T},2}/2$ resolution. The measurements are reported in units of (pb/GeV). The inclusive n -jet event samples include the events with number of jets $\geq n$. In the present thesis, the measurements are performed for $n = 2$ giving inclusive 2-jet events ($n_j \geq 2$) and for $n = 3$ giving inclusive 3-jet events ($n_j \geq 3$). The cross-section ratio R_{32} , defined in Eq. 5.2 is obtained by dividing the differential cross-sections of inclusive 3-jet events to that of inclusive 2-jet one, for each bin in $H_{\text{T},2}/2$.

$$R_{32} = \frac{\frac{d\sigma_{3\text{-jet}}}{d(H_{\text{T},2}/2)}}{\frac{d\sigma_{2\text{-jet}}}{d(H_{\text{T},2}/2)}} \quad (5.2)$$

For inclusive 2-jet events sufficient data are available up to $H_{\text{T},2}/2 < 2000$ GeV, while for inclusive 3-jet events and the ratio R_{32} , the accessible range is limited to $H_{\text{T},2}/2 < 1680$ GeV.

5.1 Data Samples

This is Plural.

we're

This measurement uses the data which was collected at the center-of-mass energy of 8 TeV by the CMS experiment in the 2012 run period of the LHC. The 2012 data is taken in four periods A, B, C, D and the data sets are divided into samples according to the run period. Further each sample is grouped into subsets based on the trigger decision. For runs B-D, the JetMon stream data sets contain prescaled low trigger threshold paths (HLTPFJet40, 80, 140, 200 and 260) while the JetHT stream data sets contain unprescaled high threshold trigger paths (HLT PFJet320 and 400). For run A, the Jet stream contains all the above mentioned trigger paths. The data

to be used in physics analysis must satisfy a certain criteria according to which it should fulfill the validation requirements of the data quality monitoring procedure.

Many

CMS uses JSON (Java Script Object Notation) format files to store the range of good lumi sections within a run. In the current analysis, the applied certification file¹⁹ is based on the final event reconstruction of the 2012 data sets. The data sets used in the current study are mentioned in the Table 5.1 along with the luminosity of each data set which increases with period. Full 2012 data sample corresponds to an integrated luminosity of 19.71 fb^{-1} .

Table 5.1: Run range and luminosity of the proton-proton collisions data collected at the center-of-mass energy of 8 TeV by the CMS experiment in the year of 2012 in different run periods.

Run	Run range	Data set	Luminosity fb^{-1}
A	190456-193621	/Jet/Run2012A-22Jan2013-v1/AOD	0.88
B	193834-196531	/Jet[Mon,HT]/Run2012B-22Jan2013-v1/AOD	4.41
C	198022-203742	/Jet[Mon,HT]/Run2012C-22Jan2013-v1/AOD	7.06
D	203777-208686	/Jet[Mon,HT]/Run2012D-22Jan2013-v1/AOD	7.37

5.1.1 Monte Carlo Samples

To have a comparison of results obtained from the data events with those from the simulated events, the MADGRAPH5 Monte Carlo (MC) event generator has been used. It has been interfaced to PYTHIA6 by the LHE event record, to generate the rest of the higher-order effects using the Parton Showering (PS) model, with tune Z2* to model the underlying event. The MC samples are processed through the complete CMS detector simulation to allow studies of the detector response and comparison to measured data on detector level.

The cross-section measured as a function of the transverse momentum p_T or the scalar sum of the transverse momentum of all jets H_T , falls steeply with the in-

¹⁹Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON

creasing p_T . So in the reasonable time, it is not possible to generate a large number of high p_T events. Hence, the events are generated in the different phase-space regions binned in H_T or the leading jet p_T . Later on, the different phase-space regions are added together in the data analyses by taking into account the cross-section of the respective different phase-space regions. The official CMS MADGRAPH5+PYTHIA6 (MG5+P6) MC samples used in this analysis are generated as slices in the H_T phase-space, as tabulated in Table 5.2 along with their cross-sections and number of events generated.

Table 5.2: The official Monte Carlo samples are produced in phase space slices in H_T with the generator MADGRAPH5 and interfaced to PYTHIA6 for the parton shower and hadronization of the events. The cross-section and number of events generated are mentioned for each sample.

Generator	Sample	Events	Cross-section pb
MADGRAPH5 + PYTHIA6	/QCD_HT-100To250_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	50129518	1.036×10^7
	/QCD_HT-250To500_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	27062078	2.760×10^5
	/QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	30599292	8.426×10^3
	/QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	13843863	2.040×10^2

5.2 Event Selection

The events are selected according to several quality criteria which ensure the high purity and high selection efficiency of the sample to be studied. This event selection also reduces beam induced background, detector-level noise and jets arising from the fake calorimeter energy deposits.

~~nonphysical measurements.~~

5.2.1 Trigger Selection

CMS implements a two-level trigger system to reduce the amount of recorded events to a sustainable rate. In this analysis the jets are the final objects to study. So single

jet trigger paths with varying thresholds are used to select events in the data. It consists of one L1 trigger seed and multiple HLT filters. The L1 jet trigger uses transverse energy sums computed by using both HCAL and ECAL in the central region ($|\eta| < 3.0$) or HF in the forward region ($|\eta| > 3.0$). The single jet triggers (HLT_PFJetX), same as the ones used for other CMS 8 TeV measurements [26, 99], are used in the current study and are tabulated in Table 5.3. A single jet trigger selects an event in which at least one jet has the transverse momentum above the threshold. HLT_PFJetX implies that there is at least one jet in the event, whose $p_T > X$ (GeV). The L1 trigger has a lower threshold to ensure full efficiency versus p_T of the HLT trigger. The p_T spectrum is steeply falling and hence the rates for low- p_T jets are very high. So it is not feasible to use a single unprescaled trigger for the selection of all required events. To collect sufficient data in the lower part of the p_T spectrum, five prescaled low- p_T trigger paths, each with different prescale value, are used. Also, one unprescaled trigger i.e. HLT_Jet320 is used in the high p_T region, in which the rate is sufficiently small to collect and store all events. During the reconstruction of the spectrum, the prescales have been taken into the account.

Table 5.3: The single jet HLT trigger paths used in the analysis. The column $H_{T,2}/2$, 99% indicates the value of $H_{T,2}/2$ at which each trigger exhibits an efficiency larger than 99%. The last column gives the effective luminosity seen by each trigger which when divided by the total integrated luminosity of 19.71 fb^{-1} , gives the effective prescale applied on a trigger over the whole run period.

Trigger Path	L1 threshold GeV	HLT threshold GeV	$H_{T,2}/2$, 99% GeV	Eff. Lumi fb^{-1}
HLT_PFJet80	36	80	120.0	0.0021
HLT_PFJet140	68	140	187.5	0.056
HLT_PFJet200	92	200	262.5	0.26
HLT_PFJet260	128	260	345.0	1.06
HLT_PFJet320	128	320	405.0	19.71

The efficiency of each trigger path as a function of $H_{T,2}/2$ is described by the turn-on curves with a rising part where the trigger is partly inefficient, until a plateau

region where the trigger is fully efficient. Hence it is important to determine the threshold above which a trigger becomes fully efficient. The threshold is the value at which the trigger efficiency exceeds 99%. The trigger efficiency for HLT_PFJetY is given by Eq. 5.3 where HLT_PFJetX is the reference trigger and is assumed to be fully efficient in the considered phase space region. The value of X is chosen previous to that of Y in p_T ordering from the trigger list so that the higher trigger condition can be emulated from the lower trigger path.

$$\epsilon_{\text{HLT_PFJetY}} = \frac{H_{T,2}/2 \left(\text{HLT_PFJetX} + (\text{L1Object}_p > Z) + (\text{HLTOBJECT}_p > Y) \right)}{H_{T,2}/2(\text{HLT_PFJetX})} \quad (5.3)$$

where Y is the p_T threshold of HLT_PFJetY and Z is the L1 seed value corresponding to the trigger path HLT_PFJetY. The denominator represents the number of events for which the reference trigger path HLT_PFJetX has been fired. The numerator is the number of events for which HLT_PFJetX has been fired along the p_T of L1Object \geq Z and the p_T of HLTOBJECT \geq Y. For example, to obtain turn-on curve for HLT_PFJet260, HLT_PFJet200 is the reference HLT path. The p_T cut on L1Object is 128 GeV and p_T cut on HLTOBJECT is 260 GeV. The threshold point at which the trigger efficiency is larger than 99% is determined by fitting the turn-on distribution with a sigmoid function described in Eq. 5.4. The trigger turn-on curves as a function of $H_{T,2}/2$ can be seen in Fig. 5.1 which are described by a sigmoid function (blue line). The error bars give the uncertainty on the efficiency which is calculated using Clopper-Pearson confidence intervals [100].

$$f_{fit}(x) = \frac{1}{2} \left(1 + \text{erf} \left(\frac{x - \mu}{\sqrt{2}\sigma} \right) \right) \quad (5.4)$$

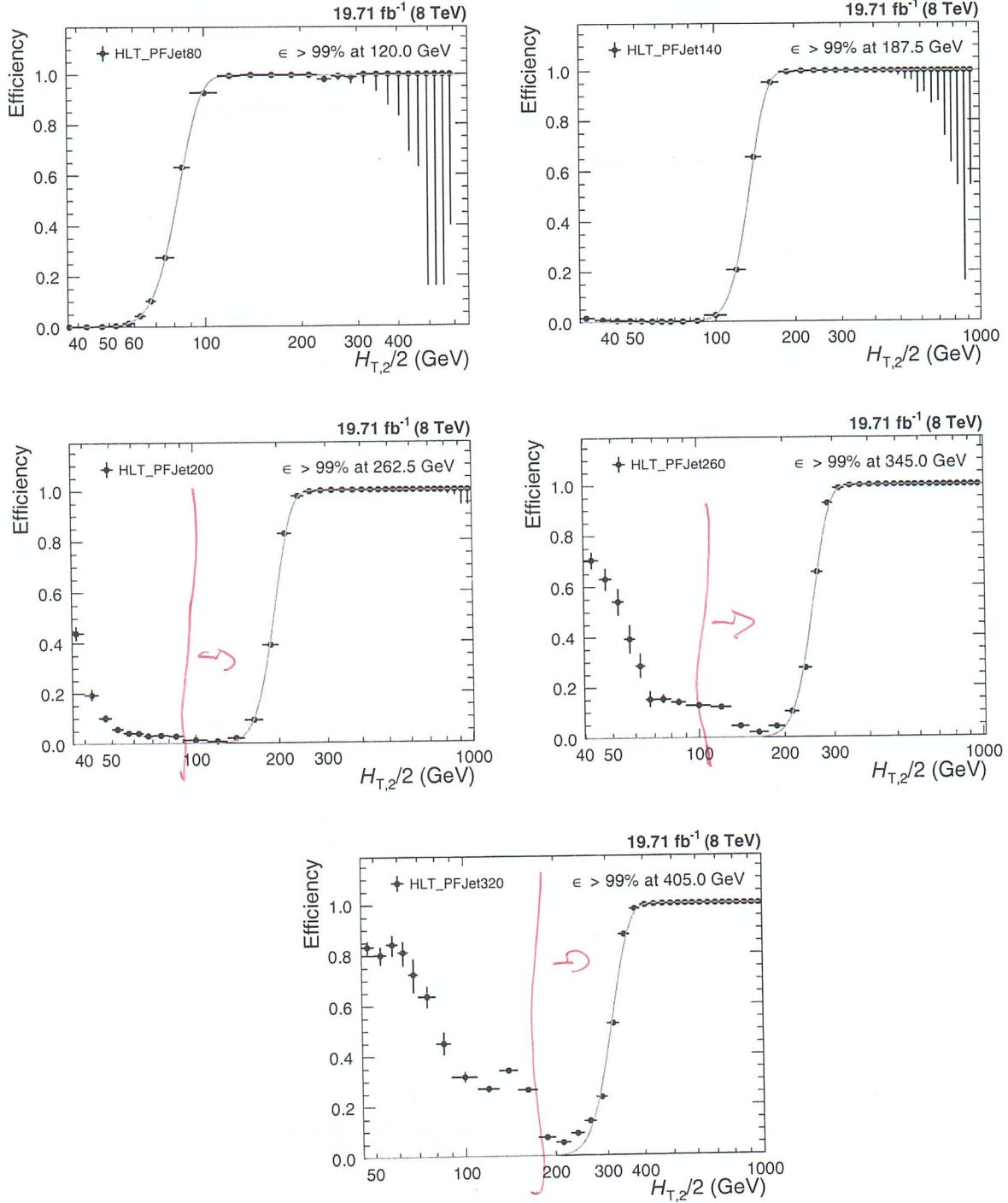


Figure 5.1: Trigger efficiencies turn-on curves for the single jet HLT trigger paths are fitted with a sigmoid function (blue line) to obtain the 99% efficiency threshold. The error bars give the uncertainty on the efficiency which is calculated using Clopper-Pearson confidence intervals [100].

5.2.2 Primary Vertex Selection

The reconstructed tracks, number of strip and pixel hits and the normalized track χ^2 identify the primary vertex (PV). The tracks are clustered according to the z-coordinate of their point of closest approach to the beam axis. A selection criteria ~~the following~~ ^{~~are imposed~~} for primary vertex should be followed which helps to identify and reject ~~the~~ beam background events. At least one good primary vertex reconstructed from at least four tracks within a distance of $|z(PV)| < 24$ cm to the nominal interaction point in a collision is required in each event. The radial distance in x-y plane, $\rho(PV)$ should not be greater than 2 cm. The number of degrees of freedom in fitting for the position of each vertex using its associated tracks should be at-least four in number.

5.2.3 Missing Transverse Energy

In an ideal detector where all particles could be identified and perfectly measured, the transverse momentum of all particles would sum up to zero. But ~~the~~ neutral weakly interacting particles, such as neutrinos, escape from typical collider detectors and do not produce any direct response in the detector ~~elements~~. The imbalance of total momentum of all visible particles can give ~~the~~ hints of the presence of such particles. The vector momentum imbalance in the plane perpendicular to the beam direction is known as missing transverse momentum or energy (E_T^{miss}). It is one of the most important observables for discriminating leptonic decays of W bosons and top quarks from background events which do not contain ^{high-energetic} neutrinos, such as multijet and Drell-Yan events.

The ratio of missing transverse energy to the total transverse energy $E_T^{\text{miss}}/\sum E_T$, shown in Fig. 5.2 for $n_j \geq 2$ (left) and $n_j \geq 3$ (right) events, shows a discrepancy between the data (black solid circles) and simulated MC (blue histogram), at the tail part of the distribution. This is because of a finite contribution

from $Z(\rightarrow \nu\bar{\nu}) + \text{jet}$ events which gives rise to non-zero E_T in the events in the data. Such events are absent in QCD simulated events in MC. Hence $E_T^{\text{miss}} / \sum E_T$ is required to be less than 0.3 to reject events with high E_T^{miss} .

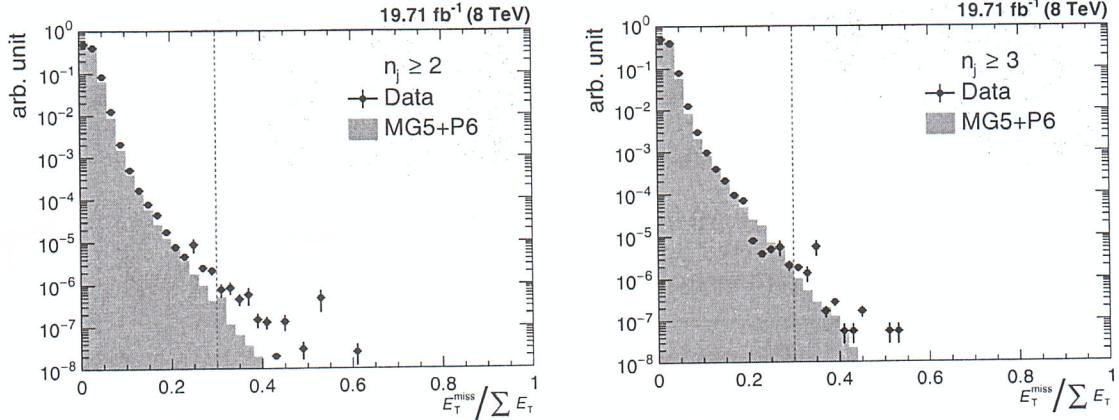


Figure 5.2: Missing transverse energy fraction of the total transverse energy per event in the data (black solid circles) and simulated Monte Carlo events (blue histogram) in inclusive 2-jet (left) and 3-jet (right) events. To remove background and noise, events with a fraction exceeding a certain threshold, here indicated with the red dashed line, are rejected.

5.2.4 Jet Identification

nonphysical

In order to suppress ~~fake~~ jets, arising from detector noise or mis-reconstructed particles, jet identification criteria (ID) have been applied. ~~Instead of applying it event-wise, it is applied to each jet.~~ The algorithm works on reconstructed jets using information of the clustered particle candidates. The official tight jet ID [101], recommended by JETMET group [102] is used. Due to pileup and electronic noise the jet constituent fractions may vary from event to event. In order to reject the noisy jets, ~~some~~ jet selection criteria are optimized to select only good quality jets. The selection criteria are implemented as selection cut on jet fractions. Table 5.4 summarizes the properties of the reconstructed jets and their respective cuts. Each jet should contain at least two particles, one of which should be a charged hadron. The cut on the fraction of neutral hadrons and photons removes HCAL noise and ECAL noise, respectively. Muons that are falsely identified and clustered as jets are

removed by the muon fraction criterion. Based on information of the tracker, additional selection cuts are enforced in the region $|\eta| < 2.4$. The charged electromagnetic fraction-cut removes the jets clustered from misidentified electrons. Furthermore, the fraction of charged hadrons in the jet must be larger than zero and jets without any charged hadrons are very likely to be pileup jets. The Figs. 5.3 and 5.4 show the distributions of the jet constituents observed in the data (black solid circles) and simulated MC events (blue histogram) for $n_j \geq 2$ and $n_j \geq 3$, respectively.

Table 5.4: The jet identification criteria (ID) removes noise and fake jets based on the properties of the reconstructed jets and the clustered particle candidates. All the tight ID selection cuts, recommended by the JETMET group are applied [102]. *Slang*

	Property	Tight ID cut
Whole η region	neutral hadron fraction	< 0.90
	neutral EM fraction	< 0.90
	number of constituents	> 1
	muon fraction	< 0.80
only $ \eta < 2.4$	charged hadron fraction	> 0
	charged multiplicity	> 0
	charged EM fraction	< 0.90

5.2.4.1 Jet ID Efficiency

The efficiency of the jet ID as a function of $H_{T,2}/2$ is studied using a tag-and-probe technique with dijet events. The two leading jets are required to be back-to-back in the azimuthal plane such that $|\Delta\phi - \pi| < 0.3$. One of the dijets is selected randomly as a “tag” jet which is required to fulfill the tight jet ID criteria. The other jet is called “probe” jet for which it is examined, whether it also passes the tight jet ID. The ID efficiency is defined as the ratio of events where the probe jet passes the ID requirements, over the total number of dijet events. It is shown as function of $H_{T,2}/2$ in Fig. 5.5 and as expected, it is always greater than 99%. The QCD cross-section decreases as a function of $H_{T,2}/2$ and hence the number of events decrease *or when* moving to higher $H_{T,2}/2$. Consequently the statistical fluctuations for ID efficiency

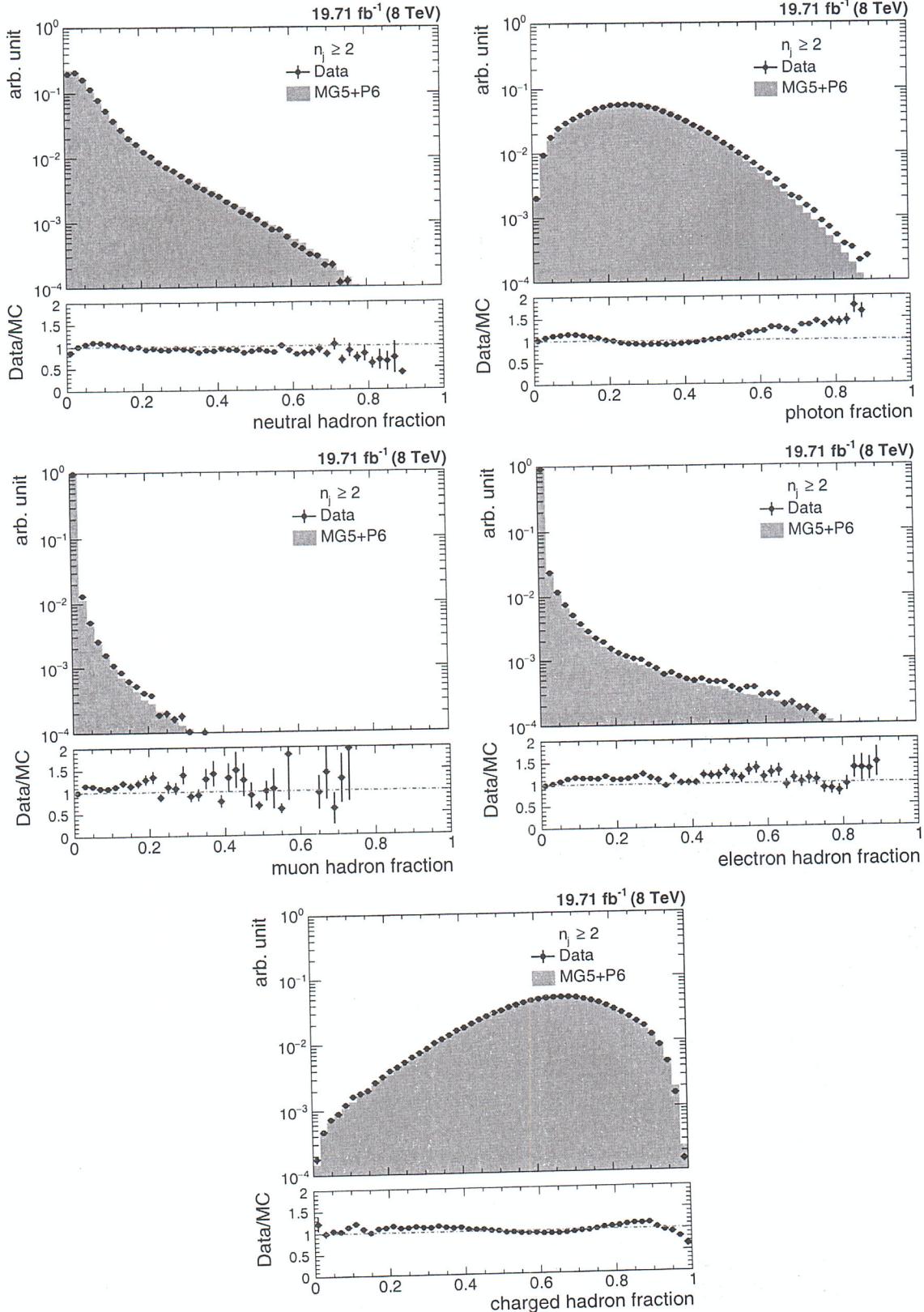


Figure 5.3: The fractions of jet constituents as observed in the data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 2-jet events. The data and simulations are normalized to the same number of events. The distributions are shown after the application of the jet ID.

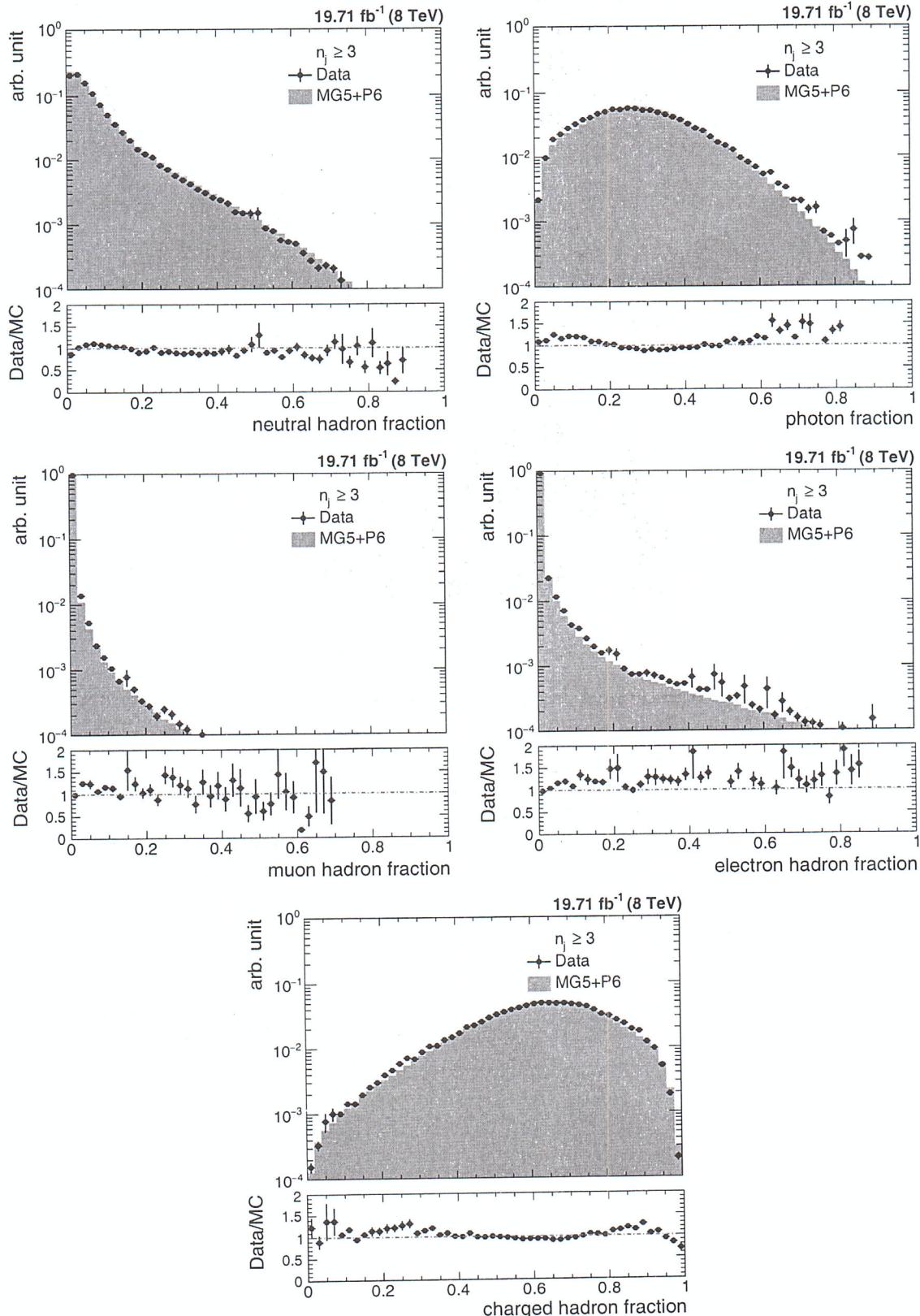


Figure 5.4: The fractions of jet constituents as observed in the data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 3-jet events. The data and simulations are normalized to the same number of events. The distributions are shown after the application of the jet ID.

are larger at higher $H_{\mathrm{T},2}/2$.

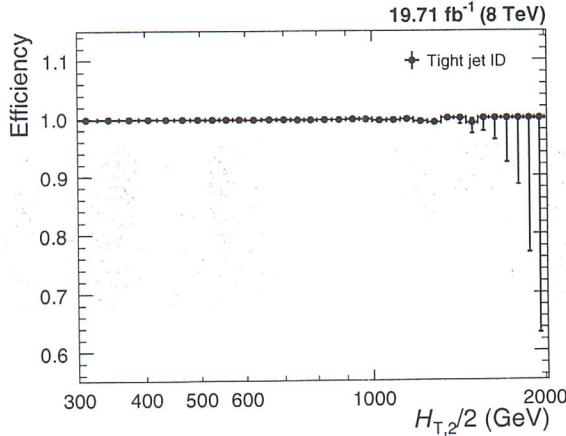


Figure 5.5: The jet ID efficiency is studied as a function of $H_{\mathrm{T},2}/2$ with tag-and-probe technique using dijet event topologies and it always exceeds 99%.

5.2.5 Jet Selection

The measurement of differential cross-sections and their ratio uses jets clustered from particle flow candidates using the anti- k_t jet algorithm with a size parameter, $R = 0.7$. The energy scale of the jets is corrected with the CMS recommended jet energy corrections, described in Sec. 4.3.2. These corrections are applied to jets in both the data²⁰ as well as in simulated events²¹. As a convention, the jets in one event are ordered in decreasing p_{T} , with the first (leading) jet being the jet with highest p_{T} . The jet selection, based on phase space cuts on transverse momentum and rapidity of jets in an event, is as follows :

- All jets having $p_{\mathrm{T}} > 150$ GeV and $|y| < 5.0$ are selected.
- Events with at least two jets are selected.
- The two leading jets should have $|y| < 2.5$ and further jets are counted only, if they lie within the same central rapidity range of $|y| < 2.5$.

²⁰Winter14_V8 jet energy corrections

²¹START53-V27 jet energy corrections

and good comparability to pQCD.

These cuts assure high detector acceptance. For a consistent comparison, the exactly same selection criteria ~~are~~ is applied in the data and simulated events as well as in theoretical calculations.

5.3 Comparison with Simulation

5.3.1 Pileup Reweighting

While generating the official Monte-Carlo samples, the number of pileup interactions ~~are~~ describing the conditions expected for each data-taking period are taken ~~care~~ ^{into account} of. But the number of pileup events implemented in the simulation $N_{\text{MC}}(N_{\text{PU,truth}})$, does not match exactly ~~with~~ the one measured in the data $N_{\text{data}}(N_{\text{PU,est.}})$. To match the pileup distributions in the data, a reweighting factor w_{PU} , as given by Eq. 5.5 is applied to the simulated events. In Fig. 5.6 the number of reconstructed vertices are shown before (left) and after pileup reweighting (right). It is observed that before pileup reweighting there was a significant mismatch of the pileup distributions in the data (black solid circles) and simulated MC events (blue histogram), which completely vanishes after reweighting.

$$w_{\text{PU}} = \frac{N_{\text{data}}(N_{\text{PU,est.}}) / \sum N_{\text{data}}}{N_{\text{MC}}(N_{\text{PU,truth}}) / \sum N_{\text{MC}}} \quad (5.5)$$

5.3.2 Comparison of Cross-sections and their Ratio

The measured data distribution of differential cross-section at the detector level is compared to the predictions of Monte Carlo simulation using MADGRAPH5 generator interfaced with PYTHIA6 (MG5+P6) including the detector simulation as well as to a fixed-order theory prediction obtained using CT10-NLO PDF set. Figure 5.7 shows the comparison of differential cross-section as a function of $H_{\text{T},2}/2$ for

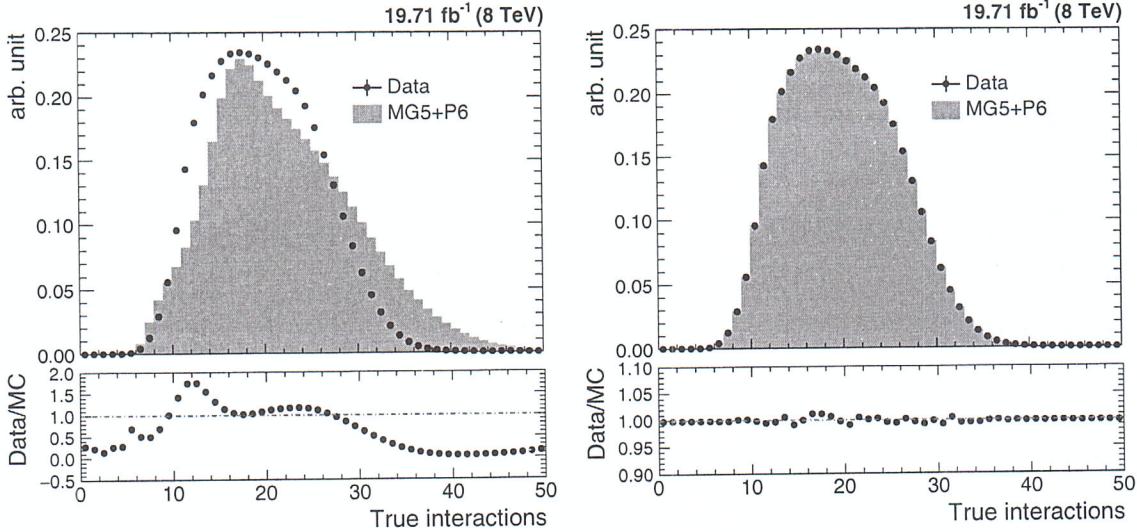


Figure 5.6: Number of reconstructed vertices in the data (black solid circles) and simulated Monte Carlo events (blue histogram) before (left) and after (right) the pileup reweighting.

$n_j \geq 2$ (left) and $n_j \geq 3$ (right) events, for the data (black solid circles), MG5+P6 MC (red empty circles) and CT10-NLO (blue histogram). The bottom panel in each plot shows the ratio of the data to the MC predictions (red line) as well as to the CT10-NLO theory predictions (blue line). The NLO predictions on parton level are not corrected for non-perturbative effects. Still the NLO predictions describe the data better as compared to the LO MC simulations which roughly describe the spectrum on detector level. ~~The S~~ufficient data for $n_j \geq 2$ and $n_j \geq 3$ events are available up to $H_{T,2}/2 < 2000$ GeV and 1680 GeV, respectively. Due to ~~some~~ kinematical constraints, the minimum cut on $H_{T,2}/2$ is 300 GeV (explained in Sec. 6.1.1). Hence the differential cross-sections are studied in the range $300 \text{ GeV} < H_{T,2}/2 < 2000$ GeV for $n_j \geq 2$ and $300 \text{ GeV} < H_{T,2}/2 < 1680 \text{ GeV}$ for $n_j \geq 3$ events.

The ratio of differential cross-sections, R_{32} as a function of $H_{T,2}/2$, is extracted by dividing the cross-section of selected inclusive 3-jet events to that of inclusive 2-jet events at any given bin size of $H_{T,2}/2$. In the cross-section ratios, the numerator and denominator are not independent samples. So to calculate the statistical uncertainty

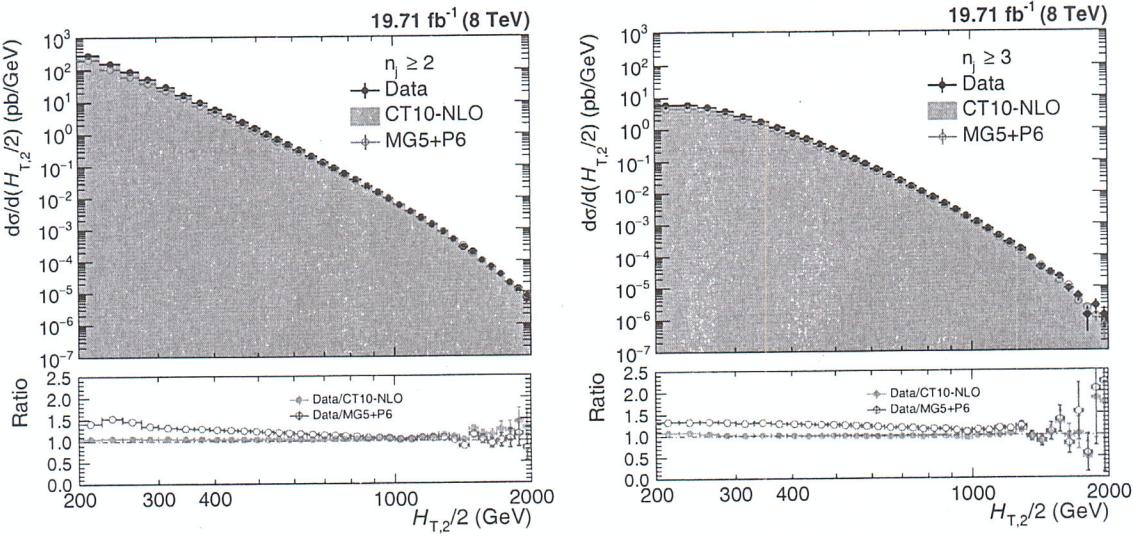


Figure 5.7: The reconstructed level differential cross-sections are compared for the data (black solid circles) and LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo (red empty circles) simulations with CT10-NLO theory predictions (blue histogram), as a function of $H_{T,2}/2$ for inclusive 2-jet (left) and 3-jet (right) events. Ratios of the data to the Monte Carlo predictions (red line) as well as to the CT10-NLO predictions (blue line) are shown in bottom panel of each plot.

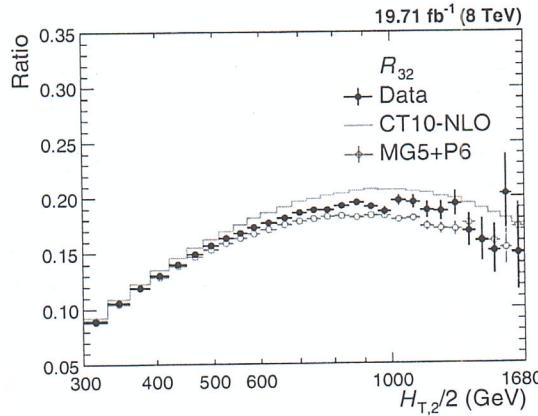


Figure 5.8: Comparison of the reconstructed level cross-section ratio R_{32} as a function of $H_{T,2}/2$, for the data (black solid circles) and LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo (red open circles) with CT10-NLO theory predictions (blue line). The error bars give the asymmetrical statistical uncertainty, calculated by the Wilson score interval method which takes into the account the correlation between the numerator and denominator.

for the cross-section ratios at reconstructed level, the Wilson score interval method is used which takes into account the correlation between the numerator and the denominator and gives asymmetric errors. Figure 5.8 shows the comparison of the

cross-section ratio R_{32} as a function of $H_{T,2}/2$, for the data (black solid circles) and LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo (red open circles), at reconstructed level with CT10-NLO theory predictions (blue line). Since in events with $n_j \geq 3$, the enough statistics for differential cross-section is available only up to 1680 GeV of $H_{T,2}/2$, R_{32} is also studied in the range $300 \text{ GeV} < H_{T,2}/2 < 1680 \text{ GeV}$. The bin-wise inclusive 2-jet and 3-jet events differential cross-sections as well as their ratio R_{32} , evaluated at detector level, along with statistical uncertainty (in %) are tabulated in Table A.1.

5.4 Jet Energy Resolution (JER)

In an ideal experiment, the value of a physical quantity would be determined exactly with an infinite precision, e.g. whenever a particle with energy E passes through an ideal calorimeter having infinite resolution, the measured energy should always be equal to E . But in real world, the measured energy of the above mentioned particle might differ from the value E . This difference of the measured quantity from its true value may be due to detector noise, uncertainties in the calibration, non-linearity of the response etc. Hence this results in the finite value of the resolution of the detector known as jet energy resolution (JER). In such a case, the measured values of energy of different particles, passing through the same detector with same energy E , will be different. Such measurements are described by a Gaussian distribution, centered around the true value of the measured quantity and its width is generally interpreted as detector resolution. Hence the importance of the detector resolution lies in the fact that it indicates how much the measured value of the observable differs from the true one i.e. how precisely a physical observable can be measured. The narrower the distribution, the higher the resolution is and hence more efficient is the detector.

Due to finite resolution of the CMS detector, the measured transverse momenta

accurate?
precise?

of jets get smeared. Since the observable in this study i.e. $H_{T,2}/2$ is the average sum of transverse momentum of the leading and sub-leading jets, the resolution of the detector has to be studied in terms of this observable. The CMS detector simulation based on MG5+P6 MC event generators is used to determine the resolution as both the particle and reconstructed level information is available. The jets clustered from stable generator particles called Gen jets as well as from particle flow candidates reconstructed from the simulated detector output called Reco jets, are used. The studies of the JETMET working group at CMS has shown that the jet energy resolution in the data is actually worse than in simulation [103]. So the reconstructed jet transverse momentum needs to be smeared additionally to match the resolution in the data. Table 5.5 shows the scaling factors (c) which need to be applied on the transverse momentum of simulated reconstructed jets. The scaling factor depends on the absolute η of the jet and are provided by JETMET working group at CMS [[103]]. The uncertainty on these measured scaling factors ($c_{central}$) needs to be taken into account in a physics analysis. This is done by smearing the reconstructed jets with two additional sets of scaling factors, c_{up} and c_{down} , that correspond to varying the factors up and down respectively, by one sigma and evaluating the impact of these new sets.

Table 5.5: JETMET working group at the CMS has shown that the jet energy resolution in the data is actually worse than in simulation [103]. To match the resolution in the data, the reconstructed jet transverse momentum in simulated events need to be smeared by applying the scale factors. The uncertainty on the resolution is given by an upwards and downwards variation c_{up} and c_{down} of the measured scaling factor $c_{central}$.

η	0.0 - 0.5	0.5 - 1.1	1.1 - 1.7	1.7 - 2.3	2.3 - 2.8
$c_{central}$	1.079	1.099	1.121	1.208	1.254
c_{down}	1.053	1.071	1.092	1.162	1.192
c_{up}	1.105	1.127	1.150	1.254	1.316

The reconstructed jet p_T is smeared randomly using a Gaussian function, $f(p_T)$

with a width widened by the scaling factor ($c_{central}$) :

$$f(p_T) = a \times \exp\left(-\frac{1}{2}\left(\frac{p_T - \mu}{\sigma}\right)^2\right) \quad (5.6)$$

where a is a constant, mean $\mu = 0$, width $\sigma = \sqrt{c_{central}^2 - 1} \cdot \text{JER}(p_T) \times p_T$ and $\text{JER}(p_T)$ is the resolution determined as a function of jet p_T using MG5+P6 MC simulated events. After smearing transverse momentum of each reconstructed jet, $H_{T,2}/2$ is calculated from both generator particle jets (Gen $H_{T,2}/2$) as well as the particle flow or reconstructed jets (Reco $H_{T,2}/2$). Then the response is calculated as defined in the Eq. 5.7.

$$R = \frac{\text{Reco } H_{T,2}/2}{\text{Gen } H_{T,2}/2} \quad (5.7)$$

The width of the response distribution in a given Gen $H_{T,2}/2$ bin is interpreted as the resolution which in good approximation can be described by 1σ of a Gaussian fit of the response distribution. A double-sided Crystal Ball function takes into account the non-Gaussian tails of the jet response distribution. The resolution as a function of $H_{T,2}/2$ is calculated separately for both $n_j \geq 2$ and $n_j \geq 3$ events. A fit example for one Gen $H_{T,2}/2$ bin is shown in Fig. 5.9 for $n_j \geq 2$ (left) and $n_j \geq 3$ (right) events. Here the black dots represent the jet response distribution and the double-sided Crystal Ball fit (blue line) is overlayed by the Gaussian fit (red line). The resolution in each Gen $H_{T,2}/2$ bin is then plotted as a function of Gen $H_{T,2}/2$.

As expected, it has been observed from Fig. 5.10 that the Crystal Ball function (blue solid circles) describes the measured distributions better as compared to Gaussian function fit (red solid circles), especially in the low- $H_{T,2}/2$ region where the non-Gaussian tails are more pronounced. Hence JER is determined using Crys-

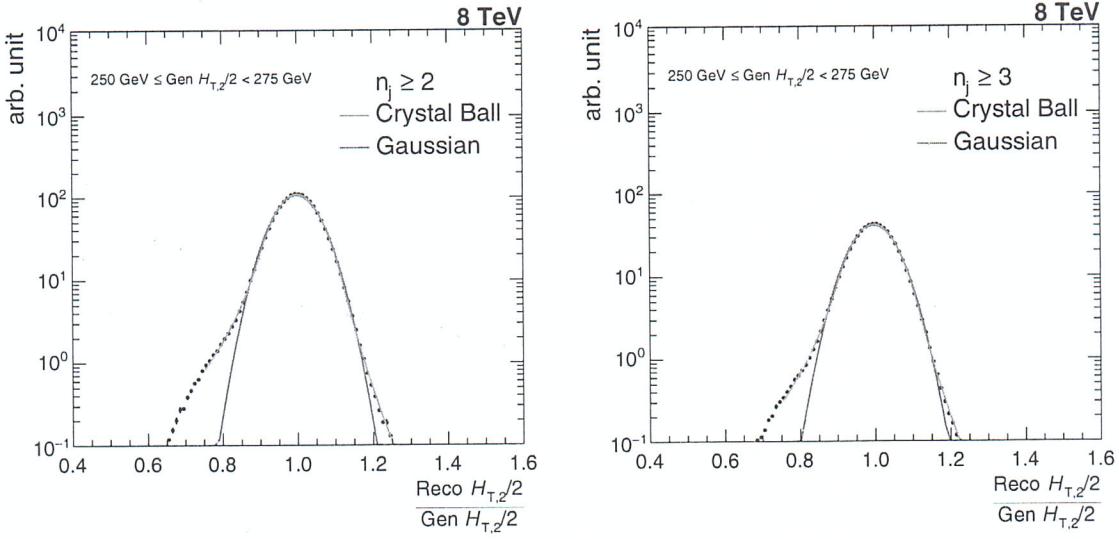


Figure 5.9: Fitting of the jet energy resolution distribution, obtained using LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo simulated events, as a function of $H_{T,2}/2$ for inclusive 2-jet (left) and 3-jet (right) events. The blue line shows the double-sided Crystal Ball function fit of $\frac{\text{Reco } H_{T,2}/2}{\text{Gen } H_{T,2}/2}$ in each Gen $H_{T,2}/2$ bin, overlayed by Gaussian fitting the core of the resolution (red line).

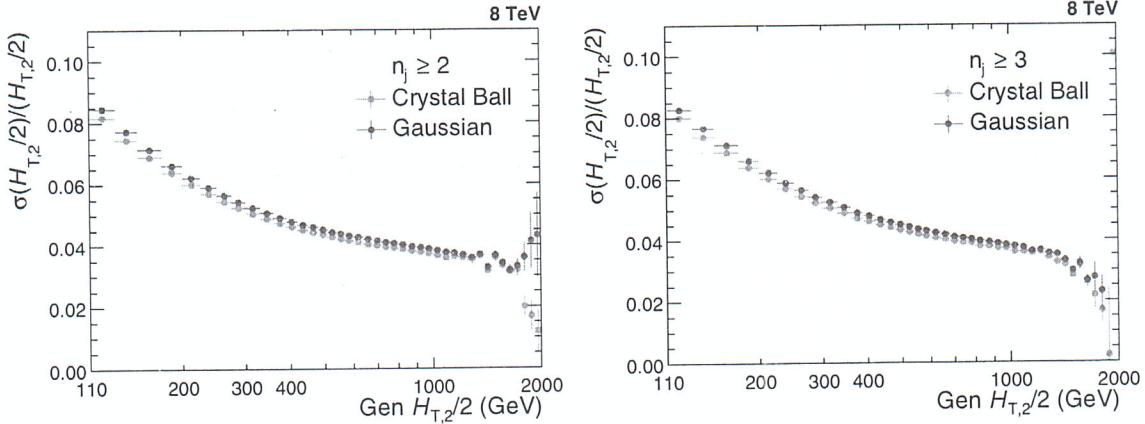


Figure 5.10: Comparison of jet energy resolution calculated using Crystal Ball fit function (blue solid circles) and Gaussian fit function (red solid circles) for inclusive 2-jet (left) and 3-jet (right) events.

tal Ball function fit. Figure 5.11 shows the final relative jet energy resolution (JER) which is described by a modified version of the NSC formula (blue solid line) [104], as mentioned in Equation 5.8. To consider the migration to lower as well as higher bins and to obtain the resolution with reasonable statistics over the full range of Gen $H_{T,2}/2$, the fit function is extrapolated to 80 GeV and up to 2000 GeV, as shown

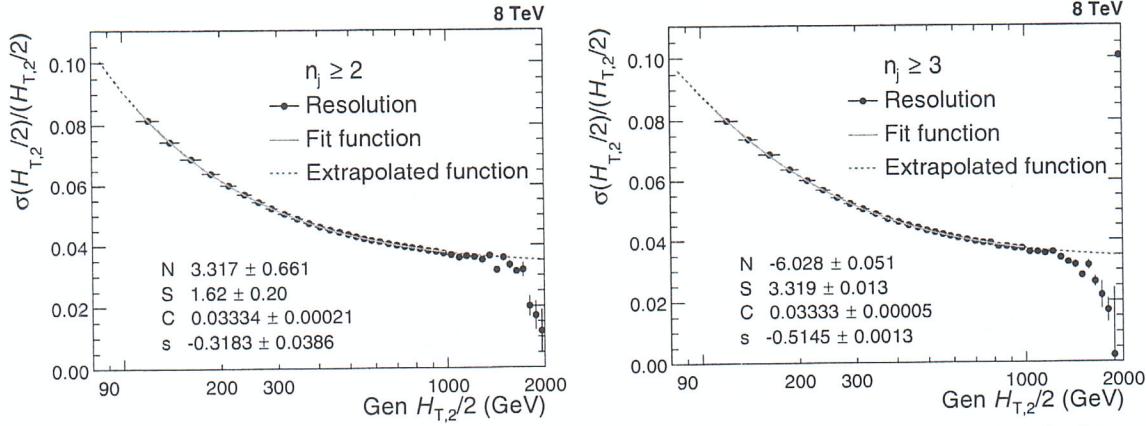


Figure 5.11: Jet energy resolution (JER) is shown as a function of Gen $H_{T,2}/2$ for inclusive 2-jet (left) and 3-jet (right) events. JER (black solid circles) is fitted by using the modified NSC-formula (blue solid line) which is extrapolated to 80 GeV and up to 2000 GeV (red dashed line) to consider the migration into lower as well as higher bins.

by red dashed line. The fit formula used here is basically the usual NSC formula which describes the resolution in terms of noise N originating due to electronic and pileup noise, a stochastic component S due to sampling fluctuation and EM fraction fluctuation per hadrons and a constant term C due to presence of dead material, magnetic field and calorimeter cell to cell fluctuations. In the low $H_{T,2}/2$ region the tracking has ~~a non-negligible influence on the resolution due to the particle flow algorithm, so the additional parameter s is introduced to obtain slightly better fits.~~ ^{improves} ~~describe this effect.~~ The parameters obtained after fitting the relative resolution using the above mentioned NSC formula are tabulated in Table 5.6 for $n_j \geq 2$ and $n_j \geq 3$ events. This calculated JER is used in unfolding procedure to smear the generated truth spectrum which is used as input in getting the response matrices and is explained in detail in Sec. 5.5.1. Since JER in $n_j \geq 2$ events is similar to that one in $n_j \geq 3$ events, N , S and C fit parameters obtained for $n_j \geq 3$ events are used for unfolding R_{32} .

$$\frac{\sigma(x)}{x} = \sqrt{\text{sign}(N) \cdot \frac{N^2}{x^2} + S^2 \cdot x^{s-1} + C^2} \quad \text{where } x = H_{T,2}/2 \quad (5.8)$$

Table 5.6: The parameters obtained by fitting the relative resolution as a function of $H_{T,2}/2$, using the modified NSC formula, for inclusive 2-jet and 3-jet events.

	N	S	C	s
Inclusive 2-jet	3.32	1.62	0.0333	-0.318
Inclusive 3-jet	-6.03	3.32	0.0333	-0.515

Since the JER is calculated using MG5+P6 Reco and Gen $H_{T,2}/2$ distributions, so it is expected that if Gen $H_{T,2}/2$ is smeared using this JER, it should match the Reco $H_{T,2}/2$. But this extracted JER in one large rapidity bin, smears the Gen $H_{T,2}/2$ too much because Smeared-Gen/Gen ratio (red line) shows a discrepancy from simulated-Reco/Gen ratio (blue line), as observed in Fig. 5.12 for $n_j \geq 2$ (left) and $n_j \geq 3$ (right) events. Some shortcomings in the detector simulation of the

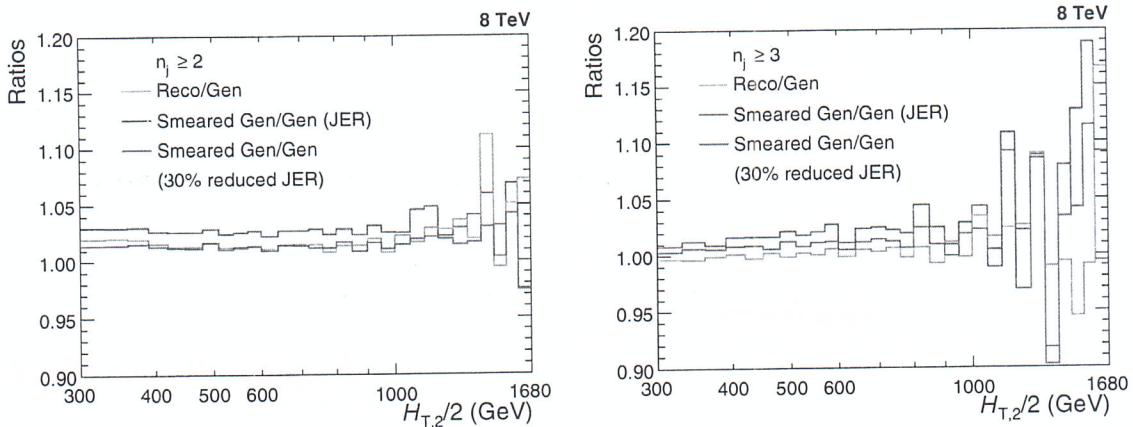


Figure 5.12: MADGRAPH5+PYTHIA6 (MG5+P6) Gen smeared using extracted jet energy resolution (JER) shows a discrepancy from simulated-Reco as Smeared-Gen/Gen ratio (red line) does not match with Reco/Gen ratio (blue line), for both inclusive 2-jet (left) and 3-jet (right) events. Smeared-Gen/Gen ratio (pink line) where Gen is smeared using 30% reduced JER matches with simulated-Reco/Gen ratio (blue line) within the statistical fluctuations. Hence an additional unfolding uncertainty is attributed by comparison to 30% reduced JER.

theory spectra lead to these small nonclosures. When the 30% reduced JER is used to smear Gen, then the ratio Smeared-Gen/Gen (pink line) matches with simulated-Reco/Gen ratio (blue line) within the statistical fluctuations. Hence an additional unfolding uncertainty is attributed by comparison to 30% reduced JER for both $n_j \geq 2$ and $n_j \geq 3$ events. Due to high statistical fluctuations at high $H_{T,2}/2$, range

up to 1680 GeV only is presented.

5.5 Unfolding

One of the main goals in an experimental measurement is to do the comparison of the data with theory predictions or with the results obtained from other experiments. But the finite resolution of a detector and the steeply falling jet p_T spectrum distorts the physical quantities. As a result, the measured observables are different from their corresponding true values. Each p_T bin content contains the migrated events from neighbouring bins along with the original events. So an unfolding process of the data should be followed in order to remove detector effects. In this analysis, the measurements are corrected for detector smearing effects and unfolded to stable particle level by using the iterative D'Agostini Bayesian algorithm as implemented in RooUnfold software package [105]. The details of the algorithm can be found in Refs. [106, 107]. In this algorithm, the number of iterations regularize the unfolding process. The obtained distribution in one iteration is taken as the input in the next one. χ^2 between two successive iterations is given by Eq. 5.9. The number of iterations stop when χ^2/N_{bins} is < 1 . A reduced χ^2 is obtained by a higher number of iterations but this will also increase the uncertainty and there are larger bin-by-bin fluctuations and correlations. So the optimization of number of iterations is very important. In the current analysis, unfolding done with “four” iterations gives the best results with low χ^2 and low bin-by-bin correlations.

$$\chi^2 = \sum_{i=1}^{N_{bins}} \left(\frac{n_i^{j+} - n_i^j}{\sqrt{n_i^j}} \right)^2 \quad (5.9)$$

where n_i^j number of events in i -th bin for j -th iteration.

The measured differential cross-sections as a function of $H_{T,2}/2$, are unfolded

separately for $n_j \geq 2$ and $n_j \geq 3$ events. The measured cross-section ratio R_{32} is also corrected for detector smearing effects and unfolded to particle level. There can be two ways to obtain unfolded cross-section ratio :

- **Method I** : First unfold separately the inclusive 2-jet and 3-jet measured cross-sections and then construct the ratio R_{32} .
- **Method II** : Unfold directly the cross-section ratio R_{32} .

In further analysis, unfolded cross-section ratio R_{32} and its systematic uncertainties are calculated using Method I, whereas Method II is used only to propagate the statistical uncertainties including bin-by-bin correlations and statistical correlations between the inclusive 3-jet and 2-jet events ~~and~~ cross-sections. Unfolding takes the response matrix as an input which is explained in the next section.

5.5.1 Response Matrices

The response matrix is a two dimensional mapping between the true and measured distributions and is used to unfold the measured data spectrum. The response matrix can be constructed by two methods :

*probably defined much earlier,
but never used consistently*

Monte Carlo Method - In this method, the response matrix is usually derived using simulated Monte Carlo (MC) samples. The true $H_{T,2}/2$ obtained from MC simulations is taken as an input and is smeared by taking into account the detector resolution. This gives the measured $H_{T,2}/2$ distributions, required to construct the response matrix. But there are several drawbacks of constructing response matrix using this method. In some phase space regions, the shape of the distribution is not well described by the LO predictions. Also, the limited number of events in the MC samples at high transverse momenta introduces high statistical fluctuations in the response matrix.

Toy Monte Carlo Method - To overcome the short comings of the above method, there is an indirect way of constructing the response matrix which uses a custom Toy Monte Carlo method. In this method, the particle level or true $H_{T,2}/2$ spectrum is obtained by fitting the theoretically predicted NLO spectrum. Then this distribution is smeared with forward smearing technique, using the extracted jet energy resolution (JER) to obtain the reconstructed level or measured $H_{T,2}/2$ spectrum. After that, the response matrix constructed from these two distributions is used for the unfolding procedure.

We have constructed the response matrices using the indirect method, separately for cross-sections and cross-section ratio which are explained in the coming ~~next~~ sections.

5.5.1.1 Inclusive Cross-sections

The NLO spectrum of the differential cross-sections for $n_j \geq 2$ and $n_j \geq 3$ events obtained using CT10-NLO PDF set are fitted with the following two different functions defined in Eq. 5.10 and 5.13. These functions describe the shape as well as normalization of the distribution.

- **Function I :**

$$f(H_{T,2}/2) = N(x_T)^{-a} (1 - x_T)^b \times \underline{\exp}(-c/x_T) \quad (5.10)$$

where N is normalization factor and a, b, c are fit parameters. The function is derived from function given below [108] :

$$f(p_T; \alpha, \beta, \gamma) = N_0(p_T)^{-\alpha} \left(1 - \frac{2 p_T \cosh(y_{min})}{\sqrt{s}}\right)^\beta \times \exp(-\gamma/p_T) \quad (5.11)$$

using

$$\alpha = a, \quad \beta = b, \quad \gamma = c * \sqrt{s}/2, \quad x_T = \frac{2 * H_{T,2}/2 * \cosh(y_{min})}{\sqrt{s}} = \frac{2 * H_{T,2}/2}{\sqrt{s}}$$
(5.12)

where transverse scaling variable x_T corresponds to the proton fractional momentum x for dijets with rapidity $y = 0$, $\sqrt{s} = 8000$ GeV and y_{min} is low-edge of the rapidity bin y under consideration (here y_{min} is taken equal to 0)

- **Function II :**

$$f(H_{T,2}/2) = A_0 \left(1 - \frac{H_{T,2}/2}{A_6}\right)^{A_7} \times 10^{F(H_{T,2}/2)}, \text{ where } F(x) = \sum_{i=1}^5 A_i \left(\log\left(\frac{x}{A_6}\right)\right)^i$$
(5.13)

where the parameter A_6 is fixed to $\frac{\sqrt{s}}{2 \cosh(y_{min})}$, $\sqrt{s} = 8000$ GeV and y_{min} is the minimum rapidity. The other parameters are derived from the fitting.

Figure 5.13 shows the fitted CT10-NLO spectrum of differential cross-section as a function of $H_{T,2}/2$ (green solid circles) using Function I (top) and using Function II (bottom) : for inclusive 2-jet (left) and 3-jet (right) events. Function I is used primarily to generate response matrices and perform the closure tests and Function II is used as an alternative function to calculate unfolding uncertainty, described in Sec. 5.6.3. To include the migration of data points to lower bins, the fit functions described by red lines are extrapolated to 80 GeV (blue dashed lines).

A flat $H_{T,2}/2$ spectrum is generated by using toy Monte Carlo events and the fit parameters obtained from the NLO spectrum using function I (as shown in Fig. 5.13) provides weights to the flat spectrum. A total of ten million events are generated randomly (in $H_{T,2}/2$ range 80-2000 GeV). These generated values are then smeared with a Gaussian function, where σ of the Gaussian is determined from the relative

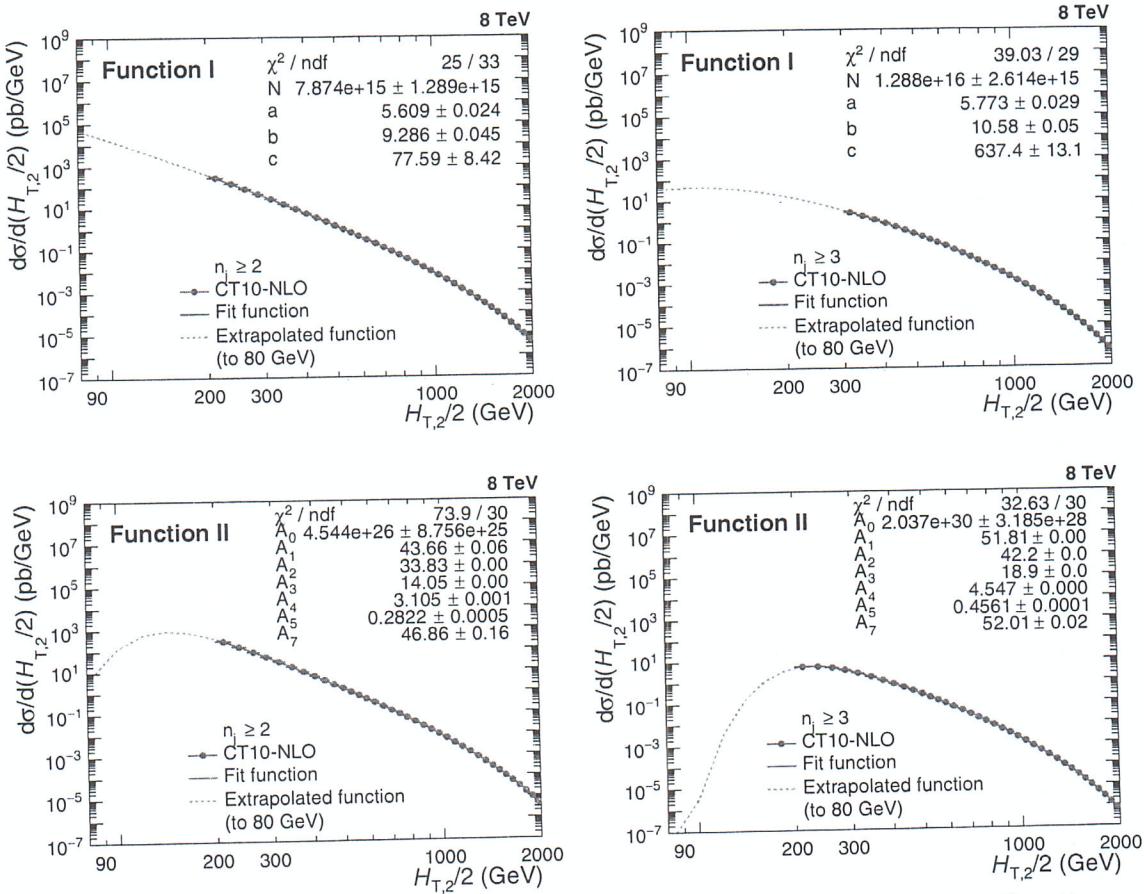


Figure 5.13: Fitted CT10-NLO spectrum of differential cross-section as a function of $H_{T,2}/2$ (green solid circles) using Function I (top) defined in Eq. 5.10 and using Function II (bottom) given by Eq. 5.13, for inclusive 2-jet (left) and 3-jet (right) events. To consider the migration to lower $H_{T,2}/2$ bins, the fit functions described by red lines are extrapolated to 80 GeV (blue dashed lines).

resolution parametrization as a function of $H_{T,2}/2$ calculated from NSC formula mentioned in equation 5.8. The parameters N, S, C used for smearing are taken from Table 5.6. These randomly generated (Gen_{Toy}) and smeared ($\text{Measured}_{\text{Toy}}$) values are used to fill the response matrices. Figure 5.14 shows the response matrices derived using the Toy MC for $n_j \geq 2$ (left) and $n_j \geq 3$ (right) events. The matrices are normalized to the number of events in each column. The response matrices are diagonal as the migrations in off-diagonal bins are much smaller than the bins along the diagonal.

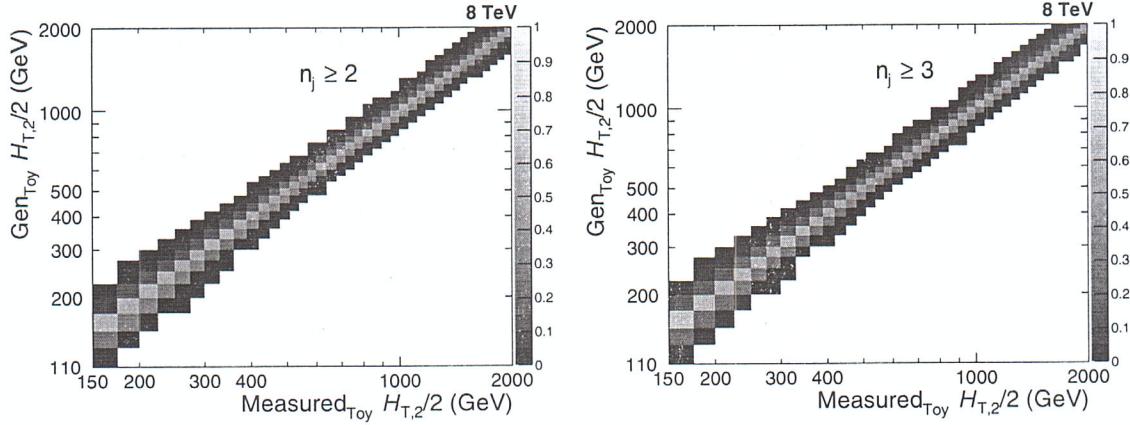


Figure 5.14: The response matrices are derived using the Toy Monte Carlo and forward smearing method, for inclusive 2-jet (left) and 3-jet (right) events. The matrices are normalized to the number of events in each column and are diagonal with small off-diagonal migrations between close-by $H_{T,2}/2$ bins.

5.5.1.2 Cross-section Ratio, R_{32}

To obtain the statistical uncertainty on the unfolded cross-section ratio R_{32} , Method II is used. In this method, the response matrix is constructed using Toy MC method as done in Sec. 5.5.1.1 for differential cross-sections. To obtain the true spectrum for R_{32} , the ratio of cross-section spectrum described by Eq. 5.10 for inclusive 3-jet to that of 2-jet events is taken. This ratio is shown by green solid circles in Fig. 5.15 (left) which is fitted using a polynomial function of degree 8 (red line). Then as explained in above section, response matrix is derived for R_{32} using the Toy Monte Carlo and forward smearing method which is shown in Fig. 5.15 (right). The matrix is normalized to the number of events in each column and is diagonal with small off-diagonal migrations between close-by $H_{T,2}/2$ bins.

5.5.2 Closure Test

A closure test has been performed to confirm the working of the unfolding procedure. In this test, $\text{Measured}_{\text{Toy}}$ spectrum is unfolded using the constructed response matrices shown in Figure 5.14. It is expected that the same Gen_{Toy} spectrum should be re-obtained after unfolding. Figure 5.16 confirms that the unfolded $\text{Measured}_{\text{Toy}}$

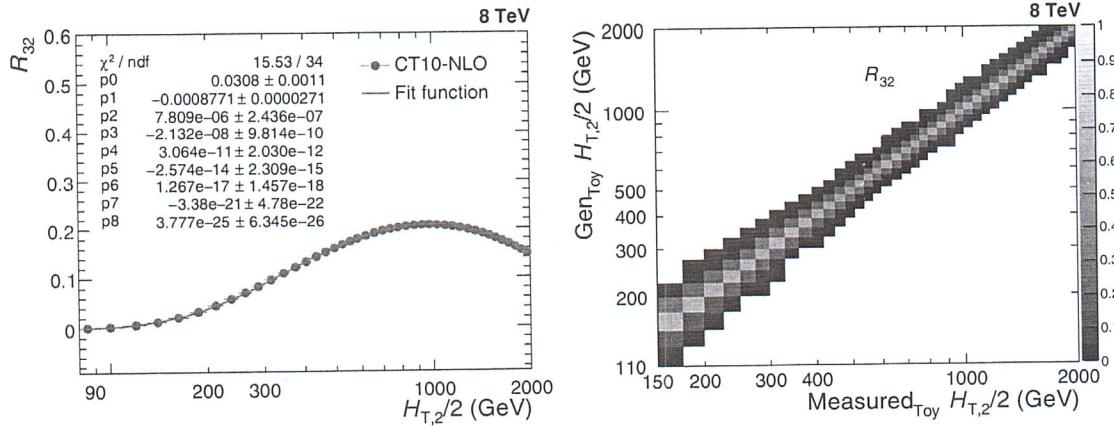


Figure 5.15: Left : The ratio of cross-sections described by Eq. 5.10 for inclusive 3-jet to that of 2-jet events is shown as a function of $H_{T,2}/2$ (green solid circles). It is fit using a polynomial function of degree 8 (red line). Right : The response matrix is derived using the Toy Monte Carlo and forward smearing method, for the cross-section ratio R_{32} . The matrix is normalized to the number of events in each column and is diagonal with small off-diagonal migrations between close-by $H_{T,2}/2$ bins.

spectrum matches exactly with Gen_{Toy} spectrum as the ratio of these distributions is perfectly flat at one for both $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events cross-sections as well as the cross-section ratio R_{32} (bottom).

For another closure test, Reco MG5+P6 MC differential cross-section distribution is unfolded using the above constructed response matrices using JER for forward smearing the randomly generated spectrum. While taking ratio of the unfolded distribution to that of Gen MG5+P6 MC, it is observed that a good closure is not obtained. This is represented by blue line in Fig. 5.17 for $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events. As observed in Fig. 5.12 in Sec. 5.4, if Reco MG5+P6 MC is unfolded using the response matrices obtained using 30% reduced JER, then the good closure is obtained as shown by red line in Fig. 5.17. Since unfolded cross-section ratio R_{32} is the ratio of unfolded differential cross-sections (Method I), same behaviour is observed for R_{32} (bottom).

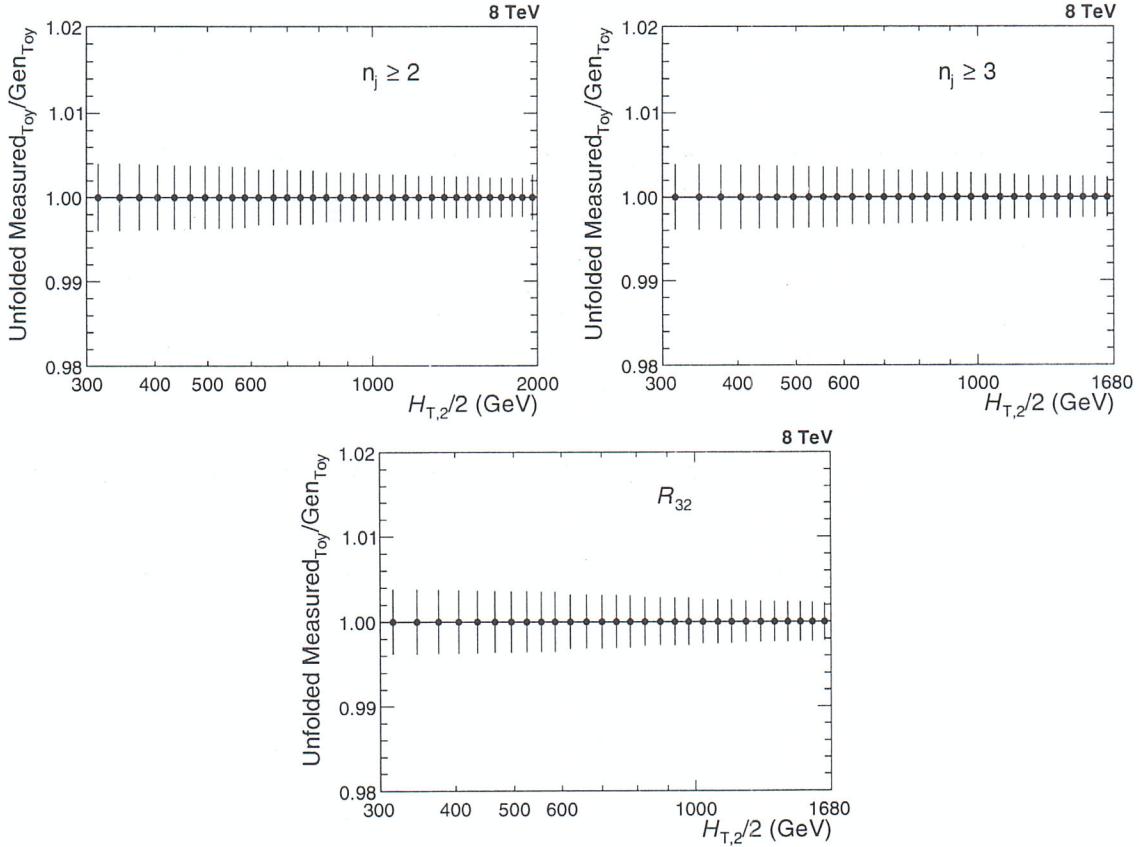


Figure 5.16: Closure test of the unfolding technique where the smeared spectrum obtained from Toy Monte Carlo method (Measured_{Toy}), is unfolded using the constructed response matrices (obtained by forward smearing the randomly generated spectrum (Gen_{Toy}) using extracted jet energy resolution (JER)). As expected, the unfolded measured_{Toy} spectrum matches exactly with Gen_{Toy} spectrum as the ratio of these distributions is perfectly flat at one for both inclusive 2-jet (top left) and 3-jet (top right) events cross-sections as well as the cross-section ratio R_{32} (bottom).

5.5.3 Unfolding of the Measurement

After validating the unfolding method, the measured differential cross-sections as well as R_{32} are unfolded using the above reconstructed response matrices. The unfolded data spectrum is compared to that of measured one in Fig. 5.18 for $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events cross-sections and for the cross-section ratio R_{32} (bottom). As already discussed that 30% reduced JER gives better closures than JER, so the unfolding of the data is done with response matrices using JER (blue solid circles) as well as 30% reduced JER (red solid circles) for smearing.

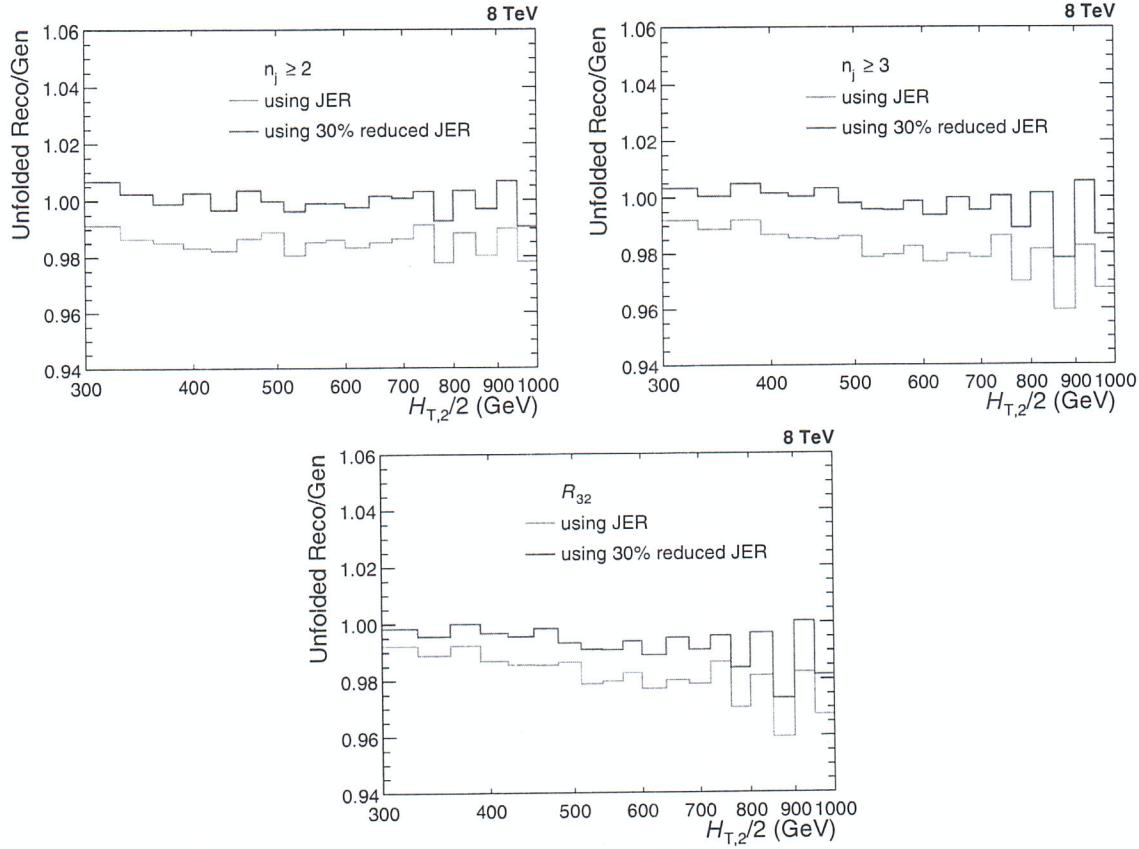


Figure 5.17: Reco MADGRAPH5+PYTHIA6 Monte Carlo (MG5+P6 MC) differential cross-section distributions unfolded with the response matrices (obtained by forward smearing the randomly generated spectrum (Gen) using extracted jet energy resolution (JER)), does not give a good closure with Gen MG5+P6 MC (blue line), for inclusive 2-jet (top left) and 3-jet (top right) events. After performing the unfolding using 30% reduced JER, a good closure is obtained (red line). Since unfolded the cross-section ratio R_{32} is the ratio of unfolded differential cross-sections, same behaviour is observed for R_{32} (bottom).

The difference between both is taken as an additional uncertainty on the unfolded measurement.

5.6 Experimental Uncertainties

In an experimental measurement of any physical observable, the uncertainties play a key role and hence are important to study in a physics analysis. The uncertainties can be categorized into two types : statistical and systematic. The statistical

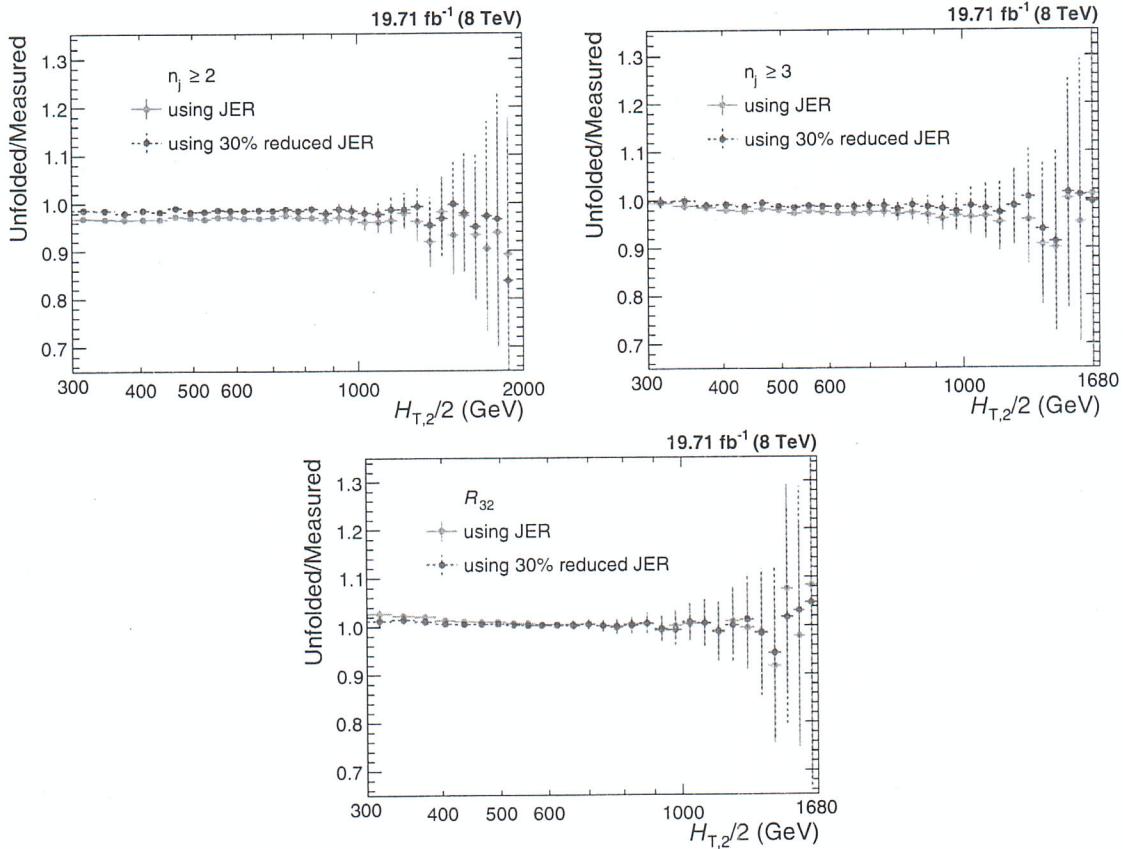


Figure 5.18: The measured differential cross-sections as well as the cross-section ratio R_{32} are unfolded as a function of $H_{T,2}/2$ using the response matrices derived using the Toy Monte Carlo and forward smearing method. The unfolded spectrum are compared with that of the measured one for inclusive 2-jet (top left) and 3-jet (top right) events cross-sections as well as for R_{32} (bottom). The unfolding is done with response matrices using JER (blue solid circles) as well as 30% reduced JER (red solid circles) for smearing. The difference between both is taken as an additional uncertainty on the unfolded measurement.

uncertainties arise due to random fluctuations depending on the number of events.
higher the smaller
The more the number of events, lesser is the statistical uncertainty. The systematic uncertainties may be due to known detector effects, model dependence, assumptions made or various corrections applied. In general, if the statistical and systematic uncertainties are uncorrelated, these can be added in quadrature to obtain the total uncertainty on the measurement. In this section, all the experimental uncertainties affecting the measurement of cross-sections and the cross-section ratio R_{32} are described. The systematic experimental uncertainties for R_{32} are propagated from the cross-sections to the ratio taking into account correlations. Due to this, the

systematic uncertainties may cancel for R_{32} completely or partially as compared to those for the individual cross-sections.

5.6.1 Statistical Uncertainty

Statistical uncertainty on the measurement is obtained through the unfolding procedure using a toy MC method. The measured data points are smeared within their statistical uncertainties to get the smeared spectrum. One million of such smeared spectra are produced and the unfolding is performed multiple times for each of the smeared spectra. The differences between the unfolded spectra and the measured one give the statistical uncertainty. The unfolding procedure introduces more statistical fluctuations which can be observed in Fig. 5.19. Here the fractional statistical uncertainties of the unfolded data (red line) are compared with those of the measured one (blue line) for $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events cross-sections as well as for the cross-section ratio R_{32} (bottom).

After the unfolding, the final statistical uncertainties become correlated among the bins such that the size of these correlations varies between 10 and 20%. The correlation (anti-) is more significant for neighbouring bins in $H_{T,2}/2$ as compared to the far off ones. In Fig. 5.20, the correlations of the statistical uncertainty after the unfolding can be seen for $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events cross-sections and for the cross-section ratio R_{32} (bottom). These correlations must be considered while performing the fits to extract the value of the strong coupling constant, α_s .

5.6.2 Jet Energy Corrections Uncertainty

As explained in Sec. 4.3.2, the measured jet energy is corrected for a variety of detector effects by using jet energy corrections (JEC). This procedure introduces uncertainties in the final corrected jet energy. There are 25 mutually independent sources which contribute to JEC. Each source presents a 1σ shift and is fully corre-

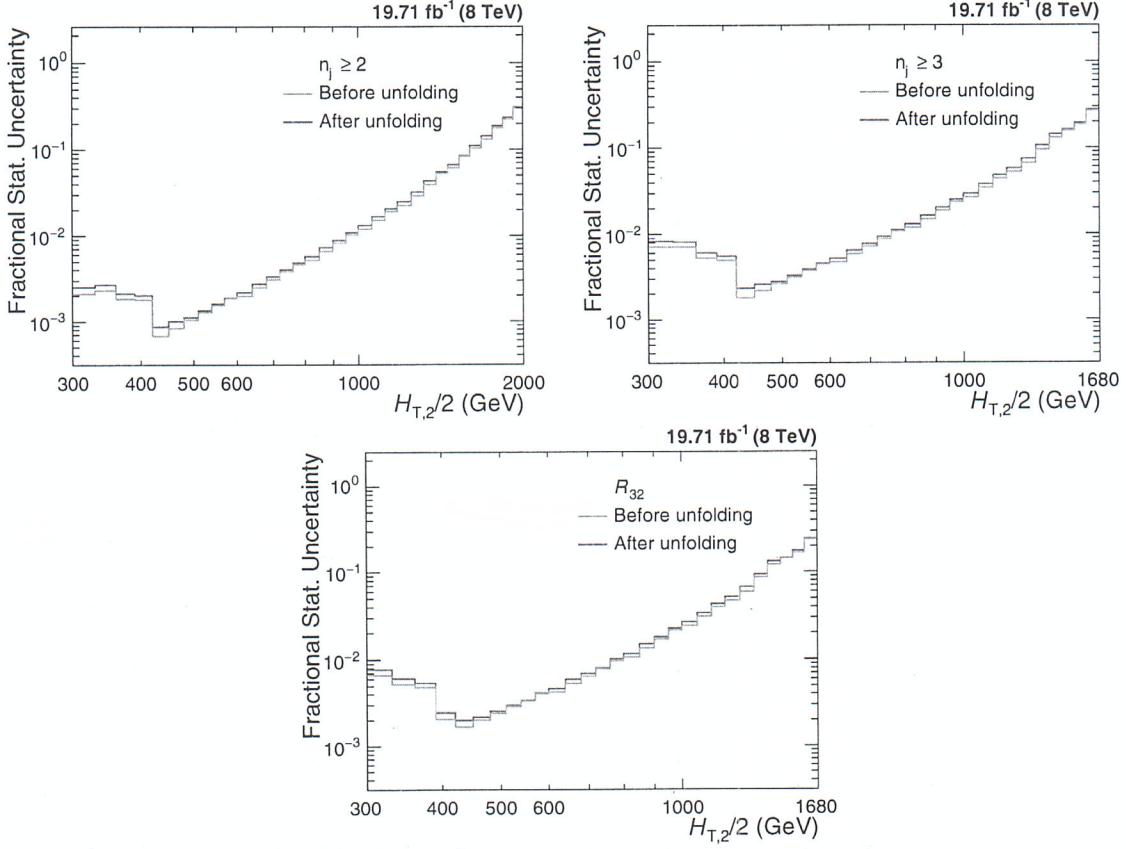


Figure 5.19: The fractional statistical uncertainties of the unfolded data (red line) are compared with those of the measured one (blue line) for inclusive 2-jet (top left) and 3-jet (top right) events cross-sections as well as for the cross-section ratio R_{32} (bottom). After unfolding, the statistical uncertainty increases slightly.

lated in p_T and η but uncorrelated to all other sources. The observable is studied with the nominal values of the jet energy which gives nominal distributions as well as by varying up and down the energy of all jets by an amount equivalent to the uncertainty. The differences between the nominal distributions and the ones obtained by varying the jet energy gives the uncertainties from each source. The Figs. A.1-A.3 show the JEC uncertainty from each source separately for inclusive 2-jet (top) and 3-jet (middle) events cross-sections and for cross-section R_{32} (bottom). The JEC uncertainties can be asymmetric in nature which leads to separate treatment of upwards and downwards variation of each source. The sum in quadrature of uncertainties from all sources gives the total JEC uncertainty. In the current anal-

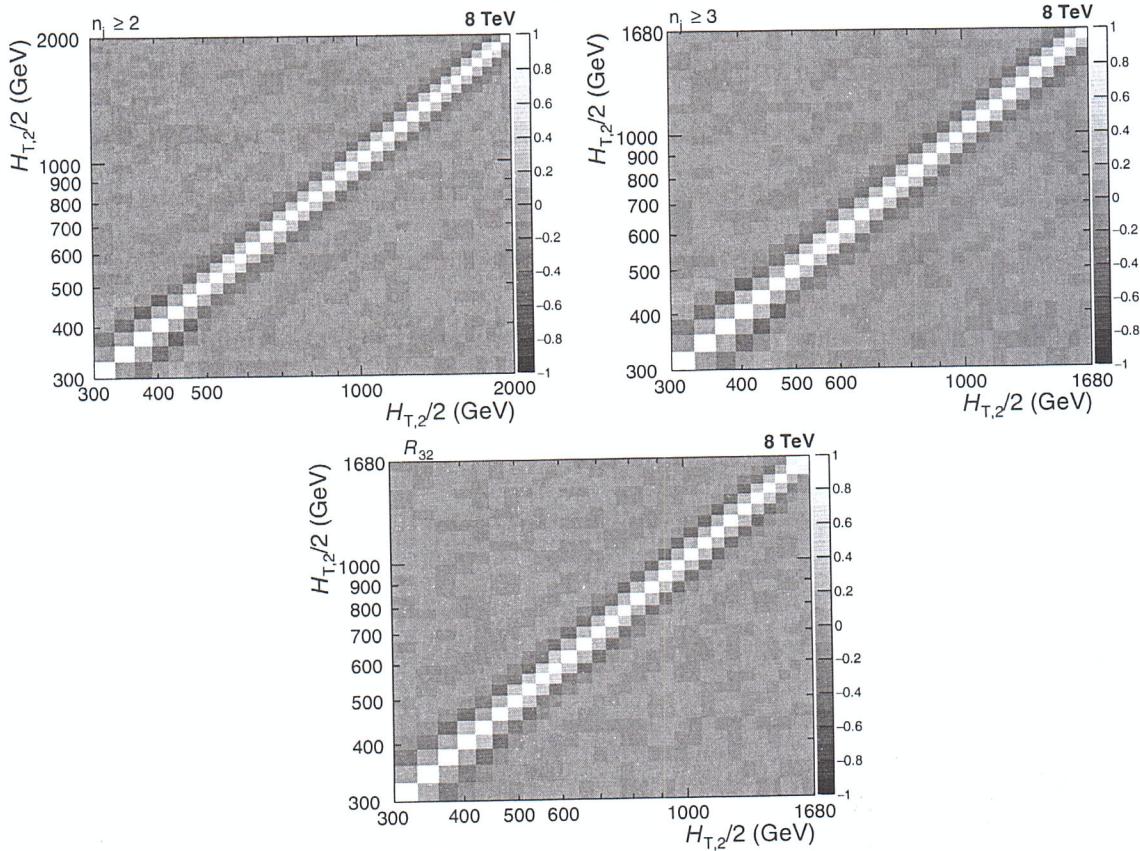


Figure 5.20: The unfolding procedure introduces the correlations of the statistical uncertainty through bin migrations which are shown here for inclusive 2-jet (top left) and ~~and~~ 3-jet (top right) events cross-sections as well as for the cross-section ratio R_{32} (bottom). The correlation (~~anti-~~) is more significant between neighbouring bins than far-ff ones.

ysis, JEC uncertainties are ~~the~~ a dominant source of experimental uncertainty at low $H_{T,2}/2$. The JEC uncertainty ranges from 3% to 10% for $n_j \geq 2$ and from 3% to 8% for $n_j \geq 3$ events cross-sections. To calculate JEC uncertainty for ratio R_{32} , the inclusive 2-jet and 3-jet events cross-sections are measured as a function of $H_{T,2}/2$ by shifting the jet p_T according to the JEC uncertainty for each source of JEC separately. Then the ratio of these cross-sections is taken and the difference of these from the central ratio R_{32} , gives the JEC uncertainty for R_{32} . As expected, JEC uncertainty for R_{32} is small as compared to that for individual cross-sections and is about 1 to 2% over all $H_{T,2}/2$ bins. Depending on the origin of sources of JEC, they are categorized into four groups which are described below in brief :

1. **Pileup** - This uncertainty originates from the differences in the transverse momentum between the true offset and the Random Cone method (i.e. essentially difference of pileup inside and outside of jets), in simulated events. This uncertainty is derived from Z/γ +jet, dijet and multijet data using fit procedure to estimate the residual pileup uncertainty after the calibration.
2. **Relative** - The forward jets are calibrated by the relative η -dependent corrections using dijet events. The main contribution to the uncertainty comes from jet energy resolution (JER), derived by varying JER scale factors up and down by quoted uncertainties and the initial and final state radiation bias corrections.
3. **Absolute** - A global fit to Z/γ +jet and multijet events gives the absolute calibration of the jet energy scale. The uncertainties are related to the lepton momentum scale for muons in $Z (\rightarrow \mu\mu)$ +jet and the single pion response in the HCAL.
4. **Flavor** - Flavor response differences are studied from simulation by cross-checking the results with quark- and gluon-tagged γ +jet and Z +jet events. These uncertainties are based on PYTHIA6.4 and HERWIG++2.3 differences propagated through the data-based calibration method.

More details of the jet energy corrections and uncertainties can be found in [109].

5.6.3 Unfolding Uncertainty

The unfolding uncertainty is comprised of three uncertainties which are explained as follows :

1. **Jet Energy Resolution** - The calculation of the jet energy resolution (JER) using simulated MG5+P6 Monte Carlo events is already explained in Sec. 5.4.

As mentioned before, the measured jet transverse momentum (p_T) in simulated MC events needs to be smeared additionally to match the resolution in the data. This smearing is done by using measured scale factors ($c_{central}$) mentioned in Table 5.5. It is recommended by JETMET group that the uncertainty on these measured scaling factors must be taken into account in a physics analysis. Since JER is used in constructing the response matrix which is an input in unfolding procedure, so the uncertainty on scale factors accounts for the unfolding uncertainty. To calculate JER uncertainty, p_T is smeared with two additional sets of scale factors corresponding to varying the factors up and down by one sigma, and corresponding $H_{T,2}/2$ is calculated. Then again JER is calculated as a function of $H_{T,2}/2$ using these upwards (c_{up}) and downwards (c_{down}) variations of the scaling factors. Alternative response matrices are built using the JER with above variations and the unfolding is performed again. The differences of the obtained unfolded spectra to the nominal ones accounts for a systematic JER uncertainty.

2. **Model Dependence** - It is explained in Sec. 5.5.1 that to obtain the true $H_{T,2}/2$ spectrum to be used in constructing response matrix using Toy MC method, the fitting of the CT10-NLO predictions is performed with the Function I described in Eq. 5.10. Using the alternative function, Function II given by Eq. 5.13, for this fitting and then constructing different response matrix, gives the model dependence of the true $H_{T,2}/2$ spectrum. The differences in unfolded distributions using the above mentioned two different response matrices give the model dependence uncertainty.
3. **Additional Uncertainty** - Small nonclosures observed in Fig. 5.12 introduces a supplementary uncertainty which is attributed by comparison of distributions unfolded using response matrices constructed using JER from simulation with that obtained with a 30% reduced JER.

All the three above mentioned uncertainties are added in quadrature to get the total unfolding uncertainty which increases from about 1% at low $H_{T,2}/2$ up to 2% at the high $H_{T,2}/2$ ends of the cross-sections for both $n_j \geq 2$ and $n_j \geq 3$ events. This uncertainty account for about less than 1% for R_{32} .

5.6.4 Luminosity Measurement Uncertainty

As discussed in Sec. 3.1.1, the luminosity delivered to CMS detector by LHC in the proton-proton collisions in the year of 2012 is measured by using the silicon pixel cluster counting method [60]. The uncertainty related to the integrated luminosity measurement is estimated to be 2.5% (syst.) and 0.5% (stat.). This uncertainty propagates directly to any absolute cross-section measurement. Hence, a total systematic uncertainty of 2.6% is considered across all the $H_{T,2}/2$ bins. At low $H_{T,2}/2$, it is similar in size as the one from JEC. This uncertainty cancels completely for R_{32} .

5.6.5 Residual Uncertainty

The small trigger and jet identification inefficiencies account for smaller than 1% uncertainties on the cross-section measurements [99, 110]. Hence, an uncorrelated residual uncertainty of 1% is assumed across all $H_{T,2}/2$ bins for both $n_j \geq 2$ and $n_j \geq 3$ events cross-sections whereas for R_{32} , it gets cancel completely.

5.6.6 Total Experimental Uncertainty

After calculating the uncertainties from all the above mentioned sources, the total experimental uncertainty on measurement of cross-sections as well as cross-section ratio R_{32} , is obtained by adding in quadrature the uncertainties from individual sources. Figure 5.21 shows the experimental uncertainties, from different sources

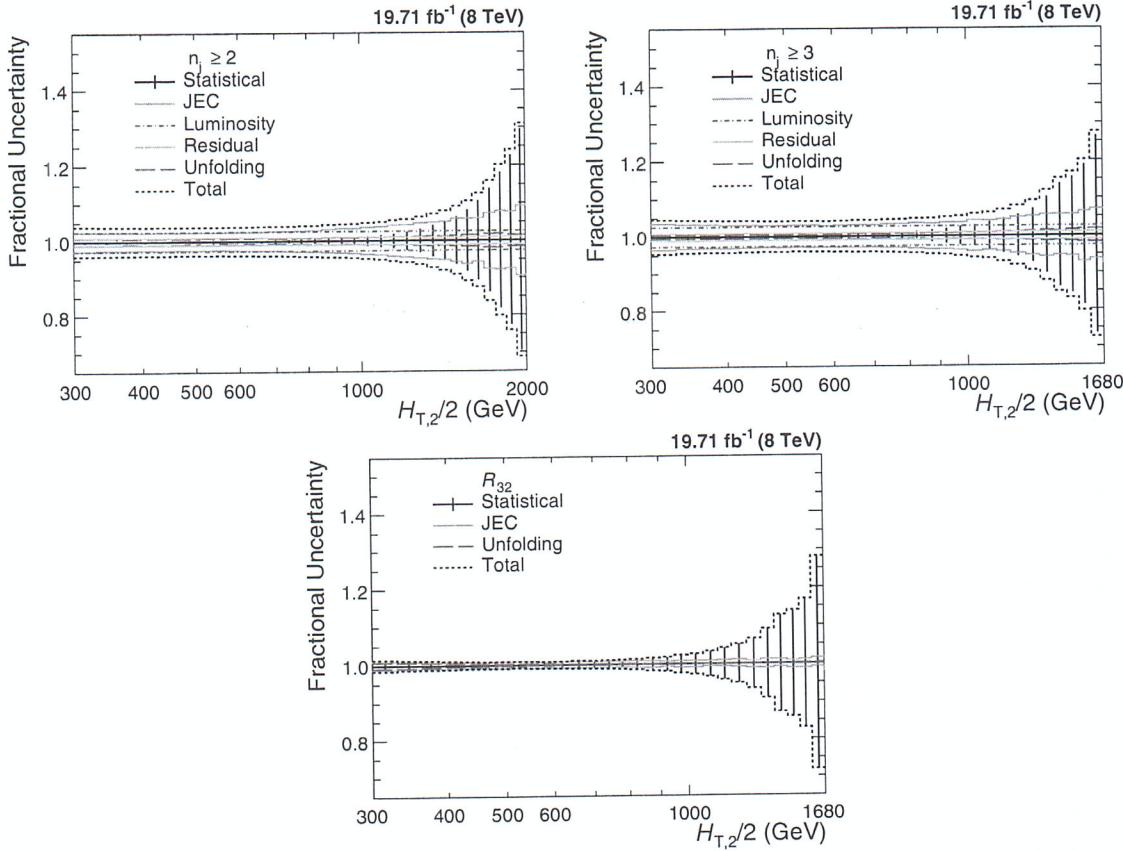


Figure 5.21: Experimental uncertainties from different sources affecting the measurement of cross-sections for inclusive 2-jet (top left) and 3-jet (top right) events and the cross-section ratio R_{32} (bottom). The error bars represent the statistical uncertainty after unfolding. The systematic uncertainties due to jet energy corrections (JEC by blue line), luminosity (red dashed line), unfolding (green dashed line) and residual effects (light purple line) are also presented. The uncertainties due to luminosity and residual effects cancel completely in R_{32} . The total uncertainty (black dashed line) is the quadrature sum of the individual sources of uncertainty.

as well as the total uncertainty, affecting the measurement of $n_j \geq 2$ (top left) and $n_j \geq 3$ (top right) events cross-sections and the cross-section ratio R_{32} (bottom). The error bars represent the statistical uncertainty obtained after unfolding. The systematic uncertainties due to jet energy corrections (JEC by blue line), luminosity (red dashed line), unfolding (green dashed line) and residual effects (light purple line) are also presented. The uncertainties due to luminosity and residual effects cancel completely in R_{32} . The total uncertainty (black dashed line) on the measurements is asymmetric in nature and dominated by the uncertainty due to the jet energy

corrections (JEC) at lower $H_{T,2}/2$ values and by statistical uncertainty at higher $H_{T,2}/2$ values.

The experimental uncertainties from each source as well as total uncertainty are also quoted in Table 5.7. The values of uncertainties (in %) from each source as well as total uncertainty, for each $H_{T,2}/2$ bin, are tabulated in Tables A.2, A.3 and A.4 for $n_j \geq 2$ and $n_j \geq 3$ events cross-sections and the cross-section ratio R_{32} , respectively.

Table 5.7: An overview of all experimental uncertainties affecting the measurement of cross-sections for inclusive 2-jet (left) and 3-jet (middle) events and the cross-section ratio R_{32} (right). The uncertainties due to luminosity and residual effects cancel completely in R_{32} . The total uncertainty is the quadrature sum of the individual sources of uncertainty.

Uncertainty Source	Inclusive 2-jet	Inclusive 3-jet	R_{32}
Statistical	< 1 to 30%	< 1 to 27%	< 1 to 28%
Jet energy corrections (JEC)	3 to 10%	3 to 8%	1 to 2%
Unfolding	1 to 2%	1 to 2%	< 1%
Luminosity	2.6%	2.6%	cancels
Residual	1%	1%	cancels
Total	4 to 32%	4 to 28%	1 to 28%

The complete data analysis of the differential inclusive 2-jet and 3-jet events cross-sections as well as their ratio R_{32} has been presented as a function of $H_{T,2}/2$. The measured spectra after correcting for detector effects through the unfolding procedure, are compared with the next-to-leading order (NLO) pQCD calculations in the next chapter.

Chapter 6

Theoretical Calculations

Measurements are not validated by theory.
It's the other way round!

The differential inclusive multijet event cross-sections measured as a function of $H_{T,2}/2$, described in the previous chapter, are validated by doing the comparison with the perturbative QCD (pQCD) theoretical calculations. The lowest order (LO) calculations roughly describe the measured cross-section distributions due to the dependence on the unphysical renormalization (μ_r) and factorization (μ_f) scales. The next-to-leading order (NLO) calculations improve the precision by reducing the dependence on μ_r and μ_f scales. This makes the NLO calculations an essential feature in the determination of fundamental parameters such as α_S and the parton distribution functions (PDFs). This chapter describes the NLO pQCD calculations used for comparison with the cross-section measurements in terms of $H_{T,2}/2$. The NLO pQCD calculations need to be corrected for the multi-parton interactions (MPI) and hadronization effects by applying non-perturbative (NP) corrections and also for the electroweak interactions (EW).

6.1 Fixed Order NLO Calculations

The NLO predictions for the differential inclusive jet event cross-sections in pQCD are computed with the NLOJET++ program version 4.1.3 [67, 68]. As explained

in Sec. 3.3.3, the interfacing of NLOJET++ program with FASTNLO [69, 70] framework is preferred over the direct calculation with NLOJET++ because with FASTNLO the calculations of the cross-sections can be repeated ~~several times with for~~ different PDFs and scale choices required for calculating the PDF and scale uncertainties. Here, FASTNLO version 2.3 framework has been used. The PDFs are accessed through the LHAPDF6 library [71, 72]. The factorization and renormalization scales are chosen equal to $H_{T,2}/2$, i.e. $\mu_f = \mu_r = H_{T,2}/2$.

In the current study, different PDF sets available for a series of different assumptions on the strong coupling constant, at the scale of the Z boson mass $\alpha_s(M_Z)$ are used for NLO calculations. Table 6.1 summarizes the existing PDF sets in LHC Run 1 (upper rows) and the newer PDF sets for Run 2 (lower rows). The different columns list the number of flavours N_f , the assumed masses M_t and M_Z of the top quark and the Z boson, respectively, the default values of $\alpha_s(M_Z)$, and the range in $\alpha_s(M_Z)$ variation available for fits with different PDF sets. All sets use a variable-flavour number scheme with at most five or six flavours apart from the ABM11 PDF, which employs a fixed-flavour number scheme with $N_F = 5$. Out of these eight PDF sets the following three are not considered further because of the below mentioned reasons :

- At NLO, predictions based on ABM11 do not describe LHC jet data at small jet rapidity [111–114].
- The HERAPDF2.0 set exclusively fits HERA DIS data with only weak constraints on the gluon PDF.
- The range of values available for $\alpha_s(M_Z)$ is too limited for the NNPDF3.0 set.

Mainly CT10 PDF set is considered for comparison between data and theory predictions as well as for calculating theoretical uncertainties.

Table 6.1: NLO PDF sets are available via LHAPDF6 with various assumptions on the value of $\alpha_s(M_Z)$. The upper rows list the existing sets in LHC Run 1 and newer ones for Run 2 are listed in lower rows, along with the corresponding number of flavours N_f , the assumed masses M_t and M_Z of the top quark and the Z boson, respectively, the default values of $\alpha_s(M_Z)$, and the range in $\alpha_s(M_Z)$ variation available for fits.

Base set	N_F	M_t (GeV)	M_Z (GeV)	$\alpha_s(M_Z)$	$\alpha_s(M_Z)$ range
ABM11 [30]	5	180	91.174	0.1180	0.110 - 0.130
CT10 [115]	≤ 5	172	91.188	0.1180	0.112 - 0.127
MSTW2008 [116, 117]	≤ 5	10^{10}	91.1876	0.1202	0.110 - 0.130
NNPDF2.3 [118]	≤ 6	175	91.1876	0.1180	0.114–0.124
CT14 [27]	≤ 5	172	91.1876	0.1180	0.111–0.123
HERAPDF2.0 [31]	≤ 5	173	91.1876	0.1180	0.110–0.130
MMHT2014 [28]	≤ 5	10^{10}	91.1876	0.1200	0.108–0.128
NNPDF3.0 [29]	≤ 5	173	91.2	0.1180	0.115–0.121

6.1.1 NLO Correction Factors

The differences between LO predictions and NLO predictions give the effect of the higher-order contributions to the pQCD predictions. These are described by an NLO correction factor, k-factor, which is derived as the ratio of cross-sections as :

$$\text{k-factor} = \frac{\sigma_{\text{NLO}}}{\sigma_{\text{LO}}} \quad (6.1)$$

The impact of the higher-order corrections is determined by the size of k-factor. The small size of k-factor indicates that the cross-section predictions are precisely described at the LO whereas the larger size hints the contributions from NLO. Figure 6.1 shows the k-factors of the NLOJET++ calculations, for inclusive 2-jet and 3-jet event cross-sections and their ratio R_{32} , using five different PDF sets. k-factor for R_{32} is obtained by taking the ratio of k-factors for inclusive 3-jet event cross-sections to that of inclusive 2-jet. The k-factors are similar for all the PDF sets in the lower region, but the differences increase in regions with larger $H_{T,2}/2$. It is observed that for inclusive 3-jet event cross-sections, k-factor jumps at the lowest $H_{T,2}/2$. This is because some jet configurations are kinematically forbidden near

the p_T cut bin i.e. 150 GeV. Since the first few bins in $H_{T,2}/2$ (below 225 GeV) still suffer from these kinematical ~~constraints~~ effects, the minimum value of $H_{T,2}/2$ studied is ~~from~~ 300 GeV.

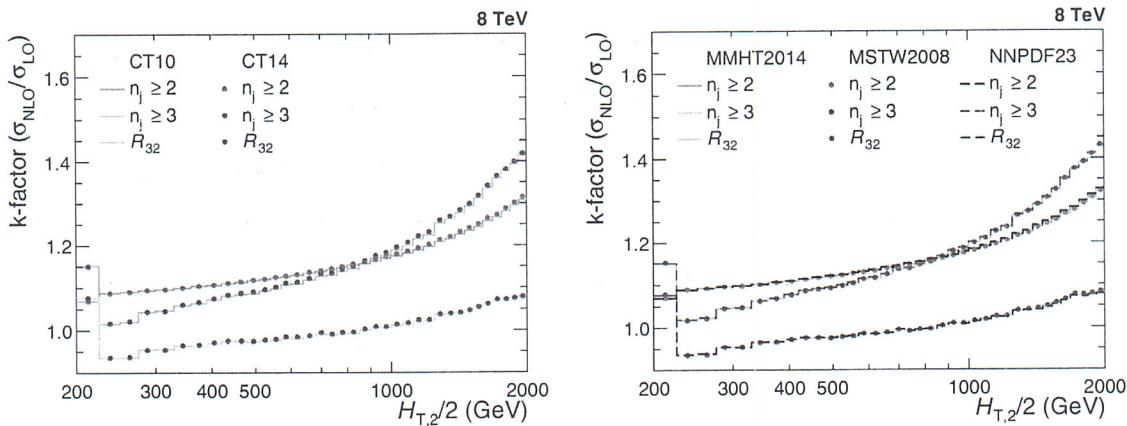


Figure 6.1: The k-factors of the NLOJET++ calculations, for inclusive 2-jet and 3-jet event cross-sections and their ratio R_{32} , using five different PDF sets.

6.1.2 Non-perturbative Corrections

The fixed-order pQCD NLO calculations predict the parton-level cross-section but lacks~~s~~ accuracy due to several effects. The partons which are emitted close to each other in phase space are not handled well in the lower order perturbation theories and hence require a parton shower (PS) correction. The scattering phenomena between partons within a colliding proton, other than the hard scattering, give rise to multi-parton interactions (MPI). The partons of the hard scattering form colorless bound states called hadrons through a process of hadronization (HAD). The MPI and hadronization cannot be modelled well within the perturbative framework. Since the fixed-order NLO calculations do not include these additional soft QCD effects, these calculations cannot be compared directly to the unfolded data. So the corrections for non-perturbative effects (NP) should be taken into account in NLO calculations. The ratio of cross-sections predicted with a nominal event generation, interfaced to the simulation of UE contributions and to the one without hadronization and MPI

effects, gives the NP correction factors which are defined as :

$$C^{\text{NP}} = \frac{\sigma^{\text{PS+HAD+MPI}}}{\sigma^{\text{PS}}} \quad (6.2)$$

In the current study, the NP effects are estimated by using samples obtained from various MC event generators with a simulation of parton shower and underlying-event (UE) contributions. The leading order (LO), HERWIG++ with the default tune of version 2.3 and PYTHIA6 with tune Z2*, and the NLO POWHEG MC event generators are considered. The matrix-element calculation is performed with POWHEG interfaced to PYTHIA8 with tune CUETS1 for the UE simulation. The ratio, defined in Eq. 6.2, is obtained for each MC generator and is fitted by a power-law function defined in Eq. 6.3. Since this ratio obtained from different MC generators have large differences, so the average of the envelope, which covers all the differences, is taken as the correction factor which is then applied as bin-by-bin multiplicative factor to the parton-level NLO cross-section. ~~The~~ ~~H~~ half of the envelope is taken as the uncertainty on the NP correction factor.

$$f(H_{\text{T},2}/2) = a \cdot (H_{\text{T},2}/2)^b + c \quad (6.3)$$

The NP correction factors, $C_{3\text{-jet}}^{\text{NP}}$ and $C_{2\text{-jet}}^{\text{NP}}$ are calculated for $n_j \geq 3$ and $n_j \geq 2$ event cross-sections, respectively and then their ratio gives the correction factor for R_{32} . The correction factors are shown in Fig. 6.2 for the inclusive 2-jet (top left) and 3-jet (top right) event cross-sections, and for the cross-section ratio R_{32} (bottom). At $H_{\text{T},2}/2 \sim 300$ GeV, the NP corrections amount to $\sim 4\text{-}5\%$ for inclusive 2-jet and 3-jet event cross-sections and $\sim 1\%$ for R_{32} , and decrease rapidly for increasing $H_{\text{T},2}/2$. On comparing the NP correction factors of R_{32} with that for individual cross-sections, it has been observed that the non-perturbative effects get reduced in R_{32} .

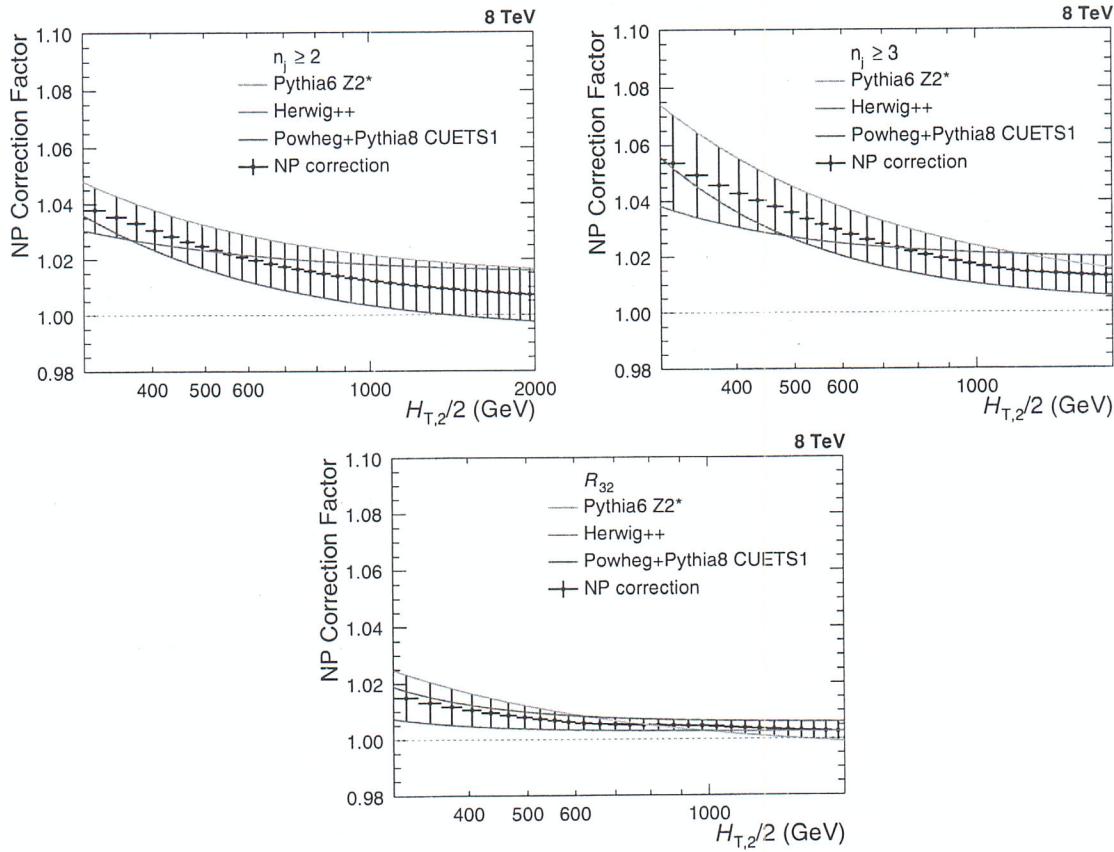


Figure 6.2: The nonperturbative (NP) corrections are presented as a function of $H_{T,2}/2$ for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections, as well as their ratio R_{32} . These corrections are calculated from the leading order HERWIG++ with the default tune of version 2.3 (red line) and PYTHIA6 with tune Z2* (blue line); and the next-to-leading order POWHEG interfaced to PYTHIA8 with tune CUETS1 (green line) Monte Carlo event generators. The black solid circles give the average NP correction factor along with the uncertainty shown by the error bars.

6.1.3 Electroweak Corrections

At LHC, the center-of-mass energy of proton-proton collisions is well beyond the electroweak (EW) scale $\sim \mathcal{O}(100 \text{ GeV})$. At such a high energy, the impact of higher order EW corrections is ~~not any more negligible~~ much more with respect to QCD effects [119] and affect the jet cross-sections at large $H_{T,2}/2$. The quark-quark scattering processes involving virtual exchanges of massive W and Z bosons contribute to electroweak (EW) corrections. The fixed-order QCD calculations do not include EW corrections and hence the NLO theory calculations are corrected for EW effects. The EW correc-

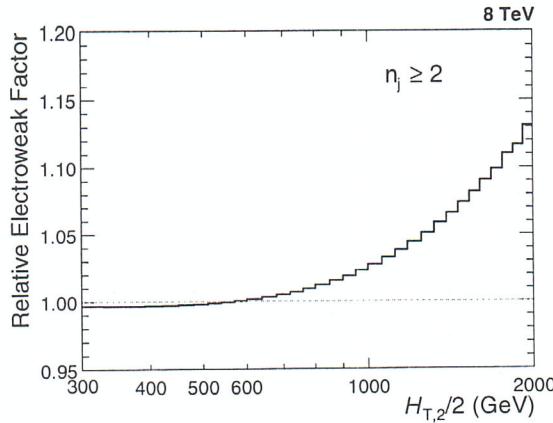


Figure 6.3: The electroweak (EW) corrections [120] in the phase space of the measurement are shown as a function of $H_{T,2}/2$ for inclusive 2-jet event cross-sections. These corrections are applied as a bin-by-bin correction factor to the fixed-order calculation of NLOJET++ as well as the MC predictions of MADGRAPH5+PYTHIA6. The EW correction factor increases up to 13% at high end of $H_{T,2}/2$ and significantly improves the agreement between data and prediction.

tions have been calculated for inclusive 1-jet and 2-jet case, in Ref. [120]. The EW correction factors in the phase space of the measurement are shown as a function of $H_{T,2}/2$ in Fig. 6.3 for inclusive 2-jet event cross-sections. These correction factor increases up to 13% at high end of $H_{T,2}/2$ which are applied as bin-by-bin correction factors to the fixed-order NLOJET++ calculations. To see the effects of EW corrections, a ratio of the data to theory predictions obtained using CT10-NLO PDF set and corrected with NP effects without including EW corrections (left) and including EW corrections (right) is plotted for inclusive 2-jet event cross-sections in Fig. 6.4. On comparing both ~~the~~ figures, it is observed that the EW corrections significantly improve the agreement between data and prediction in the high $H_{T,2}/2$ region. EW corrections are not available yet for inclusive 3-jet production and hence not applied for inclusive 3-jet event cross-sections. The guess from theory side is that EW for inclusive 2-jet and 3-jet will be similar, so for R_{32} , it is assumed to be equal to a factor of 1. Since the EW effects are not taken care of in MC simulations, these corrections are also applied to the MC predictions.

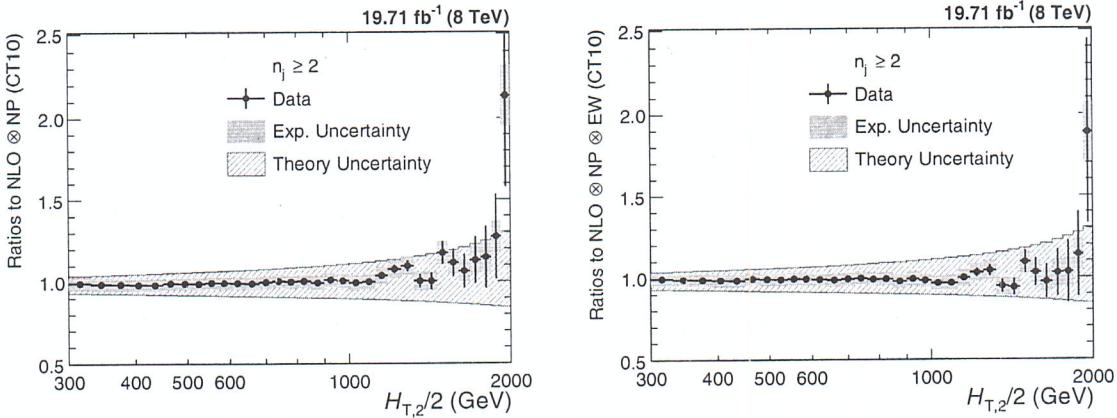


Figure 6.4: Ratio of the data over theory obtained using the CT10-NLO PDF set and corrected with non-perturbative effects (NP) without including electroweak (EW) corrections (left) and including EW corrections (right) is shown for inclusive 2-jet event cross-sections. The error bars represents the statistical uncertainty of the data and the shaded rectangles represents the total experimental systematic uncertainty. The shaded band around unity indicate the total uncertainty of the theory. The EW corrections significantly improve the agreement between data and prediction in the high $H_{T,2}/2$ region.

6.2 Theoretical Uncertainties

The measured differential inclusive multijet event cross-sections are not only sensitive to experimental uncertainties but also to the theoretical uncertainties. The renormalization and factorization scale variations, PDF uncertainties and the non-perturbative corrections contribute to theoretical uncertainties. These are described below :

6.2.1 Scale Uncertainty

In perturbative QCD calculations of cross-sections, one has to choose a renormalization (μ_r) and factorization (μ_f) scale. The dependence on scales is negligible if these calculations are performed for all orders of the perturbative series, given by Eq. 2.8. Since the NLO describes this series up to second power in α_S , it introduces a scale dependence of the measurement which is covered by systematic uncertainty known as scale uncertainty. The scale uncertainty is evaluated with the conventional recipe of varying the default scale $H_{T,2}/2$ chosen for μ_r and μ_f independently in the

following six combinations: $(\mu_r/H_{T,2}/2, \mu_f/H_{T,2}/2) = (1/2,1/2), (1/2,1), (1,1/2), (1,2), (2,1)$ and $(2,2)$. The maximal upwards and downwards deviations in cross-section from the central prediction, give the scale uncertainty. To calculate the scale uncertainty for cross-section ratio R_{32} , firstly R_{32} is obtained for each of the above mentioned scale choices and then its difference from central R_{32} is taken. The scale uncertainty calculated using CT10-NLO PDF set ranges from 5% to 13% and 11% to 17% for inclusive 2-jet and 3-jet events cross-sections respectively, and from 6% to 8% for R_{32} .

6.2.2 PDF Uncertainty

The calculation of jet cross-sections in proton-proton collisions relies upon the knowledge of PDFs. These PDF sets are obtained by global fits to all the available deep inelastic scattering (DIS) and related hard scattering data from different experiments. The various sources affect the PDFs such as the theory model, input parameters like the strong coupling constant α_S , the quark masses and the statistical and systematic uncertainty sources of the data included in the PDF fit. These sources contribute to PDF uncertainty which is evaluated according to the prescriptions given for each PDF set. The CT10-NLO PDF set [115, 121] employs the eigenvector method to evaluate the PDF uncertainties. The CT10-PDF set consists of $N_{\text{ev}} = 26$ eigenvectors with two PDF members per eigenvector k , which are varied upwards and downwards to generate a set of eigenvector pairs. The asymmetric uncertainties, ΔX^+ and ΔX^- , of a quantity X are given by Eq. 6.4 where X_0 is the central prediction, X_k^+ and X_k^- are the predictions using the upwards and downwards variation

of each eigenvector k .

$$\Delta X^+ = \sqrt{\sum_{k=1}^{N_{\text{ev}}} [\max(X_k^+ - X^0, X_k^- - X^0, 0)]^2} \quad (6.4)$$

$$\Delta X^- = \sqrt{\sum_{k=1}^{N_{\text{ev}}} [\min(X_k^+ - X^0, X_k^- - X^0, 0)]^2}$$

The symmetric uncertainty (ΔX^\pm) is given by half the difference of the upwards and downwards variations :

$$\Delta X^\pm = \sqrt{\sum_{k=1}^{N_{\text{ev}}} \left[\frac{X_k^+ - X_k^-}{2} \right]^2} \quad (6.5)$$

The CT10-NLO PDF set uncertainties are downscaled by a factor of 1.64 in order to have the uncertainties at the 68.3% confidence level $\text{CL}(1\sigma)$ instead of 90% $\text{CL}(2\sigma)$ such that to have a uniform treatment with respect to other PDF sets. The PDF uncertainty as derived with the CT10-NLO PDF set is the dominant source of uncertainty and ranges from 3% to 30% for inclusive 2-jet and from 4% to 32% for 3-jet cross-sections. For R_{32} , the ratio of predictions for inclusive 3-jet to that of 2-jet is taken for each eigen vector with upwards and downwards variations separately and then PDF uncertainty is calculated as done for individual cross-sections. The PDF uncertainty ranges and from 2% to 10% for cross-section ratio R_{32} .

6.2.3 Non-perturbative Uncertainty

As discussed in 6.1.2, the differences in the non-perturbative (NP) corrections calculated from various Monte Carlo event generators introduce the NP uncertainty which is of the order of 1% and 1 to 2% for inclusive 2-jet and 3-jet event cross-sections respectively, and < 1% for cross-section ratio R_{32} .

6.2.4 Total Theoretical Uncertainty

The total systematic theoretical uncertainties are obtained as the quadratic sum of the scale, PDF and NP uncertainties. Figure 6.5 presents the systematic theoretical uncertainties affecting the cross-section measurement for inclusive 2-jet (top left) and 3-jet (top right) events and the cross-section ratio R_{32} (bottom), using CT10-NLO PDF set. The scale (red dashed line), PDF (green line) and NP (blue dashed line) uncertainties as well as total theoretical uncertainty (black dashed line) are shown. The total theoretical uncertainty is asymmetric and dominated by PDF uncertainty which grows in magnitude with increasing value of $H_{T,2}/2$. Table 6.2

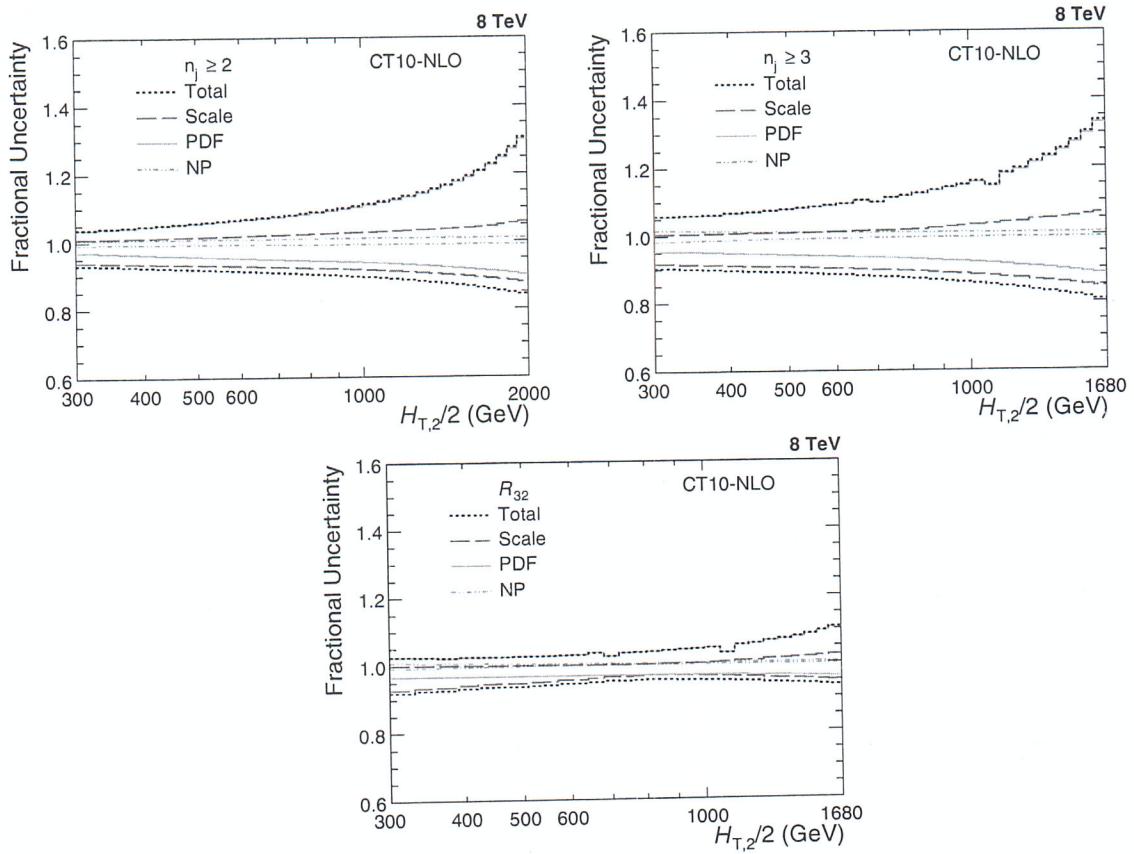


Figure 6.5: The systematic theoretical uncertainties affecting the cross-section measurement for inclusive 2-jet (top left) and 3-jet (top right) events and their ratio R_{32} (bottom). The scale (red dashed line), PDF (green line) and NP (blue dashed line) uncertainties as well as total uncertainty (black dashed line) obtained using line) uncertainties as well as total uncertainty (black dashed line) obtained using CT10-NLO PDF set are shown. The total theoretical uncertainty is asymmetric and dominated by PDF uncertainty.

quotes the values of the theoretical uncertainty from each source as well as total uncertainty affecting the measurements. The bin-wise values of uncertainties (in %) from each source as well as total uncertainty are shown in Tables A.5, A.6 and A.7 for $n_j \geq 2$ and $n_j \geq 3$ event cross-sections and the cross-section ratio R_{32} , respectively. The computation of the NLO predictions with NLOJET++ is also subject to statistical fluctuations from the complex numerical integrations. For the inclusive 2-jet event cross-sections this uncertainty is smaller than about a per mille, while for the inclusive 3-jet event cross-section it amounts to 1-9 per mille. Hence the statistical uncertainty is not considered in the total theoretical uncertainty. The small dips at ~ 700 and 1000 GeV in the PDF uncertainty for inclusive 3-jet events cross-sections and the cross-section ratio R_{32} is a feature of the CT10-NLO PDF set.

Table 6.2: Overview of all systematic theoretical uncertainties, obtained using CT10-NLO PDF set, affecting the measurement of cross-sections for inclusive 2-jet (left) and 3-jet (middle) events and the cross-section ratio R_{32} (right).

Uncertainty Source	Inclusive 2-jet	Inclusive 3-jet	R_{32}
Scale	5 to 13%	11 to 17%	6 to 8%
PDF	3 to 30%	4 to 32%	2 to 10%
Non-perturbative (NP)	1%	1 to 2%	< 1%
Total	3 to 30%	5 to 34%	3 to 11%

It is
considered
in the fit!
It is the only
uncorrelated
source from
theory!

6.3 Comparison of Theory to Data

After correcting the measurement for detector effects as well as NLO pQCD calculations for non-perturbative (NP) and electroweak (EW) effects, it is now feasible to compare the measured cross-sections with the theory predictions. Figure 6.6 shows the measured differential inclusive 2-jet and 3-jet event cross-sections as a function of $H_{T,2}/2$ after unfolding for detector effects. On the left, the measurements (points) are compared to the NLOJET++ predictions using the CT10-NLO PDF set (line),

corrected for NP effects and in addition for EW effects in the 2-jet case. On the right, the comparison is made to the predictions from MADGRAPH5+PYTHIA6 (MG+P6) with tune Z2* (line), corrected for EW effects in the 2-jet case. The error bars give the total experimental uncertainty, given by the quadrature sum of the statistical and systematic uncertainties. On a logarithmic scale, the data are in good agreement with the NLO predictions over the whole range of $H_{T,2}/2$ from 300 GeV up to 2000 (2-jet) and 1680 GeV (3-jet) respectively.

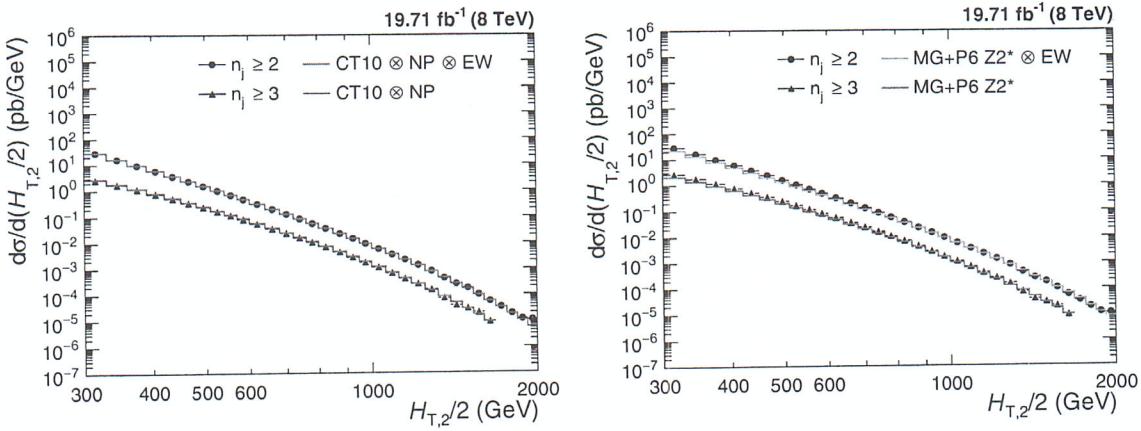


Figure 6.6: Comparison of the measured differential inclusive 2-jet and 3-jet event cross-sections as a function of $H_{T,2}/2$ to theoretical predictions. On the left, the data (points) are shown together with NLOJET++ predictions (line) using the CT10-NLO PDF set, corrected for non-perturbative (NP) and electroweak (EW) effects (2-jet) or only NP effects (3-jet). On the (right), the data (points) are compared to predictions from MADGRAPH5+PYTHIA6 (MG+P6) with tune Z2* (line), corrected for EW effects in the 2-jet case. The error bars give the total experimental uncertainty, given by the quadrature sum of the statistical and systematic uncertainties.

Figure 6.7 shows the cross-section ratio R_{32} obtained from unfolded data (solid circles) in comparison to that from NLO pQCD predictions obtained using the CT10-NLO PDF set corrected with NP corrections (line). The error bars here represent the total experimental uncertainty derived as quadratic sum from all uncertainty sources. The deviations of measured R_{32} from the predicted theoretical value can be explained by the electroweak effects which are not considered yet because of their unavailability for inclusive 3-jet event cross-sections.

For better comparisons, the ratios of the data over the theory at NLO are also shown from CT10!

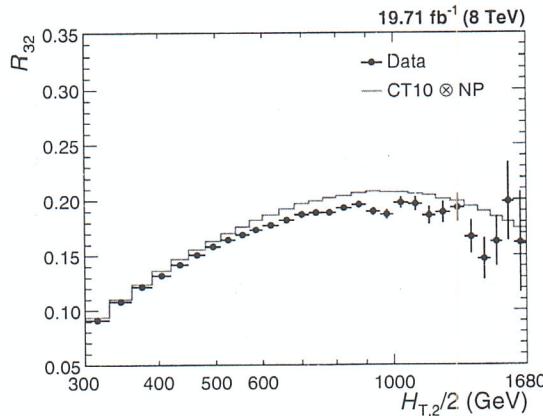


Figure 6.7: Cross-section ratio R_{32} as a function of $H_{T,2}/2$ calculated from data (solid circles) in comparison to that from NLO pQCD predictions obtained using the CT10-NLO PDF set corrected with non-perturbative (NP) corrections (line). The error bars correspond to the total experimental uncertainty derived as quadratic sum from all uncertainty sources.

studied in detail. In Fig. 6.8, the ratios of the data over NLO JET++ predictions using the CT10-NLO PDF set are shown for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections as well as their ratio R_{32} (bottom). The data are well described by the predictions within their uncertainties, which are dominated at large $H_{T,2}/2$ by PDF effects in the upwards and by scale variations in the downwards direction. A trend towards an increasing systematic excess of the 2-jet data with respect to theory, starting at about 1 TeV in $H_{T,2}/2$, is remedied by the inclusion of EWK corrections. In the 3-jet case the statistical precision of the data and the reach in $H_{T,2}/2$ is insufficient to observe any effect. The alternative PDF sets MSTW2008 and NNPDF2.3 exhibit a small underestimation of the cross-sections at high $H_{T,2}/2$.

The POWHEG framework providing the NLO dijet calculation matched to the parton showers of PYTHIA8 employed with the CUETS1 and CUETM1 tunes [78] is also used for a comparison. The ratios of the data w.r.t. theory from POWHEG+PYTHIA8 with tune CUETS1 are shown for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections as well as their ratio R_{32} (bottom) in Fig. 6.9. For comparison, the LO prediction from PYTHIA6 with tune Z2*, the tree-level multi-leg improved prediction by MADGRAPH5+PYTHIA6 with tune Z2*, and the matched

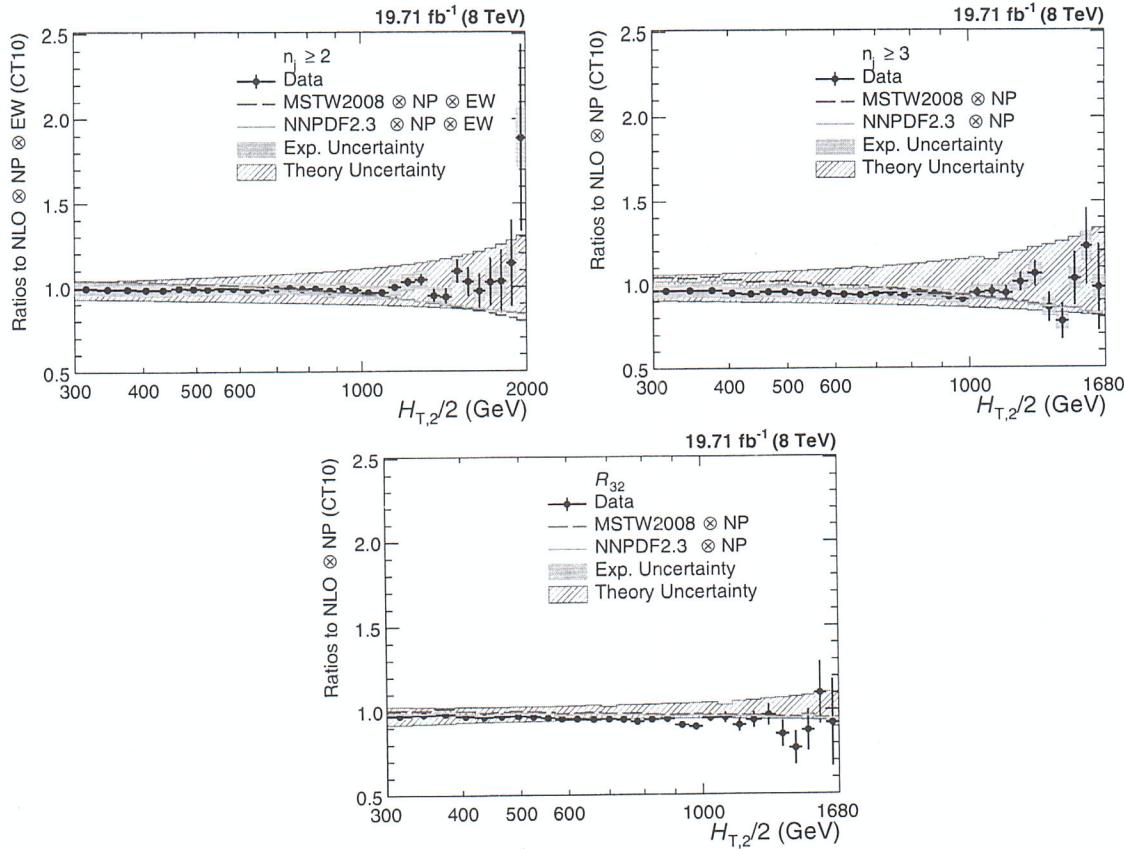


Figure 6.8: Ratio of the data over theory using the CT10-NLO PDF set for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections and their ratio R_{32} (bottom). The theory predictions are corrected for non-perturbative effects (NP) and also for electroweak effects (EW) for inclusive 2-jet only. For comparison predictions employing two other PDF sets, MSTW2008 and NNPDF2.3, are also shown. The error bars represent the statistical uncertainty of the data and the shaded rectangles represents the total experimental systematic uncertainty. The shaded band around unity indicate the total uncertainty of the theory.

NLO prediction from POWHEG+PYTHIA8 with tune CUETM1 are shown as well. EW corrections have been accounted for in this comparison for the 2-jet case only. Significant discrepancies, which are cancelled to a large extent in the ratio R_{32} , are visible in the comparison with the LO prediction from MADGRAPH5+PYTHIA6 with tune Z2*, in particular for small $H_{T,2}/2$. In contrast, the employed dijet MC POWHEG+PYTHIA8 better describe the 2-jet event cross-section, but fail for the 3-jet case.

The jet measurements at hadron colliders can be used to extract the strong coupling constant α_S , which is discussed in the next chapter.

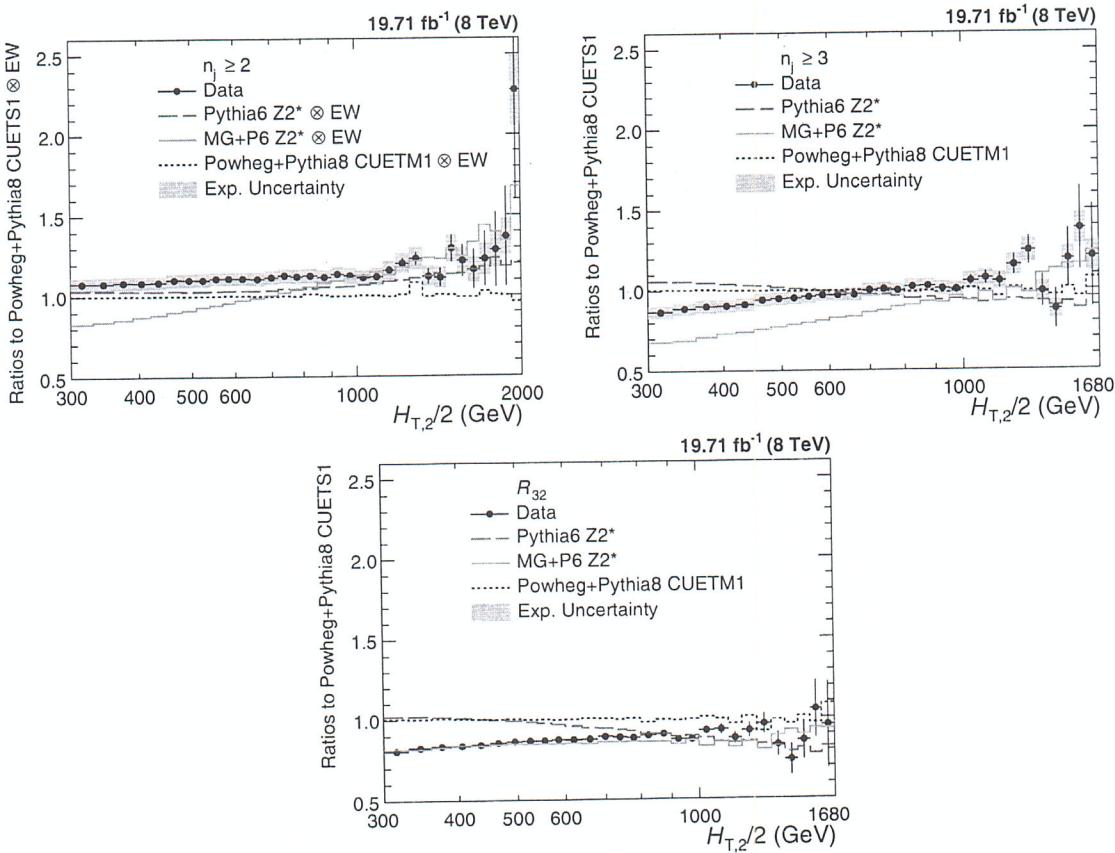


Figure 6.9: Ratio of the data over the predictions from POWHEG+PYTHIA8 with tune CUETS1 are presented for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections as well as their ratio R_{32} (bottom). For comparison the alternative tune CUETM1 of POWHEG+PYTHIA8, the tree-level multi-leg improved prediction by MADGRAPH5+PYTHIA6 with tune $Z2^*$, and the LO MC predictions from PYTHIA6 tune $Z2^*$ are shown as well. The error bars correspond to the statistical uncertainty of the data and the shaded rectangles to the total experimental systematic uncertainty. EW corrections have been accounted for in this comparison for the 2-jet case only.

Chapter 7

Determination of the Strong Coupling Constant

The inclusive jet production cross-section at hadron colliders mainly depends on the strong coupling constant α_S for a given center-of-mass energy. Hence the measurements of the inclusive jet cross-section and jet properties provide a direct probe to measure the strong coupling constant. The measurement of α_S has been already done by various experiments such as CMS [1, 99, 113, 122, 123], ATLAS [124], D0 [125, 126], H1 [127, 128], and ZEUS [129]. In this thesis, the value of the strong coupling constant at the scale of the Z boson mass $\alpha_s(M_Z)$, is extracted using the measurements of differential inclusive 2-jet and 3-jet event cross-sections as well as the cross-section ratio R_{32} , as a function of $H_{\text{T},2}/2$. The differential inclusive jet production cross-section up to NLO is given by [130] :

$$\frac{d\sigma}{d(H_{\text{T},2}/2)} = \alpha_S^2(\mu_r) \hat{X}^{(0)}(\mu_f, H_{\text{T},2}/2) [1 + \alpha_S(\mu_r) K1(\mu_r, \mu_f, H_{\text{T},2}/2)] \quad (7.1)$$

where $\frac{d\sigma}{d(H_{\text{T},2}/2)}$ is the differential inclusive jet production cross-section as a function of $H_{\text{T},2}/2$, μ_r and μ_f are the renormalization and factorization scales set equal to $H_{\text{T},2}/2$, $\alpha_S^2(\mu_r) \hat{X}^{(0)}(\mu_f, H_{\text{T},2}/2)$ is the leading order

(LO) contribution to the differential inclusive jet production cross-section and $\alpha_S^3(\mu_r)\hat{X}^{(0)}(\mu_f, H_{T,2}/2)K1(\mu_r, \mu_f, H_{T,2}/2)$ is the next-to-leading order (NLO) contribution. Equation 7.1 shows how the inclusive jet production cross-section varies with $\alpha_s(\mu_r)$.

7.1 Sensitivity of Measurements to $\alpha_s(M_Z)$

For a fixed choice of μ_r and μ_f , different input values of $\alpha_s(M_Z)$ to a PDF set will lead to different theory predictions for the differential cross-section distribution. This will give an estimate of the sensitivity of the theory predictions to the varying input value of $\alpha_s(M_Z)$. A comparison of the measured spectrum with the theory predictions obtained using all $\alpha_s(M_Z)$ inputs will give a hint of the input value of $\alpha_s(M_Z)$ for which the theory distribution has the closest matching with data. In this section, the sensitivity of the differential inclusive jet event cross-sections and the cross-section ratio, R_{32} to varying input values of $\alpha_s(M_Z)$ for different PDF sets is demonstrated by plotting the ratios of the data over theory predictions with central value of $\alpha_s(M_Z)$.

Figures 7.1, 7.2 and 7.3 present the ratio of the data to the theory predictions, corrected for NP effects, for all variations in $\alpha_s(M_Z)$ available for the PDF sets CT10, CT14, MSTW2008, MMHT2014 and NNPDF2.3 at NLO evolution order as specified in Table 6.1, for inclusive 2-jet event cross-sections, inclusive 3-jet events cross-sections and ratio R_{32} respectively. The $\alpha_s(M_Z)$ value is varied in the range 0.112-0.127, 0.111-123, 0.110-0.130, 0.108-0.128 and 0.114-0.124 in steps of 0.001 for the CT10, CT14, MSTW2008, MMHT2014 and NNPDF2.3PDF sets, respectively. The error bars correspond to the total experimental uncertainty derived as quadratic sum from all uncertainty sources. The theory predictions are also corrected for EW effects for inclusive 2-jet events cross-section. A small slope increasing with $H_{T,2}/2$ is visible for most PDFs in both cross-sections. This effect is largely cancelled in

the cross-section ratio. R_{32} exhibits a flat behaviour with respect to the predictions for all five PDF sets in the whole range of $H_{T,2}/2$ up to 1680 GeV. Therefore, these data can be used to determine the strong coupling constant, although only up to 1 TeV for the cross-sections as long as electroweak corrections are not taken into account.

Moreover, in Figs. 7.1–7.3, the different sensitivity to $\alpha_s(M_Z)$ caused by the leading power in α_S in the expansion of the 2-jet inclusive ($\propto \alpha_S^2$) and the 3-jet inclusive ($\propto \alpha_S^3$) cross-sections, and their ratio ($\propto \alpha_S^1$), is clearly visible from the spread between the calculations for the smallest and largest value of $\alpha_s(M_Z)$ within the same PDF set. This also demonstrates the potential of cross-section ratios R_{mn} with $(m-n) > 1$.

7.2 Determination of $\alpha_s(M_Z)$

As discussed in the previous section, the measured inclusive 2-jet and 3-jet event cross-sections and their ratio R_{32} can be used for a determination of the strong coupling constant $\alpha_s(M_Z)$. To extract the value of $\alpha_s(M_Z)$, a general fit procedure [1, 99, 123] has been followed as described in the following section.

7.2.1 Fitting Procedure

The value of $\alpha_s(M_Z)$ is determined by minimizing the chi-square (χ^2) between the experimental measurements and the theoretical predictions. The χ^2 is given by the following equation :

$$\chi^2 = M^T C^{-1} M \quad (7.2)$$

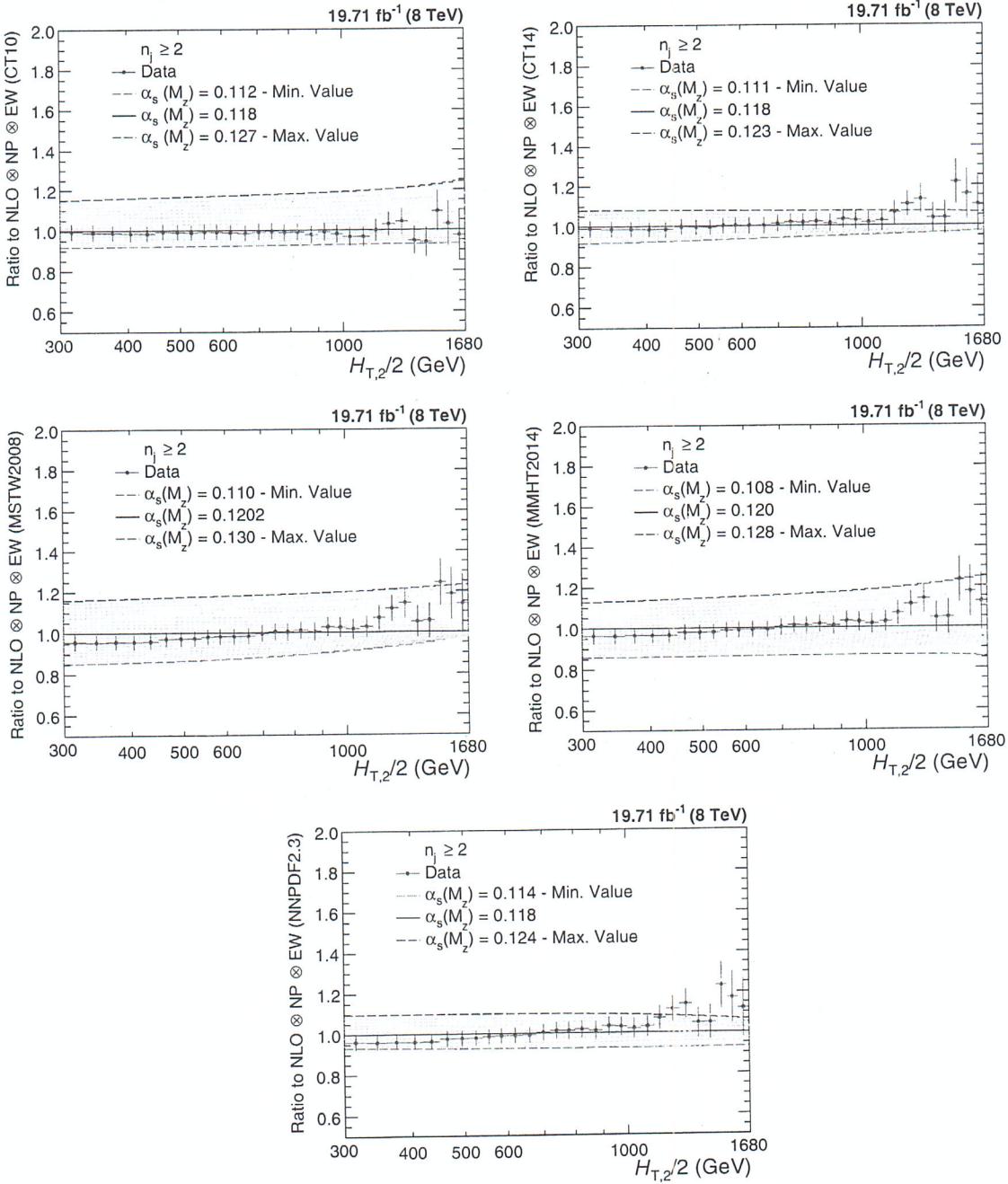


Figure 7.1: Ratio of the measured inclusive 2-jet differential cross-section to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for a series of values of $\alpha_s(M_Z)$. The $\alpha_s(M_Z)$ value is varied in the range 0.112-0.127, 0.111-0.123, 0.110-0.130, 0.108-0.128 and 0.114-0.124 in steps of 0.001 for the CT10, CT14, MSTW2008, MMHT2014 and NNPDF2.3 NLO PDF sets, respectively. The theory predictions are corrected for non-perturbative (NP) and electroweak (EW) effects. The error bars correspond to the total experimental uncertainty.

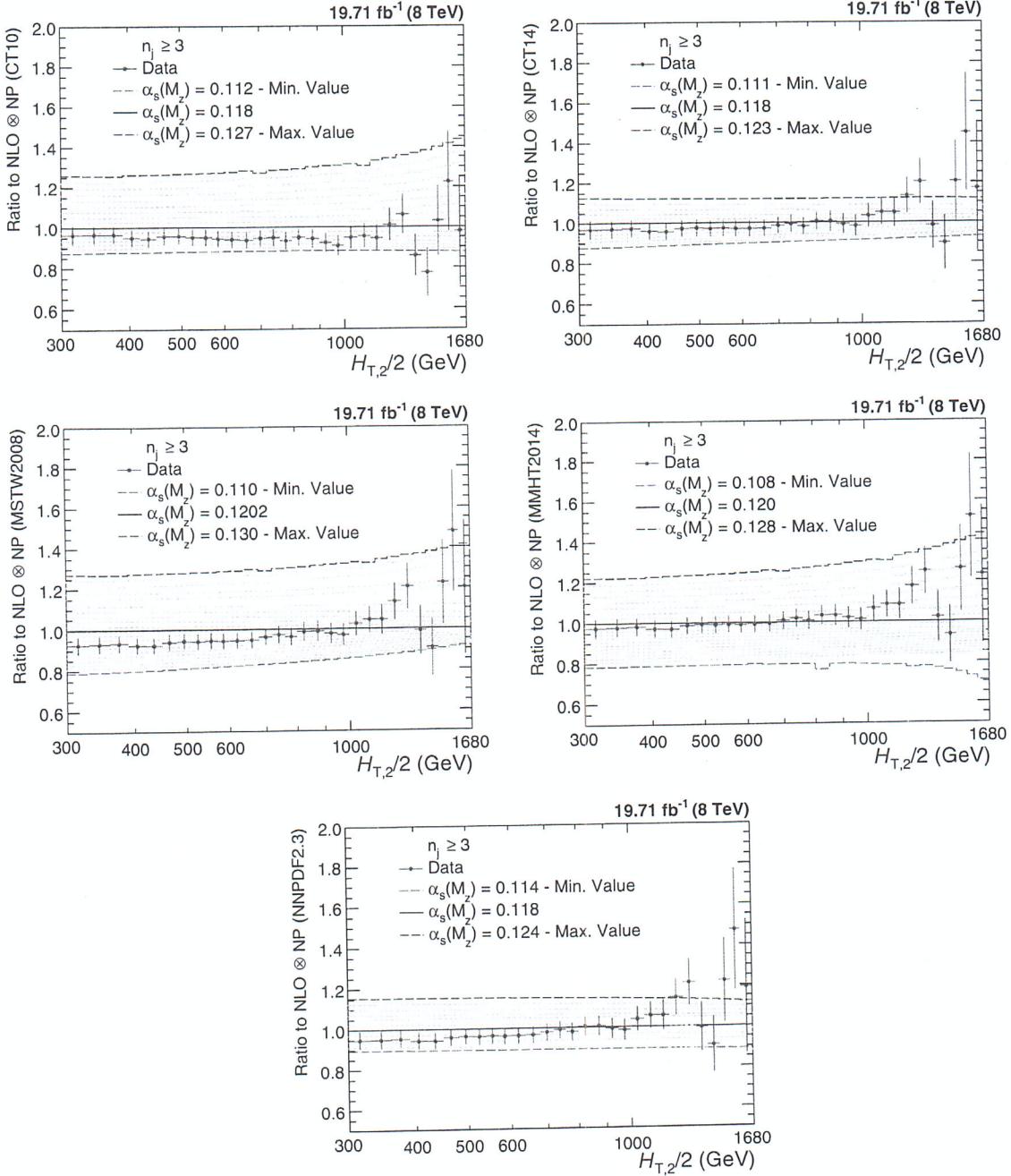


Figure 7.2: Ratio of the measured inclusive 3-jet differential cross-section to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for $\alpha_s(M_Z)$. The $\alpha_s(M_Z)$ value is varied in the range 0.112-0.127, a series of values of $\alpha_s(M_Z)$ are used: 0.112-0.127 for the CT10, 0.111-0.123, 0.110-0.130, 0.108-0.128 and 0.114-0.124 in steps of 0.001 for the CT14, MSTW2008, MMHT2014 and NNPDF2.3 NLO PDF sets, respectively. The theory predictions error bars correspond to the total experimental uncertainty. The theory predictions are corrected for non-perturbative (NP) effects.

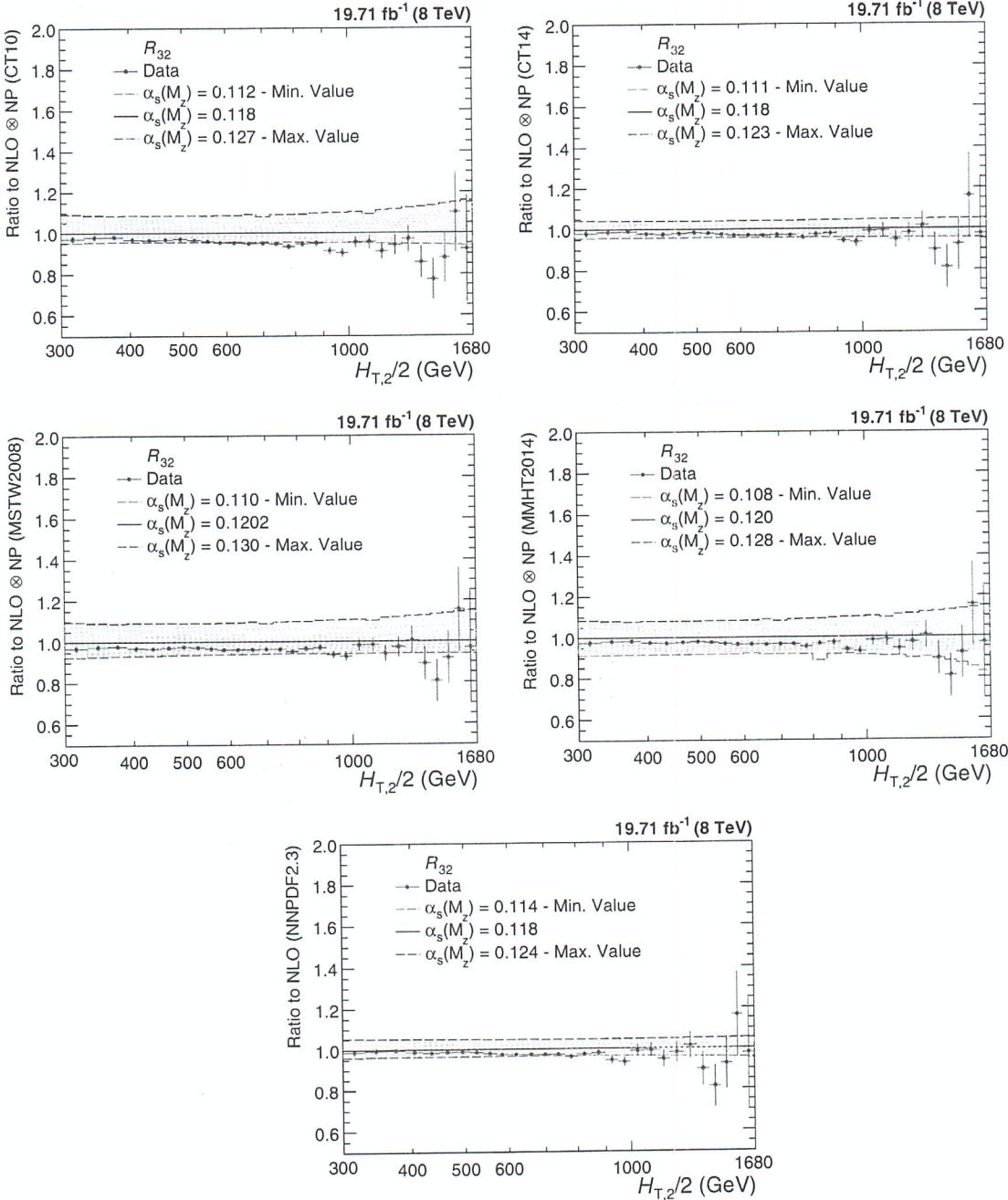


Figure 7.3: Ratio of the measured cross-section ratio, R_{32} to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for a series of values of $\alpha_s(M_Z)$. The $\alpha_s(M_Z)$ value is varied in the range 0.112-0.127, 0.111-0.123, 0.110-0.130, 0.108-0.128 and 0.114-0.124 in steps of 0.001 for the CT10, CT14, MSTW2008, MMHT2014 and NNPDF2.3 NLO PDF sets, respectively. The error bars correspond to the total experimental uncertainty. The theory predictions are corrected for non-perturbative (NP) effects.

where M is the vector of the differences between the data (D^i) and the theoretical values (T^i) in each bin i ,

$$M^i = D^i - T^i \quad (7.3)$$

and C is the covariance matrix including all experimental uncertainties as described in Sec. 5.6 and some theoretical uncertainties discussed in Sec. 6.2. The covariance matrix $C = C_{\text{exp}} + C_{\text{theo}}$ is defined as the sum of covariances of experimental and theoretical sources of uncertainty as follows :

$$C_{\text{exp}} = \text{Cov}^{\text{ExpStat}} + \sum \text{Cov}^{\text{JEC}} + \text{Cov}^{\text{Unfolding}} + \text{Cov}^{\text{Lumi}} + \text{Cov}^{\text{Residual}} \quad (7.4)$$

$$C_{\text{theo}} = \text{Cov}^{\text{TheoStat}} + \text{Cov}^{\text{NP}} + \text{Cov}^{\text{PDF}} \quad (7.5)$$

where the labelled covariance matrices account for the following effects:

- $\text{Cov}^{\text{ExpStat}}$: statistical uncertainty of the data including correlations introduced by the unfolding.
- Cov^{JEC} : the jet energy corrections (JEC) systematic uncertainty.
- $\text{Cov}^{\text{Unfolding}}$: the unfolding systematic uncertainty including the resolution (JER) and model dependence.
- Cov^{Lumi} : the luminosity uncertainty.
- $\text{Cov}^{\text{Residual}}$: a residual uncorrelated systematic uncertainty summarizing individual causes such as small trigger and identification inefficiencies, time dependence of the jet p_T resolution, and uncertainty on the trigger prescale factors.
- $\text{Cov}^{\text{TheoStat}}$: statistical uncertainty caused by numerical integrations in the cross-section computations.

- Cov^{NP} : the systematic uncertainty of the non-perturbative (NP) corrections.
- Cov^{PDF} : the PDF uncertainties.

While taking the differences between theory and data, the treatment of experimental and theoretical systematic uncertainties is crucial. The Unfolding, JEC, Lumi, and PDF and NP systematic uncertainties are treated as 100% correlated among $H_{T,2}/2$ bins. If δ_i is the total uncertainty on the differential cross-section, for the i -th $H_{T,2}/2$ bin, for any of these fully correlated sources, then the (i,j) -th element of the corresponding covariance matrix is given by $\text{COV}_{ij} = \delta_i \times \delta_j$. The JEC, unfolding, and luminosity uncertainties are treated as multiplicative to avoid the statistical bias that arises when estimating uncertainties from data. In fits of the ratio R_{32} , the luminosity and residual uncorrelated uncertainties cancel completely. Partial cancellations between the other sources of uncertainty are taken into account in the fit.

The evaluation of PDF uncertainty depends on the individual PDF set as already discussed in Sec. 6.2.2. The PDF covariance matrix construction varies among different PDF sets. The CT10, CT14, MMHT2014 and MSTW2008 NLO PDF sets employ the eigenvector method to evaluate the PDF uncertainties as explained in Sec. 6.2.2. The number of eigenvectors (N_{ev}) with two PDF members per eigenvector for CT10, CT14, MMHT2014 and MSTW2008 NLO PDF sets are 26, 28, 25 and 20, respectively. The NNPDF2.3 PDF set comes with hundred different replicas (N_{rep}) instead of different eigenvectors, as for CT10 or CT14 PDF sets. The mean uncertainty is calculated as average uncertainty from 100 different replicas. Following the prescription given in [131], the PDF uncertainty is calculated as :

$$(\Delta X)^2 = \frac{1}{N_{\text{rep}} - 1} \sum_{k=1}^{N_{\text{rep}}} [X_k - \langle X_k \rangle]^2 \quad (7.6)$$

where ΔX is the uncertainty on predicted differential cross-section, X_k is the differ-

ential cross-section for k -th replica and $\langle X_k \rangle$ is the average differential cross-section from all the replicas.

Scale uncertainties of the pQCD predictions are taken into account by employing the offset method, i.e. by performing separate fits with varying scale factors as described in the Sec. 6.2.1. The largest upwards and downwards deviations from the default factors are defined as the uncertainty. At NLO such scale variations predominantly lead to smaller cross-sections and also a smaller ratio R_{32} as visible in Fig. 6.5. As a consequence the scale uncertainty in fits is equally asymmetric, where smaller cross-sections or ratios are compensated by an increase in the fitted value for $\alpha_s(M_Z)$.

7.2.2 Fit Results

To determine the value of the strong coupling constant at the scale of the Z boson mass $\alpha_s(M_Z)$, fits to the differential inclusive 2-jet and 3-jet events cross-sections are performed using five different NLO PDF sets : CT10, CT14, MSTW2008, MMHT2014 and NNPDF2.3. The range in $H_{T,2}/2$ is restricted to be between 300 GeV and 1 TeV to avoid the region close to the minimal p_T threshold of 150 GeV for each jet at low p_T and the onset of electroweak effects at high $H_{T,2}/2$, which are available for the dijet case only. The $\alpha_s(M_Z)$ results obtained from a simultaneous fit to all 19 $H_{T,2}/2$ bins in the above mentioned range are reported in Table 7.1. For comparison, a simultaneous fit to both cross-sections ignoring any correlations, and a fit to the cross-section ratio R_{32} , fully accounting for correlations is also performed and the results are tabulated in Table 7.2. The electroweak effects are assumed to cancel in the ratio as do the luminosity and the uncorrelated uncertainty.

All cross-section fits give compatible values for $\alpha_s(M_Z)$ in the range of 0.115–0.118 whereas for the ratio R_{32} somewhat smaller values are obtained. But for individual cross-sections, χ^2/n_{dof} values are small as compared to the cross-section

ratio R_{32} . A possible explanation is an overestimation of the residual uncorrelated uncertainty of 1% that is cancelled for R_{32} . If the fits are repeated with an assumed uncertainty of 0.25% instead, the χ^2/n_{dof} values lie around unity while the $\alpha_s(M_Z)$ values are still compatible with the previous results but with slightly reduced uncertainties.

Table 7.1: Determination of $\alpha_s(M_Z)$ from the inclusive 2-jet and 3-jet event cross-sections using five PDF sets at NLO. Only total uncertainties without scale variations are quoted. The results are obtained from a simultaneous fit to all 19 $H_{\text{T},2}/2$ bins in the restricted range of $300 < H_{\text{T},2}/2 < 1000$ GeV.

PDF set	Inclusive 2-jets			Inclusive 3-jets		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}
CT10	0.1174	0.0032	3.0/18	0.1169	0.0027	5.4/18
CT14	0.1160	0.0035	3.5/18	0.1159	0.0031	6.1/18
MSTW2008	0.1159	0.0025	5.3/18	0.1161	0.0021	6.7/18
MMHT2014	0.1165	0.0034	5.9/18	0.1166	0.0025	7.1/18
NNPDF2.3	0.1183	0.0025	9.7/18	0.1179	0.0021	9.1/18

Table 7.2: Determination of $\alpha_s(M_Z)$ from the inclusive 2-jet and 3-jet event cross-sections simultaneously and from their ratio R_{32} using five PDF sets at NLO. Only total uncertainties without scale variations are quoted. The results are obtained from a simultaneous fit to all 38 (left) and 19 (right) $H_{\text{T},2}/2$ bins in the restricted range of $300 < H_{\text{T},2}/2 < 1000$ GeV. For comparison, correlations between the two cross-sections are neglected in the simultaneous fit on the left, but fully taken into account in the ratio fit on the right.

PDF set	Inclusive 2- and 3-jets			R_{32}		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}
CT10	0.1170	0.0026	8.2/37	0.1141	0.0028	19./18
CT14	0.1161	0.0029	9.1/37	0.1139	0.0032	15./18
MSTW2008	0.1161	0.0021	11./37	0.1150	0.0023	21./18
MMHT2014	0.1168	0.0025	11./37	0.1142	0.0022	19./18
NNPDF2.3	0.1188	0.0019	15./37	0.1184	0.0021	12./18

To investigate how the electroweak (EW) corrections affect the fit results for $\alpha_s(M_Z)$, the range in $H_{\text{T},2}/2$ is extended to $300 < H_{\text{T},2}/2 < 1680$ GeV. $\alpha_s(M_Z)$ values are obtained from fits to the inclusive 2-jet event cross-section in this range with or without EW correction factors and the results are presented in Table 7.3. The largest impact is a reduction in χ^2/n_{dof} , which indicates a better agreement

when EW effects are included. In addition, a tendency to slightly smaller $\alpha_s(M_Z)$ values is observed without the EW corrections. For the ratio R_{32} , it is expected that these effects are much reduced.

Table 7.3: Determination of $\alpha_s(M_Z)$ from the inclusive 2-jet event cross-section using five PDF sets at NLO without (left) and with (right) electroweak (EW) corrections. Only total uncertainties without scale variations are quoted. The results are obtained from a simultaneous fit to all 29 $H_{T,2}/2$ bins in the range of $300 < H_{T,2}/2 < 1680$ GeV.

PDF set	without EW			with EW		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	χ^2/n_{dof}
CT10	0.1163	0.0034	15./28	0.1165	0.0032	14./28
CT14	0.1137	0.0033	24./28	0.1144	0.0033	17./28
MSTW2008	0.1093	0.0028	27./28	0.1133	0.0023	19./28
MMHT2014	0.1127	0.0032	32./28	0.1141	0.0032	21./28
NNPDF2.3	0.1162	0.0024	31./28	0.1168	0.0024	23./28

From Fig. 7.3 follows that only the PDF sets MSTW2008 and MMHT2014 provide a large enough range in $\alpha_s(M_Z)$ values to ensure fits without extrapolation. The other three PDF sets are at the limit such that reliable fits cannot be performed for all scale settings and/or bins in scale $Q = H_{T,2}/2$. Since many systematic uncertainties cancel completely or partially in the cross-section ratio R_{32} as compared to the individual cross-sections, R_{32} is used mainly to determine the value of $\alpha_s(M_Z)$. Table 7.4 gives the complete results for MSTW2008 and MMHT2014 for the full range in $H_{T,2}/2$ of 300 GeV up to 1.68 TeV along with the corresponding components of PDF, scale, NP and total experimental except scale uncertainties. In contrast to fits at NLO using cross-sections where the scale uncertainty recipe usually leads to a very asymmetric behaviour with larger downward uncertainties in the case, this is inverted for the fits to the cross-section ratio R_{32} . The scale uncertainty is the most dominant source of total uncertainty on $\alpha_s(M_Z)$. These values are determined with the central renormalization and factorization scales i.e. $\mu_r = \mu_f = H_{T,2}/2$. The values are also determined for the six scale factor combinations for the two PDF sets MSTW2008 and MMHT2014 and the results are shown in Table 7.5.

The uncertainty decomposition for $\alpha_s(M_Z)$ determined from cross-section ratio R_{32} is performed in four sub-ranges of $H_{T,2}/2$ and the results are shown in Table 7.6. The statistical uncertainty of the NLO computation is negligible in comparison to any of the other sources of uncertainty. Electroweak corrections, significant only at high $H_{T,2}/2$, are assumed to cancel between the numerator and denominator.

Finally, the values of the strong coupling constant at the scale of mass of Z boson $\alpha_s(M_Z)$, determined from the ratio R_{32} using the two most compatible PDF sets, are :

- Using the MSTW2008 PDF set, which dates from before the LHC start :

$$\begin{aligned} \alpha_s(M_Z) &= 0.1150 \pm 0.0010 \text{ (exp)} \pm 0.0013 \text{ (PDF)} \pm 0.0015 \text{ (NP)} {}^{+0.0050}_{-0.0000} \text{ (scale)} \\ &= 0.1150 \pm 0.0023 \text{ (all except scale)} {}^{+0.0050}_{-0.0000} \text{ (scale)} \end{aligned} \quad (7.7)$$

- Using the MMHT2014 PDF set, which uses the LHC jet data to determine the PDF parameters :

$$\begin{aligned} \alpha_s(M_Z) &= 0.1142 \pm 0.0010 \text{ (exp)} \pm 0.0013 \text{ (PDF)} \pm 0.0014 \text{ (NP)} {}^{+0.0049}_{-0.0006} \text{ (scale)} \\ &= 0.1142 \pm 0.0022 \text{ (all except scale)} {}^{+0.0049}_{-0.0006} \text{ (scale)} \end{aligned} \quad (7.8)$$

7.3 Running of the Strong Coupling Constant

The value of the strong coupling constant α_S depends on the energy scale Q and it decreases with the increase of scale Q . To study this dependence, the determination of α_S is carried out at different energies. The procedure to extract $\alpha_S(Q)$ is same as the one followed for the $\alpha_s(M_Z)$. To have different energy scales, the fitted $H_{T,2}/2$

Table 7.4: Determination of $\alpha_s(M_Z)$ from the ratio R_{32} using the two most compatible PDF sets MSTW2008 and MMHT2014 at NLO along with the corresponding components of PDF, scale, NP and total (except scale) experimental uncertainties. The results are obtained from a simultaneous fit to all 29 $H_{T,2}/2$ bins in the full range of $300 < H_{T,2}/2 < 1680$ GeV.

PDF set	$\alpha_s(M_Z)$	exp	PDF	NP	all exc.	scale	χ^2/n_{dof}
MSTW2008	0.1150	± 0.0010	± 0.0013	± 0.0015	± 0.0023	$^{+0.0050}_{-0.0000}$	26./28
MMHT2014	0.1142	± 0.0010	± 0.0013	± 0.0014	± 0.0022	$^{+0.0049}_{-0.0006}$	24./28

Table 7.5: Determination of $\alpha_s(M_Z)$ from the ratio R_{32} in the $H_{T,2}/2$ range from 300 up to 1680 GeV at the central scale and for the six scale factor combinations for the two PDF sets MSTW2008 and MMHT2014.

$\mu_r/H_{T,2}/2$	$\mu_f/H_{T,2}/2$	MSTW2008		MMHT2014	
		$\alpha_s(M_Z)$	χ^2/n_{dof}	$\alpha_s(M_Z)$	χ^2/n_{dof}
1	1	0.1150	26./28	0.1142	24./28
1/2	1/2	0.1165	77./28	0.1160	73./28
2	2	0.1200	18./28	0.1191	18./28
1/2	1	0.1150	53./28	0.1136	48./28
1	1/2	0.1150	30./28	0.1142	28./28
1	2	0.1155	23./28	0.1147	22./28
2	1	0.1180	19./28	0.1175	19./28

Table 7.6: Uncertainty decomposition for $\alpha_s(M_Z)$ from the determination of α_S from the jet event rate R_{32} in bins of $H_{T,2}/2$. The statistical uncertainty of the NLO computation is negligible in comparison to any of the other sources of uncertainty. Electroweak corrections, significant only at high $H_{T,2}/2$, are assumed to cancel between the numerator and denominator.

$H_{T,2}/2$ (GeV)	MSTW2008					MMHT2014				
	$\alpha_s(M_Z)$	exp	PDF	NP	scale	$\alpha_s(M_Z)$	exp	PDF	NP	scale
300-420	0.1157	± 0.0015	± 0.0014	± 0.0019	$^{+0.0053}_{-0.0000}$	0.1158	± 0.0014	± 0.0010	± 0.0019	$^{+0.0052}_{-0.0000}$
420-600	0.1153	± 0.0011	± 0.0014	± 0.0018	$^{+0.0057}_{-0.0000}$	0.1154	± 0.0011	± 0.0012	± 0.0017	$^{+0.0056}_{-0.0000}$
600-1000	0.1134	± 0.0013	± 0.0016	± 0.0019	$^{+0.0052}_{-0.0000}$	0.1140	± 0.0012	± 0.0012	± 0.0018	$^{+0.0045}_{-0.0000}$
1000-1680	0.1147	± 0.0029	± 0.0017	± 0.0018	$^{+0.0063}_{-0.0011}$	0.1154	± 0.0025	± 0.0014	± 0.0015	$^{+0.0056}_{-0.0011}$
300-1680	0.1150	± 0.0010	± 0.0013	± 0.0015	$^{+0.0050}_{-0.0000}$	0.1142	± 0.0010	± 0.0013	± 0.0014	$^{+0.0049}_{-0.0006}$

range 300 - 1680 GeV is divided into four different sub-ranges as shown by the first column in Table 7.7. Each of the $H_{T,2}/2$ range is associated with a scale Q , which is the differential cross-section weighted average $H_{T,2}/2$ scale from the inclusive 2-jet calculations and integrated over all the measured $H_{T,2}/2$ bins contributing to that given $H_{T,2}/2$ range. Let N_{bin}^j be the total number of measured $H_{T,2}/2$ bins contributing to the j -th $H_{T,2}/2$ range, then the corresponding scale Q_j , shown in second column of Table 7.7, is calculated as :

$$Q_j = \frac{\sum_{i=1}^{N_{bin}^j} H_{T,2}^i \left[\frac{d\sigma}{d(H_{T,2}/2)} \right]^i}{\sum_{i=1}^{N_{bin}^j} \left[\frac{d\sigma}{d(H_{T,2}/2)} \right]^i} \quad (7.9)$$

The value of $\alpha_s(M_Z)$ is extracted in each $H_{T,2}/2$ range. These extracted $\alpha_s(M_Z)$ values are evolved to the corresponding values $\alpha_s(Q)$ and are quoted in Table 7.7 along with the extracted $\alpha_s(M_Z)$ values and the total uncertainty. The evolution is performed for five flavours at 2-loop order with the RUNDEC program [132, 133]. The obtained $\alpha_s(Q)$ points (black solid circles) are shown as a function of scale Q in Fig. 7.4. The black solid line and the yellow uncertainty band are evolved using $\alpha_s(M_Z) = 0.1150 \pm 0.0023$ (all except scale) $^{+0.0050}_{-0.0000}$ (scale) obtained using MSTW2008 NLO PDF set. The world average [21] (dashed line) and results from other measurements of the CMS [1, 99, 113, 122, 123], ATLAS [124], D0 [125, 126], H1 [127, 128], and ZEUS [129] experiments are also imposed. The current measurement is in very good agreement within the uncertainty with other results obtained by previous experiments as well as with the world average value of $\alpha_s(M_Z) = 0.1181 \pm 0.0011$ derived in Ref. [21].

Table 7.7: Evolution of the strong coupling constant between the scale of the Z boson mass and the cross-section averaged $H_{T,2}/2$ scale $\langle Q \rangle$ for the separate determinations in each respective fit range. The evolution is performed for five flavours at 2-loop order with the RUNDEC program [132, 133].

$H_{T,2}/2$ (GeV)	$\langle Q \rangle$ (GeV)	$\alpha_s(M_Z)$	$\alpha_s(Q)$	No. of data points	χ^2/n_{dof}
300-420	340	$0.1157^{+0.0060}_{-0.0030}$	$0.0969^{+0.0041}_{-0.0021}$	4	2.8/3
420-600	476	$0.1153^{+0.0062}_{-0.0025}$	$0.0928^{+0.0039}_{-0.0016}$	6	6.1/5
600-1000	685	$0.1134^{+0.0059}_{-0.0028}$	$0.0879^{+0.0035}_{-0.0017}$	9	7.1/8
1000-1680	1114	$0.1147^{+0.0074}_{-0.0040}$	$0.0841^{+0.0039}_{-0.0021}$	10	5.4/9

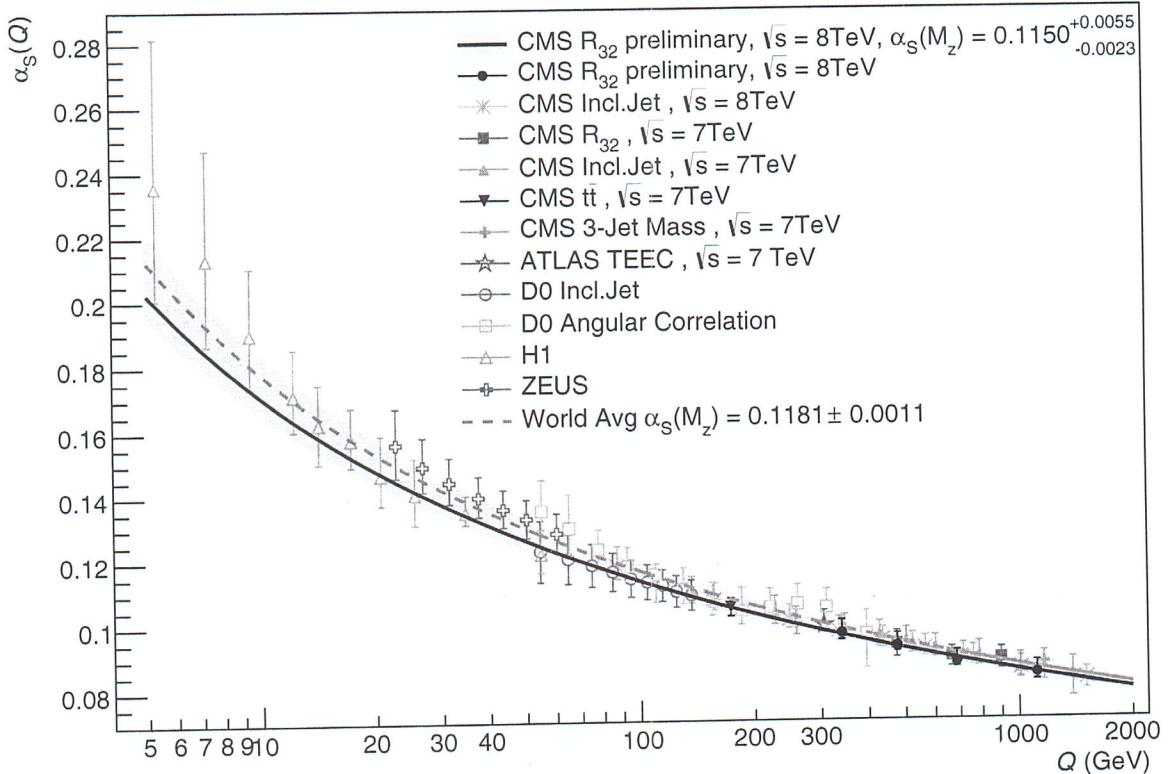


Figure 7.4: The running $\alpha_s(Q)$ as a function of the energy scale Q is shown as obtained by using the MSTW2008 NLO PDF set. The solid line and the uncertainty band are drawn by evolving the extracted $\alpha_s(M_Z)$ values using the 2-loop 5-flavour renormalization group equations as implemented in RUNDEC [132, 133]. The dashed line represents the evolution of the world average [21] and the black circles correspond to the $\alpha_s(Q)$ determinations presented in Table 7.7. Results from other measurements of CMS [1, 99, 113, 122, 123], ATLAS [124], D0 [125, 126], H1 [127, 128], and ZEUS [129] are superimposed.

Chapter 8

Summary

Inclusive multijet production cross-section measured precisely in terms of jet transverse momentum is one of the important observables in understanding physics at hadron colliders. It provides the essential information about the structure of parton through parton distribution functions (PDFs) and the precise measurement of the strong coupling constant α_s . The value of the strong coupling constant at the scale of the Z boson mass $\alpha_s(M_Z)$ can be determined using cross-section ratio instead of individual cross-sections because many uncertainties of theoretical and experimental origin cancel in the ratio which reduces the dependence on PDFs, renormalization and factorization scales, luminosity etc.

In this thesis, a measurement of the inclusive 2-jet and 3-jet event cross-sections as well as the cross-section ratio R_{32} has been presented. The data sample has been collected from proton-proton collisions recorded with the CMS detector at a centre-of-mass energy of 8 TeV and corresponds to an integrated luminosity of 19.7 fb^{-1} . The jets are reconstructed with the anti- k_t clustering algorithm for a jet size parameter $R = 0.7$. The inclusive 2-jet and 3-jet event cross-sections are measured differentially as a function of the average transverse momentum of the two leading jets, referred as $H_{\text{T},2}/2$. The ratio R_{32} is obtained by dividing the differential cross-sections of inclusive 3-jet events to that of inclusive 2-jet one in

Physical review

each bin of $H_{T,2}/2$. An appropriate selection criteria has been designed for choosing the best events for analysis. The measurements are performed at a central rapidity of $|y| < 2.5$ in a range of $300 < H_{T,2}/2 < 2000$ GeV for inclusive 2-jet event cross-sections and $300 < H_{T,2}/2 < 1680$ GeV for inclusive 3-jet event cross-sections and ratio R_{32} .

The measured cross-sections after correcting for detector effects by using an iterative unfolding procedure are compared to the perturbative QCD predictions computed, using NLOJET++ program, at next-to-leading order (NLO) accuracy and complemented with non-perturbative (NP) corrections that are important at low $H_{T,2}/2$. The data are found to be well described by NLO calculations. The upwards trend observed in the inclusive 2-jet and 3-jet data at high $H_{T,2}/2$ in comparison to the prediction at NLO QCD, is explained by the onset of electroweak (EW) corrections in the 2-jet case. For the 3-jet event cross-sections these corrections have not yet been computed theoretically. In the 3-jet to 2-jet cross-section ratio R_{32} , the EW corrections are assumed to cancel. In fact, NLO QCD provides an adequate description of R_{32} in the accessible range of $H_{T,2}/2$. In contrast, leading order (LO) tree-level Monte Carlo (MC) predictions obtained using MADGRAPH5 event generator interfaced to PYTHIA6 exhibit significant deviations. The sources of experimental and theoretical uncertainties are studied in detail. The experimental uncertainty ranges from 4 to 32% for inclusive 2-jet event cross-sections, from 4 to 28% for 3-jet event cross-sections and from 1 to 28% for cross-section ratio R_{32} . It is dominated by the uncertainty due to the jet energy corrections (JEC) at lower $H_{T,2}/2$ values and by statistical uncertainty at higher $H_{T,2}/2$ values. The theoretical uncertainty ranges from 3 to 30% and 5 to 34% for inclusive 2-jet and 3-jet event cross-sections respectively and from 3 to 11% for ratio R_{32} . The PDF uncertainty derived with the CT10-NLO PDF set is the dominant source of theoretical uncertainty.

The inclusive multijet cross-sections being proportional to the powers of the strong coupling constant α_S ($\sigma_{n\text{-jet}} \propto \alpha_S^n$) are used to extract the value of the strong coupling constant at the scale of the Z boson mass $\alpha_s(M_Z)$. In cross-section ratio R_{32} which is proportional to α_S , many uncertainties and PDF dependencies largely cancel and hence becomes the better tool to extract the value of $\alpha_s(M_Z)$. In this thesis, a fit of the ratio of the inclusive 3-jet event cross-section to that of 2-jet, R_{32} in the range $300 < H_{T,2}/2 < 1680$ GeV, using the MSTW2008 PDF set gives :

$$\begin{aligned}\alpha_s(M_Z) &= 0.1150 \pm 0.0010 \text{ (exp)} \pm 0.0013 \text{ (PDF)} \pm 0.0015 \text{ (NP)} {}^{+0.0050}_{-0.0000} \text{ (scale)} \\ &= 0.1150 \pm 0.0023 \text{ (all except scale)} {}^{+0.0050}_{-0.0000} \text{ (scale)}\end{aligned}$$

and using the MMHT2014 PDF set gives :

$$\begin{aligned}\alpha_s(M_Z) &= 0.1142 \pm 0.0010 \text{ (exp)} \pm 0.0013 \text{ (PDF)} \pm 0.0014 \text{ (NP)} {}^{+0.0049}_{-0.0006} \text{ (scale)} \\ &= 0.1142 \pm 0.0022 \text{ (all except scale)} {}^{+0.0049}_{-0.0006} \text{ (scale)}\end{aligned}$$

The equally compatible values of $\alpha_s(M_Z)$ are determined with separate fits to the inclusive 2-jet and 3-jet event cross-sections provided the range in $H_{T,2}/2$ is restricted to $300 < H_{T,2}/2 < 1000$ GeV. The extracted $\alpha_s(M_Z)$ values in sub-ranges of $H_{T,2}/2$ are evolved to corresponding $\alpha_S(Q)$ along with the error bars at different scales Q . The current measurement of $\alpha_s(M_Z)$ and the running of $\alpha_S(Q)$ as a function of Q is in well agreement within uncertainties with the world average value of $\alpha_s(M_Z) = 0.1181 \pm 0.0011$ [21] and already existing determinations performed by the CMS and other experiments. The results on α_S reported here are consistent with the energy dependence predicted by the renormalization group equation (RGE) [18] which states that the strong force becomes weaker at short distances corresponding to large momentum transfers.

The inclusion of the EW corrections in inclusive 2-jet event cross-sections become relevant at $H_{T,2}/2$ beyond 1 TeV. Their availability for 3-jet one and hence cross-section ratio R_{32} can improve the precision of the measurement of $\alpha_s(M_Z)$. Also when the theoretical calculations will become available for inclusive 4-jet event

cross-sections, the various cross-section ratios such as $R_{43} \propto \alpha_S^1$ and $R_{42} \propto \alpha_S^2$ can be measured to extract the value of the strong coupling constant more precisely. Currently LHC is running at high center-of-mass energy of 13 TeV delivering a higher instantaneous luminosity and this makes possible to access the extended phase space and perform the measurements with improved accuracy.