

**MEASUREMENT OF MULTIJET CROSS-SECTION RATIOS  
IN PROTON-PROTON COLLISIONS WITH THE CMS  
DETECTOR AT THE LHC**

**A THESIS**

**Submitted to the  
FACULTY OF SCIENCE  
PANJAB UNIVERSITY, CHANDIGARH  
for the degree of**

**DOCTOR OF PHILOSOPHY**

**2018**

**ANTERPREET KAUR**

**DEPARTMENT OF PHYSICS  
CENTRE OF ADVANCED STUDY IN PHYSICS  
PANJAB UNIVERSITY  
CHANDIGARH, INDIA**



*Dedicated to  
my Parents*



## *Acknowledgement*



# Abstract

The hadrons colliding at very high center-of-mass energies provide a direct probe to the nature of the underlying parton-parton scattering physics. The scattering of the elementary quarks and gluons, constituents of the incoming hadron beams, produces partons carrying high momenta, which then fragment and hadronize producing a spray of particles. These particles get clustered in the form of jets. The jets being the final structures observed in the detector, preserve the energy and direction of the initial partons. Hence jets can serve as a direct test of theory of strong interactions called Quantum Chromodynamics. The inclusive multijet production cross-section is an important observable which provides the details of parton distribution functions (PDF) of the colliding hadrons and the precise measurement of the strong coupling constant  $\alpha_s$ . Instead of individual cross-sections, the ratio of cross-sections is a better tool to determine the value of  $\alpha_s$  as many theoretical and experimental uncertainties cancel in the ratio.

A measurement of inclusive multijet event cross-sections and the cross-section ratio is presented using the data from the proton-proton collisions collected with the CMS detector at a center-of-mass energy of 8 TeV corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$ . 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 as well as the ratio of the 3-jet over 2-jet event cross-section ( $R_{32}$ ) are measured as a function of the average transverse momenta  $p_T$  of the two leading jets in a phase space region ranging up to jet  $p_T$  of 2.0 TeV and an absolute rapidity of  $|y| = 2.5$ . The measurements, after correcting for detector effects are well described by predictions at next-to-leading order in perturbative quantum chromodynamics and additionally compared to several Monte Carlo event generators. The strong coupling constant at the scale of the  $Z$  boson mass is extracted from a fit of the measured  $R_{32}$  which gives  $\alpha_s(M_Z) = 0.1150 \pm 0.0010 \text{ (exp)} \pm 0.0013 \text{ (PDF)} \pm 0.0015 \text{ (NP)} {}^{+0.0050}_{-0.0000} \text{ (scale)}$  using MSTW2008 PDF set. This measurement agrees well with the world average

value of  $\alpha_s(M_Z) = 0.1181 \pm 0.0011$  as well as previous measurements.

# Contents

<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xxi</b>
<hr/>	
<b>1 Introduction</b>	<b>1</b>
<b>2 Theoretical Background</b>	<b>7</b>
2.1 Standard Model . . . . .	7
2.2 Quantum Chromodynamics . . . . .	10
2.2.1 Perturbative Quantum Chromodynamics . . . . .	14
2.2.2 Renormalization and Running of the Strong Coupling . . . . .	16
2.3 Hadronic Collisions . . . . .	18
2.3.1 Parton Shower and Hadronization . . . . .	21
2.3.2 Underlying Event . . . . .	24
2.4 Jets . . . . .	25
2.4.1 Jet Algorithms . . . . .	26
<b>3 Experimental Setup</b>	<b>33</b>
3.1 The Large Hadron Collider . . . . .	34
3.1.1 Luminosity Measurement . . . . .	37

---

3.1.2	Pileup Interactions . . . . .	38
3.2	The Compact Muon Solenoid . . . . .	40
3.2.1	Coordinate System . . . . .	41
3.2.2	Inner Tracker System . . . . .	43
3.2.3	Electromagnetic Calorimeter . . . . .	46
3.2.4	Hadronic Calorimeter . . . . .	48
3.2.5	Superconducting Magnet . . . . .	51
3.2.6	Muon System . . . . .	51
3.2.7	Trigger and Data Acquisition System . . . . .	53
3.2.7.1	Jet Triggers . . . . .	55
3.2.7.2	Data Acquisition System . . . . .	56
3.2.8	Data Management . . . . .	56
3.3	Software Tools . . . . .	57
3.3.1	CMSSW Framework . . . . .	58
3.3.2	ROOT . . . . .	59
3.3.3	NLOJET++ and FASTNLO . . . . .	59
<b>4</b>	<b>Event Generation, Simulation and Reconstruction</b> . . . . .	<b>61</b>
4.1	Event Generation and Simulation Software . . . . .	61
4.1.1	PYTHIA . . . . .	63
4.1.2	MADGRAPH5 . . . . .	63
4.1.3	HERWIG . . . . .	64
4.1.4	POWHEG . . . . .	64
4.2	Detector Simulation . . . . .	64
4.3	Event Reconstruction . . . . .	66
4.3.1	Particle Flow Algorithm . . . . .	66

4.3.2	Jet Energy Corrections . . . . .	71
<b>5</b>	<b>Measurement of the Differential Inclusive Multijet Cross-sections and their Ratio</b>	<b>75</b>
5.1	Data Samples . . . . .	76
5.1.1	Monte Carlo Samples . . . . .	77
5.2	Event Selection . . . . .	78
5.2.1	Trigger Selection . . . . .	78
5.2.2	Primary Vertex Selection . . . . .	82
5.2.3	Missing Transverse Energy . . . . .	82
5.2.4	Jet Identification . . . . .	83
5.2.4.1	Jet ID Efficiency . . . . .	84
5.2.5	Jet Selection . . . . .	87
5.3	Comparison with Simulation . . . . .	88
5.3.1	Pileup Reweighting . . . . .	88
5.3.2	Comparison of Cross-sections and their Ratio . . . . .	88
5.4	Jet Energy Resolution (JER) . . . . .	91
5.5	Unfolding . . . . .	97
5.5.1	Response Matrices . . . . .	98
5.5.1.1	Inclusive Cross-sections . . . . .	99
5.5.1.2	Cross-section Ratio, $R_{32}$ . . . . .	102
5.5.2	Closure Test . . . . .	102
5.5.3	Unfolding of the Measurement . . . . .	104
5.6	Experimental Uncertainties . . . . .	105
5.6.1	Statistical Uncertainty . . . . .	107
5.6.2	Jet Energy Corrections Uncertainty . . . . .	107

5.6.3	Unfolding Uncertainty . . . . .	110
5.6.4	Luminosity Measurement Uncertainty . . . . .	112
5.6.5	Residual Uncertainty . . . . .	112
5.6.6	Total Experimental Uncertainty . . . . .	112
<b>6</b>	<b>Theoretical Calculations</b>	<b>115</b>
6.1	Fixed Order NLO Calculations . . . . .	115
6.1.1	NLO Correction Factors . . . . .	117
6.1.2	Non-perturbative Corrections . . . . .	118
6.1.3	Electroweak Corrections . . . . .	120
6.2	Theoretical Uncertainties . . . . .	122
6.2.1	Scale Uncertainty . . . . .	122
6.2.2	PDF Uncertainty . . . . .	123
6.2.3	Non-perturbative Uncertainty . . . . .	124
6.2.4	Total Theoretical Uncertainty . . . . .	125
6.3	Comparison of Theory to Data . . . . .	126
<b>7</b>	<b>Determination of the Strong Coupling Constant</b>	<b>131</b>
7.1	Sensitivity of Measurements to $\alpha_s(M_Z)$ . . . . .	132
7.2	Determination of $\alpha_s(M_Z)$ . . . . .	133
7.2.1	Fitting Procedure . . . . .	133
7.2.2	Fit Results . . . . .	139
7.3	Running of the Strong Coupling Constant . . . . .	142
<b>8</b>	<b>Summary</b>	<b>147</b>
<b>9</b>	<b>Hardware and Software Activities Undertaken</b>	<b>151</b>

9.1 Silicon Photo-Multipliers . . . . .	152
9.2 MicroTCA . . . . .	156
9.3 Other Activities . . . . .	161
<b>Appendix A</b>	<b>163</b>
A.1 Cross-section Ratio, $R_{32}$ . . . . .	164
A.2 Individual Sources of Jet Energy Correction Uncertainties . . . . .	165
A.3 Experimental Uncertainties . . . . .	169
A.4 Theoretical Uncertainties . . . . .	172
A.5 Definitions . . . . .	175
<b>Bibliography</b>	<b>177</b>
<b>List of Publications</b>	<b>193</b>
<b>Reprints</b>	<b>199</b>



# List of Figures

2.1	The Standard Model summarizing the properties of elementary particles and their forces of interaction. . . . .	8
2.2	The fundamental Feynman rules for different processes of quantum chromodynamics. . . . .	13
2.3	The Feynman diagrams of leading-order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) processes in quantum chromodynamics. . . . .	15
2.4	Running of the strong coupling constant evolved at the energy scale $Q$ as a function of $Q$ . . . . .	19
2.5	Schematic illustration of the factorization theorem in a collision of two protons. . . . .	21
2.6	Illustration of the hadronization process in Lund string model. . . . .	23
2.7	A proton-proton collision involving the main hard scattering process along with the low momentum transfer underlying event (UE) contributions. . . . .	25
2.8	Formation of a jet in a proton-proton collision. . . . .	26
2.9	Illustration of infrared and collinear unsafe behaviour of jet algorithms. .	27
2.10	The clustering of particles into jets using different jet algorithms. . .	30

---

3.1	An overview of the different experiments of the Large Hadron Collider (LHC), a complex particle accelerator and collider located at CERN.	35
3.2	The integrated luminosity delivered by stable beams to CMS during proton-proton collisions. . . . .	38
3.3	In a proton-proton collision, the particles produced from the hard interaction are clustered into a jet. The hard interaction corresponds to the main vertex. The particles produced in the interactions other than the hard one, form a pileup jet. . . . .	39
3.4	The three dimensional view of the CMS detector along with its sub-detector components. . . . .	41
3.5	Front view of the CMS detector along with its various components. .	42
3.6	The right-handed coordinate system used by the CMS detector. . .	43
3.7	A longitudinal view of the CMS detector is shown in the $y$ - $z$ plane. .	44
3.8	A longitudinal view of the inner tracking system is shown in $r$ - $z$ plane.	45
3.9	A geometric view of one quarter of the electromagnetic calorimeter (ECAL) in $y$ - $z$ plane. . . . .	47
3.10	Longitudinal section of one quarter of the hadronic calorimeter (HCAL) in $r$ - $\eta$ plane. . . . .	49
3.11	A longitudinal view of the CMS muon system showing the location of the three gaseous particle detectors. . . . .	53
3.12	Work flow of the L1 trigger system consisting of local, regional and global components. . . . .	54
3.13	Architecture of the CMS Data Acquisition (DAQ) system. . . . .	57
3.14	The schematic overview of the CMS computing grid. . . . .	58

4.1	The comparison between Monte Carlo (MC) simulations generated by event generators and the real data produced by the particle collisions and observed in the detectors. . . . .	62
4.2	The Particle Flow (PF) algorithm is used by the CMS to identify and reconstruct the particles. The PF converts the sub-detector measurements back to physical particle objects. . . . .	69
4.3	Formation of jets in a proton-proton collision at different levels. . . . .	71
4.4	A schematic diagram of the factorized jet energy corrections (JEC). .	72
5.1	Trigger efficiencies turn-on curves for the single jet HLT trigger paths.	81
5.2	Missing transverse energy fraction of the total transverse energy per event in the data and simulated Monte Carlo events. . . . .	83
5.3	The fractions of jet constituents for different types of PF candidates for inclusive 2-jet events. . . . .	85
5.4	The fractions of jet constituents for different types of PF candidates for inclusive 3-jet events. . . . .	86
5.5	The jet ID efficiency is studied as a function of $H_{T,2}/2$ with tag-and-probe technique using dijet event topologies and it always exceeds 99%. . . . .	87
5.6	Number of reconstructed vertices before and after the pileup reweighting. . . . .	89
5.7	Comparison of differential cross-sections for the data with simulated events and CT10-NLO theory predictions. . . . .	90
5.8	Comparison of the cross-section ratio for the data with simulated events and CT10-NLO theory predictions. . . . .	90
5.9	Fitting of the jet energy resolution distribution as a function of $H_{T,2}/2$ . .	94

---

5.10 Comparison of jet energy resolution calculated using Crystal Ball fit function and Gaussian fit function . . . . .	94
5.11 Jet energy resolution (JER) is shown as a function of Gen $H_{T,2}/2$ . . . . .	95
5.12 Additional unfolding uncertainty. . . . .	96
5.13 Fitted CT10-NLO spectrum of differential cross-section as a function of $H_{T,2}/2$ . . . . .	101
5.14 The response matrices are derived using the Toy Monte Carlo and forward smearing method. . . . .	102
5.15 Left : The ratio of cross-sections for inclusive 3-jet to that of 2-jet events as a function of $H_{T,2}/2$ . Right : The response matrix is derived using the Toy Monte Carlo and forward smearing method, for the cross-section ratio $R_{32}$ . . . . .	103
5.16 Closure test of the unfolding technique . . . . .	104
5.17 Reco differential cross-section distributions unfolded with the response matrices. . . . .	105
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. . . . .	106
5.19 The fractional statistical uncertainties of the unfolded data are compared with those of the measured one. . . . .	108
5.20 The unfolding procedure introduces the correlations of the statistical uncertainty through bin migrations. . . . .	109
5.21 Experimental uncertainties from different sources affecting the measurement of cross-sections and the cross-section ratio. . . . .	113
6.1 The k-factors using five different PDF sets. . . . .	118

6.2	The nonperturbative (NP) corrections are presented as a function of $H_{T,2}/2$ . . . . .	120
6.3	The electroweak (EW) corrections as a function of $H_{T,2}/2$ . . . . .	121
6.4	Ratio of the data over theory obtained using the CT10-NLO PDF set. . . . .	122
6.5	The systematic theoretical uncertainties affecting the cross-section measurement and the cross-section ratio. . . . .	125
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. . . . .	127
6.7	Cross-section ratio as a function of $H_{T,2}/2$ calculated from data in comparison to that from NLO pQCD predictions obtained using the CT10-NLO PDF set. . . . .	128
6.8	Ratio of the data over theory using the CT10-NLO PDF set. . . . .	129
6.9	Ratio of the data over the predictions from Monte Carlo simulations. . . . .	130
7.1	Ratio of the measured inclusive 2-jet differential cross-section to theory predictions using different PDF sets. . . . .	134
7.2	Ratio of the measured inclusive 3-jet differential cross-section to theory predictions using different PDF sets. . . . .	135
7.3	Ratio of the measured cross-section ratio to theory predictions using different PDF sets. . . . .	136
7.4	The running $\alpha_S(Q)$ as a function of the energy scale $Q$ . . . . .	145
9.1	The arrangement of Silicon Photo-Multipliers (SiPMs) on the Mounting Board (MB). . . . .	153
9.2	The breakdown voltage (BV) is estimated using LED method and its variation is shown over time for 18 channels of one readout module (RM). . . . .	155

---

9.3	The relative variation of the SiPM gain is presented over time for a single RM with 18 channels. The gain is stable over a time from the middle of February to the beginning of March in 2014 and the relative variation of the gain lies within 2% . . . . .	156
9.4	The distribution of the relative variations in gain for all the installed SiPMs is fitted with a Gaussian function. It has a width of only 0.5 % and all gain variations are within 3%. . . . .	157
9.5	$\mu$ TCA crate showing the different slots. . . . .	158
9.6	A test-stand designed to monitor the working of Power Mezzanines/Auxiliary Power Mezzanines (PMs/APMs) through stability tests. . . . .	159
9.7	A test-stand installed at Department of Physics, Panjab University, Chandigarh to perform the stability tests for monitoring the working of Power Mezzanines/Auxiliary Power Mezzanines (PMs/APMs). . .	160
A.1	The fractional jet energy correction (JEC) uncertainties from individual sources (Part I). . . . .	166
A.2	The fractional jet energy correction (JEC) uncertainties from individual sources (Part II). . . . .	167
A.3	The fractional jet energy correction (JEC) uncertainties from individual sources (Part III). . . . .	168

# List of Tables

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. . . . .	77
5.2	The official Monte Carlo samples produced in phase space slices in $H_T$ with the generator <b>MADGRAPH5</b> and interfaced to <b>PYTHIA6</b> . . . . .	78
5.3	The single jet HLT trigger paths used in the analysis. . . . .	79
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. . . . .	84
5.5	The jet energy resolution in the data is actually worse than in simulation. To match the resolution in the data, the reconstructed jet transverse momentum in simulated events need to be smeared by applying the scale factors. . . . .	92
5.6	The parameters obtained by fitting the relative resolution as a function of $H_{T,2}/2$ . . . . .	96
5.7	An overview of all experimental uncertainties affecting the measurement of cross-sections and the cross-section ratio. . . . .	114
6.1	NLO PDF sets available via <b>LHAPDF6</b> with various assumptions on the value of $\alpha_s(M_Z)$ . . . . .	117

6.2	Overview of all systematic theoretical uncertainties affecting the measurement of cross-sections and the cross-section ratio. . . . .	126
7.1	Determination of $\alpha_s(M_Z)$ from the inclusive 2-jet and 3-jet event cross-sections using five PDF sets. . . . .	140
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. . . . .	140
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. . . . .	141
7.4	Determination of $\alpha_s(M_Z)$ from the ratio $R_{32}$ using the two most compatible PDF sets MSTW2008 and MMHT2014. . . . .	143
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. . .	143
7.6	Uncertainty decomposition for $\alpha_s(M_Z)$ from the determination of $\alpha_S$ from the jet event rate $R_{32}$ in bins of $H_{T,2}/2$ . . . . .	143
7.7	Evolution of the strong coupling constant. . . . .	145
A.1	Differential cross-sections ( $\times 10^{-3}(\text{pb}/\text{GeV})$ ) and the cross-section ratio $R_{32}$ at detector level in each bin of $H_{T,2}/2$ , along with statistical uncertainty (in %). . . . .	164
A.2	Experimental uncertainties (in %) affecting the cross-section measurement in each bin of $H_{T,2}/2$ for inclusive 2-jet events. . . . .	169
A.3	Experimental uncertainties (in %) affecting the cross-section measurement in each bin of $H_{T,2}/2$ for inclusive 3-jet events. . . . .	170

A.4 Experimental uncertainties (in %) affecting the cross-section measurement in each bin of $H_{\mathrm{T},2}/2$ for cross-section ratio. . . . .	171
A.5 Theoretical uncertainties (in %), calculated using CT10-NLO PDF set, affecting the cross-section measurement in each bin of $H_{\mathrm{T},2}/2$ for inclusive 2-jet events. . . . .	172
A.6 Theoretical uncertainties (in %), calculated using CT10-NLO PDF set, affecting the cross-section measurement in each bin of $H_{\mathrm{T},2}/2$ for inclusive 3-jet events. . . . .	173
A.7 Theoretical uncertainties (in %), calculated using CT10-NLO PDF set, affecting the measurement of cross-section ratio in each bin of $H_{\mathrm{T},2}/2$ . . . . .	174



# 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 particle detectors constituting the real data.

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  $\alpha_S$  convoluted with PDFs. The lowest-order  $\alpha_S^2$  term represents the production of two partons in final states whereas terms of higher-order  $\alpha_S^3$ ,  $\alpha_S^4$  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  $\sigma_{i-jet}$ , given by the process

$\text{pp} \rightarrow i\text{jets} + X$ , is proportional to  $\alpha_s^i$ . 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  $\alpha_S$ . Also the ratio of cross-sections given by Eq. 1.1 is proportional to the QCD coupling constant  $\alpha_S$  and hence can be used to determine the value of  $\alpha_S$ .

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

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  $19.7 \text{ fb}^{-1}$  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  $H_{T,2}/2$  in this thesis. The strength of the strong force,  $\alpha_S$  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  $\alpha_S$  depends on the energy scale  $Q$  and it decreases with the increase of  $Q$  scale. The running of  $\alpha_S$  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)<sup>1</sup>, which precisely describes the evolution of  $\alpha_S$  at the renormalization scale of QCD.

The organization of this thesis<sup>2</sup> 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.

---

<sup>1</sup>According 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.

<sup>2</sup>The 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  $\alpha_S$  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.

### 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-

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.

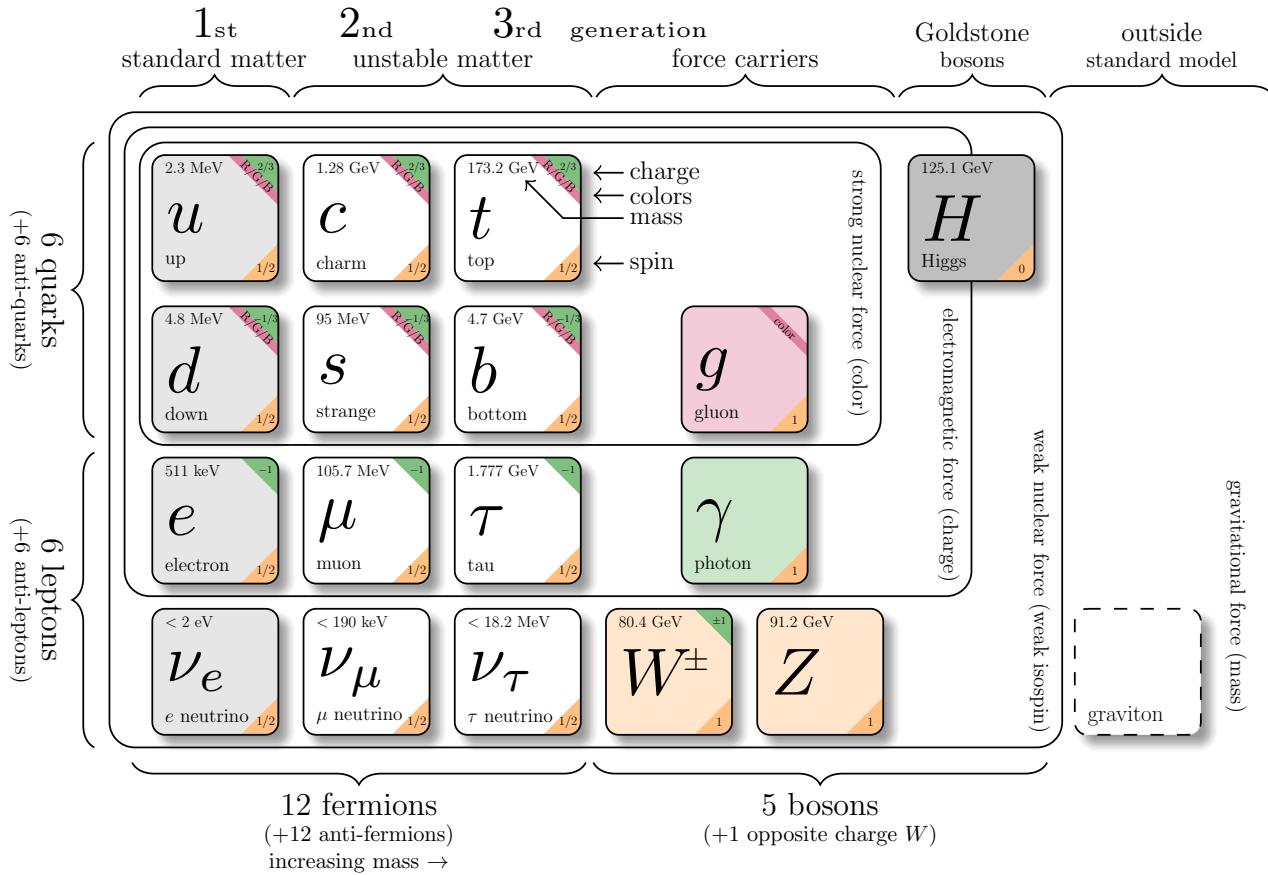


Figure 2.1: The Standard Model<sup>3</sup> summarizing the properties of elementary particles known as fermions (leptons and quarks) grouped into three generations, gauge bosons as mediators for the interactions, the scalar Higgs boson and not corporated graviton for the gravitational force.

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

Depending on how the fermions interact, these are classified into two categories - leptons ( $\ell$ ) and quarks ( $q$ ). The leptons are of six types : electron ( $e$ ), muon ( $\mu$ ) and tau ( $\tau$ ) with electric charge  $Q = -1$ <sup>4</sup> and the corresponding neutrinos : electron neutrino ( $\nu_e$ ), muon neutrino ( $\nu_\mu$ ) and tau neutrino ( $\nu_\tau$ ) 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 ( $\gamma$ ) 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.

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

---

<sup>4</sup>all charges are expressed in units of elementary charge  $e$

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 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 the 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

has a same role as the electric charge in electromagnetic interactions. However, the mediator of electromagnetic interactions i.e. 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 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  $\mathcal{A}_\mu^A$ ;  $\mathcal{L}_{quarks}$  defines the interaction between spin- $\frac{1}{2}$  quark fields  $q_a$  of mass  $m_q$  and spin-1 gluon fields  $\mathcal{A}_\mu^A$  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,

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

$$F_{\mu\nu}^A = \partial_\mu A_\nu^A - \partial_\nu A_\mu^A - g_s f_{ABC} A_\mu^B A_\nu^C \quad (2.3)$$

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  $\gamma_\mu$  are the Dirac  $\gamma$ -matrices.

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

$A_\mu^A$  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  $\lambda^A$  [16]. The generator matrices  $T^A$  follow the commutation relations :

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

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

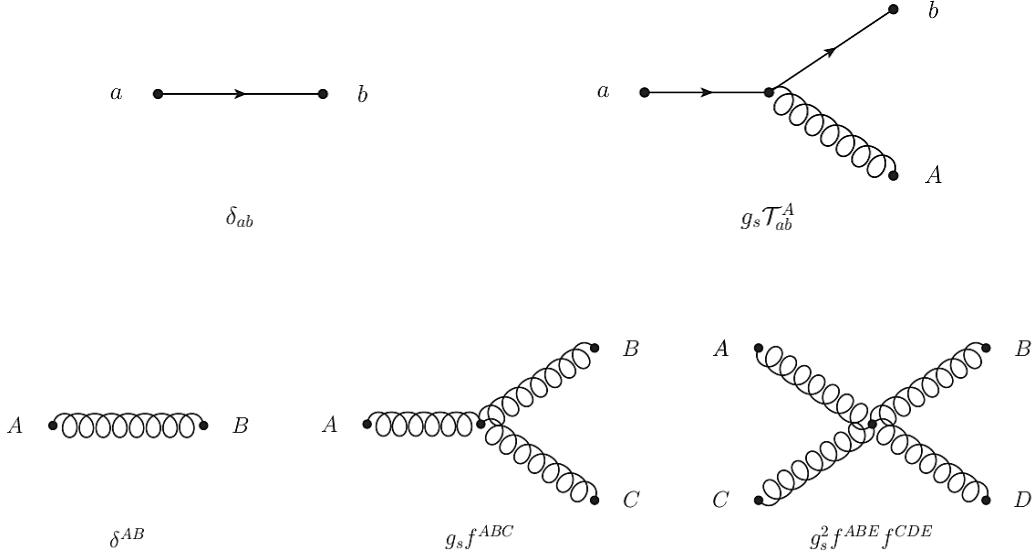


Figure 2.2: The fundamental Feynman rules of a free quark-field term (top left), quark-gluon interaction term (top right), free gluon-field term (bottom left), cubic gluon self-interaction term (bottom middle) and quartic gluon self-interaction term (bottom right). Taken from [17].

is given by

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

where  $\xi$  may be any finite constant. This choice fixes the class of covariant gauges with  $\xi$  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  $\eta^A$  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  $\alpha$ . 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

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 pair making difficult 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  $\alpha_s$  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  $\alpha_S$  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 of  $\alpha_S$ . 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

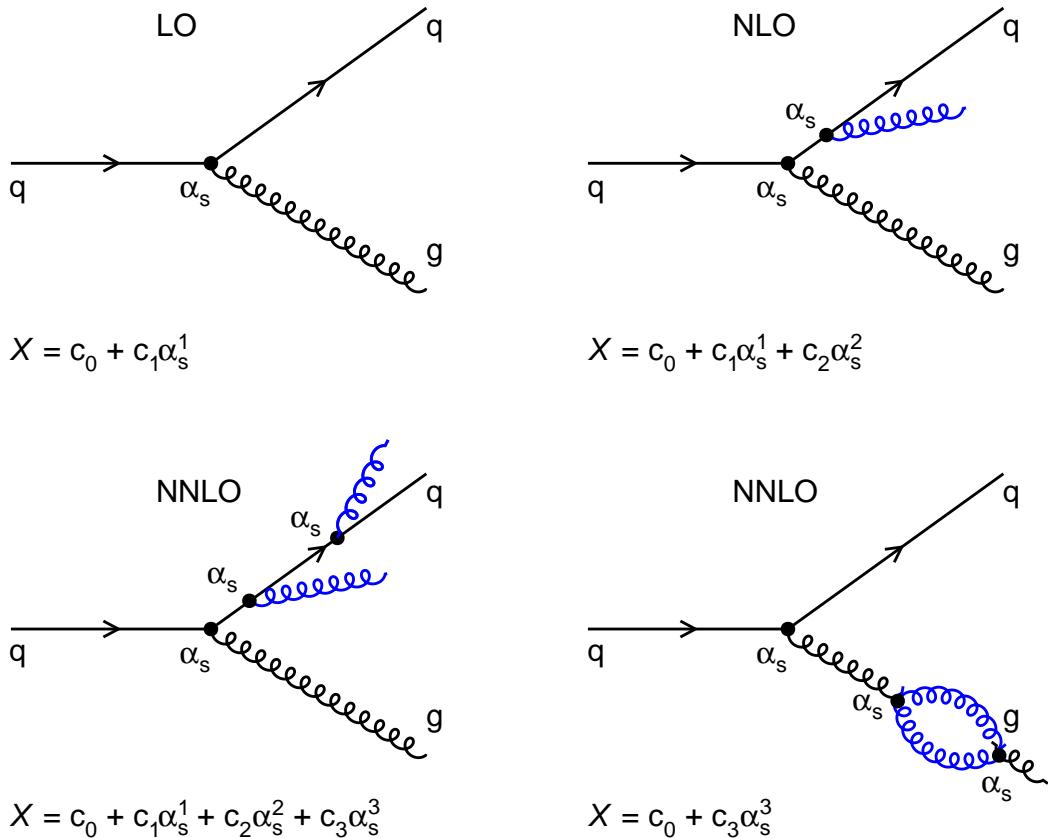


Figure 2.3: The Feynman diagrams<sup>5</sup> of leading-order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) processes in quantum chromodynamics along with the perturbative expansion of any observable  $X$  in powers of the strong coupling constant  $\alpha_S$ . At each successive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

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

<sup>5</sup>Drawn 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  $\mu_r$ . At first, the divergences are regularized temporarily by introducing a cut-off to the loop momenta at  $\mu_r$  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  $\mu_r$ . The exact dependence of  $\alpha_s(\mu_r^2)$  on  $\mu_r$  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  $\mu_r$  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  $\beta$  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  $\beta$ -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  $\alpha_S$  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  $\Lambda_{QCD}$  is the constant of integration. The perturbative coupling becomes large at the scale  $\Lambda_{QCD}$  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  $\alpha_S$  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

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

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

Since  $\alpha_S$  is 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

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

This value is derived using data from deep inelastic scattering process, electron-positron annihilation processes, hadronic  $\tau$  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  $\alpha_S$  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 ( $\sigma$ ) of a certain process which gives the probability that the two hadrons interact and give rise to that

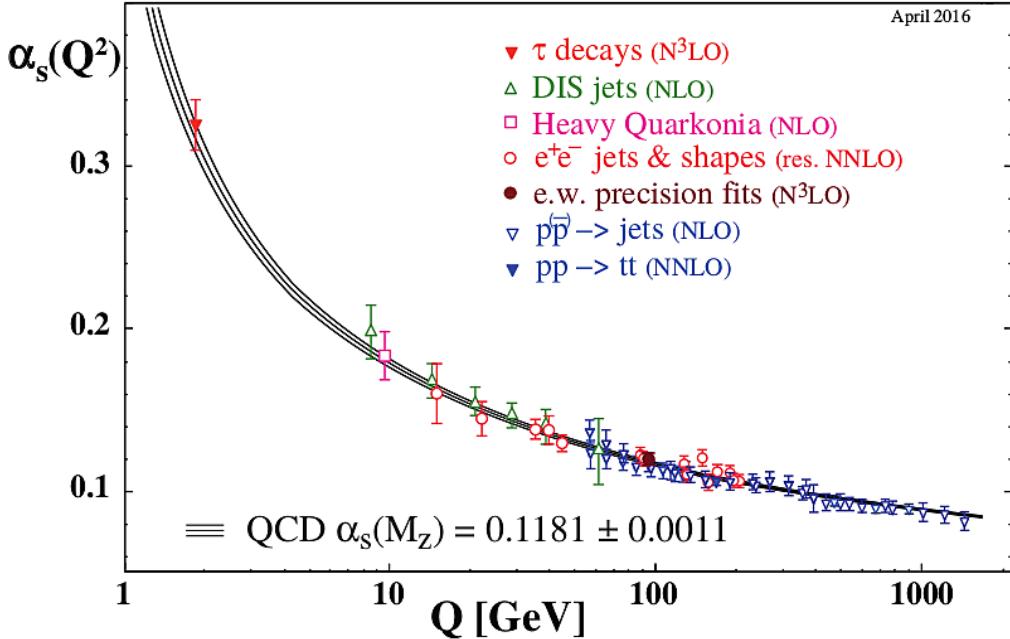


Figure 2.4: The different experimental measurements of the strong coupling constant  $\alpha_S$  evolved at the energy scale  $Q$  are shown as a function of  $Q$ . These describe the running of the  $\alpha_S$  up to the 1 TeV scale. Taken from [21].

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  $\sigma$  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.  $\mu_f$  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

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

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

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  $\mu_f$ . 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  $\mu_f$  and the renormalization scale  $\mu_r$ . 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  $\mu_f$  is given by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [23–25] equations which use  $\alpha_S$  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 which mainly differ in choice of input data sets, treatment of heavy quarks, order of perturbation theory, way of treating experimental errors and theoretical assumptions. The global PDFs are the CTEQ [27], MMHT [28], NNPDF [29], ABM [30] and HERAPDF [31] at LO, NLO and NNLO.

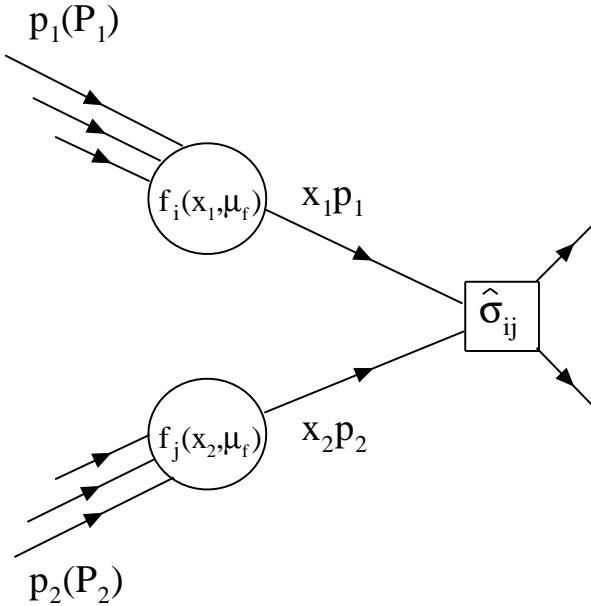


Figure 2.5: Schematic illustration<sup>6</sup> of the factorization theorem in a collision of two protons  $P_1$  and  $P_2$  having momenta  $p_1$  and  $p_2$ , respectively. In a hard-scattering process at a scale  $Q^2$ , the two partons  $x_1$  and  $x_2$  participate with momenta  $x_1 p_1$  and  $x_2 p_2$ . 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  $\mu_f$ .

### 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 lead 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-

---

<sup>6</sup>Drawn 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  $\sim 1$  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  $\alpha_S$  evolves and becomes large. 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 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)<sup>7</sup>, 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/fm<sup>2</sup> 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.

---

<sup>7</sup> 1 femtometre =  $1 \times 10^{-15}$  metres

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.

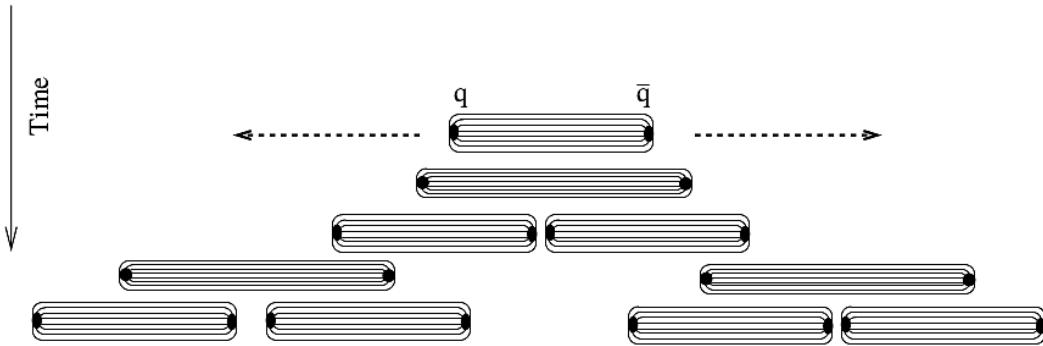


Figure 2.6: Illustration of the hadronization process in Lund string model<sup>8</sup>. When the quark  $q$  and anti-quark  $\bar{q}$  are pulled apart from each other, the potential energy of the gluonic string connecting the quarks increases. As it becomes of the order of hadron masses, the string 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

---

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

fragmentation model. However, this model has problems in dealing with the decay of very massive clusters.

### 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  $\alpha_S$ . The precise knowledge of the jet production cross-section can help to extract the value of  $\alpha_S$ .

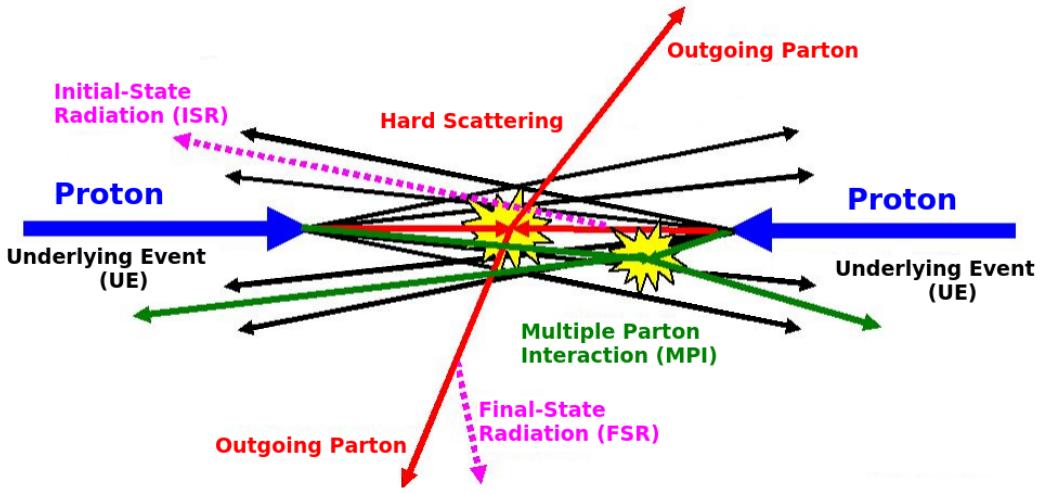


Figure 2.7: A proton-proton collision<sup>9</sup> involving the main hard scattering process along with the low momentum transfer underlying event (UE) contributions coming from initial- and final-state radiations (ISR and FSR) complemented with multiple parton interactions (MPI) and collisions from leftover partons called beam remnants.

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 search 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  $\alpha_S$ . In the next section, we focus on the definition of a jet.

## 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

<sup>9</sup>Source : [The Energy Dependence of Min-Bias and the Underlying Event at CDF](#)

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 are followed in the form of jet algorithms.

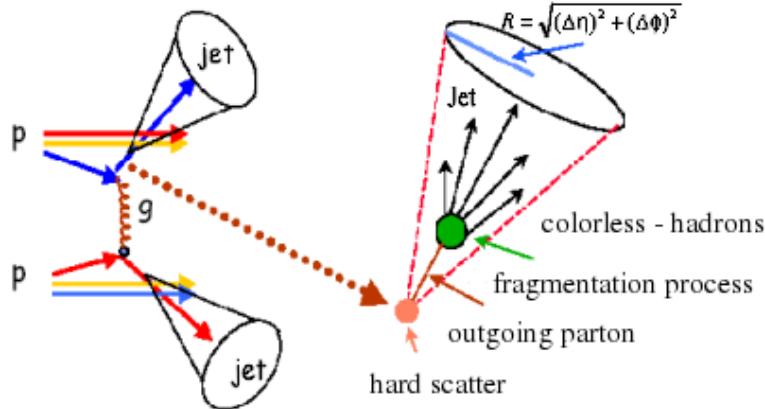


Figure 2.8: In a proton-proton collision, the outgoing partons of the hard scattering process undergo fragmentation and hadronization processes producing a shower of partons which get collimated into a conical jet with radius R.

### 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 associ-

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-section 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

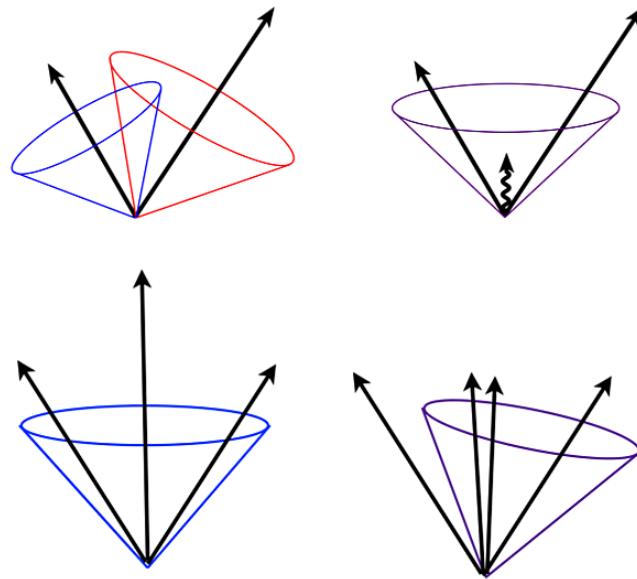


Figure 2.9: Top : Infrared unsafe behaviour of jet algorithm is illustrated where the presence of soft radiation between two jets may cause a merging of the jets that would not occur in the absence of the soft radiation. Bottom : Collinear unsafe behavior of jet algorithm is shown in which the number of jets change due to a collinear splitting<sup>10</sup>.

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

<sup>10</sup>Source : <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  $\eta\text{-}\phi$  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 an 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.

**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 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

particle to the beam are calculated.

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

(2.18)

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

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 produce jets with 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

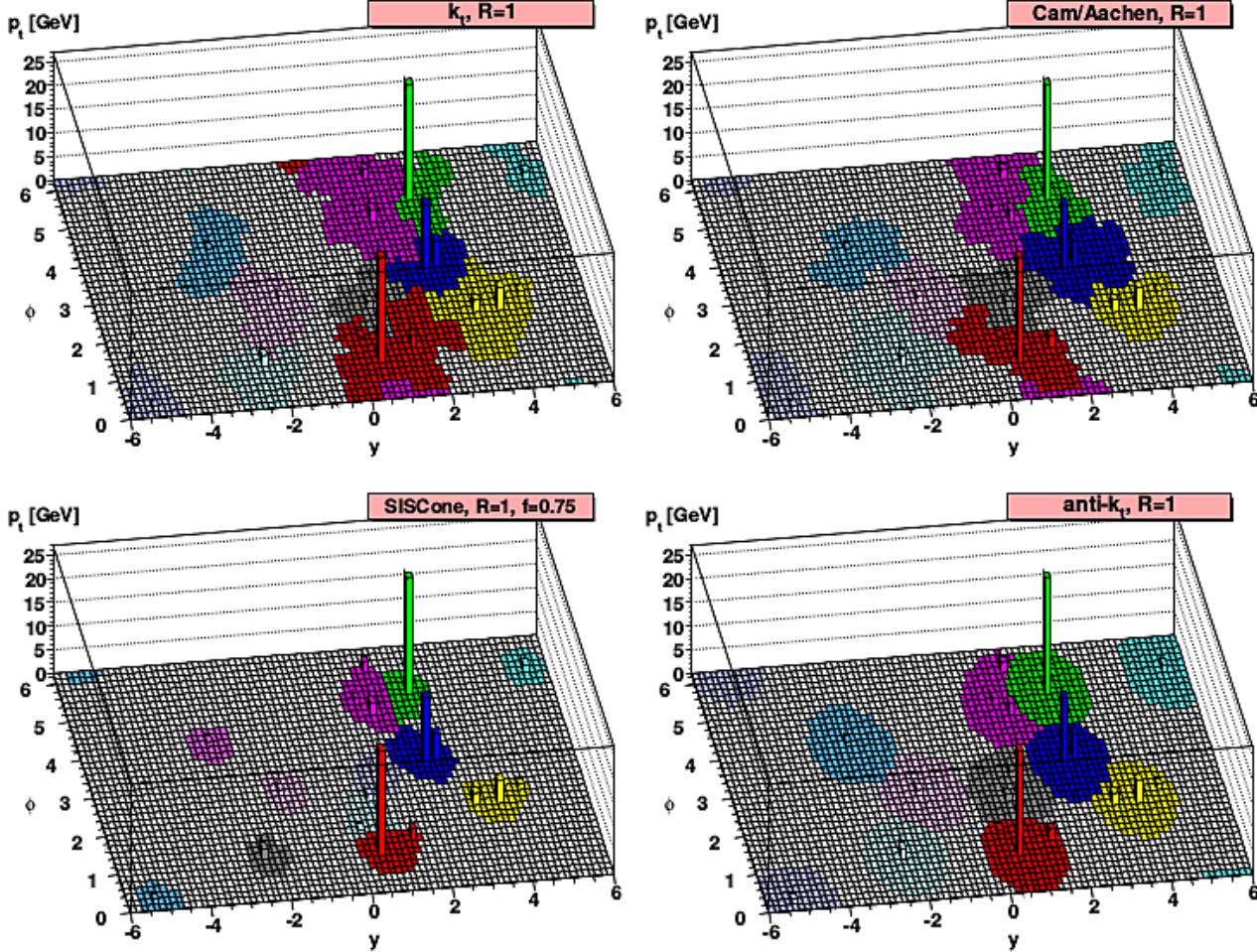


Figure 2.10: The clustering of particles, in  $y$ - $\phi$  space at the parton level, into jets clustered with the  $k_t$  (top left), Cambridge/Aachen (top right), SIScone (bottom left) and anti- $k_t$  (bottom right) algorithms with  $R = 1$ . The towers represent the jet  $p_T$ . The anti- $k_t$  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 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

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.



# Chapter 3

## Experimental Setup

The hadron colliders aim at search for elementary particles and study the dynamics of interactions between them as predicted by the Standard Model or beyond the Standard Model theories. For the same beam energy, higher center-of-mass energy can be achieved in the hadron colliders as compared to the fixed target experiments. Due to the availability of very high center-of-mass energy of the colliding hadrons, it becomes possible for the researchers to understand the fundamental structure of the universe deeply and to look back in its history. The masses of the  $Z$  and  $W$  bosons, discovered by the UA1 and UA2 experiments [48, 49], were measured precisely at the Large Electron-Positron (LEP) collider. The experiment at proton-antiproton collider, Tevatron at FNAL discovered the top quark and also measured its mass. The search for the long awaited Higgs boson was carried out by the currently running most powerful accelerator, the Large Hadron Collider (LHC). Still many questions related to the the nature of dark matter, the existence of super-symmetry (SUSY) or the extra dimensions, are yet to be answered.

The European Organization for Nuclear Research (CERN) is a world-class fundamental physics research organization founded in 1954. In the beginning, it concentrated on pure physics research to understand the inside of the atom, justifying the word “nuclear” in name. At present, the main area of research at CERN

is particle physics which studies the fundamental constituents of matter and their forces of interactions. To accomplish this task, several particle detectors have been built by CERN which probe the physics at the TeV energy scale.

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [50] is the world's biggest and the most powerful particle accelerator and collider built by CERN. It occupies the circular tunnel between the border of France and Switzerland which was previously used by LEP collider [51]. The circumference of the tunnel is 27 km and it lies underground at a depth ranging from 50 to 175 metres. The two beams of particles are accelerated in directions opposite to each other. There are 1,232 dipole magnets to maintain the beams in their circular path. The additional 392 quadrupole magnets are present to keep the beams focused to increase the probabilities of interaction between the particles. The LHC mainly collides proton beams, but it also collides proton-lead and lead-lead as well as xenon-xenon nuclei at different centre-of-mass energies. Since this thesis presents the study done using proton-proton (pp) collisions data, we will discuss the acceleration of protons only.

The protons pass through a series of accelerators which increase their energy successively before their injection into the main ring of LHC. An overview of the various accelerators and detectors comprising the complex structure of the LHC is shown in Fig. 3.1. A bottle of hydrogen gas is the source of protons. The stripping of electrons from hydrogen gas atoms using an electric field yields protons. These protons are accelerated up to 100 keV through a radiofrequency quadrupole which provides the first focusing and a further acceleration to 750 keV energy. The linear particle accelerator (LINAC2) increases the energy of protons to 50 MeV. Then these protons are injected into the Proton Synchrotron Booster (PSB) in the form of bunches where they get accelerated to 1.4 GeV energy. The Proton Synchrotron

(PS) further enhances the energy of protons to 25 GeV which is then increased to 450 GeV by the Super Proton Synchrotron (SPS). Finally the protons are injected into two beam pipelines of the main LHC ring where their energy increases to the beam energy. The total center-of-mass energy in head-on collisions between beams of same mass particles is twice the energy of the beams.

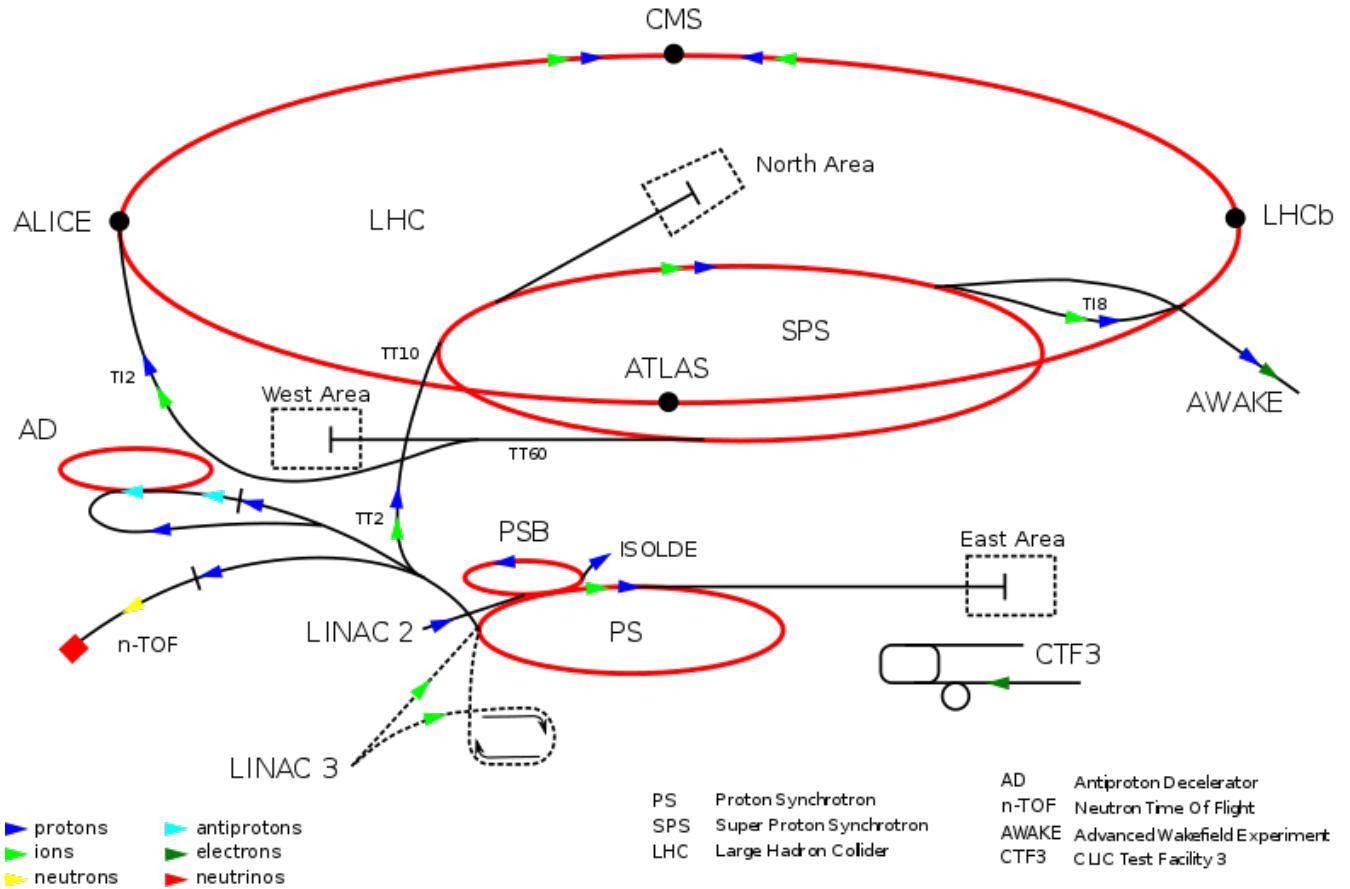


Figure 3.1: An overview of the different experiments of the Large Hadron Collider (LHC), a complex particle accelerator and collider located at CERN<sup>11</sup>.

The accelerated beams collide at four interaction points around which six detectors are located : ALICE (A Large Ion Collider Experiment) [52], ATLAS (A Toroidal LHC Apparatus) [53], CMS (Compact Muon Solenoid) [54–56], LHCb (Large Hadron Collider for Beauty) [57], LHCf (Large Hadron Collider forward) [58]

<sup>11</sup>Source : [https://en.wikipedia.org/wiki/Large\\_Hadron\\_Collider](https://en.wikipedia.org/wiki/Large_Hadron_Collider)

and TOTEM (Total, elastic and diffractive cross-section measurement) [59]. The CMS and ATLAS are two general purpose detectors dedicated to the validation of the Standard Model theory predictions, existence of super-symmetry (SUSY) and also looking for extra dimensions. The ALICE is a heavy-ion detector which studies quark-gluon plasma, a state of matter believed to be present just after the Big Bang, produced in collisions of lead ions. The LHCb experiment will explore the differences between matter and antimatter and new physics through b-quark (beauty) studies. TOTEM experiment is dedicated to cross-section measurements whereas LHCf focuses on forward physics.

The LHC successfully injected the first protons on September 10, 2008 but after few days there was magnetic quench in bending magnets which lead to a loss of  $\sim 6$  tonnes of liquid helium. After recovery from this incident, at first the low-energy beams were circulated in the tunnel on November 20, 2009 and after three days the first collisions took place in all four detectors at  $\sqrt{s} = 450$  GeV. The LHC achieved 1.18 TeV energy per beam on November 30, 2009. This made LHC the world's highest energy particle accelerator and left behind the Tevatron having record of 0.98 TeV per beam for eight years. The LHC recorded pp collisions at  $\sqrt{s} = 2.36$  TeV around December 15, 2009. After this the beam energy was ramped up to 3.5 TeV on March 19, 2010 which resulted in the first pp collisions at  $\sqrt{s} = 7$  TeV on March 30, 2010. The beam energy was kept at 3.5 TeV throughout 2011, and increased to 4 TeV in 2012. After a long shutdown for two years, the LHC restarted in 2015 and collided the proton beams at a much higher center-of-mass energy of 13 TeV and is running successfully till now. In the coming years, protons will be made to collide at a designed  $\sqrt{s} = 14$  TeV with luminosity up to  $10^{34}$   $\text{cm}^{-2}\text{s}^{-1}$ . In this thesis, work has been carried out using the pp collisions data collected by the CMS detector at  $\sqrt{s} = 8$  TeV in the year of 2012.

### 3.1.1 Luminosity Measurement

Luminosity ( $\mathcal{L}$ ) is one of the most important parameters of an accelerator which characterizes its performance. It defines the rate at which collisions occur and is given by the number of collisions produced in a detector per  $\text{cm}^2$  and per second. Cross-section ( $\sigma$ ) is a measure of the probability that an event can take place.  $\mathcal{L}$  is related to total number of events  $N$  of a process over a time period  $T$  and  $\sigma$  as :

$$N = \int_0^T \mathcal{L} \sigma dt = \mathcal{L}_{int} \sigma \quad (3.1)$$

where  $\int_0^T \mathcal{L} dt = \mathcal{L}_{int}$  gives the total integrated luminosity.  $\mathcal{L}_{int}$  is expressed in the units of area, usually in  $\text{barn}^{-1}$  and gives a direct indication of number of events produced in a process. For example, an integrated luminosity of  $10 \text{ fb}^{-1}$  means that 10 events are produced in a process having cross-section equal to  $1 \text{ fb}$ .

The luminosity depends on the particle beam parameters as :

$$\mathcal{L} = \frac{N_b N_p^2 f_{rev} \gamma F}{4\pi \epsilon_n \beta^*} \quad (3.2)$$

where  $N_b$  is the number of bunches per beam,  $N_p$  is the number of particles in each bunch,  $f_{rev}$  is the revolution frequency of the beam,  $\gamma$  is the relativistic gamma factor and  $F$  is the geometric luminosity reduction factor. The effective collision area of the two beams is related to the normalized transverse beam emittance  $\epsilon_n$  and the value of the betatron function  $\beta^*$  at the interaction point.

The CMS experiment constantly monitors the instantaneous luminosity delivered by LHC which is shown versus time in Fig. 3.2 for proton-proton collisions at nominal center-of-mass energy for the years 2010-2017. The relative instantaneous luminosity is calculated by using two methods [60] : Hadron Forward (HF) method by measuring the particle flux in the hadron forward calorimeter and Counting

method where the number of reconstructed vertices in the pixel tracker are counted. The measurement of the absolute luminosity is performed using van-der-Meer scans done in special runs of the LHC [61]. The uncertainty on the measured luminosity for 2012 data set is 2.5% (syst.) and 0.5% (stat.).

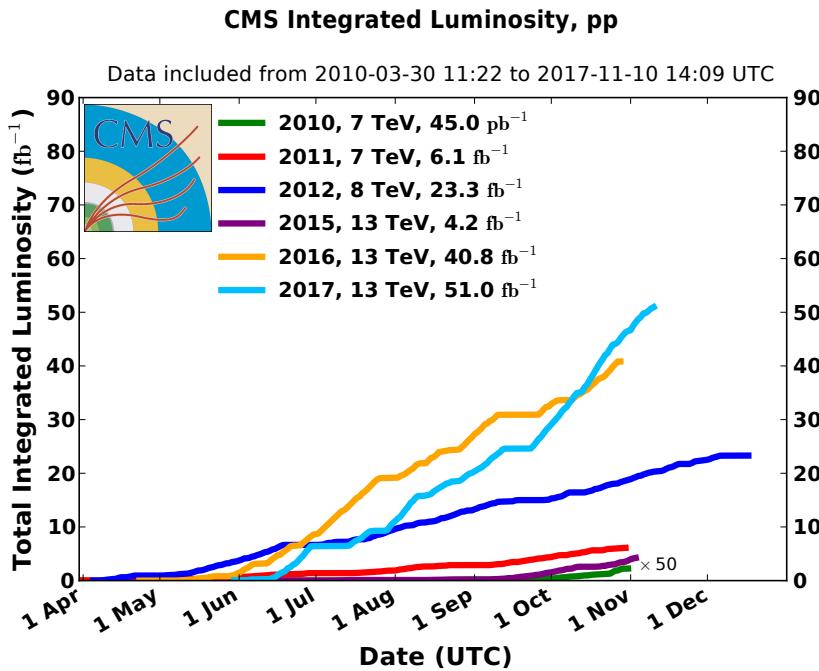


Figure 3.2: The integrated luminosity delivered by stable beams to CMS during proton-proton collisions taking place at nominal center-of-mass energy, is shown versus time for data-taking in 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange) and 2017 (light blue) run periods of the LHC<sup>12</sup>.

### 3.1.2 Pileup Interactions

To observe the extremely rare events, the event rate in a collider should be very high. This demands delivered luminosity to be high which is achieved by increasing the number of bunches or increasing the number of protons per bunch. However, this comes at the cost of multiple proton-proton interactions coming from independent hadron-hadron collisions occurring in the same bunch crossing, called pileup (PU)

<sup>12</sup>Source : <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>

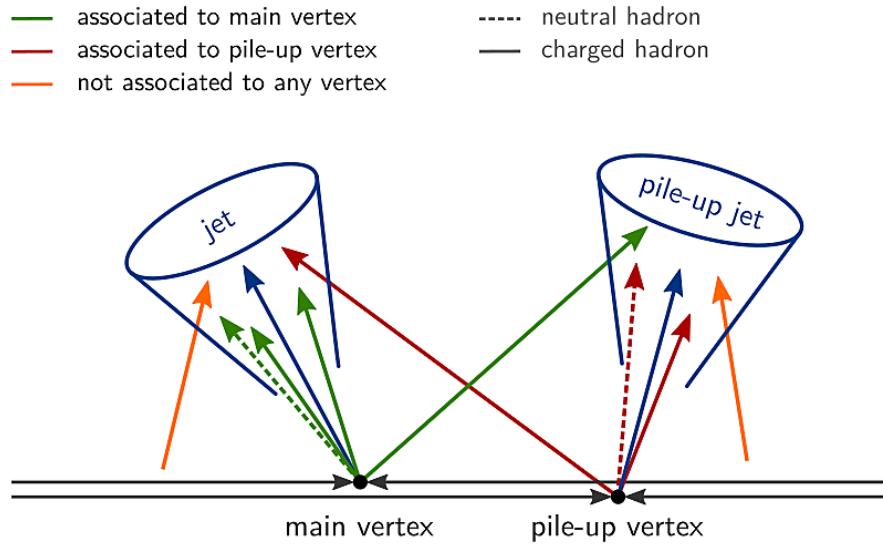


Figure 3.3: In a proton-proton collision, the particles produced from the hard interaction are clustered into a jet. The hard interaction corresponds to the main vertex. The particles produced in the interactions other than the hard one, form a pileup jet<sup>13</sup>.

interactions. The hard interaction in every event is accompanied by a large amount of PU interactions which give rise to low  $p_T$  jets. The vertex of pileup interaction is reconstructed from tracks pointing to it as shown in Fig. 3.3. The pileup due to the additional collisions within a single bunch crossing is called in-time pileup whereas pileup coming from collisions other than hard scattering in other bunch crossings is known as out-of-time pileup. The pileup itself cannot be directly measured, it can be correlated to various other directly measurable quantities. The number of primary vertices ( $N_{PV}$ ) is directly related to the amount of pileup as the pileup comes from the additional proton-proton interactions. The greater the  $N_{PV}$ , the more pileup energy is added to the jets which needs to be subtracted.

<sup>13</sup>Source : <http://cds.cern.ch/record/1747055>

## 3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) detector is a general purpose detector located at the interaction point 5 (P5) of the main LHC ring, near the village of Cessy in France. The name of CMS comes from its compact size with main emphasis on the detection of muons and enclosed within high solenoidal magnetic field. The CMS detector aims at identifying the different types of particles produced in proton-proton and heavy ion collisions and measuring their energies and momenta. This is achieved by concentric layers of different sub-detectors arranged in a cylindrical complex structure with 21.6 m length and 15 m diameter. The silicon-based tracker surrounds the the interaction point and forms the innermost layer. After the tracker, there are layers of a scintillating crystal electromagnetic calorimeter (ECAL) and a sampling hadron calorimeter (HCAL). The calorimeters are enclosed inside the superconducting solenoid. Outside the magnet lies the large muon detectors embedded inside an iron yoke. The three dimensional view of the CMS detector along with its components is presented in Fig. 3.4. The CMS was constructed in parts at ground and assembled later on in the cavern. The components are easily accessible for upgrades or repairs as the detector can be opened up into movable slices. Figure 3.5 shows the front view of the CMS detector differentiating individual components which contribute to event reconstruction. The dashed and solid lines represent the invisible and visible tracks, respectively, of the reconstructed particles. The different particles are : photons ( $\gamma$ ), muons ( $\mu^\pm$ ), electrons ( $e^-$ ), neutrons (n) and charged hadrons ( $\pi^\pm$ ).

A brief overview of the CMS detector has been presented and the details of the its design as well as physics performance are available in Ref. [55,56]. Before going into the details of each sub-detector, first the CMS coordinate system is described in the next section.

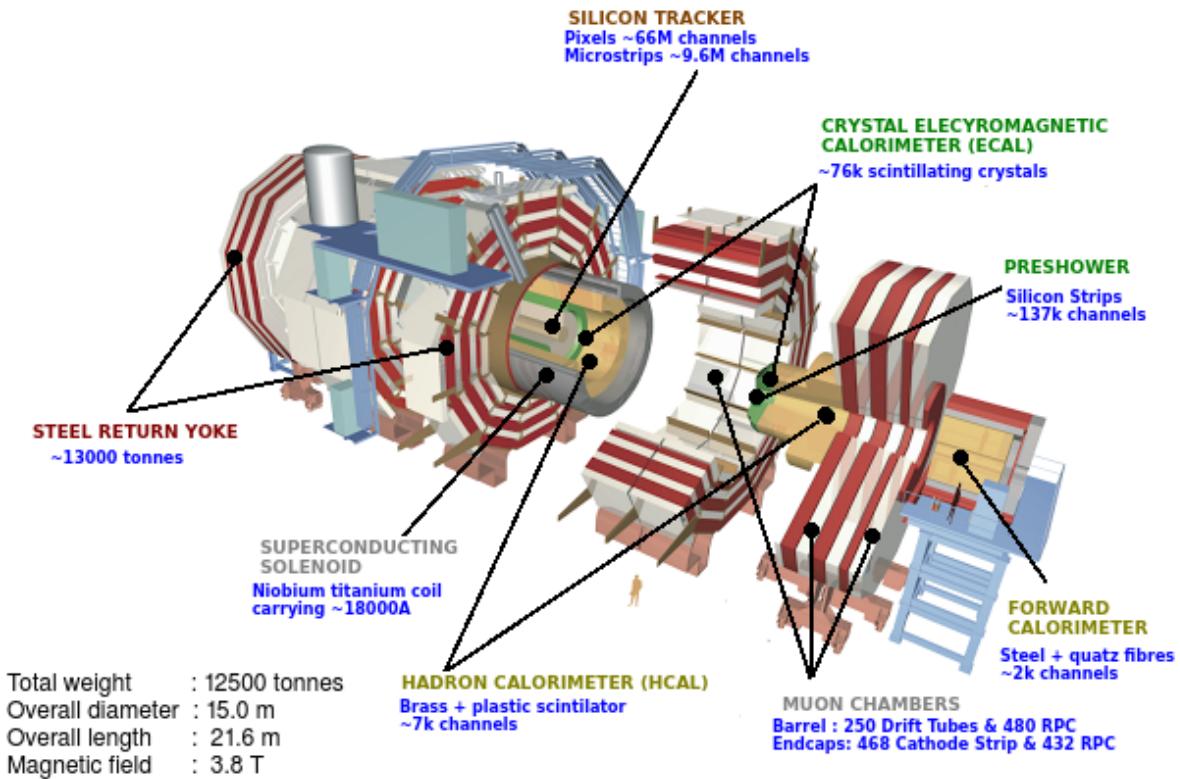


Figure 3.4: The three dimensional view of the CMS detector along with its sub-detector components<sup>14</sup>.

### 3.2.1 Coordinate System

A right-handed coordinate system, illustrated in Fig. 3.6, is used by the CMS detector. The origin of the co-ordinate system lies at the nominal interaction point (IP) of the collisions. The  $x$ -axis points horizontally from the IP and towards the center of the LHC ring. The  $y$ -axis points vertically upwards and the  $z$ -axis along the beam direction towards the Jura mountains. Following customary polar coordinate conventions : the azimuthal angle  $\phi$  is measured from the  $x$ -axis in the  $x$ - $y$  plane as  $\phi = \tan^{-1}(\frac{y}{x})$  where  $\phi = 0$  points to the  $+x$  axis and  $\phi = \pi/2$  points to the  $+y$  axis. The polar angle  $\theta$ , is calculated from the  $z$ -axis in the  $z$ - $y$  plane as  $\theta = \tan^{-1}(\sqrt{\frac{x^2 + y^2}{z^2}})$  with  $\theta = 0$  corresponding to the  $+z$  direction and  $\theta = \pi$  to the

<sup>14</sup>Source : <https://orbiterchspacenews.blogspot.in/2013/04/cern-cms-prepares-for-future.html>

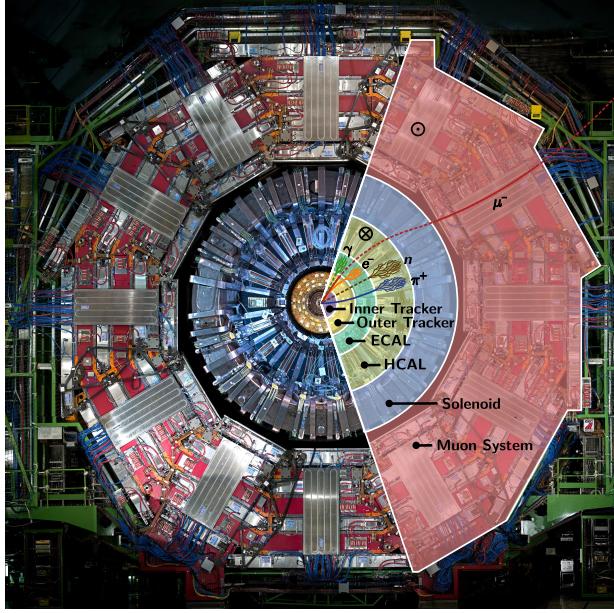


Figure 3.5: Front view of the CMS detector along with its components : inner tracker, outer tracker, electromagnetic calorimeter, hadronic calorimeter, solenoid and muon system. The path of different particles detected by dedicated sub-detectors are shown by dashed (invisible track) and solid (visible track) lines.  $\otimes$  and  $\odot$  gives the direction of magnetic field inside the solenoid and in the return yoke, respectively. Taken from [56].

$-z$  direction. The quantities rapidity  $y$  and the pseudorapidity  $\eta$  are preferred over the angles  $\theta$  and  $\phi$ . The rapidity and pseudorapidity are given by Eq. 3.3. Both the quantities are equal for massless particles.

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (3.3)$$

$$\eta = - \ln \left( \tan \left( \frac{\theta}{2} \right) \right)$$

The difference between rapidities  $\Delta y$  is invariant under longitudinal Lorentz boost whereas it does not hold for  $\eta$ . Hence  $y$  is considered in this thesis. The angular distance between the two particles is defined by  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ . The momentum component transverse to the direction of beam  $p_T$ , is computed from

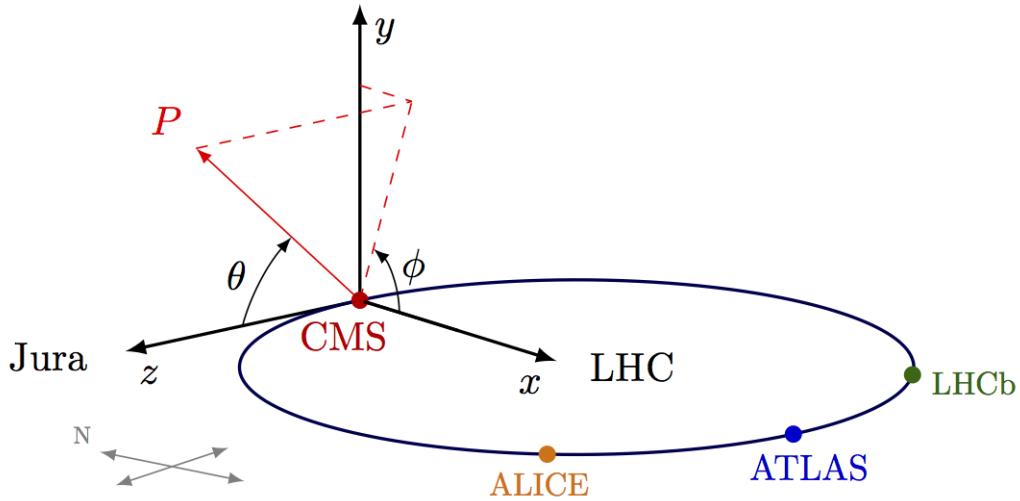


Figure 3.6: The CMS detector uses the right-handed coordinate system<sup>15</sup> having origin at the interaction point (IP). The  $x$ -axis points horizontally from the IP towards the center of the LHC ring, the  $y$ -axis points vertically upwards whereas the  $z$ -axis along the beam direction towards the Jura mountains. The azimuthal angle  $\phi$  is measured from the  $x$ -axis in the  $x$ - $y$  plane and the polar angle  $\theta$  is calculated from the  $z$ -axis in the  $z$ - $y$  plane.

the  $x$ - and  $y$ -components as  $p_T = \sqrt{p_x^2 + p_y^2}$  and the transverse energy is given by  $E_T = E \sin\theta$ . After introducing the CMS coordinate system, further the detector subsystems are described briefly in the following sections. In Fig. 3.7, a longitudinal section of the CMS detector shows the location of different sub-systems along with the superconducting solenoid, in the  $y$ - $z$  plane.

### 3.2.2 Inner Tracker System

The charged particles produced from the LHC collisions leave their trajectories as they move outward from the interaction point. The particle flux within the detector decreases as  $1/r^2$ . So the tracks of the particles need to be measured as close to the collision point as possible and in a precise manner. The innermost tracking system of the CMS consists of silicon detectors and measures the hits produced by the charged particles. It surrounds the interaction point and has a cylindrical volume

<sup>15</sup>Source : [https://wiki.physik.uzh.ch/cms/latex:example\\_spherical\\_coordinates](https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates)

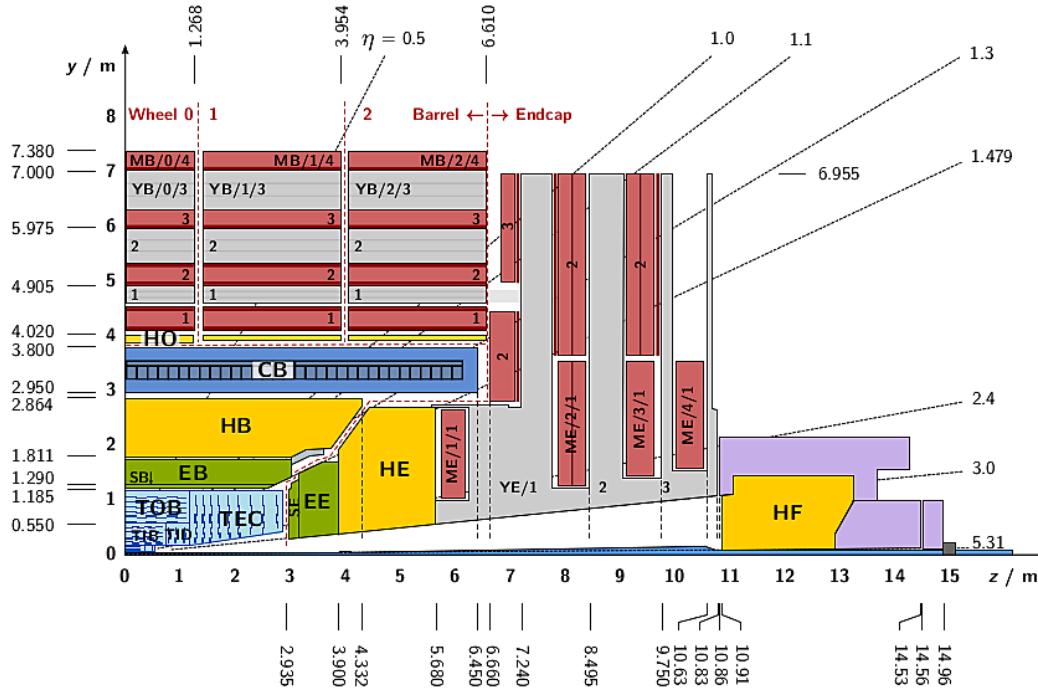


Figure 3.7: A longitudinal view of the CMS detector is shown in the  $y$ - $z$  plane<sup>16</sup>. It shows the tracking detector (TIB, TID, TOB, TEC) close to the nominal interaction point at  $(0,0)$ , the electromagnetic (EB, EE) and hadronic (HB, HE, HO, HF) calorimeters. The coil of the solenoid magnet (CB) surrounds the inner barrel region. The iron return yoke (YB, YE) is interleaved with the muon chambers (MB, ME).

of length of 5.8 m and a diameter of 2.5 m and covers a pseudorapidity range up to  $|\eta| < 2.5$ . The passage of the charged particles through the silicon detector material produces small ionization currents which get detected as hits. Such multiple hits when combined, reconstruct the track which gives the information about the direction and transverse momentum  $p_T$  of the charged particle. Silicon detectors have a much higher resolution in tracking charged particles as compared to the older ones such as cloud chambers or wire chambers. CMS inner tracking system shown in Fig. 3.8 consists of two sub-systems :

**Pixel Detector** - A pixel detector is located close to the beam pipe. It has three

<sup>16</sup>Source : <http://cds.cern.ch/record/1747055>

co-centric barrel layers lying at radii of 4.4, 7.3 and 10.2 cm from the beam pipe. It

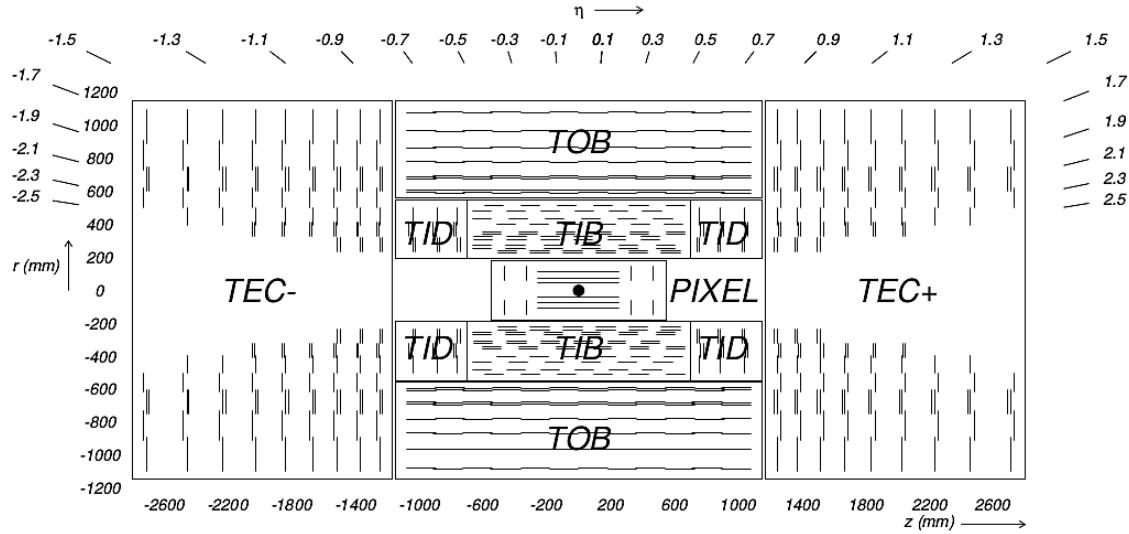


Figure 3.8: A longitudinal view of the inner tracking system is shown in  $r$ - $z$  plane. The CMS tracking system is made up of the silicon pixel detector and the silicon strip detector. The silicon strip detector has four components : The Tracker Inner Barrel (TIB) complemented by the Tracker Inner Disks (TID) which are further surrounded by the Tracker Outer Barrel (TOB) in barrel region. Tracker End Cap (TEC) covers high  $\eta$  ranges up to  $\eta = 2.5$ . Taken from [54].

has two disks of pixel modules on each side of barrel. Taking the design luminosity of LHC i.e.  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , about 1000 particles are produced from more than 20 overlapping proton-proton collisions. These particles traverse through the tracker for each bunch crossing, i.e. every 25 ns. The size of each pixel is  $100 \mu\text{m}$  (in  $r, \phi$ )  $\times$   $150 \mu\text{m}$  (in  $z$ ) which gives an average occupancy of  $10^{-4}$  per bunch crossing. Due to the large Lorentz effect, the pixel tracker has a spatial resolution of  $10 \mu\text{m}$  and  $20 \mu\text{m}$  in  $(r, \phi)$  and  $z$  plane respectively, which is beneficial for a precise determination of the primary and secondary vertices and good momentum resolution.

**Strip Detector** - After coming out of the pixel detector the charged particles traverse through ten layers of silicon strip detectors, reaching out to a radius of 130 cm. The silicon strip detector has four layers of inner barrel (TIB) assembled in shells with two inner endcaps (TID), each having three small discs. There is an outer barrel tracker (TOB) consisting of six concentric layers. Finally two endcaps

(TEC) are placed at the end of tracker. Each part of the tracker has silicon modules which are designed with dedicated functions. The strip detector performs the measurement of the particle tracks with a reduced resolution of  $23\text{ }\mu\text{m}$  which hints the smaller particle flux at larger distances from the collision point. The active silicon area of CMS tracker is about  $200\text{m}^2$  which makes it the largest silicon tracker. Along with the measurement of tracks, the energy also needs to be measured for which the calorimeters are present outside the tracker. To measure the momenta of the particles precisely, they should interact with the tracker to a minimum extent. In contrast, to measure their energy, they are required to interact with the calorimeters fully.

### 3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a homogeneous and hermetic calorimeter used to slow down the produced photons and electrons/positrons and measure their energy by absorbing them into the detector material. The barrel part of the ECAL is made up of 61200 lead tungstate ( $\text{PbWO}_4$ ) crystals and each of the two end caps has 7324 crystals.  $\text{PbWO}_4$  is a very dense material having a short radiation length of  $X_0 = 0.89\text{ cm}$  and covers the pseudorapidity up to  $|\eta| < 3.0$ . The incorporation of oxygen makes it highly transparent and enables to emit scintillation light. The small Molière radius of  $2.19\text{ cm}$  of this material, gives a fine granularity. These properties leads to compact size of ECAL so that it can be easily placed within the solenoid magnet.

When the electrons, positrons or photons produced in the collisions, hit the crystals of ECAL, they produce electromagnetic showers through the subsequent processes of bremsstrahlung and electron-positron pair production. The energy of the particles deposited by the photoelectric effect and the Compton scattering causes excitation of the material atomic state and the emission of photons. The number of emitted photons is directly proportional to the energy of the incident particles.

The emitted photons are detected by silicon avalanche photo diodes (APDs) in the barrel region and vacuum phototriodes (VPT) in the end-cap region. Figure 3.9 presents a geometric view of ECAL in the  $y$ - $z$  plane showing the arrangement of different parts of ECAL : the ECAL barrel (EB) extending up to  $|\eta| < 1.479$  using more than 60000 crystals and ECAL endcaps (EE) covering the region  $1.479 < |\eta| < 3.0$  with an additional 15000 crystals. The preshower detectors (ES) made of lead absorbers and silicon detectors are put in front of the endcaps to distinguish high energetic single photons from low energetic photon pairs originating from neutral pions decays.

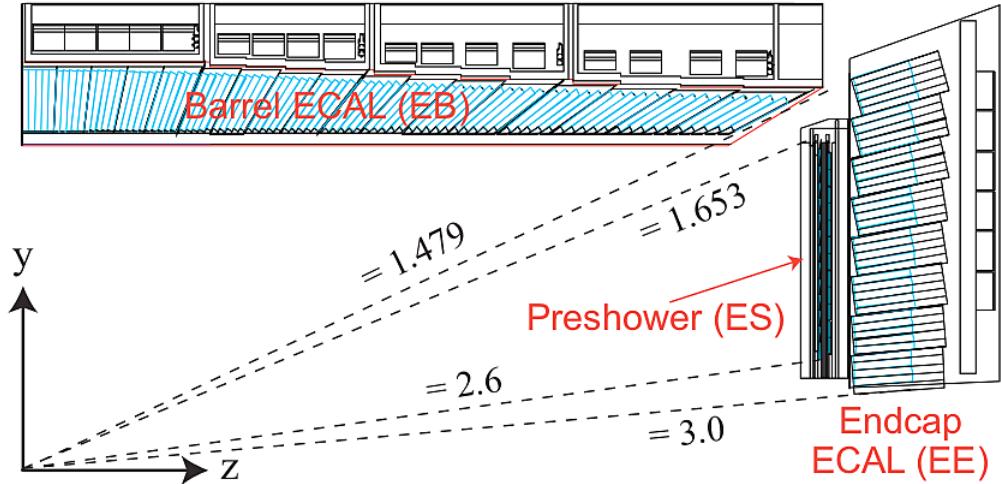


Figure 3.9: A geometric view of one quarter of the electromagnetic calorimeter (ECAL) in  $y$ - $z$  plane showing the arrangement of sub-modules covering the barrel region (EB) and the endcaps (EE). ECAL is complemented with preshower detector (ES) mounted in front of the endcaps. Taken from [55].

The relative energy resolution of the ECAL has been measured to be [62] :

$$\left( \frac{\sigma(E)}{E} \right)^2 = \left( \frac{2.8\%}{\sqrt{E}} \right)^2 + \left( \frac{12\%}{E} \right)^2 + \left( 0.30\% \right)^2 \quad (3.4)$$

where  $E$  is the energy in GeV. The first term is the stochastic component caused by fluctuations in the energy deposited in the preshower absorber and lateral shower containment. The second term is the contribution by noise and the last is the

constant term which comes from leakage of energy from the back of the crystal, inter-calibration errors and non-uniformity of the longitudinal light collection.

### 3.2.4 Hadronic Calorimeter

At CMS, the major fraction of the particles produced in proton-proton collisions is hadrons. These are usually collimated in a given direction producing conical structures called jets. The jets have both hadronic (charged and neutral) and electromagnetic components which are detected and measured by the combined CMS calorimeter system. The calorimeters are designed in a way that a particle loses all of its energy as it travels through them. The calorimeters measure the energies and directions of particle jets which indirectly give the energies and directions of quarks, gluons and neutrinos, initiating the jets. The energy deposits and the locations of these deposits are used to determine the directions and momenta of charged particles. But the neutral hadrons do not leave any track and hence their energy cannot be measured directly. The energy of neutral hadrons is measured by taking into account the missing transverse energy ( $E_T^{\text{miss}}$ ). The determination of  $E_T^{\text{miss}}$  is a crucial tool in searching the new particles and new physics phenomena. Here, the hadron calorimeter (HCAL) comes into play which can detect neutral particles with non-zero mass such as neutrons. Since the neutral hadrons carry  $\sim 10\%$  energy of the total jet energy, HCAL is an essential sub-system of the CMS detector and contributes to most of CMS's physics studies.

HCAL is a sampling calorimeter installed inside the solenoid coil. It consists of a non-magnetic brass absorber with a short interaction length of  $\lambda_I = 16$  cm and is interleaved with plastic scintillators having wavelength-shifting (WLS) fibres as readout. The highly energetic hadrons further produces a large number of pions and nucleons by inelastic interactions. The hadronic shower spreads more than the electromagnetic showers because of large transverse momentum of the secondary particles. As the energy of the particles is lower than a certain threshold, the ion-

ization and low-energy hadronic processes come into play. The active scintillation material excites and blue-violet light is emitted. The wavelength shifters connect

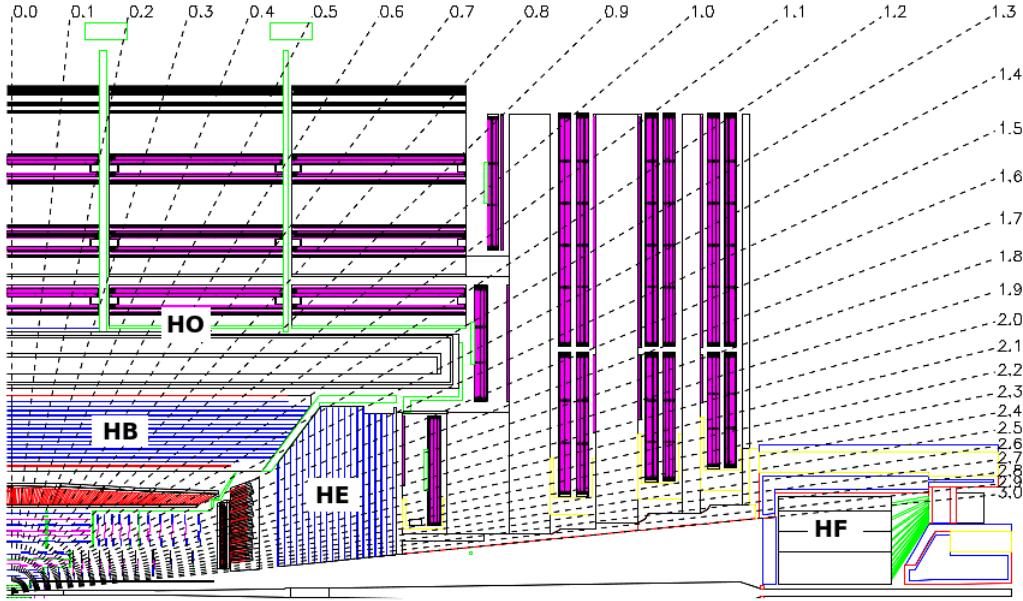


Figure 3.10: Longitudinal section of one quarter of the hadronic calorimeter (HCAL) in  $r\text{-}\eta$  plane. It consists of different parts : hadron barrel (HB), hadron outer (HO) and hadron endcap (HE) and hadron forward (HF). Taken from [54].

all scintillators to photodiodes and read out the signals and further pass them to the data acquisition system. The longitudinal view of one quarter of the HCAL presented in Fig. 3.10 shows the different parts :

**Hadron Barrel -** The hadron barrel (HB) is divided into two identical half barrel sections on either side of the interaction point. Each half barrel is made of 18 azimuthal wedges which are further divided into four azimuthal sectors each giving a granularity of  $\Delta\phi = 0.087$ . In  $z$  direction, the plastic scintillators are divided into 16 intervals of granularity  $\Delta\eta = 0.087$ . HB covers the region up to  $|\eta| < 1.305$  and overlaps with endcaps for  $1.305 \leq |\eta| \leq 1.392$ . Since HB has the highest resolution ( $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ ), it is optimal for calibration of the jet energy scale. The thickness of the HCAL amounts to 7-11 interaction lengths which are sufficient

enough to absorb most of the hadrons.

**Hadron Outer** - The total amount of material in barrel region to absorb the hadronic shower is not sufficient. This requirement is fulfilled by placing an outer hadron (HO) calorimeter as a tail catcher on top of the coil of the magnet. The HO uses the solenoid coil as an additional absorber having interaction lengths of  $1.4/\sin\theta$  and measures the tails of hadron showers penetrating the HB and the coil. Since the HO is physically located inside the muon system, it is strongly constrained by its geometry. The muon system is subdivided into 5 rings along the  $z$ -axis. Each of these rings is 2.536 m wide in  $z$ -direction and the HO is placed as first sensitive layer in these rings, with a scintillator thickness of 10 mm. The central ring ( $\eta = 0$ ) has two scintillator layers placed on each side of 19.5 cm thick iron layer.

**Hadron Endcap** - The hadron endcaps (HE) extend the pseudorapidity range up to  $|\eta| < 3.0$ . About 34% of the particles produced in the final state reach this region. The granularity in  $\Delta\eta \times \Delta\phi$  is  $0.087 \times 0.087$  up to  $|\eta| < 1.6$  and  $0.17 \times 0.17$  for  $|\eta| > 1.6$ . The main challenges faced in the construction of the HE were the use of non-magnetic material in order to not disturb the magnetic field and the close distance to the beam line. The continuous damages caused by radiations decrease the detector response which should be monitored at regular intervals.

**Hadron Forward** - The hadron forward (HF) calorimeter lies at a distance of  $z = \pm 11.2$  m from the interaction point, covering the  $2.8 < |\eta| < 5.2$  region. The HF has a cylindrical structure with an outer radius of 130.0 cm. It is azimuthally subdivided into 36,  $20^\circ$  modular wedges. The HF is made up of 5 mm thick grooved steel plates which have quartz fibers inserted into the grooves. The fibres running parallel to the beam line are bundled to form  $0.175 \times 0.175$  ( $\Delta\eta \times \Delta\phi$ ) towers. The HF detects the jets having very high  $\eta$  and also the hadronization products of the beam remnants. The iron absorbers and quartz fibers act as active material to measure the emitted Cerenkov light and to produce the signal in the photomultipliers

(PMT).

The relative hadronic energy resolution of the barrel HCAL and ECAL combination can be parametrized as :

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + b^2 \quad (3.5)$$

where  $a$  is a stochastic term and  $b$  is a constant term. These values have been measured [63] as  $a = (0.847 \pm 0.016) \sqrt{\text{GeV}}$  and  $b = 0.074 \pm 0.008$  whereas for HF the measured values are  $a = 1.98 \sqrt{\text{GeV}}$  and  $b = 0.09$ .

### 3.2.5 Superconducting Magnet

The superconducting magnet is an essential feature of the CMS detector which is 13m long and 6m in diameter. Its refrigerated superconducting high-purity aluminium-stabilized niobium-titanium coils cooled at 4 Kelvin produces a magnetic field of 4 Teslas (T). The magnet will run at 3.8 T in order to maximize its lifetime. This intense solenoidal field makes the compactness and cylindrical symmetry of the detector possible. The magnet is placed between the calorimeters and the muon system. The solenoidal magnetic field parallel to the beam bends the tracks of the high momentum charged particles in the transverse plane. The curvature of the trajectory increases with the strength of the magnetic field which make possible to determine the transverse momentum more precisely. The magnet is complemented by an iron yoke ( $\sim 10000$  tonnes) which returns the magnetic field at 2 T.

### 3.2.6 Muon System

As the name of CMS suggests, the detection of muons is of central importance in the CMS detector. Out of all the known stable particles, only the muons and neutrinos pass through the calorimeter without depositing their energies. They interact

very little with matter and can travel long distances through the dense matter. The charged muons can be detected by having an additional tracking system outside the calorimeters whereas the neutrinos are practically undetectable as they escape completely without being tracked in any of the layers of the calorimeters. Their presence can be detected from the missing energy carried by them. The CMS muon system is installed outside the calorimeters in the iron return yoke of the magnet which acts as a hadron absorber for the identification of the muons. Along with the muon identification, the muon system also measures their momenta and is also used for the triggering purposes. Good muon momentum resolution and trigger capability are enabled by the high-field solenoidal magnet and its flux-return yoke. The CMS muon system measures the momenta and charge of muons over the entire kinematic range of the LHC. The muon system shown in Fig. 3.11 consists of three types of gaseous particle detectors :

**Drift Tube** - The muon barrel (MB) detector has four concentric layers of drift tube (DT) chambers inside the iron yoke which covers the region up to  $|\eta| < 1.2$ . DT stations are distributed into 5 wheels along the  $z$  direction. Each wheel is divided into 12 sectors, each sector covering a  $30^\circ$  azimuthal angle. The DT is an aluminium tube having length of 2.5 m and area of  $4.2 \times 1.3 \text{ cm}^2$ . It is filled with a gas mixture consisting of 58% Ar + 15 % CO<sub>2</sub>.

**Cathode Strip Chambers** - In the forward region, the muon and background flux is higher. In this region, cathode strip chambers (CSC) are preferred because of their fast response time, high radiation tolerance and fine segmentation. In each end cap, four stations of CSCs are installed which cover the region of  $0.9 < |\eta| < 2.4$ . Each CSC is trapezoidal in shape and consists of 6 gas gaps. Each gap has a plane of radial cathode strips and a plane of anode wires lying in perpendicular direction to the strips.

**Resistive Plate Chambers** - Both DT and CSC are accompanied by resistive

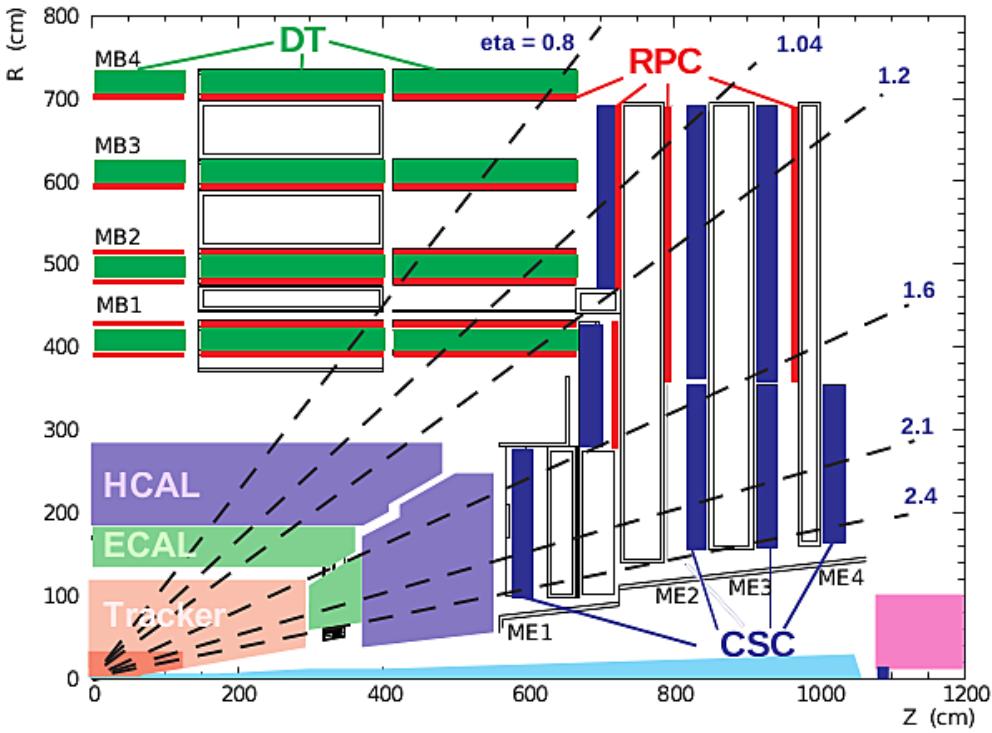


Figure 3.11: A longitudinal view of the CMS muon system showing the location of the three gaseous particle detectors : four Drift Tube (DT) stations in the barrel (MB1-MB4, green), four stations of Cathode Strip Chambers (CSC) in the endcap (ME1-ME4, blue), and the Resistive Plate Chambers (RPC) stations (red)<sup>17</sup>.

plate chambers (RPC) which are double-gap chambers. RPCs operate in avalanche mode to ensure good performance at high rates. They help to resolve ambiguities in attempting to make tracks from multiple hits in a chamber. They also provide additional points for determination of a muon trajectory and give fast response to the trigger system which is described in the following section.

### 3.2.7 Trigger and Data Acquisition System

At the LHC, the interaction rates in proton-proton collisions are very high. In the 2012 run period, the beam crossing frequency was 25 ns. At this frequency, around 40 million bunch crossings occur per second with an average of around 20 collisions

<sup>17</sup>Source : <https://arxiv.org/abs/1209.2646>

per bunch crossing. But the rate at which the information can be stored is much lower than collision rates. Hence, either the storage rate should be increased or event rates should be decreased. This is achieved by using an efficient trigger system which retains the interesting signal events and rejects the background events. An event should be accepted or rejected very quickly, based on signals of certain physics objects inside the detector. CMS has a two-level complex trigger system :

**Level-1 Trigger** - The Level-1 (L1) trigger system is based on custom electronics which stores the events at maximum rate of 100 kHz and then forward them to the next level triggers. The L1 system uses only coarsely segmented data from calorimeter and muon detectors and holds all the high-resolution data in pipeline memories in the front-end electronics. The work flow of the L1 trigger system, consisting of local, regional and global components, is shown in Fig. 3.12. The local triggers known

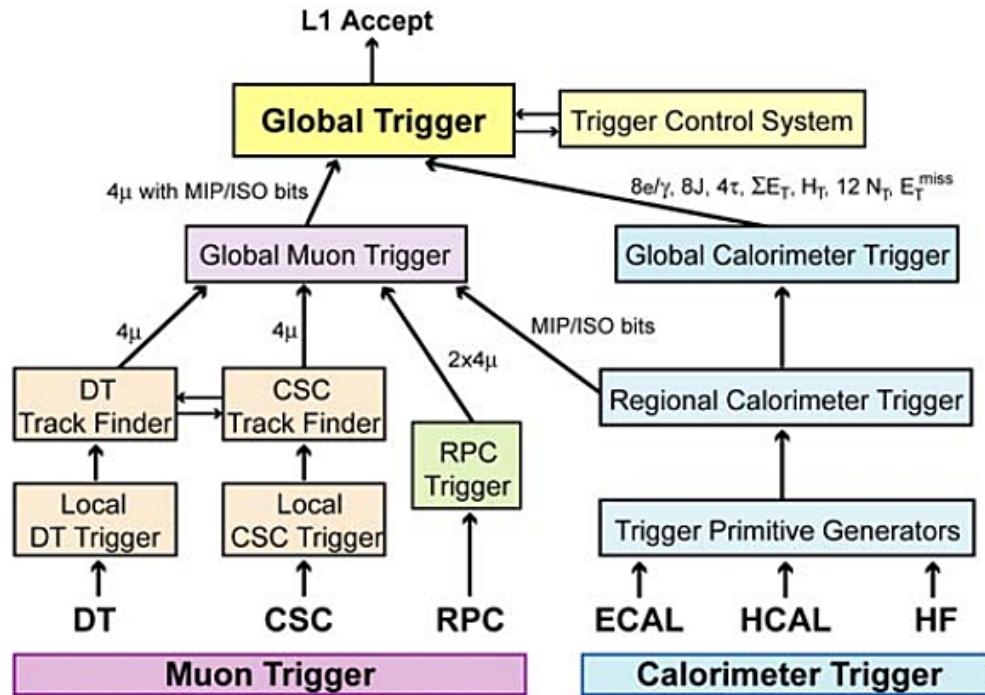


Figure 3.12: Work flow of the L1 trigger system consisting of local, regional and global components. Taken from [54].

as Trigger Primitive Generators are based on energy deposits in calorimeter trigger

towers and tracks in muon chambers. The Regional Triggers combine their information and use pattern logic to determine ranked and sorted trigger objects such as electron or muon candidates in limited spatial regions. The rank is determined as a function of energy or momentum and quality, which reflects the level of confidence attributed to the L1 parameter measurements, based on detailed knowledge of the detectors and trigger electronics and on the amount of information available. The Global Calorimeter and Global Muon Triggers determine the highest-rank calorimeter and muon objects and transfer them to the Global Trigger (GT), the top entity of the Level-1 hierarchy. The events accepted by the GT are further evaluated by the HLT.

**High Level Trigger -** At the second step, a software-based High-Level Trigger (HLT) reduces the maximum L1 accepted rate of 100 kHz to a final output rate of 100 Hz. The HLT system filters events by performing physics selections based on the offline reconstruction software. The on-line processor farm provides the HLT and a fraction of the accepted events are passed to the Data Acquisition (DAQ) system for further processing.

### 3.2.7.1 Jet Triggers

At CMS, there are various types of triggers depending on the analysis to be performed. The triggers based on jet properties and missing transverse energy ( $E_T^{\text{miss}}$ ) are important for search for new physics whereas the single-jet triggers are mainly designed to study quantum chromodynamics (QCD). This thesis uses the single-jet triggers to select the events for analysis. At L1, the single-jet triggers use information mainly from the calorimeters by looking for the highest energy deposit. The sums of transverse energy from ECAL and HCAL are computed in  $4 \times 4$  trigger towers, except in the HF region where this quantity is measured in the whole trigger tower itself. If the calculated sum is greater than a certain threshold, the event is selected at L1 and it is passed to the HLT. At HLT level, the jets are reconstructed

using the anti- $k_t$  jet clustering algorithm. The inputs to the jet algorithm are either calorimeter towers giving “CaloJet” objects, or the reconstructed particle flow objects giving “PFJet” objects. The processing of reconstruction algorithm takes a long time and hence the jet trigger paths are divided into multiple selection steps. At first, the jets are reconstructed from calorimeter towers. The PF algorithm is run only for events in which at least one calorimeter jet passes a certain  $p_T$  threshold. The jets are then again clustered again from the PF candidates. In 2012, most of the jet trigger paths take PFJets as their inputs. The rate of jet events is quite high, so PFJet trigger paths have a pre-selection based on CaloJets. The matching between CaloJets and PFJets is required in single PFJet paths. Due to the flexibility of the HLT, it is possible to apply the jet energy corrections during the HLT selection.

### 3.2.7.2 Data Acquisition System

As the L1 trigger accepts events at a rate of 100 kHz, the Data Acquisition (DAQ) system has to process the events at the same speed. It reads out the data of all detector sub-components and assembles the complete events, see Fig. 3.13. The data is subsequently passed to the HLT which further reduces the rates to a few hundred events per second. Finally, the events are merged and saved to a local storage system, from which they are continuously transferred to the Tier-0 computing center at CERN.

### 3.2.8 Data Management

Although the trigger system reduces the collision rate enough to be stored, still there is a huge amount of the data need to be analyzed. An efficient computing infrastructure and the software is required for storing and distributing the data. To meet this need, the LHC has a data storage infrastructure called the Worldwide LHC Computing Grid (WLCG) [64]. WLCG provides a hierarchical structure, as shown in Fig. 3.14, in a series of four levels or Tiers. Each Tier is made up of several

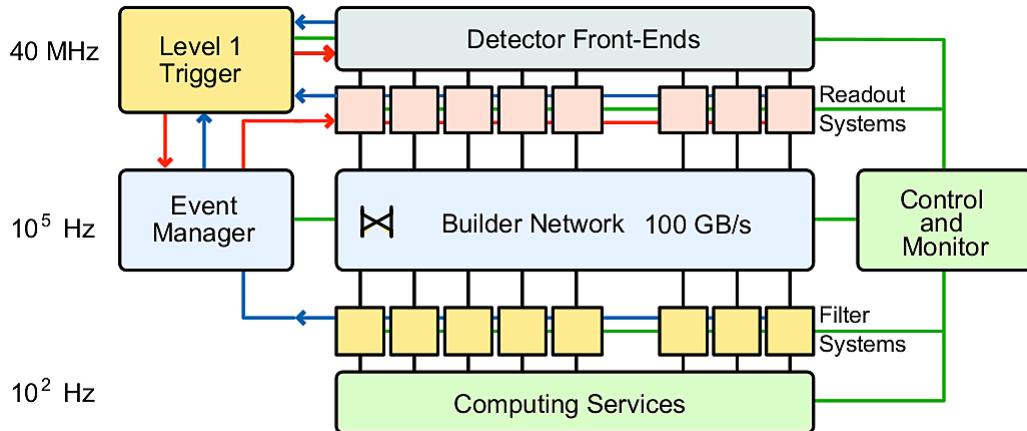


Figure 3.13: Architecture of the CMS Data Acquisition (DAQ) system. Taken from [54].

computer centres. All the raw collision data collected by CMS is converted into a format suitable for offline analysis and sorted in the form of the data sets at the Tier-0 site at CERN. This processed data is then transferred to Tier-1 centers all over the world where reconstruction algorithms are run. Further reconstructed and simulated data is distributed to Tier-2 sites, where it is available for physics analysis mainly performed on Tier-3 sites.

### 3.3 Software Tools

Every year, the CMS is recording a huge amount of collision and simulation data. This data is analyzed iteratively to improve the understanding of the detector and the measured physics. So a dedicated data structure and software tools are required for data analysis. These are included in the software framework referred to as CMSSW framework.

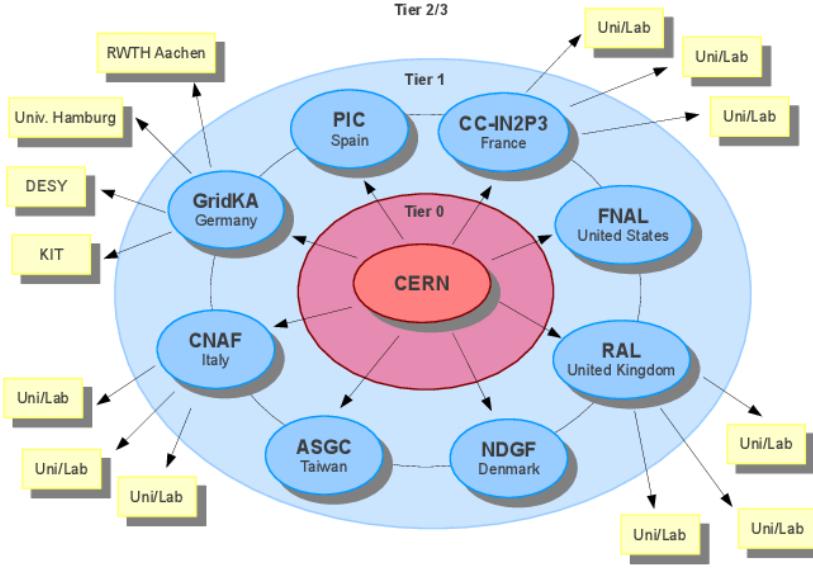


Figure 3.14: The schematic overview of the CMS computing grid. All data collected by CMS is stored at the Tier-0 site at CERN which is then transferred to Tier-1 centers all over the world. Further reconstructed and simulated data is distributed to Tier-2 sites, where it is available for physics analysis mainly performed on Tier-3 sites. Taken from [64].

### 3.3.1 CMSSW Framework

The CMS software framework (CMSSW) [65] provides all necessary tools required to perform a physics analysis. It is built on top of an event data model (EDM). It is a container for arbitrary C++ objects, e.g. recorded raw data and reconstructed physical objects or derived physical quantities of an event. The reconstruction algorithms in the CMSSW framework are divided into modules, which can be dynamically loaded and run. The event processing model in CMSSW is run by one executable, called `cmsRun`. SCRAM (Source Configuration, Release, And Management) is a configuration and management tool in the framework. It builds a runtime environment and make available all the necessary shared libraries. The shared libraries reduces memory consumption by loading only required modules during runtime. The CMSSW framework performs calibration, event generation, detector simulation, event reconstruction as well as data analysis by implementing the

codes either in C++ or Python languages. To reduce the event content, a process called skimming is performed where only necessary data is preserved.

### 3.3.2 ROOT

ROOT [66] is an open source object-oriented data analysis framework, developed by CERN. It consists of a huge C++ library provided with all the functionalities to store and analyze large amounts of the data. It provides histogramming methods in 1, 2 and 3 dimensions, curve fitting functions, minimization procedures, graphics and visualization classes. The command language of ROOT is command line interpreter (CINT), with several extensions to C++ which makes ROOT a versatile package. ROOT can be extended dynamically by linking external libraries. The events generated or analyzed in CMSSW framework are stored in a tree structure in files using ROOT libraries. In this thesis, ROOT has been used extensively for storing information of events or objects, for analyzing the events, for fitting as well as plotting purposes.

### 3.3.3 NLOJET++ and FASTNLO

The cross-sections for jet production at leading order (LO) and next-to-leading order (NLO) are evaluated using a C++ program called NLOJET++ [67, 68]. It uses the dipole subtraction method for the separation of the divergences and can calculate up to three-jet observables at NLO precision. The perturbative QCD cross-sections are calculated using Monte Carlo integration methods which are very time consuming. It makes PDF fits or estimations of uncertainties difficult where the calculations of the cross-sections are need to be repeated. So the NLOJET++ is interfaced to the FASTNLO project [69, 70] which performs fast re-evaluations of cross-sections. It stores the perturbative coefficients obtained with NLOJET++ in a way that the strong coupling constant and the PDFs can be changed afterwards without a recalculation of the perturbative coefficients.

All the event generators and cross-section calculation tools take the PDFs as an input. They are either hard coded in the generators or accessed via a standardized interface with the LHAPDF library [71, 72]. LHAPDF provides a unified and easy way to use the PDF sets by storing them in the data files. It provides interpolation routines to read the PDFs and interpolate the PDFs at all scales. It also allows access to single PDF members without needing to load whole sets. LHAPDF is supported by many MC event generators and other physics programs.

# Chapter 4

## Event Generation, Simulation and Reconstruction

Event simulation plays a significant role in the operation of any experiment. Before the real data taking, the reconstruction algorithms, efficiency of trigger paths, analysis strategies and other operational details of the experiment need to be studied and well optimized. This is achieved by simulation of the apparatus and the expected processes using the Monte Carlo (MC) method [73]. In high energy physics, the simulation of experimental data is done in two steps : event generation and detector simulation. Event generators simulate a collision starting from the proton-proton interaction up to the production of the final decay products, to be observed with the detector. The output of an event generator is taken as input in a detector simulation program which models the interactions between the generated final-state particles and the detector. This requires a sophisticated and complex simulation of the detector material and the behaviour of the interacting particles.

### 4.1 Event Generation and Simulation Software

In real world, the machine or collider produces interactions which are observed by detectors. The interesting events are stored and reconstructed afterwards for a

physics analysis. In the MC world, the role of machines is played by the event generators. The event generators generate simulated events as detailed as observed by a detector. The output produced by an event generator is in the form of “events” with the same behaviour and fluctuations as the real data. The detector simulation takes the output of the generator as an input and allows a precise prediction and verification of the entire experimental setup. The comparison of real and Monte Carlo world is presented in Fig. 4.1.

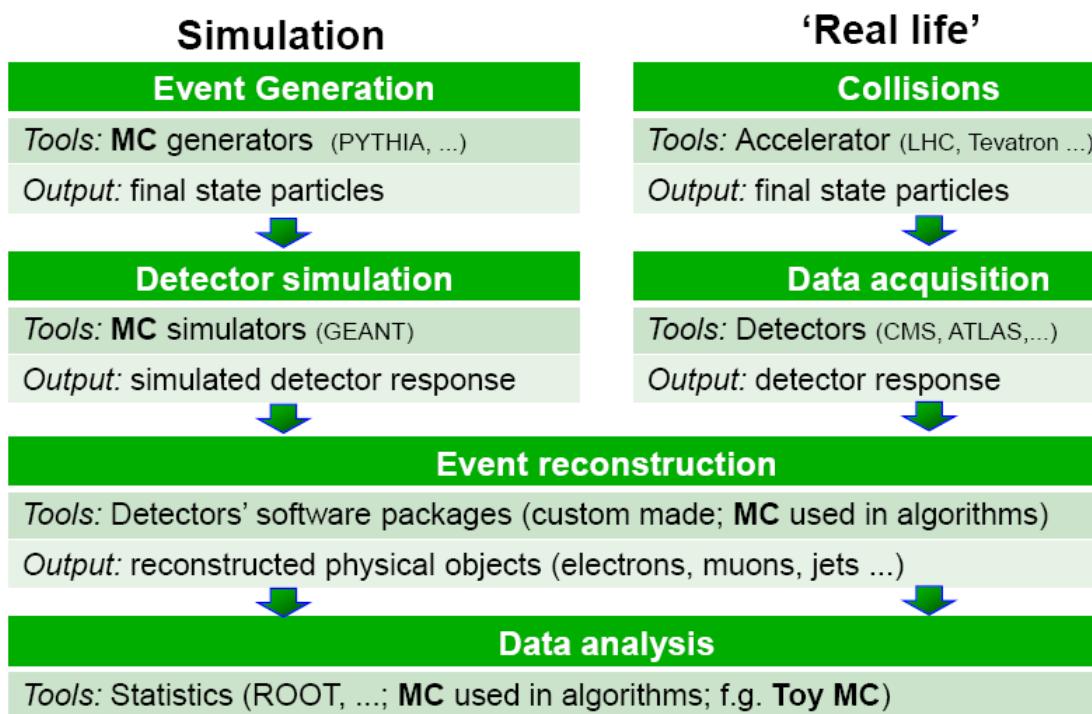


Figure 4.1: The comparison between Monte Carlo (MC) simulations generated by event generators and the real data produced by the particle collisions and observed in the detectors.

There are a variety of MC event generators which are commonly used in high energy physics. In this thesis, the three leading order (LO) generators : PYTHIA, MADGRAPH5 and HERWIG as well as the next-to-leading order (NLO) generator POWHEG are used. These generators are described one by one in the following sections.

### 4.1.1 PYTHIA

PYTHIA is the most widely used program to generate the collisions at high energies for p-p, e-e and e-p colliders. It contains theoretical models for a number of physics aspects which includes hard and soft processes, parton distribution functions, initial-state and final-state parton showers, multi-parton interactions, fragmentation and decay. It has a feature to interface with the external programs. It uses the Lund string hadronization model [74] to describe the hadronization process. PYTHIA was originally coded in FORTRAN language under the version 6 i.e. PYTHIA6 [75]. In 2004, it was rewritten in C++ and was released as PYTHIA8 [76] in 2007. The two versions differ in the description of multi-parton interactions. Both the versions use LO calculations to derive the colored partons from the hard interaction which hadronize into colorless objects like hadrons. For the studies in this thesis, PYTHIA6 with tune Z2<sup>\*</sup> [77] and PYTHIA8 with tunes CUETS1 and CUETM1 [78] have been used.

### 4.1.2 MADGRAPH5

MADGRAPH5 [79] generates matrix elements for high energy physics processes, such as decays and  $2 \rightarrow n$  scattering processes. The event information of the particles involved in the hard process such as particle ID, momenta, spin etc. is stored in the Les Houches format [80] which can be interfaced to other generators. In the current study, MADGRAPH5 has been interfaced to PYTHIA6 with tune Z2<sup>\*</sup> to handle the rest of the generation steps which involve parton showering and hadronization. Matching algorithms make sure that no double-counting occurs between the tree-level and the parton-shower-model-generated partons. MADGRAPH5+PYTHIA6 (MG5+P6) samples are used mainly for general comparisons to the data and calculating the detector resolution.

### 4.1.3 HERWIG

HERWIG (Hadron Emission Reactions With Interfering Gluons) [81] is a multi-purpose event generator which performs the LO calculations. It uses angular ordering for parton showers and cluster model for hadronization. The hard lepton-lepton, lepton-hadron and hadron-hadron scattering as well as soft hadron-hadron processes can be simulated using HERWIG generator. This generator can be interfaced to external matrix element generators. HERWIG was written in FORTRAN language and a version in C++ is also available as HERWIG++ [82]. We have used the HERWIG++ generator with the default tune of version 2.3 [83] to generate the samples which have been used to study the non-perturbative effects.

### 4.1.4 POWHEG

POWHEG (Positive Weight Hardest Emission Generator) generator performs the fixed NLO calculations merged with parton showers [84–86]. This generator used a computer framework known as POWHEG BOX [87] to implement NLO calculations in shower MC programs. It can be interfaced with all modern shower MC programs that support the Les Houches Interface format. It contains the hard matrix elements for NLO dijet production. POWHEG has been interfaced to PYTHIA8 with tunes CUETS1 and CUETM1 to include the parton shower and hadronization,

## 4.2 Detector Simulation

The particles generated by Monte Carlo event generators are passed through a computer program which does the detector simulation. It defines the detector system including its geometry, material and electronics properties. The detector simulation describes the nature of the interactions of the particles with the material of the detector. While propagating through the detector material, these particles are allowed to decay according to their known branching fractions and decay kinematics. The

particles can interact with the detector material through several physical processes, including electron bremsstrahlung, energy loss by ionization, multiple scattering, hadron showering etc., which are simulated or parametrized in the corresponding parts of the detector.

In CMS, the detector response is simulated by two approaches [55] : Full Simulation and Fast Simulation. Full Simulation is based on a C++ simulation toolkit GEANT4 (GEometry ANd Tracking) [88]. It is a successor of a FORTRAN based GEANT3 and handles the interactions of particles with matter over a wide range of energy. In GEANT4, the uniform and non-uniform electromagnetic, magnetic and electric fields can be specified. The equation of motion of the particle in the field gives the track of the particle. A physical interaction of a track in the sensitive region of a detector is called a hit. The secondary particles produced are stored in a stack with the information of their kinematic properties as well as the vertex position where the interaction has occurred. A large number of Monte Carlo events may have to be produced for a feasible physics analysis. The complete detector simulation of CMS using GEANT4 is rather time consuming. So a Fast Simulation framework [89] has been developed in the general software framework of the CMS for the fast simulation of the detector response. In Fast Simulation, detector effects are parametrized instead of simulating these from first principles as done in Full Simulation. In Fast Simulation package, the events are produced at much faster rates as compared to the Full Simulation package, while maintaining almost the same level of accuracy for physics studies. The format of the Fast Simulation data output is fully compatible with the standard Full Simulation one.

After simulating the detector response, it is then transformed into a digital signal with the help of electronics and this step is called digitization. The simulated output of the detector response needs to be as close as possible to the real data coming from the CMS detector. After this, the event reconstruction algorithms are applied to both simulated and real events.

## 4.3 Event Reconstruction

The aim of the event reconstruction is to identify the particles passing through the detector by interpreting the electrical signals produced in digitization. In event reconstruction, analysis-level objects are created by combining recorded signals from the tracker, calorimeters and muon detectors. Initially the reconstructed hits are collected which are combined to form tracks and calorimetric towers. Then higher level objects such as electrons, muons, photons and jets are reconstructed by combining the tracks and energy deposits. In CMS, all the particles are identified and reconstructed with a Particle Flow (PF) algorithm, discussed in detail in the next section.

### 4.3.1 Particle Flow Algorithm

In the CMS, the identification and reconstruction of the particles is performed using the event reconstruction technique called Particle Flow (PF) algorithm [90, 91]. The PF algorithm combines the information from the individual sub-detectors. The additional identification and reconstruction of the tracks enhances the reconstruction performance. These tracks are used to identify the primary vertices in an event.

**Track Reconstruction** - The particles produced from the collisions leave the tracks in the sub-detectors as they traverse through the CMS detector. They follow helix paths due to presence of the strong magnetic field. An efficient tracking system is needed to measure the momenta of the charged particles from the curvature of the tracks. The CMS uses an iterative tracking algorithm, the Combinatorial Track Finder (CTF) algorithm [92], which first generates the track seeds by grouping the hits in the innermost layers. Then a pattern recognition is performed using a Kalman filter [93], where hits coinciding with the predicted trajectory of the charged particle are found. Then the best parameters are estimated for all hits along the trajectory and fitting is performed. Finally, the quality criteria are applied to the

tracks to reject the badly reconstructed ones and to decrease the fake rate, defined as the ratio between fake tracks and all tracks. These all steps are performed iteratively which are ordered by the level of difficulty in identifying the tracks.

**Primary Vertex Reconstruction** - The increasing center-of-mass energy and instantaneous luminosity enhances the probabilities of multiple pp collisions per bunch crossing as well occurrence of two or more hard interactions between partons in the same pp collisions giving rise to pileup events. Hence, the identification of the primary vertex of the main hard interaction becomes important. The CMS performs the primary vertex reconstruction in two steps : First, a set of promptly produced tracks are selected and grouped together in clusters, using a deterministic annealing (DA) algorithm [94]. The tracks are grouped based on their  $z$ -coordinate at the point of closest approach to the beam line. For each track, the  $z$ -coordinate of the point of closest approach to the beam-line is referred as  $z_i$  with associated uncertainty as  $\sigma_i$ . These tracks must be assigned to an unknown number of vertices denoted by a  $z$ -coordinate of  $z_k$ . If a track is assigned to only one vertex, it is referred as a hard assignment. It is represented by values of probability  $p_{ik} = 1$  if track  $i$  is assigned to vertex  $k$  and 0 otherwise. For the soft assignments, the tracks can be associated with more than one vertex such that  $p_{ik}$  has value between 0 and 1 representing the probability of the assignment of track  $i$  to vertex  $k$  in a large ensemble of possible assignments.  $\chi^2$  is defined as :

$$\chi^2 = \sum_{ik} p_{ik} \frac{(z_i - z_k)^2}{\sigma_i^2} \quad (4.1)$$

Instead of minimizing the  $\chi^2$ , the DA algorithm finds the most likely distribution for a given value  $\chi_0^2$  which is then decreased until it finds a good reliable minimum. Once the tracks are assigned to the different vertices, a three dimensional fitting is done using the adaptive vertex fitter [95] where each track is assigned a weight ( $w$ )

between 0 and 1. After fitting, the tracks are labelled as either good with  $w = 1$  or outliers with  $w$  value close to 0. The sum of these weights gives a rough estimate of the number of tracks associated with the vertex. The number of degrees of freedom in a fit is given by  $n_{\text{dof}} = 2\sum_i w_i - 3$ , where  $w_i$  is the weight of the  $i$ th track and the sum runs over all tracks associated with the vertex. The vertices having  $n_{\text{dof}} > 4$  i.e. having at least four tracks assigned to each vertex, are considered. All the reconstructed vertices are ordered by the sum of the squared track momenta  $\sum p_T^2$  and the one having largest sum is selected as the primary vertex of interest.

The PF algorithm works independently of the vertex reconstruction. The PF event reconstruction algorithm basically converts the detector signals back to physical objects by using PF event reconstruction algorithm which is illustrated in Fig. 4.2. The transverse momenta of the final state stable particles or energies of the calorimeter towers are taken as the inputs to the PF algorithm. The PF algorithm first collects the reconstructed hits in each sub-detector independently and creates a list of reconstructed elements (referred as blocks) which consists of charged tracks in tracker, energy clusters in calorimeters and muon tracks in muon system. Then a link algorithm connects topologically compatible blocks producing PF objects. The PF objects consist of all stable particles : electrons, muons, photons, charged and neutral hadrons. The energy of the the electrons is determined from the track momentum at the main interaction vertex along with the corresponding ECAL energy deposits and the energy sum of all bremsstrahlung photons associated with the tracks. The curvature of the tracks in tracker and muon chamber is used to estimate the energy of muons. The energy of photons is obtained directly from the ECAL measurement, corrected for zero-suppression effects<sup>18</sup>. The energy of charged hadrons is calculated by combining the track momentum and corresponding energy clusters in ECAL and HCAL, corrected for zero-suppression effects as well

---

<sup>18</sup>To suppress noise in the calorimeters, only cells with energies above a given threshold are considered and this procedure is known as zero-suppression.

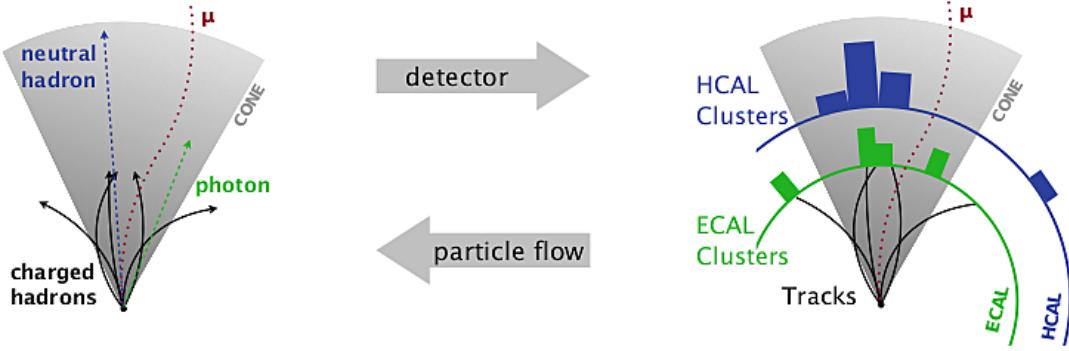


Figure 4.2: The Particle Flow (PF) algorithm is used by the CMS to identify and reconstruct the particles. The PF converts the sub-detector measurements back to physical particle objects. Taken from [17].

as calibrated for the nonlinear response of the calorimeters. The energy of neutral hadrons is obtained from the corresponding calibrated ECAL and HCAL energies only. Along with the reconstruction of the PF objects, missing transverse energy ( $E_T^{\text{miss}}$ ) is also determined using PF algorithm.  $E_T^{\text{miss}}$  is defined as the negative vector sum of transverse momenta ( $p_T$ ) of all the isolated stable particles reconstructed in an event i.e.  $E_T^{\text{miss}} = -\sum_i \vec{p}_{T,i}$ . To avoid any kind of double-counting of energy, blocks of all PF reconstructed particle objects are removed and the energy of the calorimeter clusters is recalculated. Finally, the collection of PF objects is used to reconstruct the jets by using the different jet clustering algorithms. The jets are collimated sprays of hadrons and other particles produced by the hadronization of the quarks or gluons. The detailed description of jets is given in Sec. 2.4. The jet algorithms, discussed in Sec. 2.4.1, are used for clustering of stable partons, particles or reconstructed particles. The jets are formed at different levels : parton level, particle level and detector or reconstructed level, as illustrated in Fig. 4.3. In the CMS detector, jets are the localized deposits of energy in the calorimeter cells along with the large number of tracks in the direction of the deposited energy. The typical jet energy fractions contributed by charged particles, photons and neutral

hadrons are 65%, 25% and 10%, respectively. So, the PF algorithm reconstructs about 90% of the jet constituents with good precision, whereas only 10% depend on the less accurate response of the HCAL. Depending on the type of input to the jet algorithm, jets can be categorized into following different types :

**Generator Jets** - The stable particles generated by the Monte Carlo event generators are clustered into generator jets (GenJets). At this particle level, the passage through the detector simulation has not been carried out. The objects at this level are photons, charged and neutral hadrons. Since the energy of GenJets is independent of the detector response, these are considered as reference objects for estimating the jet energy corrections, discussed in Sec. 4.3.2.

**Calorimetric Jets** - The Calorimetric jets (CaloJets) are reconstructed by taking the energy clusters deposited in the ECAL and HCAL calorimeter towers as inputs. One calorimetric tower consists of one HCAL cell surrounded by an array of  $5 \times 5$  ECAL cells. The tower's four-momenta are computed by taking the direction from the interaction point to the tower center. All towers with a transverse energy measurement above 300 MeV are considered in the clustering process. CaloJets are relatively simple objects because only calorimeter information is deployed, but they are strongly affected by the non-linearity of the calorimeter response. Since the readout of calorimeter measurements is fast, CaloJets are commonly used by the trigger system.

**Particle Flow Jets** - The clustering of particle flow candidates give detector level jets called Particle Flow jets (PFJets). The four-momenta of the particles is taken as the input. The use of the tracker system and high granularity of the ECAL gives better energy resolution calculated using the independent measurements of charged hadrons and photons clustered to form a jet. Hence PFJets perform better than CaloJets and are the standard jets used at CMS.

This thesis presents the study of the jets formed by clustering the particle flow

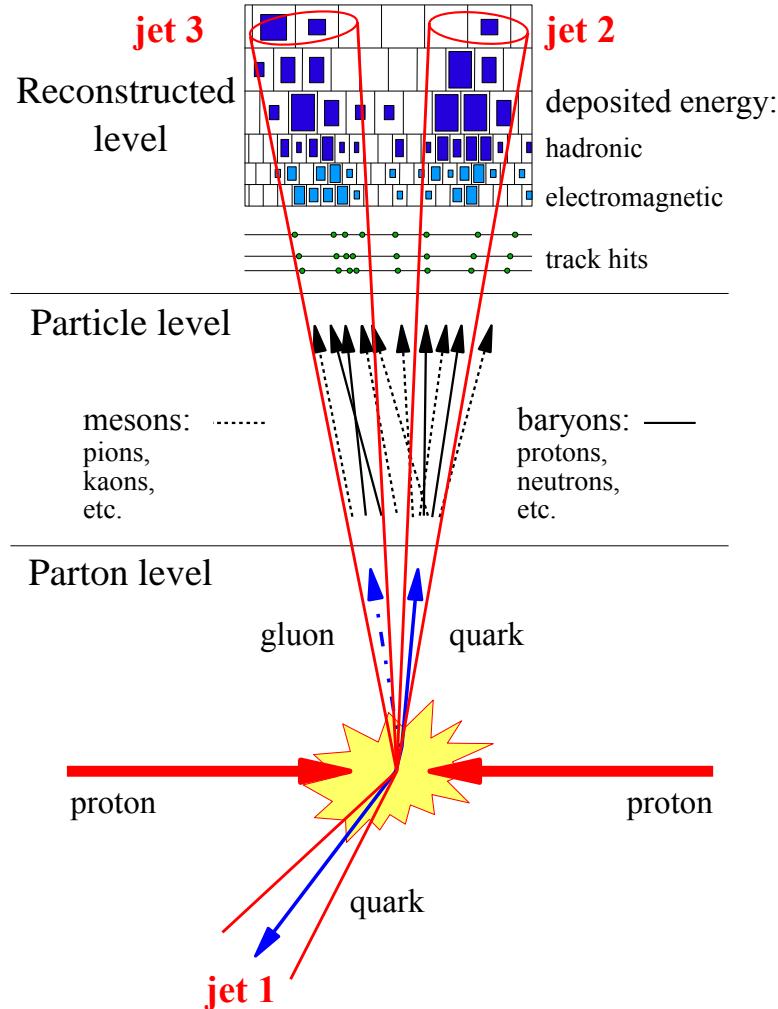


Figure 4.3: In a proton-proton collision, the hard scattered quarks and gluons fragment to produce the showers of partons which get clustered into parton jets. The produced partons undergo hadronization and produce hadrons which form particle jets. The hadrons deposit their energies into the calorimeters in the form of reconstructed jets. Taken from [96].

candidates using the anti- $k_t$  algorithm with a jet size parameter of  $R = 0.7$ .

### 4.3.2 Jet Energy Corrections

The measured energy of jets cannot be directly translated to the energy at true particle or parton level. This is because of the nonlinear and nonuniform response of the calorimeters, effects of pileup and small residual effects in the data remain-

ing after the corrections based on Monte Carlo simulations. Hence the jet energy corrections (JEC) [97, 98] are used to correct the measured jet energy and relate it to the corresponding true particle jet energy. To correct the energy of jets, the CMS follows a factorized approach, as presented in Fig. 4.4, where JEC are applied in a sequential manner with fixed order, i.e. the output of one step serves as the input for the next one. Each level of correction takes care of a different effect and is independent of each other. At each step, the jet four-momenta is scaled with a correction factor which depends on jet  $p_T$ ,  $\eta$ , flavor etc.

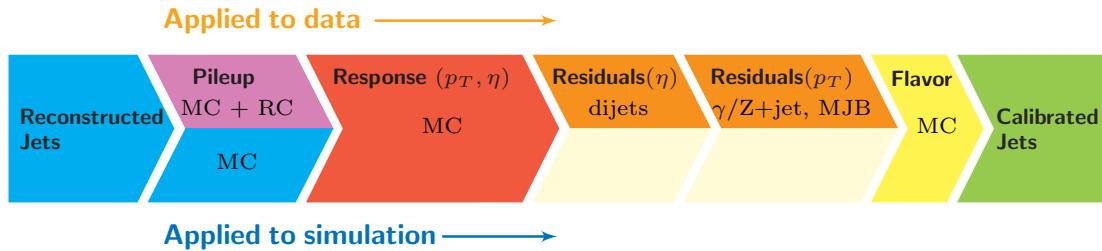


Figure 4.4: A schematic diagram of the factorized jet energy corrections (JEC) applied to the data (upper half) and simulation (lower half). The reconstructed jets are corrected for pileup effects, non-uniform  $p_T$  and  $\eta$  response and residual differences between the data and Monte Carlo simulations along with optional flavor corrections. All corrections marked with MC are derived from simulation studies, RC stands for random cone, and MJB refers to the analysis of multijet events. Taken from [98].

The corrected jet transverse momentum  $p_T^{\text{corr}}$  is obtained by applying all correction factors subsequently on raw or uncorrected jet transverse momentum  $p_T^{\text{raw}}$  as below :

$$p_T^{\text{corr}} = c_{\text{res}}(\eta, p_T'') \cdot c_{\text{mc}}(\eta, p_T') \cdot c_{\text{pileup}}(\eta, \rho, A_j, p_T^{\text{raw}}) \cdot p_T^{\text{raw}} \quad (4.2)$$

where  $p_T'$  is the transverse momentum obtained after applying the pileup correction factor  $c_{\text{pileup}}$  on  $p_T^{\text{raw}}$ ,  $p_T''$  is the transverse momentum obtained after applying the additional correction factor  $c_{\text{mc}}$  because of relative and absolute effects derived from MC. Finally, a correction factor  $c_{\text{res}}$  is applied for residual effects derived from the data. The corrections applied at each step are discussed below :

**Pileup Corrections** - The additional proton-proton collisions occur within the same bunch-crossing along with the main hard interaction and give rise to pileup events. The particles produced from the pileup events get clustered into the jets coming from the hard interaction and increase the jet energy. This extra energy needs to be subtracted from the reconstructed jet energy. This is done by applying the pileup corrections to raw jet  $p_T^{\text{raw}}$ . The pileup corrections are determined by simulating a sample of QCD dijet events with and without pileup effects. The pileup correction factor,  $c_{\text{pileup}}$  is calculated from jet area method using the pileup density  $\rho$  in the event and the jet area  $A_j$ .  $c_{\text{pileup}}$  is parametrized as a function of  $\rho$ ,  $A_j$ , jet  $p_T$  and  $\eta$ . There are corrections for residual differences between the data and detector simulation which are determined using the random cone (RC) method in zero-bias events. Hence the different pileup corrections are applied to the data and the MC simulations.

**MC Corrections** - The next correction applied to the pileup corrected jets is based on MC simulated QCD events. Due to the inefficiencies introduced by the detector simulation, the reconstructed jet  $p_T$  is not the same as that of the generated one. This difference is corrected with the factor,  $c_{\text{mc}}$  which is derived by comparing the measured jet  $p_T$  to the particle level jet  $p_T$ . The corrections are determined as a function of jet  $p_T$  and  $\eta$  which make the detector response uniform over these two variables.

**Residual Data Corrections** - The jets corrected with above mentioned corrections are further corrected for remaining small differences between the data and MC simulations. This correction is applied only to the data. The correction factor  $c_{\text{res}}$  is derived using data-driven methods. The relative residual corrections are evaluated using dijet events in which a probe jet is calibrated using a tag jet. The last correction applied is the absolute residual correction in which the precisely reconstructed  $Z$  bosons balanced to a jet are used to calibrate the jet energy.

**Flavor Corrections** - These corrections correct the jets for flavor dependence ( $b$ ,  $\tau$  etc.) and are optional. These are extracted using  $Z+jet$  and photon+jets simulated events. The flavor corrections have not been applied for 8 TeV CMS data.

The process of correction of jets by using JEC introduces uncertainties in the final corrected jet energy which are discussed in Sec. 5.6.2. After correcting the jets, the multijet event cross-sections are measured which are discussed in the following chapter.

# 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,  $H_{T,2}/2 = \frac{1}{2}(p_{T,1} + p_{T,2})$ , where  $p_{T,1}$  and  $p_{T,2}$  denote the transverse momenta of the two leading jets.

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

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

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 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. 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 file<sup>19</sup> 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  $19.71 \text{ fb}^{-1}$ .

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 $\text{fb}^{-1}$
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-

---

<sup>19</sup>Cert\_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 region 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 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	$1.036 \times 10^7$
	/QCD_HT-250To500_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	27062078	$2.760 \times 10^5$
	/QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	30599292	$8.426 \times 10^3$
	/QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	13843863	$2.040 \times 10^2$

## 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.

### 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, different 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  $19.71 \text{ fb}^{-1}$ , 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 $\text{fb}^{-1}$
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.

$$\epsilon_{\text{HLT\_PFJetY}} = \frac{H_{T,2}/2 \left( \text{HLT\_PFJetX} + (\text{L1Object\_}p_T > Z) + (\text{HLTOBJECT\_}p_T > Y) \right)}{H_{T,2}/2(\text{HLT\_PFJetX})} \quad (5.3)$$

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<sup>20</sup> [100].

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

---

<sup>20</sup>See details in Sec. A.5

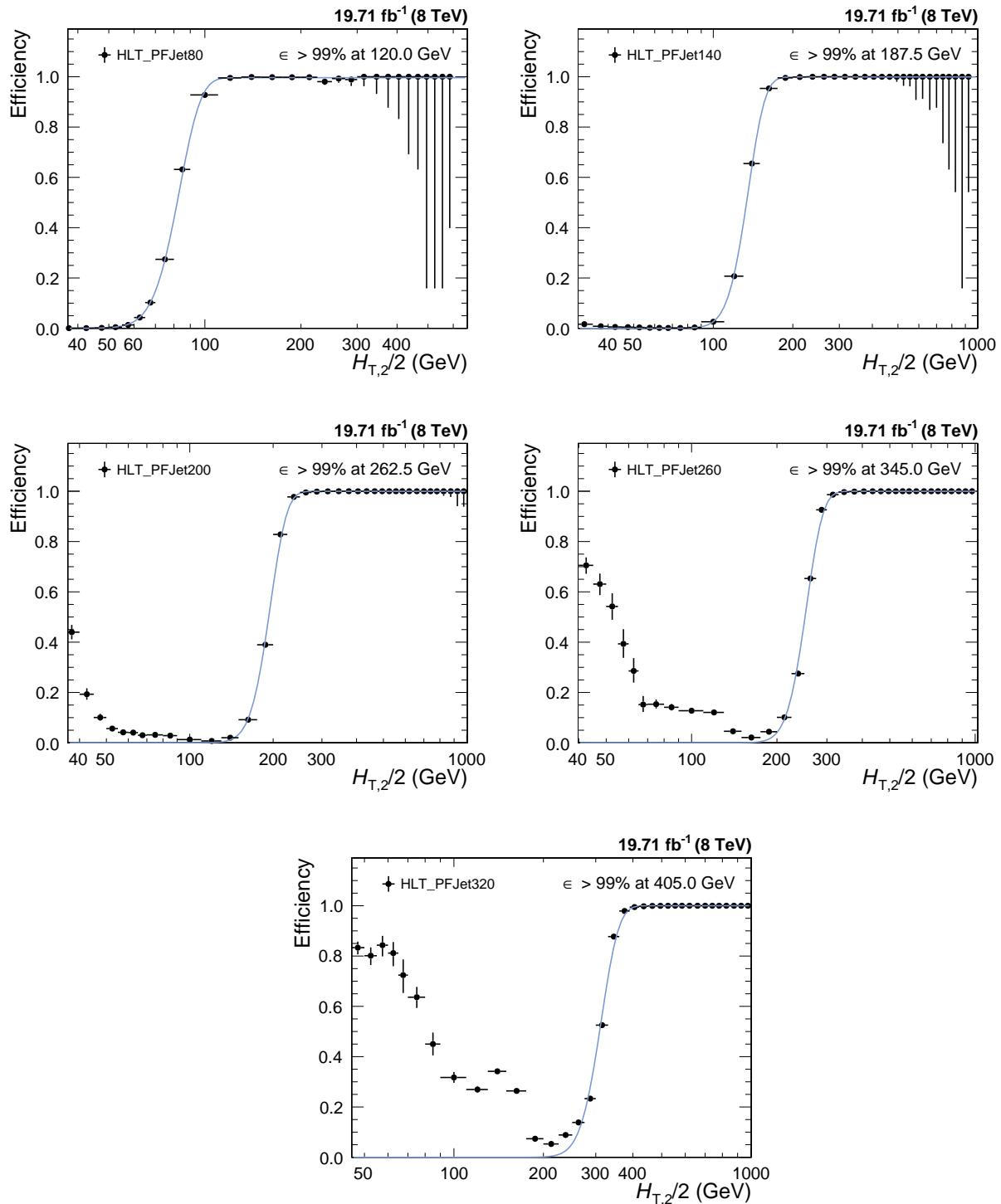


Figure 5.1: Trigger efficiencies turn-on curves for the single jet HLT trigger paths are fitted with a sigmoid function (blue line) to obtain the 99% efficiency threshold. The error bars give the uncertainty on the efficiency which is calculated using Clopper-Pearson confidence intervals [100].

### 5.2.2 Primary Vertex Selection

The reconstructed tracks, number of strip and pixel hits and the normalized track  $\chi^2$ , 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 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 ( $E_T^{\text{miss}}$ ). 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 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  $E_T^{\text{miss}}$ .

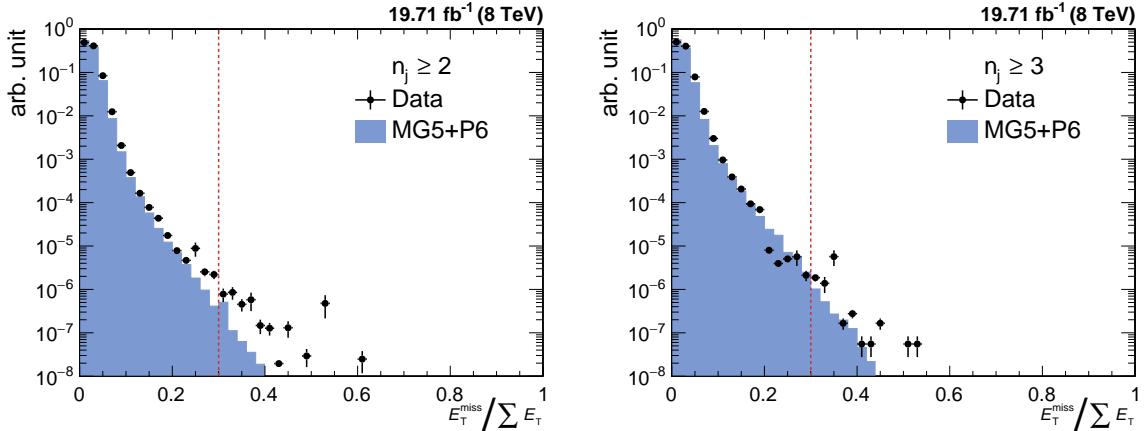


Figure 5.2: Missing transverse energy fraction of the total transverse energy per event in the data (black solid circles) and simulated Monte Carlo events (blue histogram) in inclusive 2-jet (left) and 3-jet (right) events. To remove background and noise, events with a fraction exceeding a certain threshold,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

### 5.2.4 Jet Identification

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 on 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].

	Property	Tight ID cut
Whole $\eta$ 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 on moving to higher  $H_{T,2}/2$ . Consequently the statistical fluctuations for ID efficiency

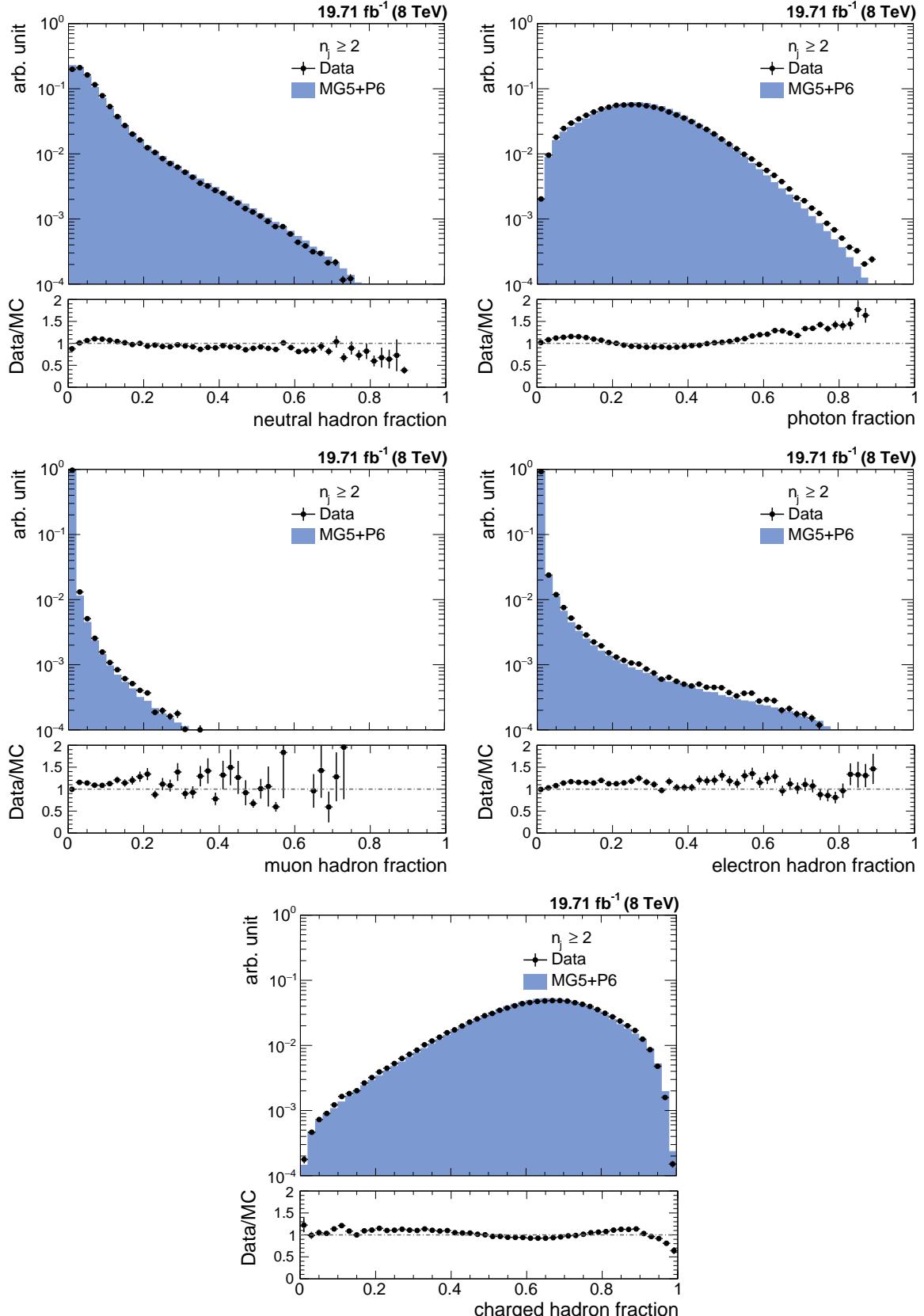


Figure 5.3: The fractions of jet constituents as observed in the data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 2-jet events. The data and simulations are normalized to the same number of events. The distributions are shown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

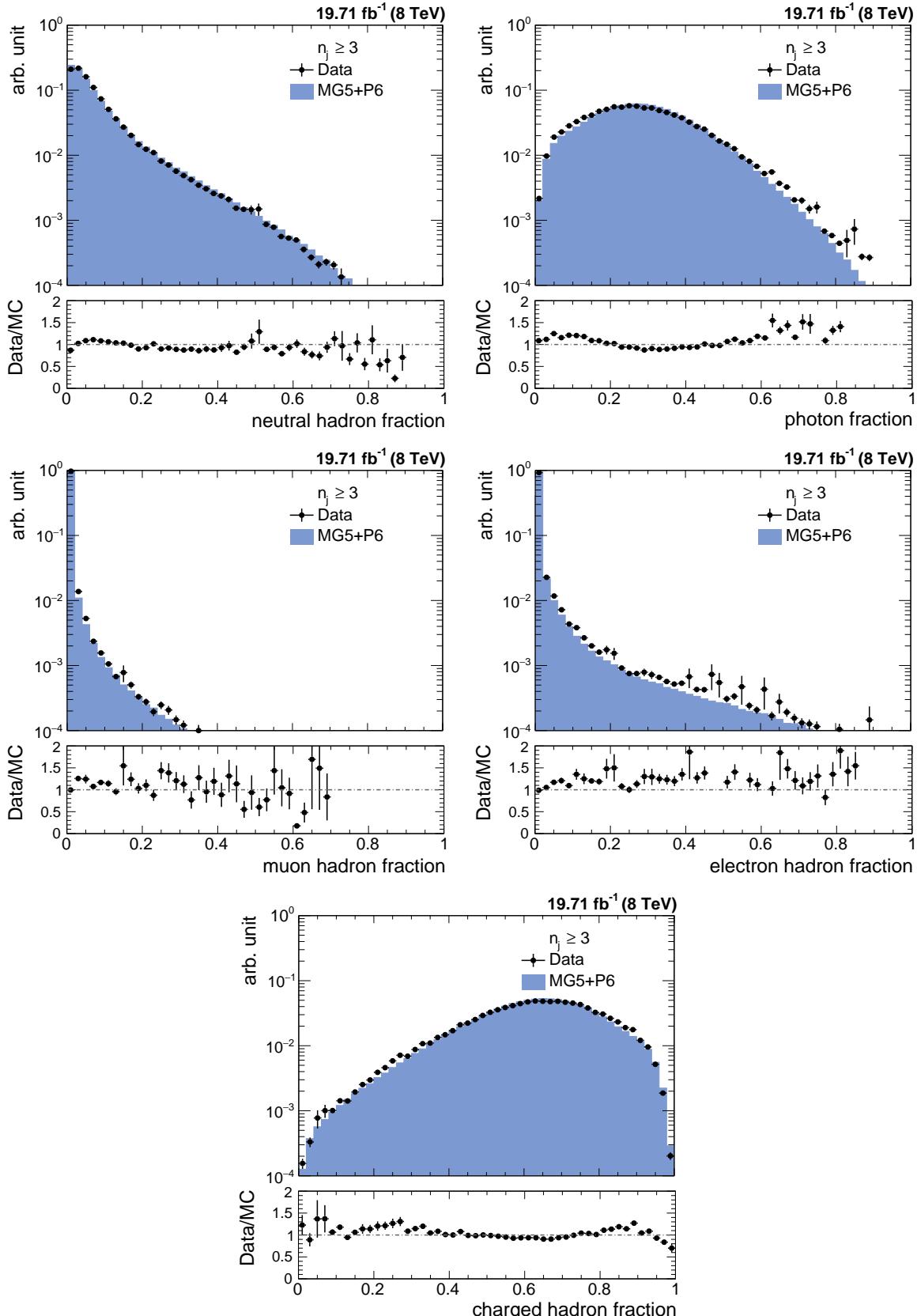


Figure 5.4: The fractions of jet constituents as observed in the data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 3-jet events. The data and simulations are normalized to the same number of events. The distributions are shown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

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

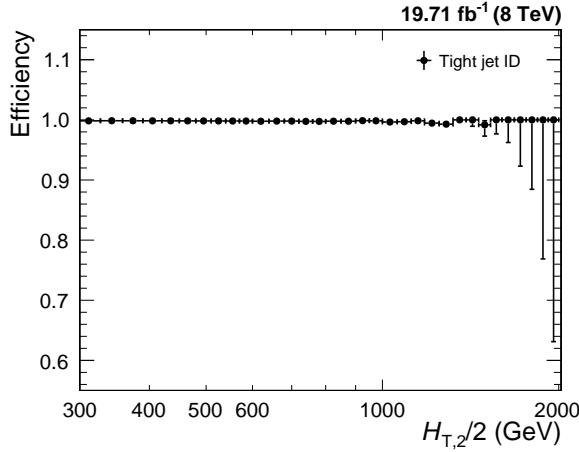


Figure 5.5: The jet ID efficiency is studied as a function of  $H_{\mathrm{T},2}/2$  with tag-and-probe technique using dijet event topologies and it always exceeds 99%.

### 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 data<sup>21</sup> as well as in simulated events<sup>22</sup>. As a convention, the jets in one event are ordered in decreasing  $p_{\mathrm{T}}$ , with the first (leading) jet being the jet with highest  $p_{\mathrm{T}}$ . 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  $|y| < 2.5$ .

---

<sup>21</sup>Winter14\_V8 jet energy corrections

<sup>22</sup>START53\_V27 jet energy corrections

These cuts assure high detector acceptance. For a consistent comparison, the exactly same selection criteria 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 describing the conditions expected for each data-taking period are taken care 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  $w_{\text{PU}}$ , 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

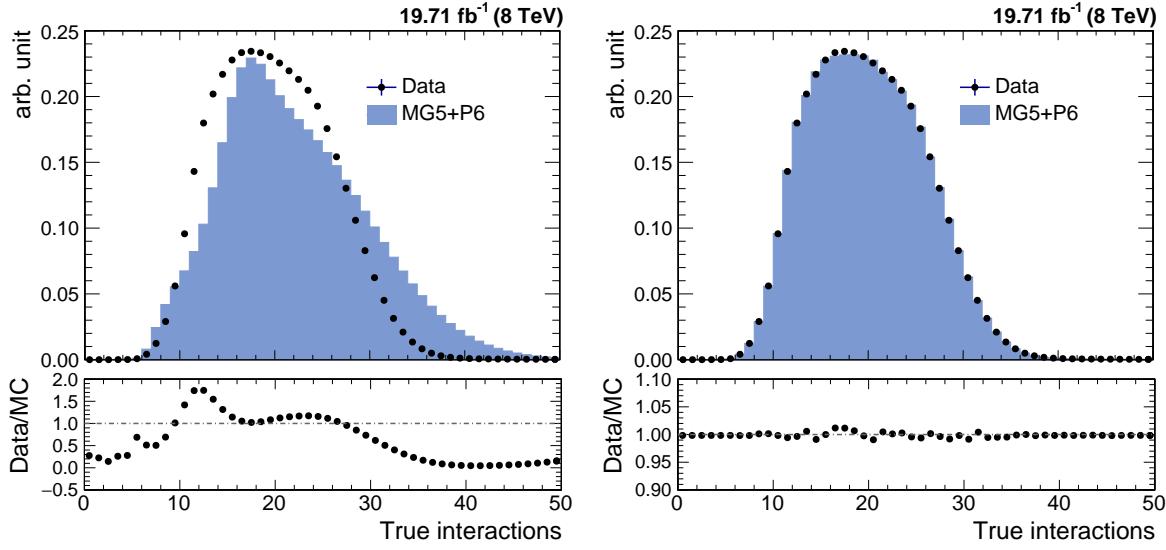


Figure 5.6: Number of reconstructed vertices in the data (black solid circles) and simulated Monte Carlo events (blue histogram) before (left) and after (right) the pileup reweighting.

$n_j \geq 2$  (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 sufficient 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  $300 \text{ GeV} < H_{T,2}/2 < 1680$  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

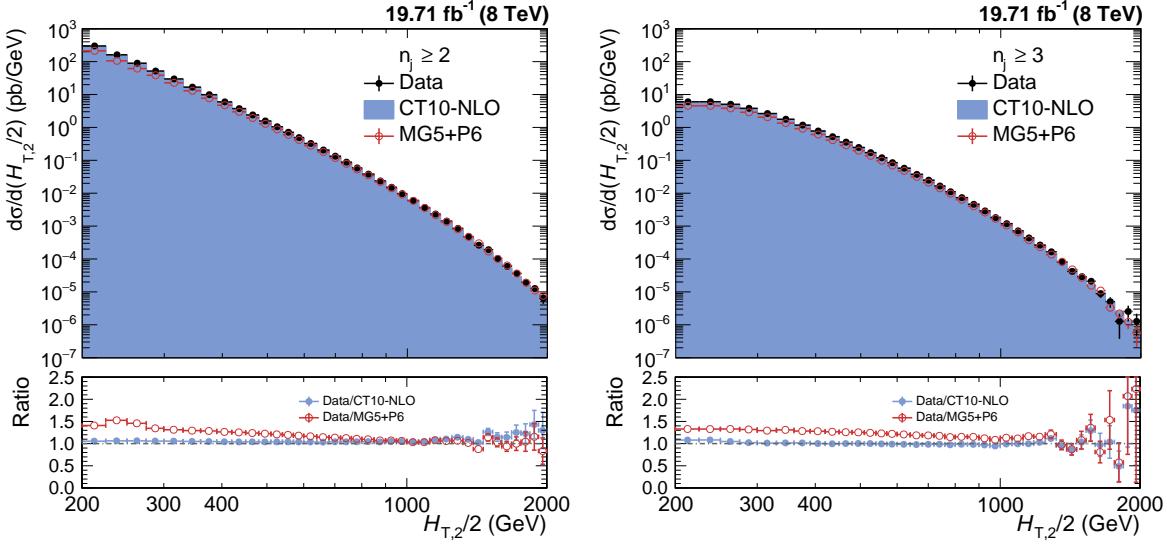


Figure 5.7: The reconstructed level differential cross-sections are compared for the data (black solid circles) and LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo (red empty circles) simulations with CT10-NLO theory predictions (blue histogram), as a function of  $H_{T,2}/2$  for inclusive 2-jet (left) and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

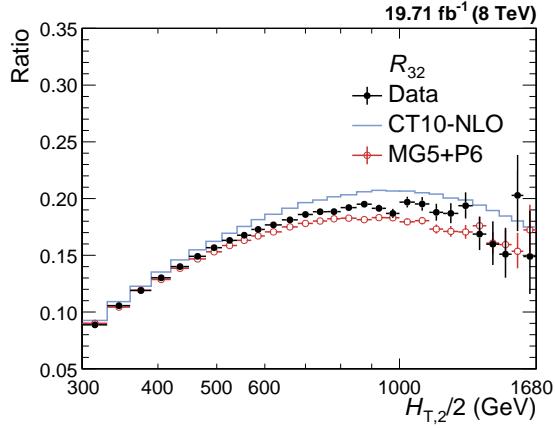


Figure 5.8: Comparison of the reconstructed level cross-section ratio  $R_{32}$  as a function of  $H_{T,2}/2$ , for the data (black solid circles) and LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo (red empty circles) with CT10-NLO theory predictions (blue line). The error bars give the asymmetrical statistical 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 empty 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  $R_{32}$ , calculated 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

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  $\eta$  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  $c_{central}$ .

$\eta$	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  $1\sigma$  of a Gaussian fit of the response distribution. A double-sided Crystal Ball function<sup>23</sup> 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

---

<sup>23</sup>See details in Sec. A.5

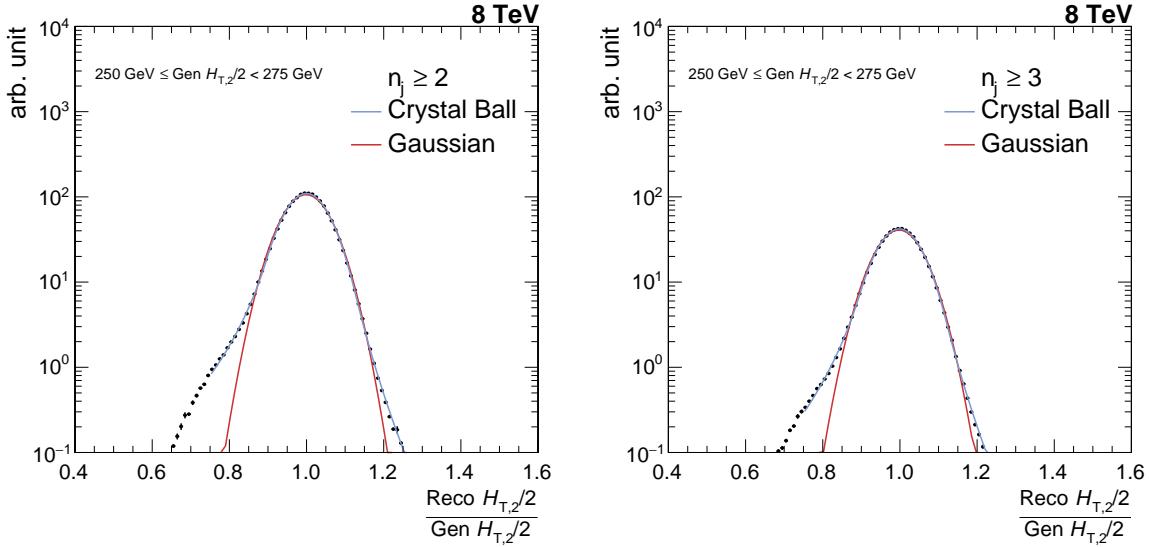


Figure 5.9: Fitting of the jet energy resolution distribution, obtained using LO MADGRAPH5+PYTHIA6 (MG5+P6) Monte Carlo simulated events, as a function of  $H_{T,2}/2$  for inclusive 2-jet (left) and 3-jet (right) events. The blue line shows the double-sided Crystal Ball function fit of  $\frac{\text{Reco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

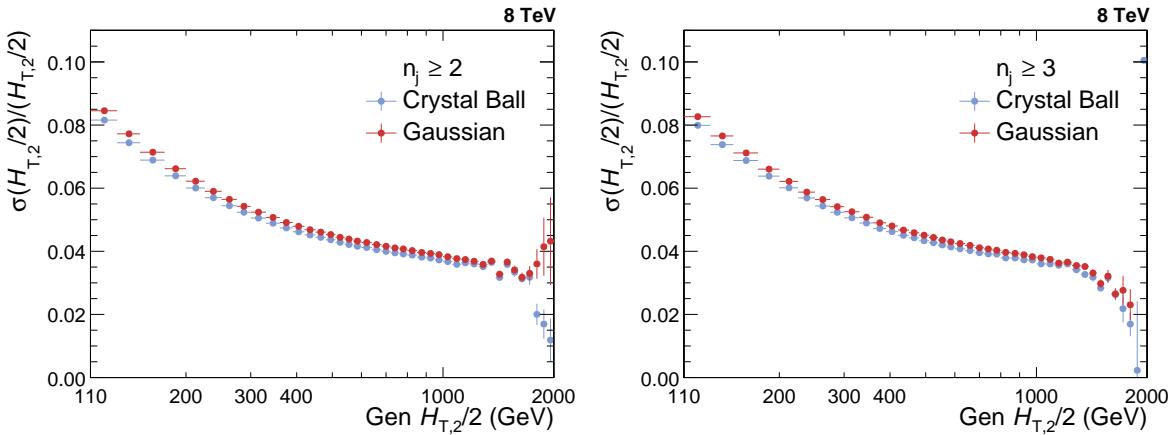


Figure 5.10: Comparison of jet energy resolution calculated using Crystal Ball fit function (blue solid circles) and Gaussian fit function (red solid circles) for inclusive 2-jet (left) and 3-jet (right) events.

the non-Gaussian tails are more pronounced. Hence JER is determined using Crystal 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 Eq. 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

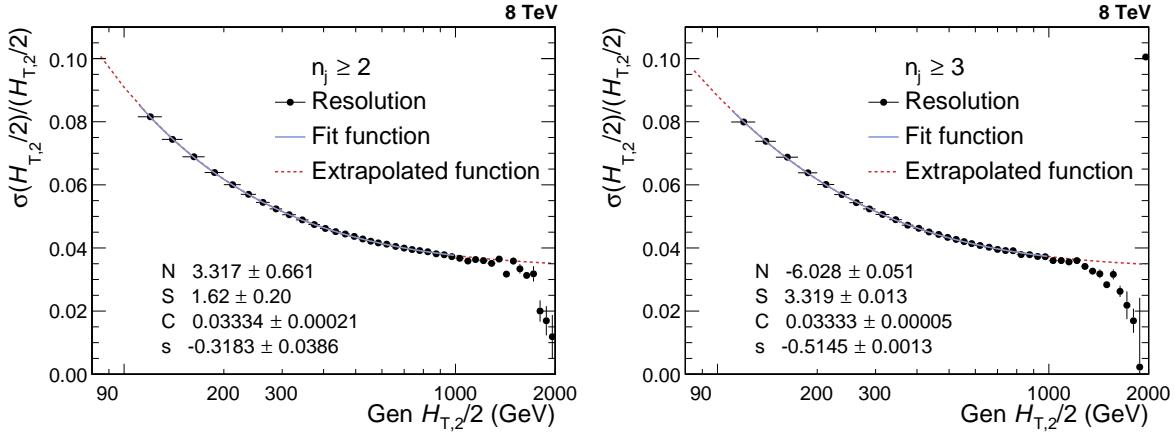


Figure 5.11: Jet energy resolution (JER) is shown as a function of Gen  $H_{T,2}/2$  for inclusive 2-jet (left) and 3-jet (right) events. JER (black solid circles) is fitted by using the modified NSC-formula (blue solid line) which is extrapolated to 80 GeV and up to 2000 GeV (red dashed line) to consider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

$H_{T,2}/2$ , the fit function is extrapolated to 80 GeV and up to 2000 GeV, as shown 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. 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

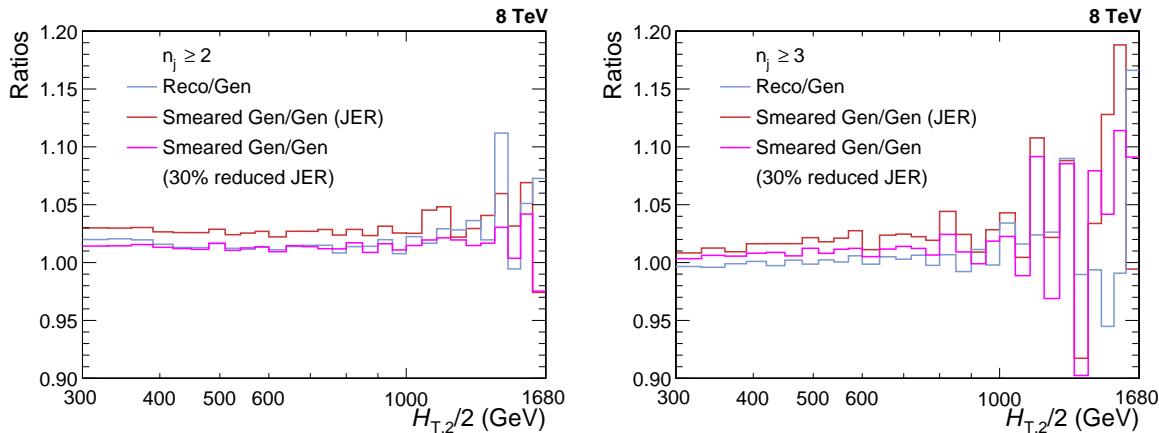


Figure 5.12: MADGRAPH5+PYTHIA6 (MG5+P6) Gen smeared using extracted jet energy resolution (JER) shows a discrepancy from simulated-Reco as Smeared-Gen/Gen ratio (red line) does not match with Reco/Gen ratio (blue line), for both inclusive 2-jet (left) and 3-jet (right) events. Smeared-Gen/Gen ratio (pink 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.  $\chi^2$  between two successive iterations is given by Eq. 5.9. The number of iterations stop when  $\chi^2/N_{bins}$  is  $< 1$ . A reduced  $\chi^2$  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  $\chi^2$  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 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 :

**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 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 :**

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

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  $\sigma$  of the Gaussian is determined from the relative

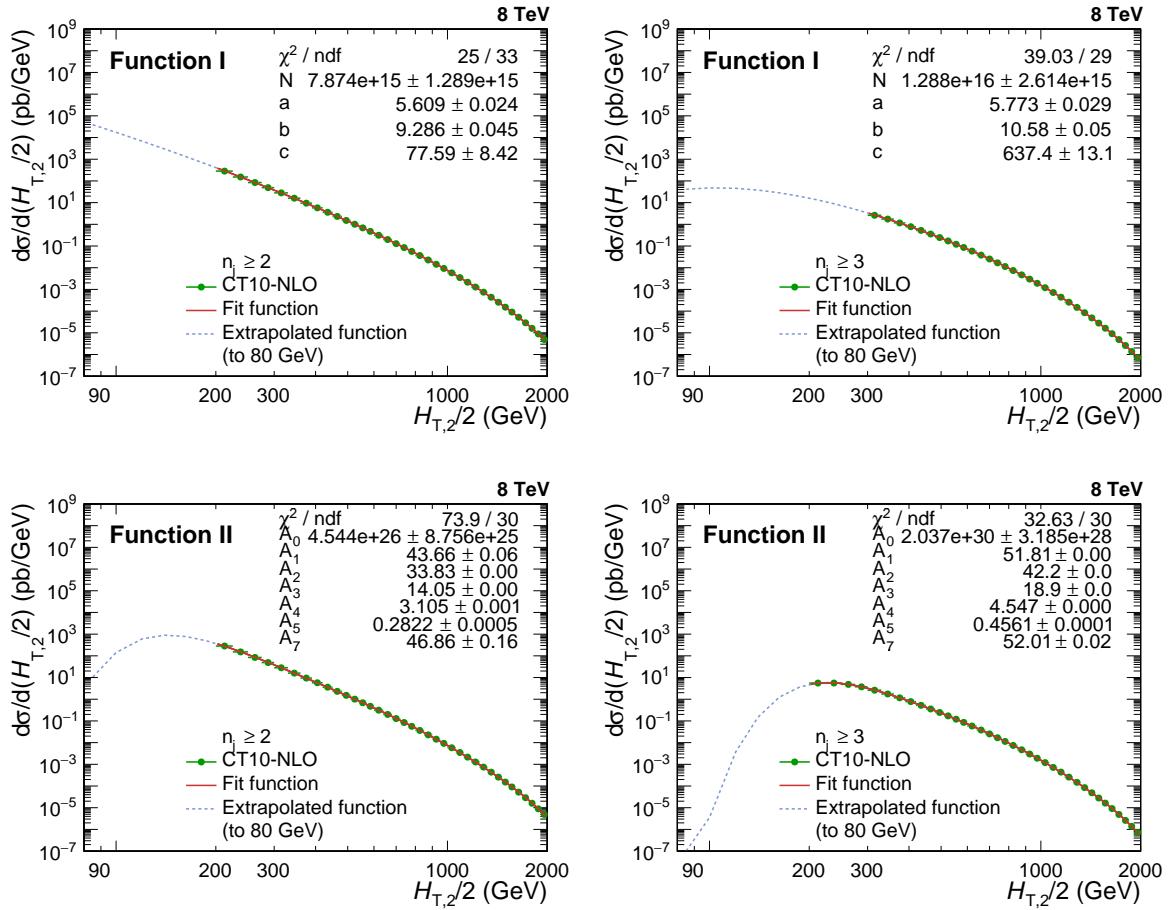


Figure 5.13: Fitted CT10-NLO spectrum of differential cross-section as a function of  $H_{T,2}/2$  (green solid circles) using Function I (top) defined in Eq. 5.10 and using Function II (bottom) given by Eq. 5.13, for inclusive 2-jet (left) and 3-jet (right) events. To consider the migration to lower  $H_{T,2}/2$  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 Eq. 5.8. The parameters  $N$ ,  $S$ ,  $C$  used for smearing are taken from Table 5.6. These randomly generated ( $\text{Gen}_{\text{Toy}}$ ) 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.

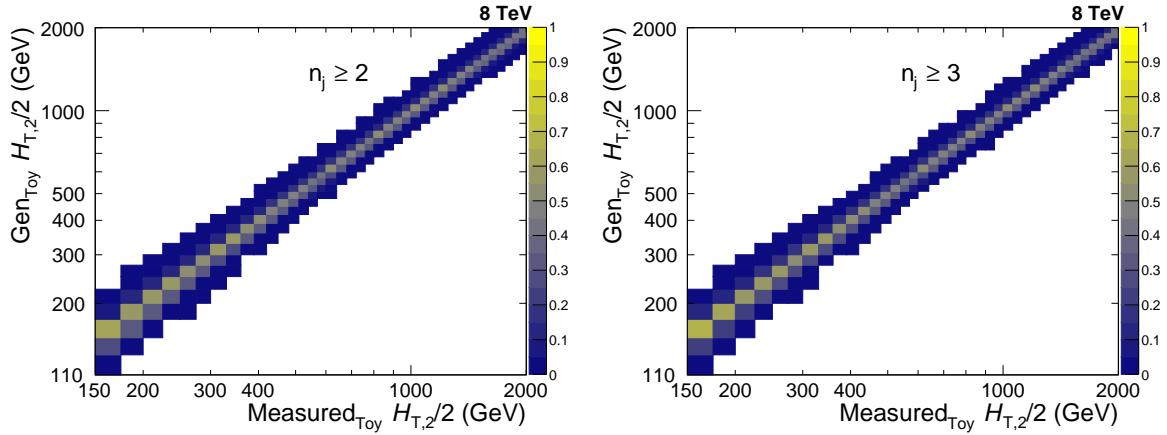


Figure 5.14: The response matrices are derived using the Toy Monte Carlo and forward smearing method, for inclusive 2-jet (left) and 3-jet (right) events. The matrices are normalized to the number of events in each column and are diagonal with small off-diagonal migrations between close-by  $H_{T,2}/2$  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  $\text{Gen}_{\text{Toy}}$  spectrum should be re-obtained after unfolding. Figure 5.16 confirms that the unfolded  $\text{Measured}_{\text{Toy}}$

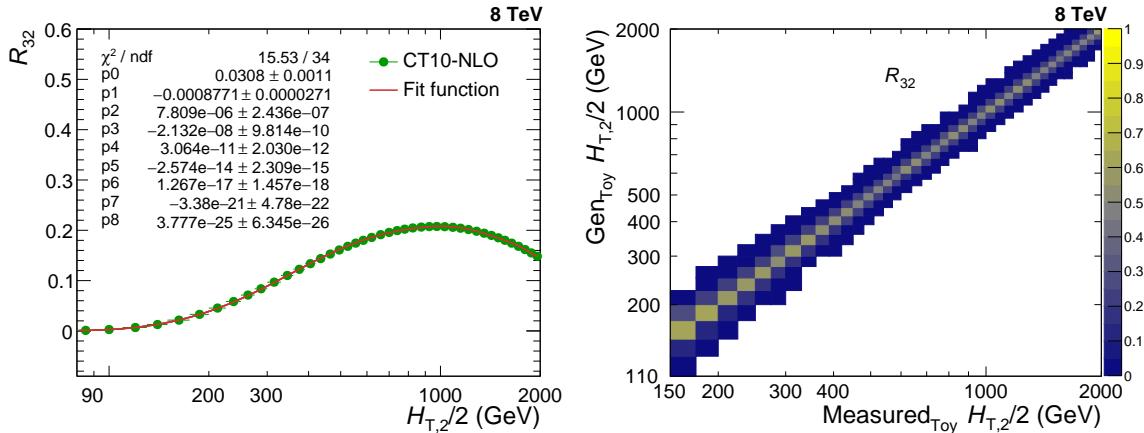


Figure 5.15: Left : The ratio of cross-sections described by Eq. 5.10 for inclusive 3-jet to that of 2-jet events is shown as a function of  $H_{T,2}/2$  (green solid circles). It is fit using a polynomial function of degree 8 (red line). Right : The response matrix is derived using the Toy Monte Carlo 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  $\text{Gen}_{\text{Toy}}$  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).

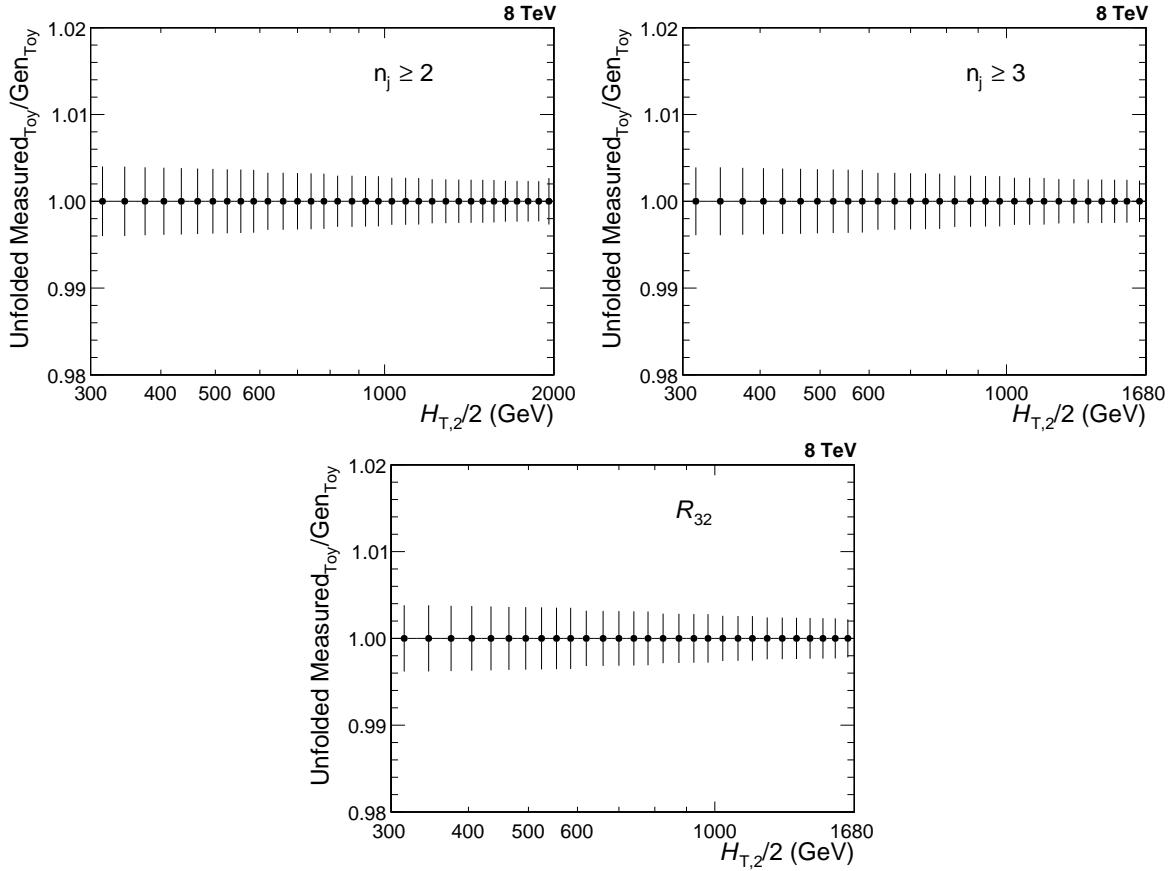


Figure 5.16: Closure test of the unfolding technique where the smeared spectrum obtained from Toy Monte Carlo method (Measured<sub>Toy</sub>), is unfolded using the constructed response matrices (obtained by forward smearing the randomly generated spectrum (Gen<sub>Toy</sub>) using extracted jet energy resolution (JER)). As expected, the unfolded measured<sub>Toy</sub> spectrum matches exactly with Gen<sub>Toy</sub> 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.

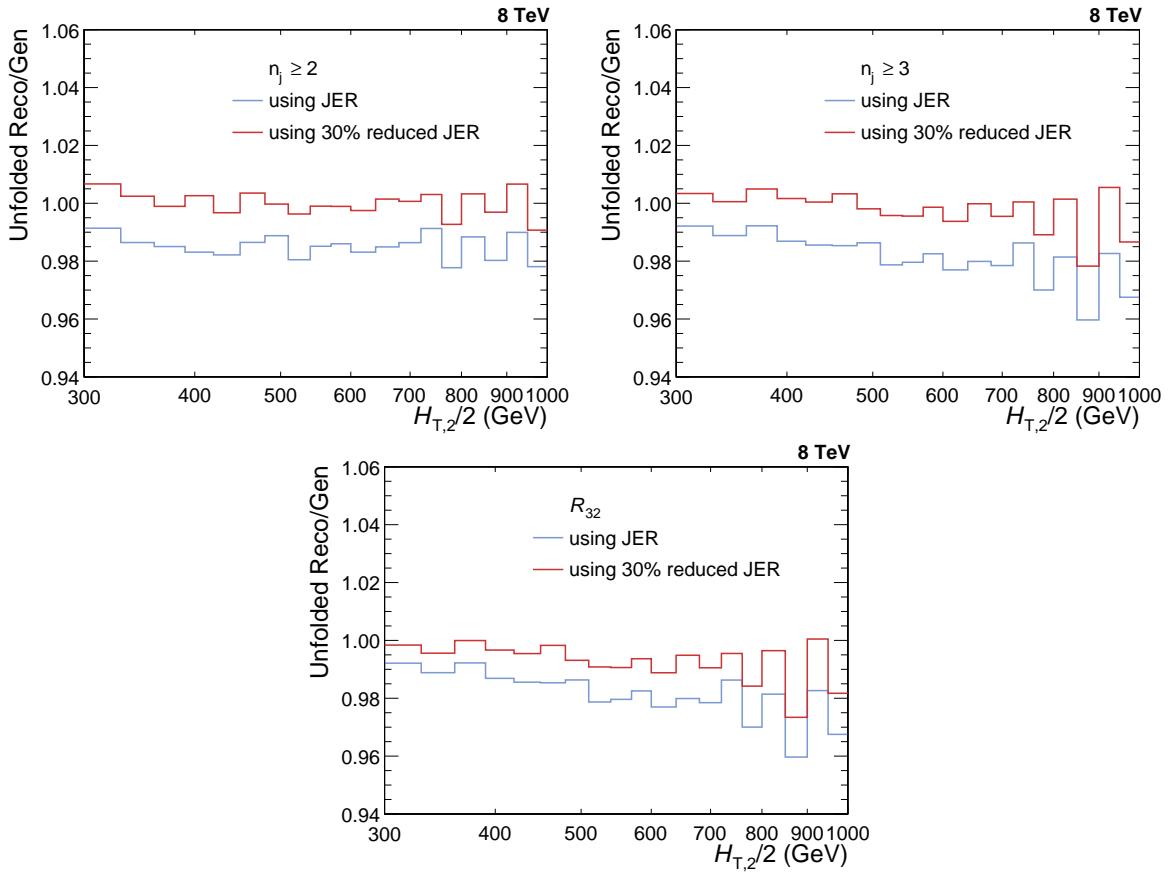


Figure 5.17: Reco MADGRAPH5+PYTHIA6 Monte Carlo (MG5+P6 MC) differential cross-section distributions unfolded with the response matrices (obtained by forward smearing the randomly generated spectrum (Gen) using extracted jet energy resolution (JER)), does not give a good closure with Gen MG5+P6 MC (blue 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

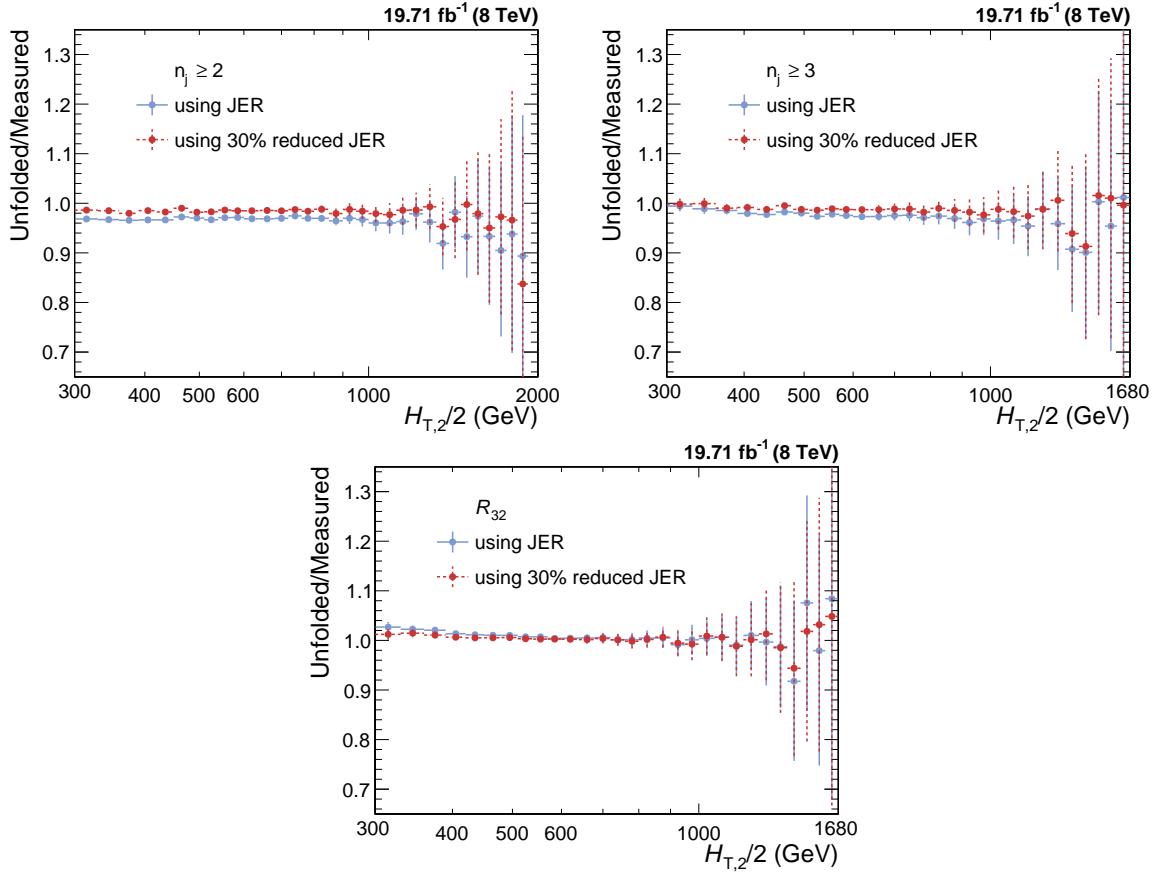


Figure 5.18: The measured differential cross-sections as well as the cross-section ratio  $R_{32}$  are unfolded as a function of  $H_{T,2}/2$  using the response matrices derived using the Toy Monte Carlo and forward smearing method. The unfolded spectrum are compared with that of the measured one for 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. 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,  $\alpha_S$ .

### 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  $1\sigma$  shift and is fully corre-

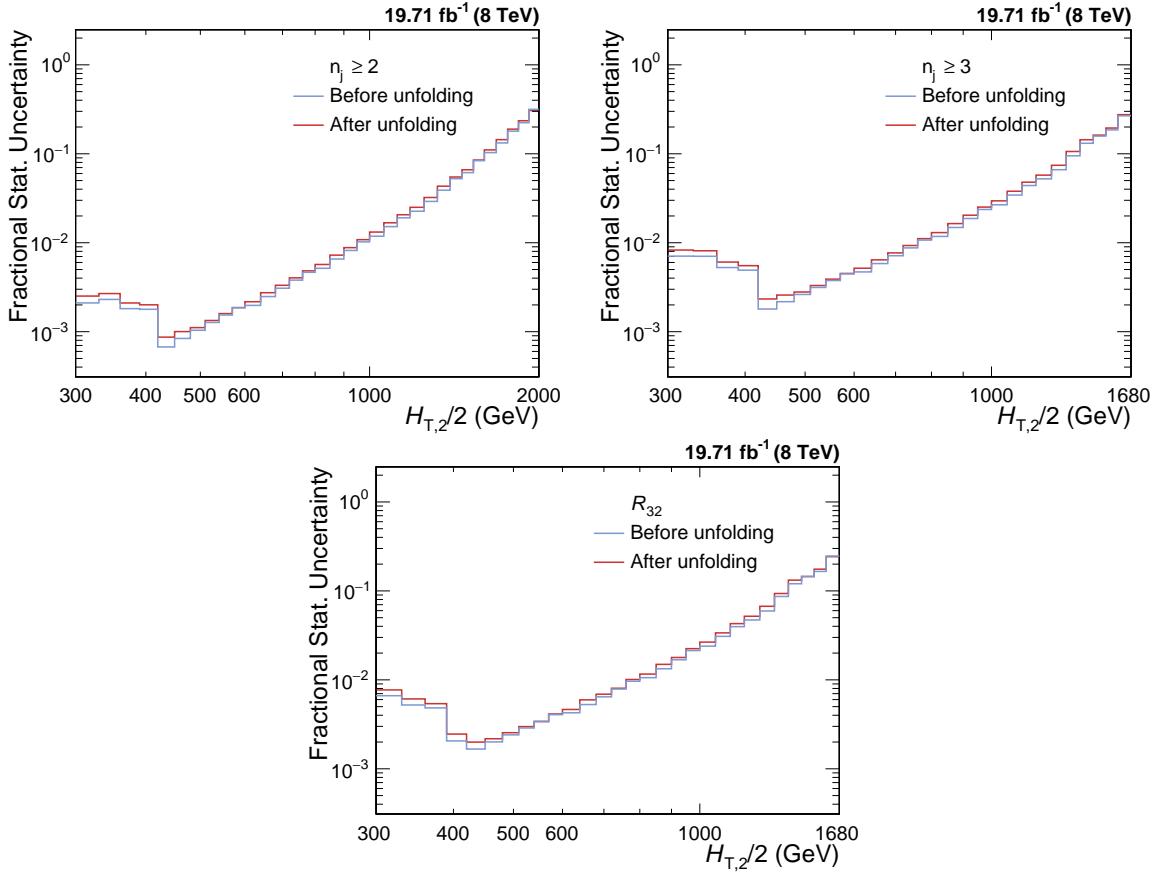


Figure 5.19: The fractional statistical uncertainties of the unfolded data (red line) are compared with those of the measured one (blue line) for inclusive 2-jet (top left) and 3-jet (top right) events cross-sections as well as for the cross-section ratio  $R_{32}$  (bottom). After unfolding, the statistical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

lated in  $p_T$  and  $\eta$  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-

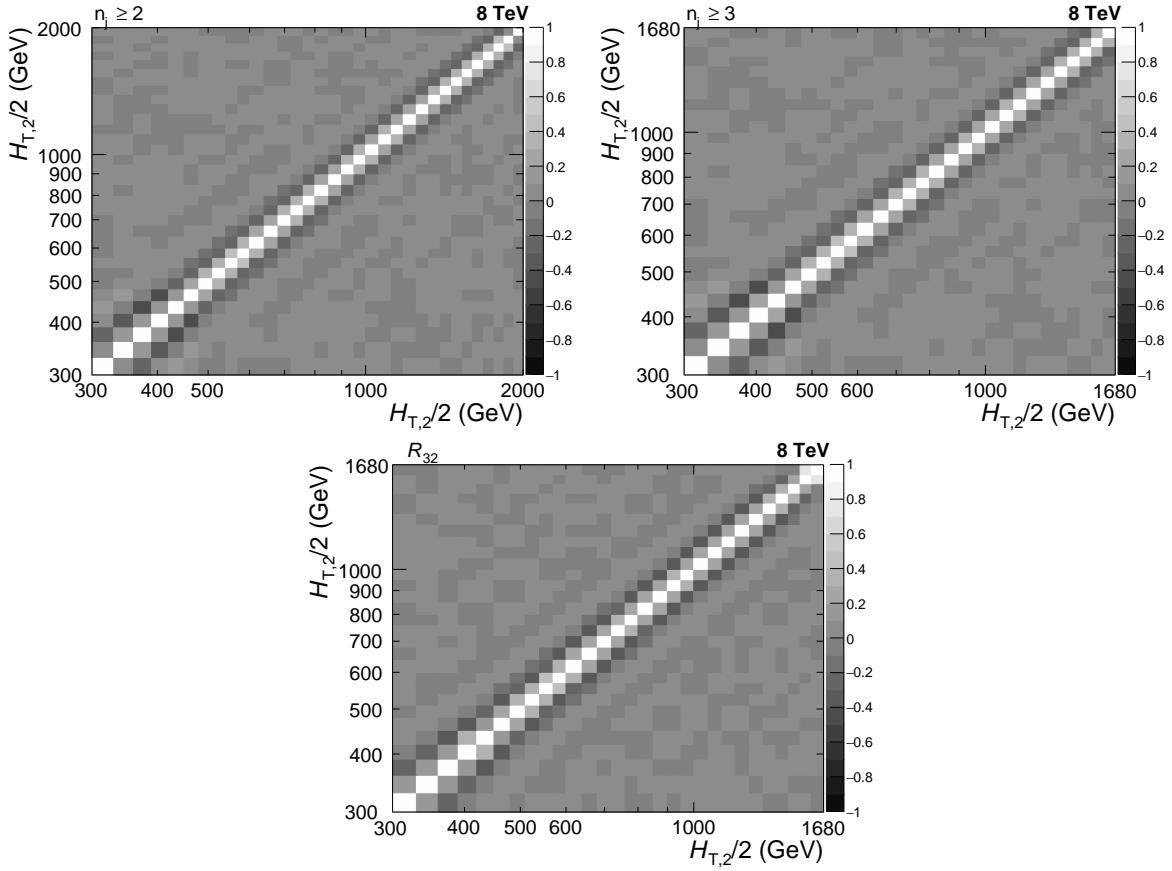


Figure 5.20: The unfolding procedure introduces the correlations of the statistical uncertainty through bin migrations which are shown here for inclusive 2-jet (top left) and and 3-jet (top right) events cross-sections as well as for the cross-section ratio  $R_{32}$  (bottom). The correlation (anti-)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

ysis, JEC uncertainties are 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  $Z/\gamma$ +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  $\eta$ -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  $Z/\gamma$ +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  $\gamma$ +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

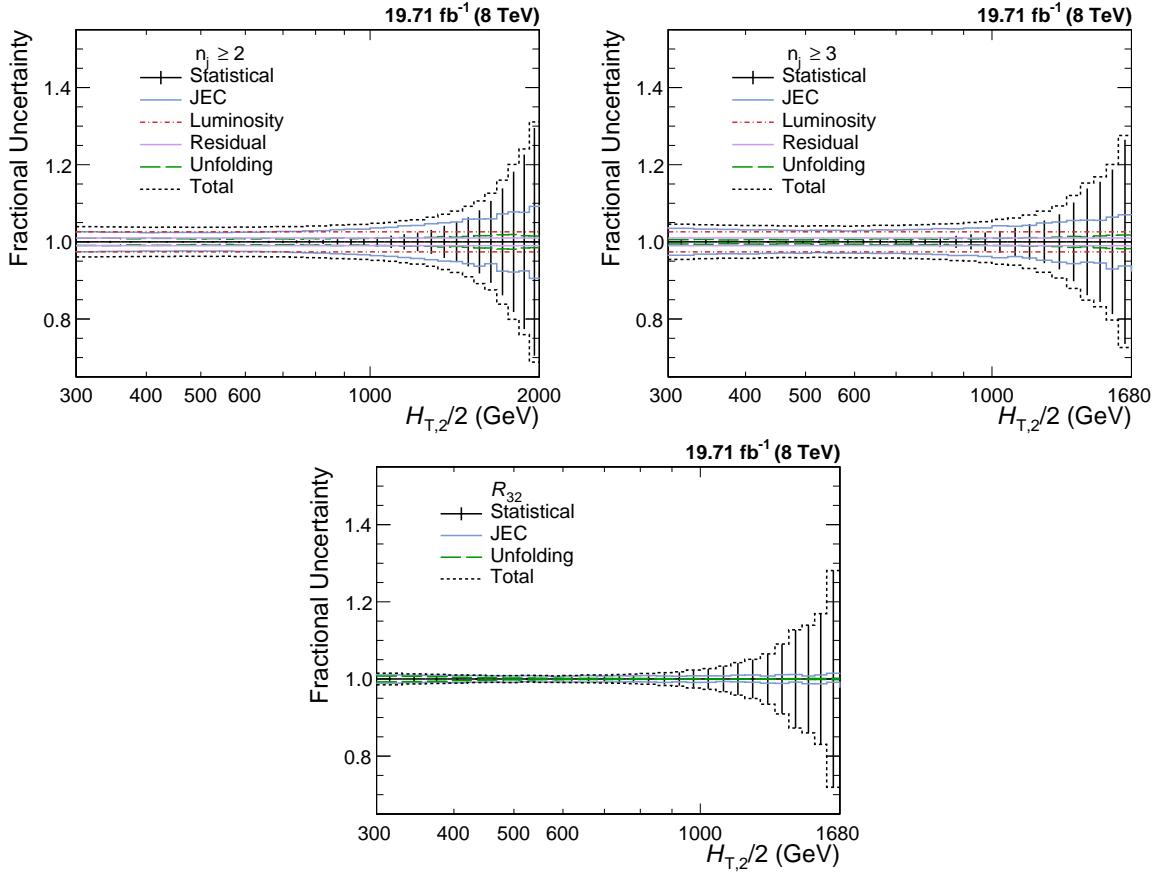


Figure 5.21: Experimental uncertainties from different sources affecting the measurement of cross-sections for inclusive 2-jet (top left) and 3-jet (top right) events and the cross-section ratio  $R_{32}$  (bottom). The error bars represent the statistical uncertainty after unfolding. The systematic 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  $H_{\mathrm{T},2}/2$  values and by statistical uncertainty at higher  $H_{\mathrm{T},2}/2$  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  $H_{\mathrm{T},2}/2$  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  $H_{\mathrm{T},2}/2$ . 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

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 ( $\mu_r$ ) and factorization ( $\mu_f$ ) scales. The next-to-leading order (NLO) calculations improve the precision by reducing the dependence on  $\mu_r$  and  $\mu_f$  scales. This makes the NLO calculations an essential feature in the determination of fundamental parameters such as  $\alpha_S$  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 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  $H_{\text{T},2}/2$ , i.e.  $\mu_f = \mu_r = H_{\text{T},2}/2$ .

In the current study, the 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 0.0001 as step size from lowest to highest value. All PDF 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]	$\leq 5$	172	91.188	0.1180	0.112 - 0.127
MSTW2008 [116, 117]	$\leq 5$	$10^{10}$	91.1876	0.1202	0.110 - 0.130
NNPDF2.3 [118]	$\leq 6$	175	91.1876	0.1180	0.114–0.124
CT14 [27]	$\leq 5$	172	91.1876	0.1180	0.111–0.123
HERAPDF2.0 [31]	$\leq 5$	173	91.1876	0.1180	0.110–0.130
MMHT2014 [28]	$\leq 5$	$10^{10}$	91.1876	0.1200	0.108–0.128
NNPDF3.0 [29]	$\leq 5$	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 effects, the minimum value of  $H_{T,2}/2$  studied is from 300 GeV.

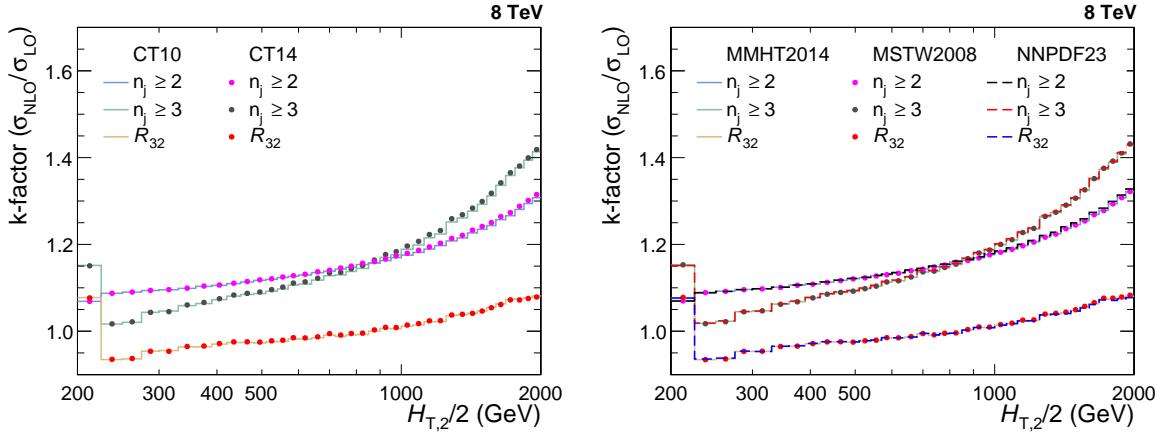


Figure 6.1: The k-factors of the NLOJET++ calculations, for inclusive 2-jet and 3-jet event cross-sections and their ratio  $R_{32}$ , using five different PDF sets.

### 6.1.2 Non-perturbative Corrections

The fixed-order pQCD NLO calculations predict the parton-level cross-section but lacks 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 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  $R_{32}$ .

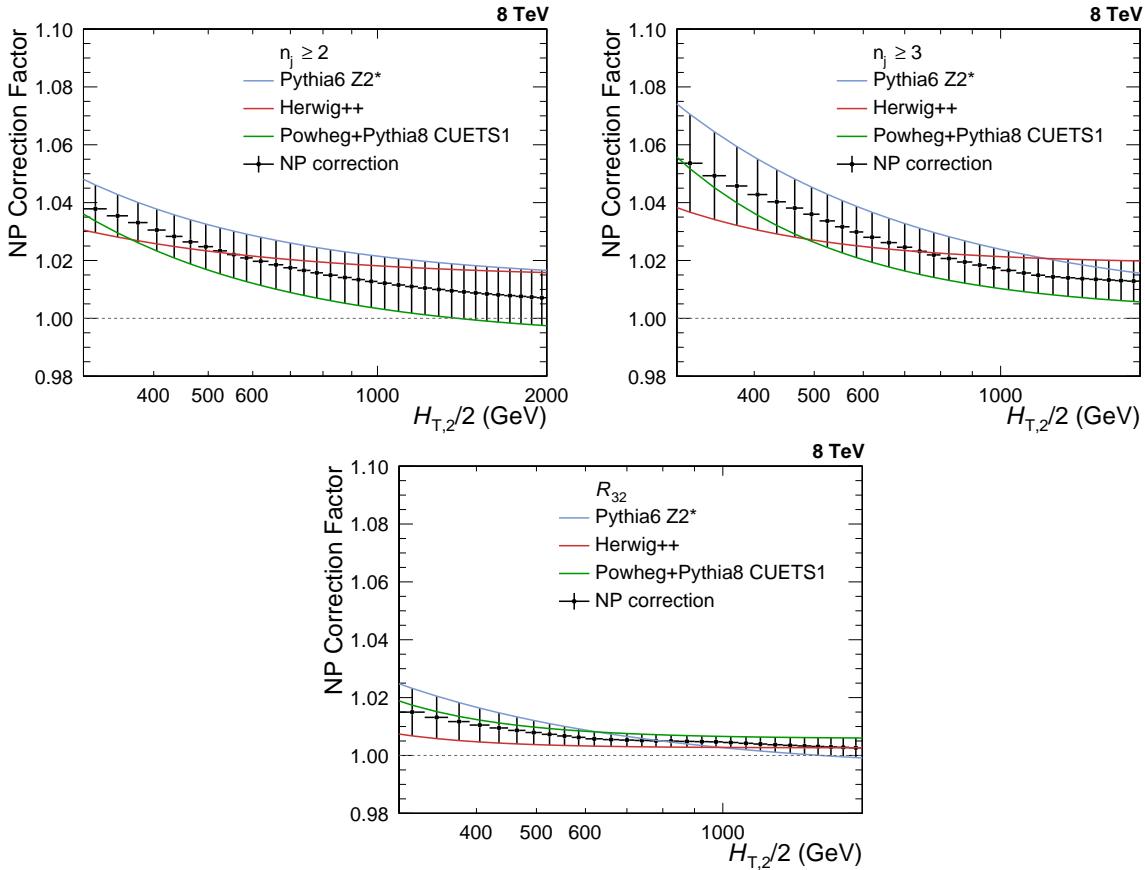


Figure 6.2: The nonperturbative (NP) corrections are presented as a function of  $H_{T,2}/2$  for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections, as well as their ratio  $R_{32}$ . These corrections are calculated from the leading order HERWIG++ with the default tune of version 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 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-

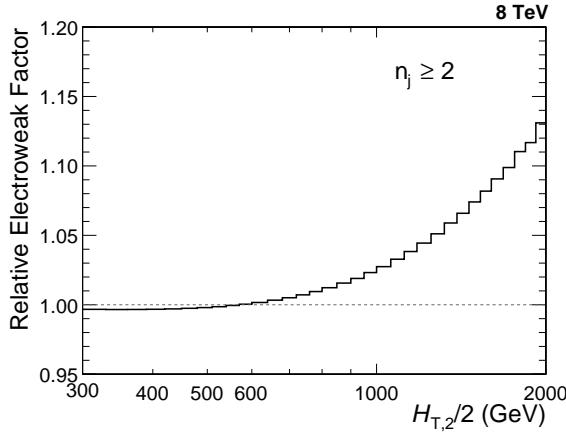


Figure 6.3: The electroweak (EW) corrections [120] in the phase space of the measurement are shown as a function of  $H_{T,2}/2$  for inclusive 2-jet event cross-sections. These corrections are applied as a bin-by-bin correction factor to the fixed-order calculation of NLOJET++ as well as the MC predictions 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.

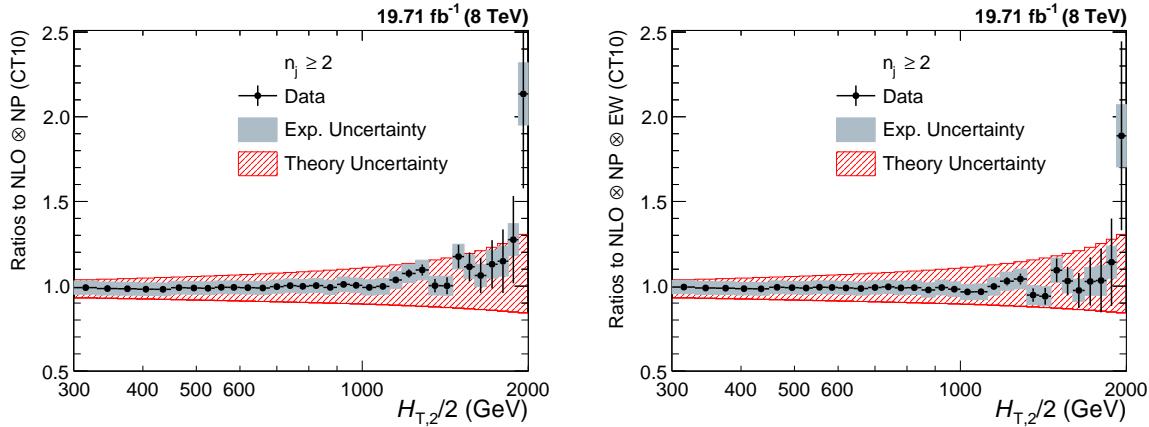


Figure 6.4: Ratio of the data over theory obtained using the CT10-NLO PDF set and corrected with non-perturbative effects (NP) without including electroweak (EW) corrections (left) and including EW corrections (right) is shown for inclusive 2-jet event cross-sections. The error bars represents the statistical 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 ( $\mu_r$ ) and factorization ( $\mu_f$ ) 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  $\alpha_S$ , 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  $\mu_r$  and  $\mu_f$  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  $\alpha_S$ , 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  $N_{ev} = 26$  eigenvectors with two PDF members per eigenvector  $k$ , which are varied upwards and downwards to generate a set of eigenvector pairs. The asymmetric uncertainties,  $\Delta X^+$  and  $\Delta X^-$ , 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 ( $\Delta X^\pm$ ) 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

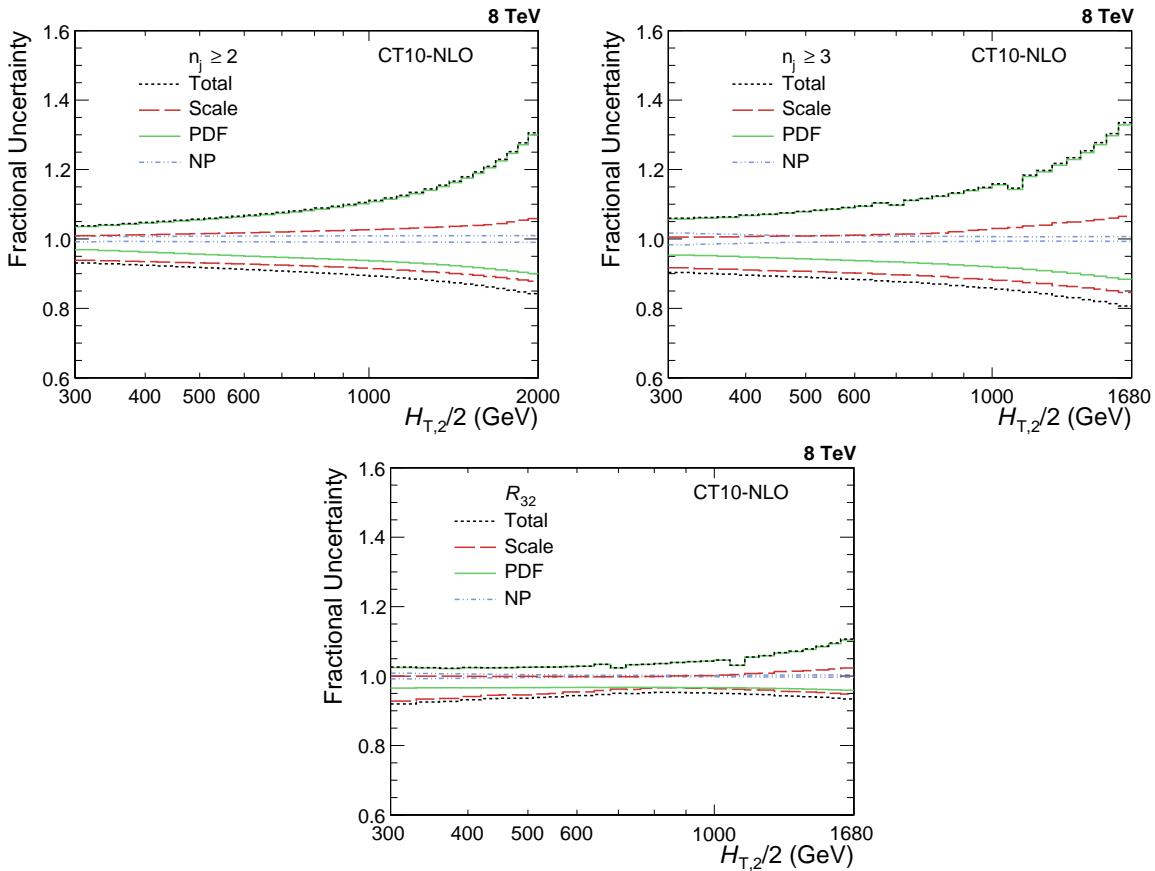


Figure 6.5: The systematic theoretical uncertainties affecting the cross-section measurement for inclusive 2-jet (top left) and 3-jet (top right) events and their ratio  $R_{32}$  (bottom). The scale (red dashed line), PDF (green line) and NP (blue dashed line) uncertainties as well as total uncertainty (black dashed line) obtained using 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  $\sim 700$  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%

### 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<sup>\*</sup> (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.

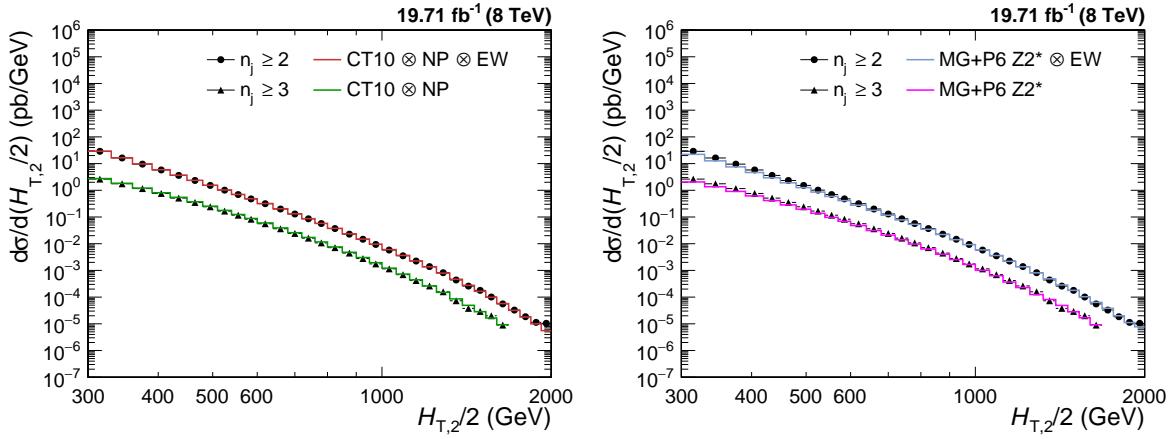


Figure 6.6: Comparison of the measured differential inclusive 2-jet and 3-jet event cross-sections as a function of  $H_{T,2}/2$  to theoretical predictions. On the left, the data (points) are shown together with NLOJET++ predictions (line) using the CT10-NLO PDF set, corrected for non-perturbative (NP) 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<sup>\*</sup> (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

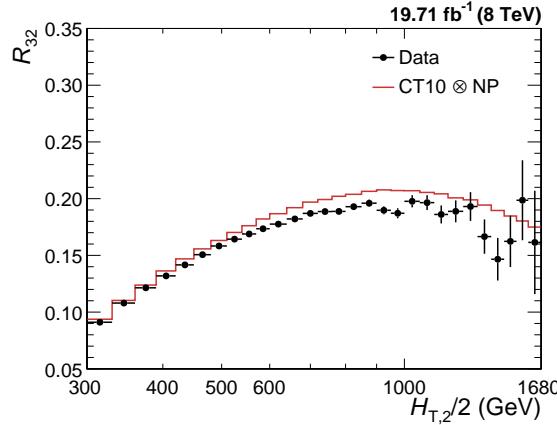


Figure 6.7: Cross-section ratio  $R_{32}$  as a function of  $H_{T,2}/2$  calculated from data (solid circles) in comparison to that from NLO pQCD predictions obtained using the CT10-NLO PDF set corrected with non-perturbative (NP) corrections (line). The error bars correspond to the total experimental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

studied in detail. In Fig. 6.8, the ratios of the data over NLOJET++ 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

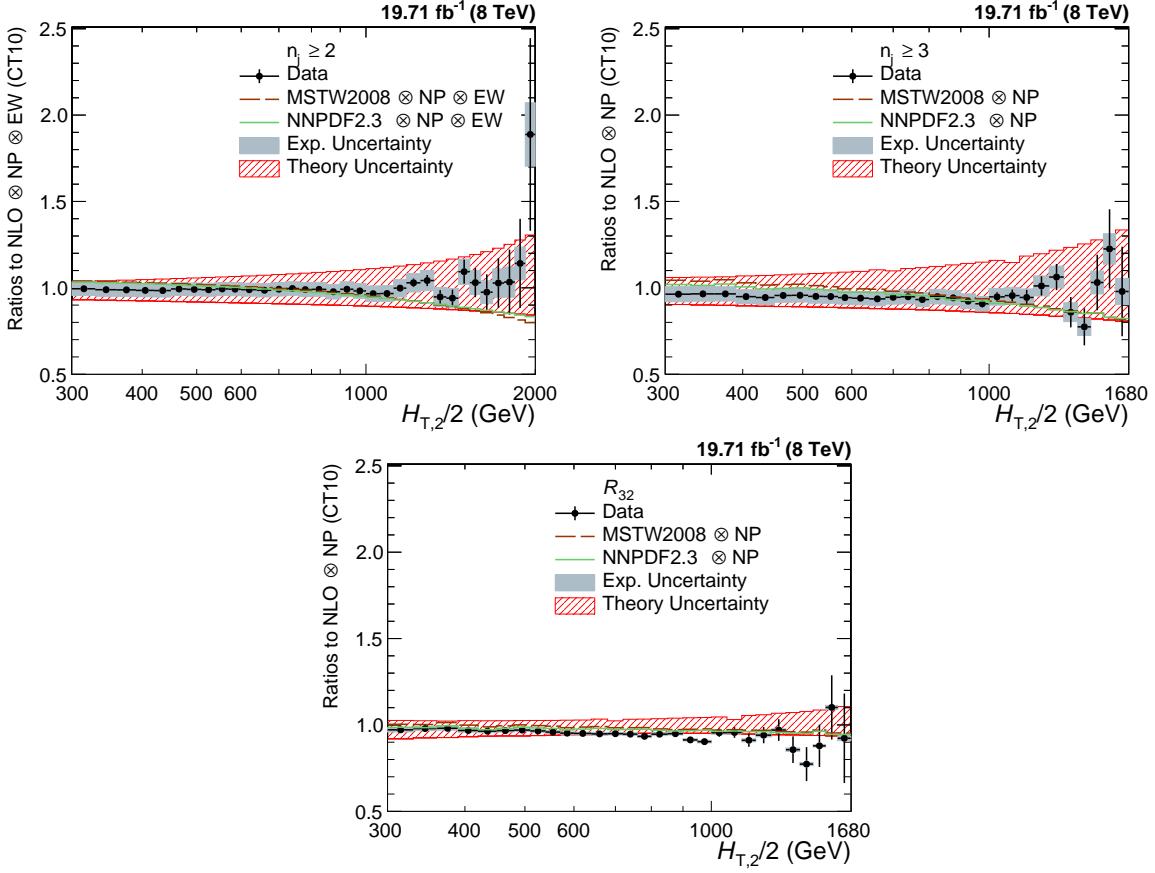


Figure 6.8: Ratio of the data over theory using the CT10-NLO PDF set for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections and their ratio  $R_{32}$  (bottom). The theory predictions are corrected for non-perturbative effects (NP) and also for electroweak effects (EW) for inclusive 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  $\alpha_S$ , which is discussed in the next chapter.

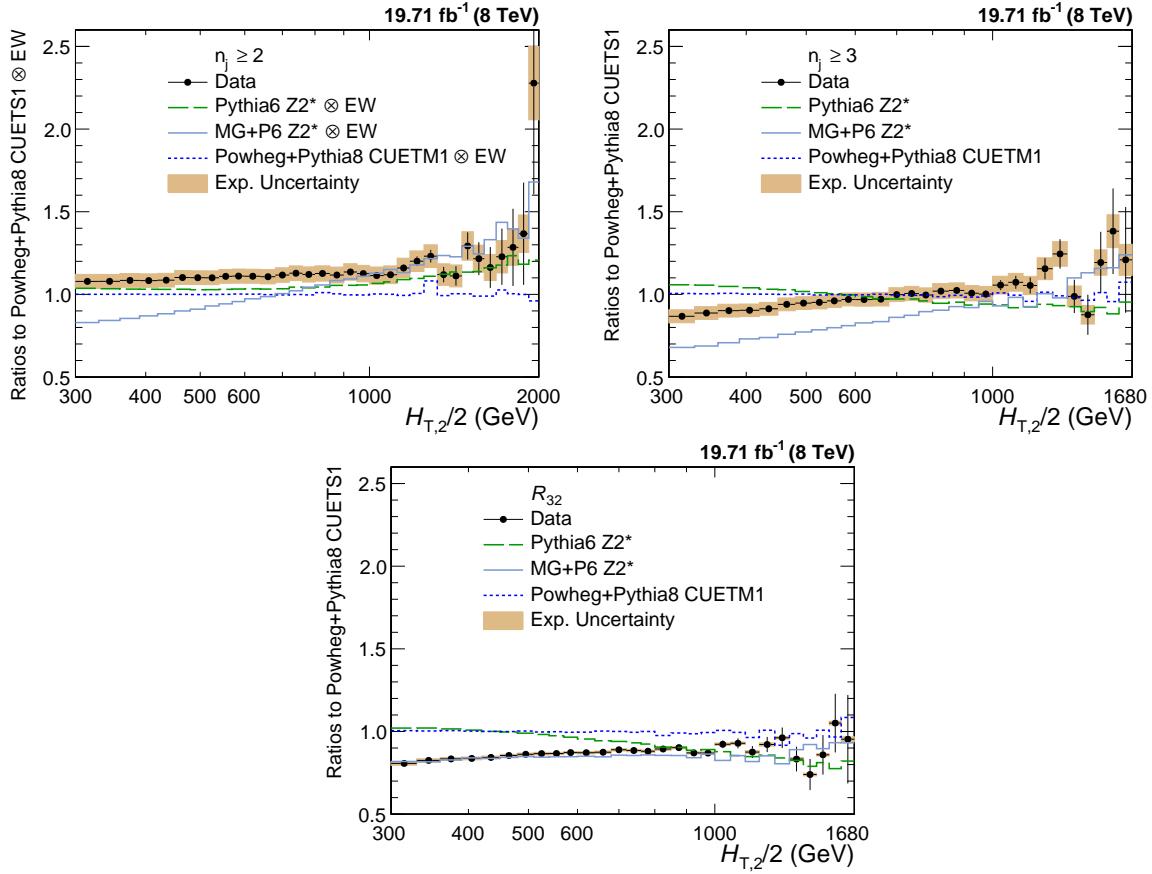


Figure 6.9: Ratio of the data over the predictions from POWHEG+PYTHIA8 with tune CUETS1 are presented for inclusive 2-jet (top left) and 3-jet (top right) event cross-sections as well as their ratio  $R_{32}$  (bottom). For comparison the alternative tune CUETM1 of POWHEG+PYTHIA8, the tree-level multi-leg 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  $\alpha_s$  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  $\alpha_s$  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  $H_{\text{T},2}/2$ ,  $\mu_r$  and  $\mu_f$  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  $\mu_r$  and  $\mu_f$ , 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  $\alpha_S$  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 ( $\chi^2$ ) between the experimental measurements and the theoretical predictions. The  $\chi^2$  is given by the following equation :

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

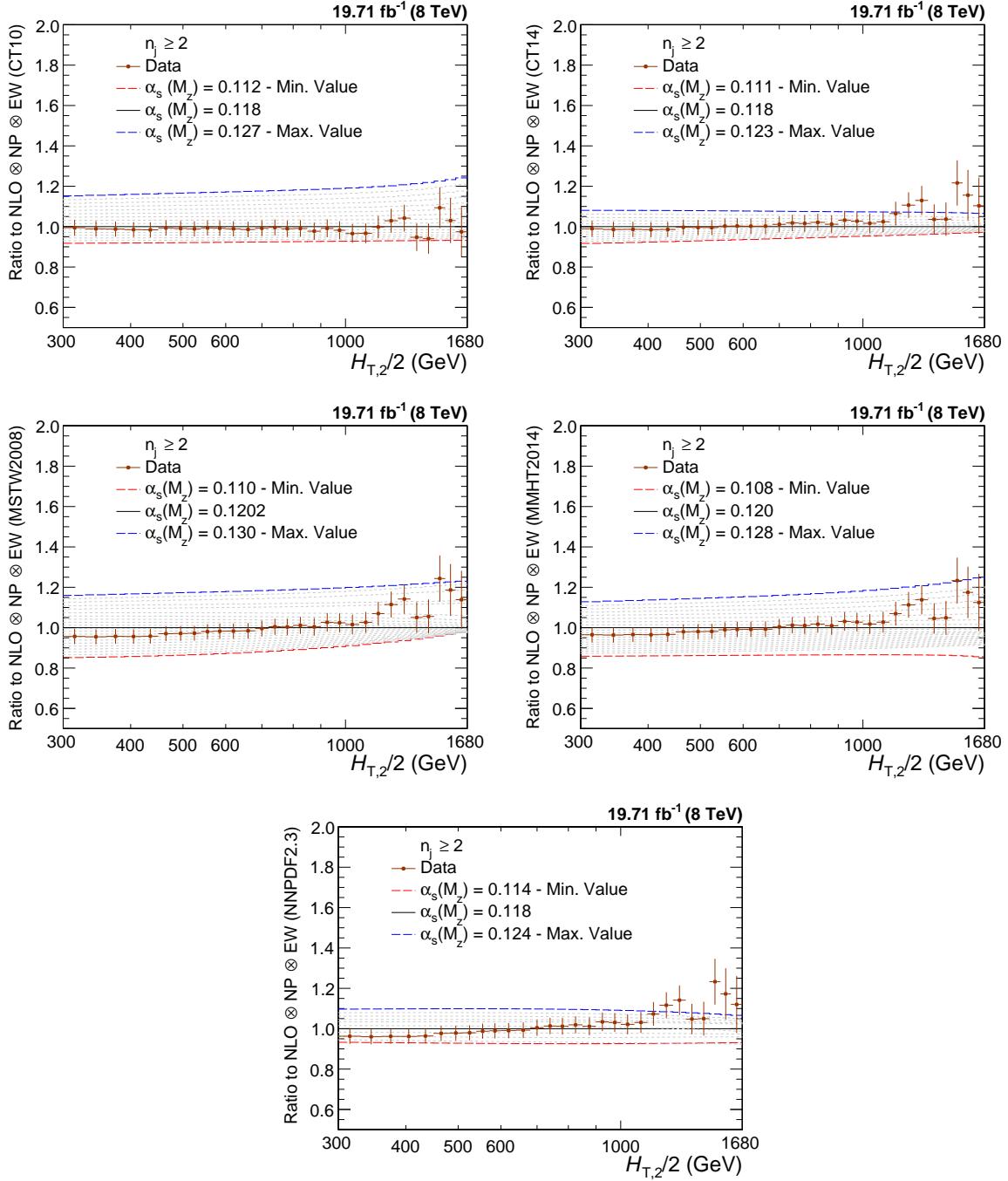


Figure 7.1: Ratio of the measured inclusive 2-jet differential cross-section to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for a series of values of  $\alpha_s(M_Z)$ . The  $\alpha_s(M_Z)$  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 error bars correspond to the total experimental uncertainty. The theory predictions are corrected for non-perturbative (NP) and electroweak (EW) effects.

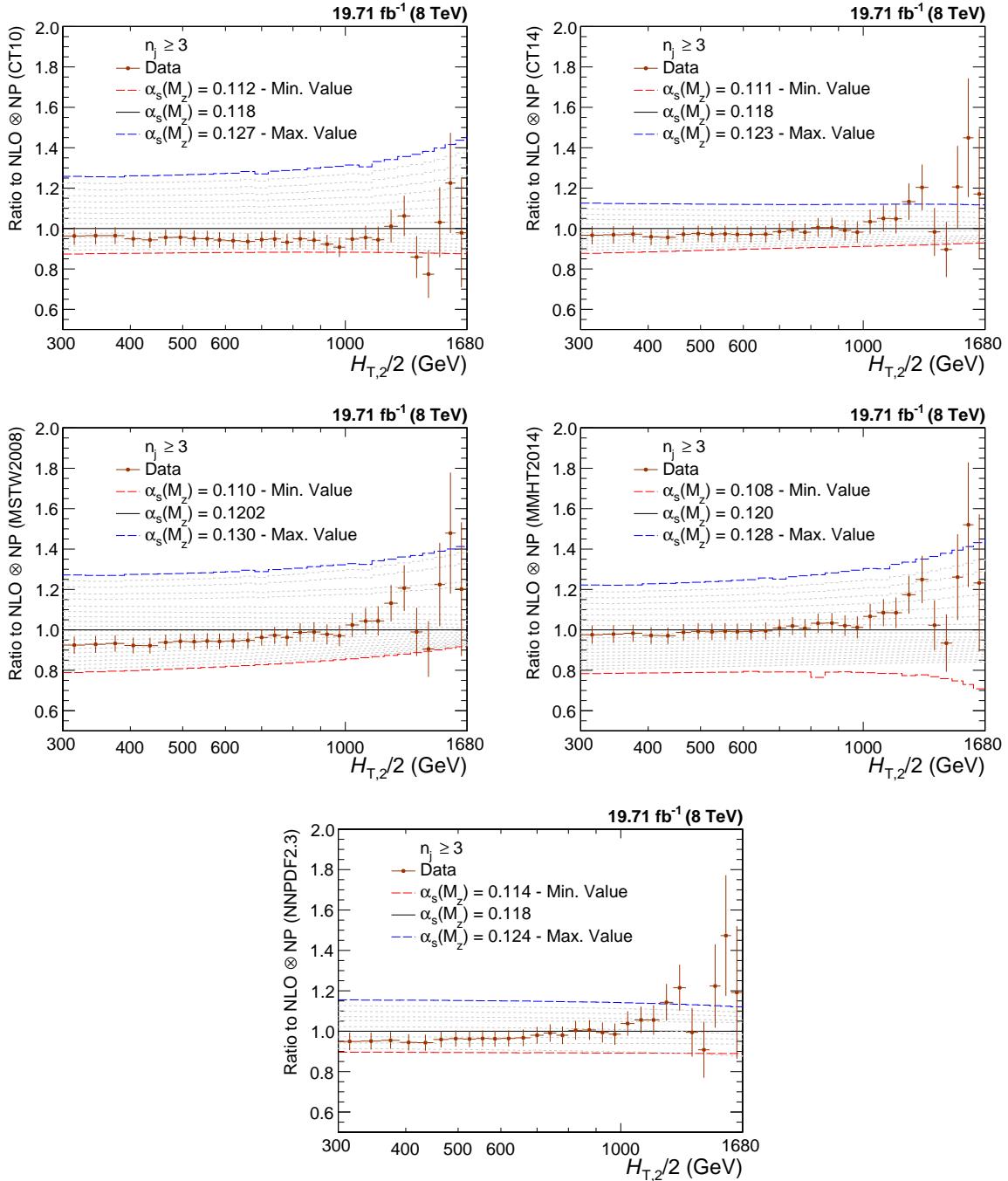


Figure 7.2: Ratio of the measured inclusive 3-jet differential cross-section to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for a series of values of  $\alpha_s(M_Z)$ . The  $\alpha_s(M_Z)$  value is varied in the range 0.112-0.127, 0.111-1.23, 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.

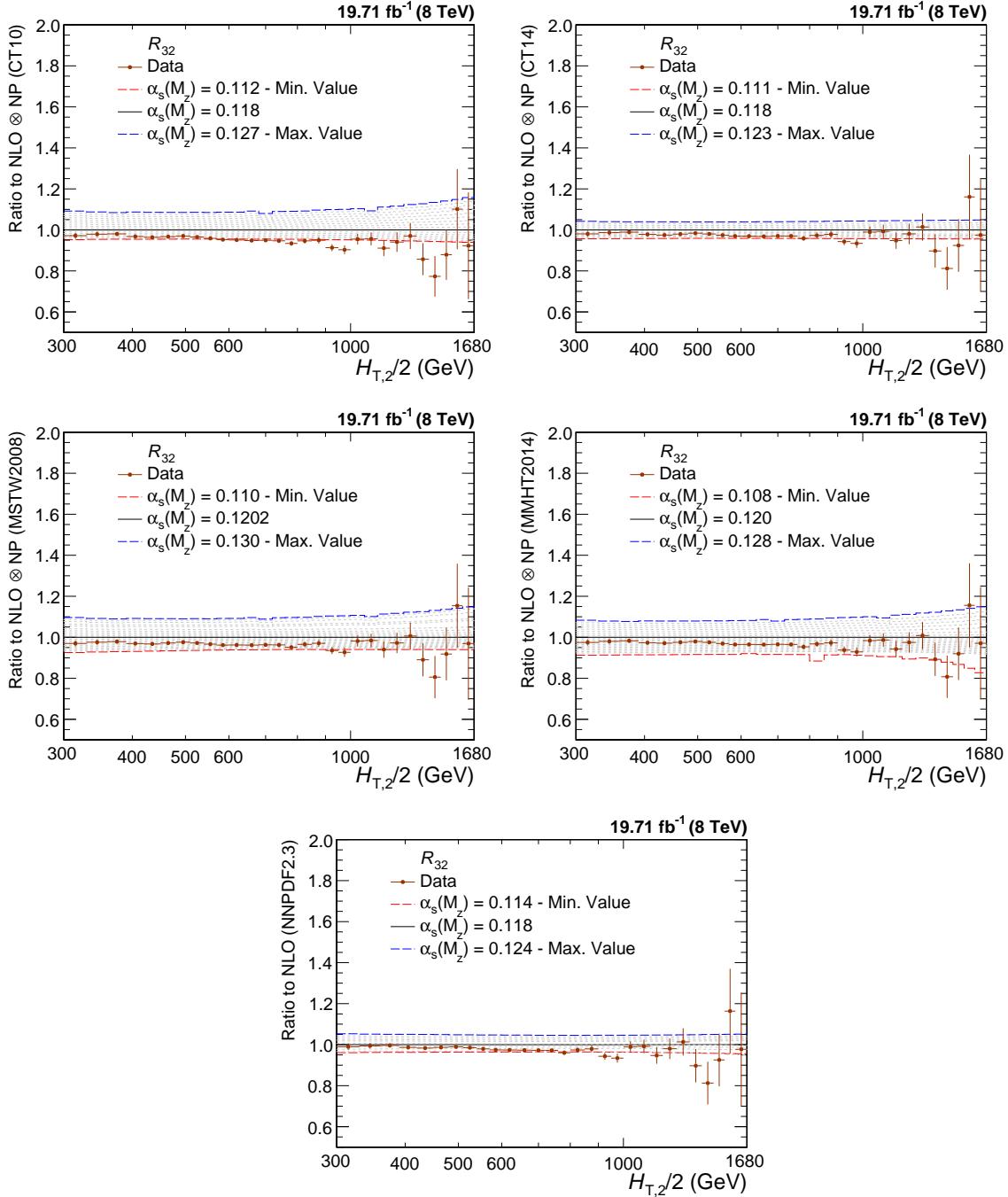


Figure 7.3: Ratio of the measured cross-section ratio,  $R_{32}$  to theory predictions using the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and NNPDF2.3 (bottom) NLO PDF sets for a series of values of  $\alpha_s(M_Z)$ . The  $\alpha_s(M_Z)$  value is varied in the range 0.112-0.127, 0.111-123, 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.
- $\text{Cov}^{\text{JEC}}$ : the jet energy corrections (JEC) systematic uncertainty.
- $\text{Cov}^{\text{Unfolding}}$ : the unfolding systematic uncertainty including the resolution (JER) and model dependence.
- $\text{Cov}^{\text{Lumi}}$ : 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.

- $\text{Cov}^{\text{NP}}$ : the systematic uncertainty of the non-perturbative (NP) corrections.
- $\text{Cov}^{\text{PDF}}$ : 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  $H_{\text{T},2}/2$  bins. If  $\delta_i$  is the total uncertainty on the differential cross-section, for the  $i$ -th  $H_{\text{T},2}/2$  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 ( $N_{\text{ev}}$ ) 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 ( $N_{\text{rep}}$ ) 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  $\Delta X$  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  $H_{\mathrm{T},2}/2$  is restricted to be between 300 GeV and 1 TeV to avoid the region close to the minimal  $p_{\mathrm{T}}$  threshold of 150 GeV for each jet at low  $p_{\mathrm{T}}$  and the onset of electroweak effects at high  $H_{\mathrm{T},2}/2$ , which are available for the dijet case only. The  $\alpha_s(M_Z)$  results obtained from a simultaneous fit to all 19  $H_{\mathrm{T},2}/2$  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,  $\chi^2/n_{\mathrm{dof}}$  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  $\chi^2/n_{\text{dof}}$  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		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$
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			$R_{32}$		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$
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  $\chi^2/n_{\text{dof}}$ , 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		
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$
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 :

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

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

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

### 7.3 Running of the Strong Coupling Constant

The value of the strong coupling constant  $\alpha_S$  depends on the energy scale  $Q$  and it decreases with the increase of scale  $Q$ . To study this dependence, the determination of  $\alpha_S$  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	$\chi^2/n_{\text{dof}}$
MSTW2008	0.1150	$\pm 0.0010$	$\pm 0.0013$	$\pm 0.0015$	$\pm 0.0023$	$^{+0.0050}_{-0.0000}$	26./28
MMHT2014	0.1142	$\pm 0.0010$	$\pm 0.0013$	$\pm 0.0014$	$\pm 0.0022$	$^{+0.0049}_{-0.0006}$	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	
		$\alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$	$\alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$
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  $\alpha_s$  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	$\pm 0.0015$	$\pm 0.0014$	$\pm 0.0019$	$^{+0.0053}_{-0.0000}$	0.1158	$\pm 0.0014$	$\pm 0.0010$	$\pm 0.0019$	$^{+0.0052}_{-0.0000}$
420-600	0.1153	$\pm 0.0011$	$\pm 0.0014$	$\pm 0.0018$	$^{+0.0057}_{-0.0000}$	0.1154	$\pm 0.0011$	$\pm 0.0012$	$\pm 0.0017$	$^{+0.0056}_{-0.0000}$
600-1000	0.1134	$\pm 0.0013$	$\pm 0.0016$	$\pm 0.0019$	$^{+0.0052}_{-0.0000}$	0.1140	$\pm 0.0012$	$\pm 0.0012$	$\pm 0.0018$	$^{+0.0045}_{-0.0000}$
1000-1680	0.1147	$\pm 0.0029$	$\pm 0.0017$	$\pm 0.0018$	$^{+0.0063}_{-0.0011}$	0.1154	$\pm 0.0025$	$\pm 0.0014$	$\pm 0.0015$	$^{+0.0056}_{-0.0011}$
300-1680	0.1150	$\pm 0.0010$	$\pm 0.0013$	$\pm 0.0015$	$^{+0.0050}_{-0.0000}$	0.1142	$\pm 0.0010$	$\pm 0.0013$	$\pm 0.0014$	$^{+0.0049}_{-0.0006}$

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	$\chi^2/n_{\text{dof}}$
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

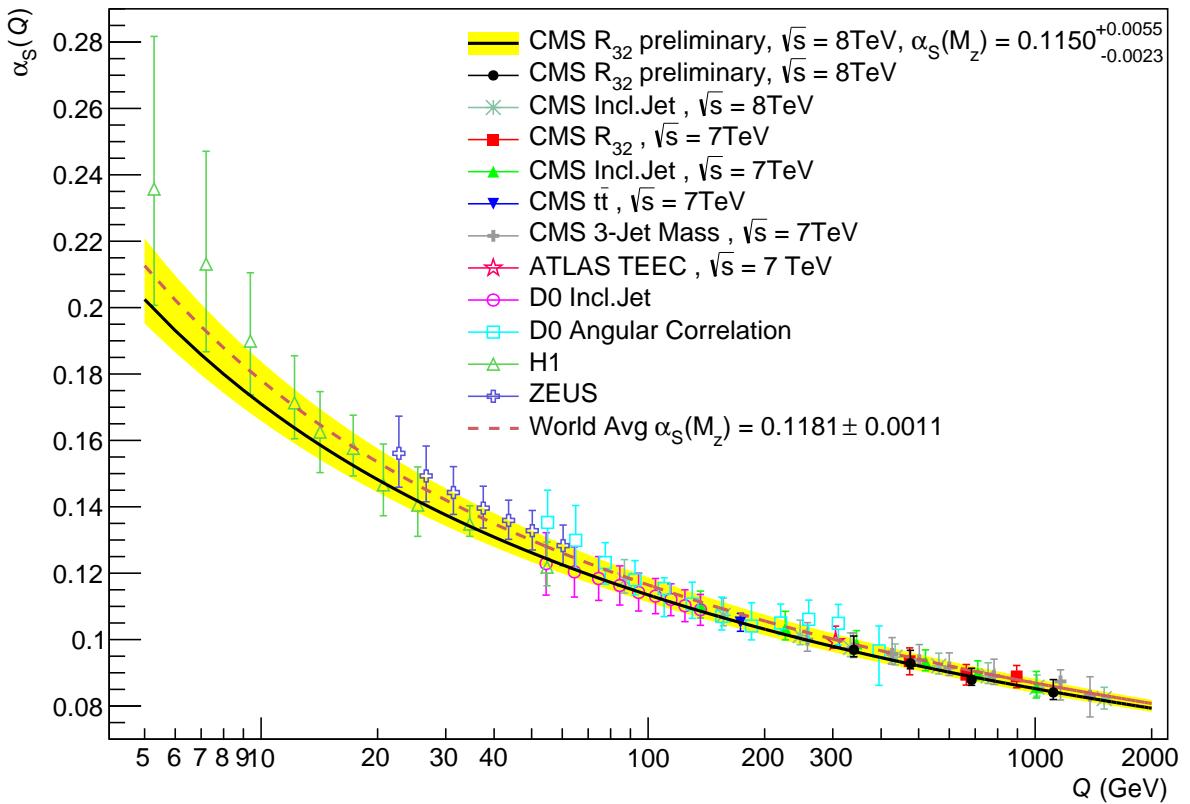


Figure 7.4: The running  $\alpha_S(Q)$  as a function of the energy scale  $Q$  is shown as obtained by using the MSTW2008 NLO PDF set. The solid line and the uncertainty band are drawn by evolving the extracted  $\alpha_s(M_Z)$  values using the 2-loop 5-flavour renormalization group equations as implemented 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  $\alpha_S$ . 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  $19.7\text{ fb}^{-1}$ . 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

each bin of  $H_{\mathrm{T},2}/2$ . 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  $300 < H_{\mathrm{T},2}/2 < 2000$  GeV for inclusive 2-jet event cross-sections and  $300 < H_{\mathrm{T},2}/2 < 1680$  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  $H_{\mathrm{T},2}/2$ . 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  $H_{\mathrm{T},2}/2$  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  $H_{\mathrm{T},2}/2$ . 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  $H_{\mathrm{T},2}/2$  values and by statistical uncertainty at higher  $H_{\mathrm{T},2}/2$  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  $\alpha_S$  ( $\sigma_{n\text{-jet}} \propto \alpha_S^n$ ) 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  $\alpha_S$ , 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 :

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

In both values, the dominant uncertainty is of theoretical origin. 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  $\alpha_S$  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.

# Chapter 9

## Hardware and Software Activities Undertaken

The Hadron Outer (HO) calorimeter of the CMS detector has been installed to measure the energy which is not contained by the barrel calorimeters. The HO improves the energy resolution of highly energetic hadrons and hence provides a better jet energy reconstruction as well as missing transverse energy resolution. HO is situated in the barrel return yokes (YB) in front of the first layer of muon chambers. The YBs consist of 5 rings of iron: YB0, YB $\pm$ 1 and YB $\pm$ 2. Rings R0, R $\pm$ 1 and R $\pm$ 2 of the HO system are located in YB0, YB $\pm$ 1 and YB $\pm$ 2 respectively. In the original design of HO, the signals of the detector were read out by hybrid photo-diodes (HPDs) by converting the wavelength shifted scintillator light into electrical charges. The electronics readout system of the HO is built using the 18-channel readout modules (RMs). The RMs combine the photo-sensors, amplifiers and analogue to digital converters (ADC) and are housed in crates (RBX). The newly developed SiPM system consists of three circuit boards : a mounting Board (MB) holding the SiPMs, a bias board generating the SiPM operation voltage and the a control board connecting the SiPMs electrically to the HCAL readout electronics and monitors the operation of the individual SiPMs. The array of 18 SiPMs is

mounted on one side of the MB, as shown in Fig. 9.1. The geometrical constraints required a total of 132 RMs to read all channels. The HPDs were chosen as photo-sensors due to its high gain and magnetic field tolerance. But in Run I conditions of the CMS, HPDs proved to be less optimal because of the discharge caused by the fringe field of the CMS magnet, low gain and photo detection efficiency, and ageing effects. Due to these inefficiencies, the HPDs needed to be replaced with multi-pixel Geiger-mode avalanche photo-diodes also known as silicon photo-multipliers (SiPMs). The SiPMs are preferred because of the low operating voltage, relatively high gain, magnetic field insensitivity and high photon-detection efficiency. This replacement was carried out during the first LHC long shutdown (LS1) in 2013-2014 [134], where the LHC was upgraded to higher luminosity ( $5 \times 10^{34} \text{ cm}^2 \text{s}^{-1}$ ) and center-of-mass energy of proton-proton collisions was increased from 4 TeV to 6.5 TeV per beam. During this up-gradation, the replacement of HPDs took place in two steps : first the existing RMs with HPDs were taken out from the detector and then were rebuilt with SiPMs. After verifying the working of the RMs, they were re-installed in the CMS detector. In collaboration with HO group of the CMS, we participated in the re-installation of the RMs with SiPMs in place of HPDs specially in the sectors YB+1 and YB+2 of HCAL during the visit to CERN in March-April, 2014.

## 9.1 Silicon Photo-Multipliers

A silicon photo-multiplier (SiPM) used in HO is a Hamamatsu Multi-Pixel Photon Counter (MPPC) in a surface mounted device housing. It has a cell pitch of 50  $\mu\text{m}$  with an active area of  $3 \times 3 \text{ mm}^2$ . The operating voltage required is  $\sim 70 \text{ V}$  with a gain  $\sim 6 \times 10^5 \text{ fC/photo-electron}$  when operated at a voltage greater than the breakdown voltage by 1 V. At this over-voltage which is given by the bias voltage subtracted from the breakdown voltage, the typical dark current rate is of the order

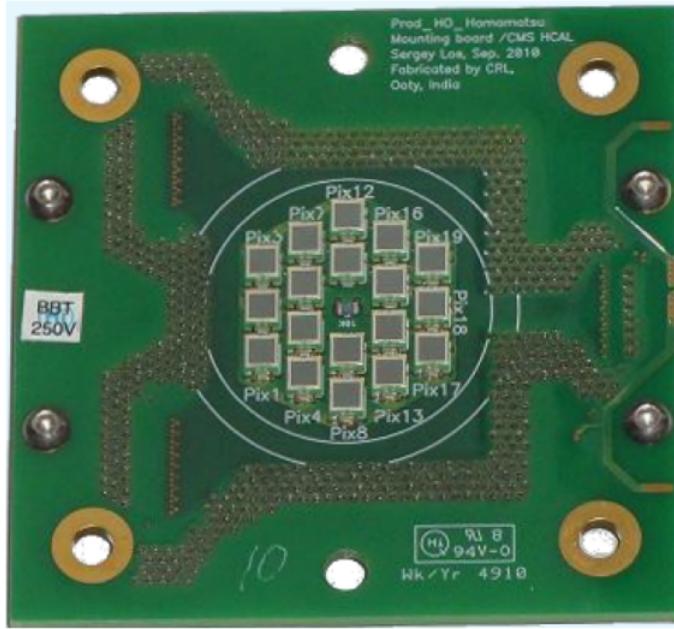


Figure 9.1: The arrangement of Silicon Photo-Multipliers (SiPMs) on the Mounting Board (MB). Taken from [137]

of a few hundred kHz. Along with the installation of SiPMs, the commissioning of the upgraded parts also took place. The commissioning includes the quick identification of the problems with the new and existing hardware, validation of the installation and repairs, in case of any malfunctions, during the access to the hardware. To monitor and optimize the operational parameters, two types of the data were used : the signals from SiPMs in the absence of light referred as pedestal events (PED) and the charge distributions collected by illuminating the SiPMs with a calibrated light emitting diodes referred as LED events. While performing the Quality Control (QC) analysis, the PED and LED events or runs were taken through an online software created by HO CMS group. These runs were used to study and analyze the properties of SiPMs.

In the first step of commissioning, a communication test was performed with the readout system along with the verification of slow control operation and channel response. After that, the measurements and optimization of SiPM operational variables were carried out. One of the important quantities of the SiPMs to study is the

change of breakdown voltage (BV) and calibration factor gain (G) with the temperature. BV gives the threshold voltage after which the diodes switch to avalanche mode and gain corresponds to the charge produced by SiPM for a single photo-electron (SPE). The gain of SiPMs depends linearly on the temperature with a relative dependence of 8% gain shift per K at an operating point of 1.5 V over-voltage [135]. As the gain depends linearly on the over-voltage, the change of the breakdown voltage with temperature translates into a change of the gain with temperature. This dependence requires an active control of the temperature of SiPMs with better than 0.1 K stability. So a peltier element is mounted on the back of the MB for cooling purposes in order to stabilize the temperature. We mainly studied the variations of BV and gain with temperature :

**Breakdown Voltage** - The BV can be determined either by using the pedestal spectrum of the SiPMs or the signal of a short LED pulse. In the first method, the gain is estimated by performing a scan of the bias voltage. The dependence of the measured gain on the bias voltage is fitted with a linear function. The extrapolation of fit function to zero gain gives the value of the breakdown voltage. The second method uses the relative change of the measured signal (S) when pulsing an LED and varying the bias voltage (V). The distribution of  $dS/SdV$  as a function of V whose maximum provides the BV. This distribution is fitted by a simple Gaussian function to obtain the value of BV. The variation of the breakdown voltage, obtained using the LED method, over time is shown for one RM (18 channels) in Fig. 9.2. When the detector is operated in stable conditions, it is observed that the BV measurements also stay stable within 50 mV which illustrate the reliability of the BV determination. This variation is studied over a period from the end of January, 2014 to the beginning of March, 2014 which indicates that if any changes are observed in gain or breakdown voltage, they are not caused by changes in the temperature.

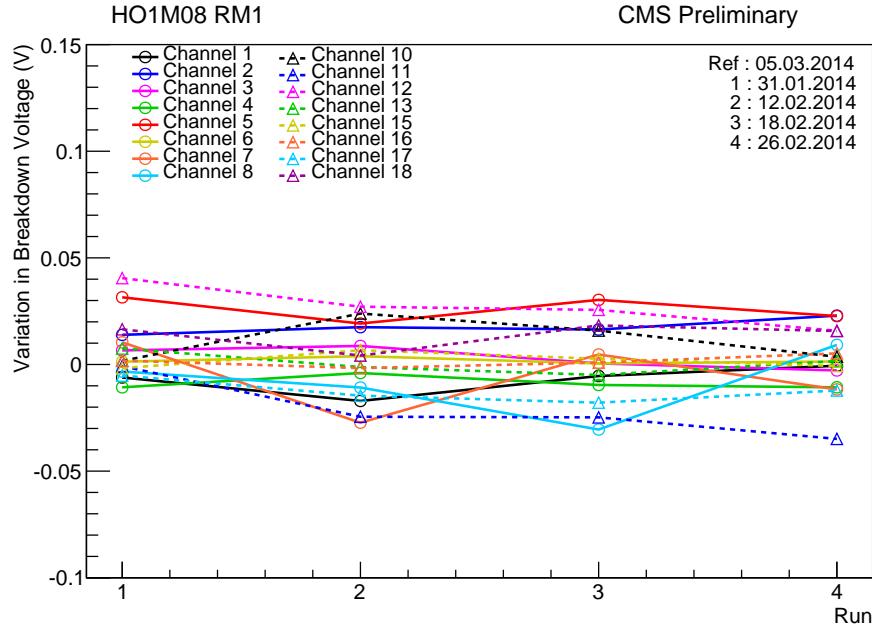


Figure 9.2: The breakdown voltage (BV) is estimated using LED method and its variation is shown over time for 18 channels of one readout module (RM). When the detector is operated in stable conditions, the BV measurements also are stable within 50 mV illustrating the reliability of the BV determination.

**Gain** - The gain is defined as the factor by which Geiger-mode avalanche (whether initiated by the photoelectric effect or thermal carrier excitation) multiplies the initiating electron to form the MPPC's output charge per avalanche. The gain is determined by generating short light pulses with an LED onto the SiPM. Assuming the intensity of the light pulse not too large and  $N$  as number of photons reaching the SiPM, the signal should be equal to  $N \times$  gain. According to Poisson statistics, the sigma of the photon number is  $\sqrt{N}$  and the sigma of the measured signal is  $\sqrt{N} \times$  gain. Dividing the variance of the signal by the mean of the signal one gets  $\sigma^2/\text{mean} = \text{gain}$ . Figure 9.3 shows the relative variation of the gain versus time for a single SiPM mounting board. We observed that gain is stable over a time from the middle of February to the beginning of March in 2014 and the relative variation of the gain lies within 2%.

The relative gain variation with time is plotted for all installed SiPMs as presented

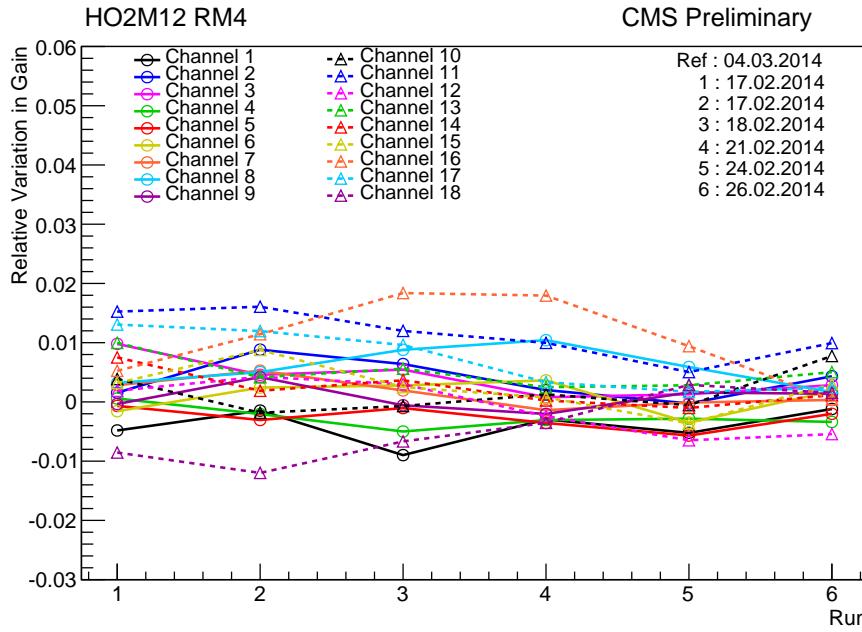


Figure 9.3: The relative variation of the SiPM gain is presented over time for a single RM with 18 channels. The gain is stable over a time from the middle of February to the beginning of March in 2014 and the relative variation of the gain lies within 2%

in Fig. 9.4 which is fitted using a Gaussian function. The distribution has a width of only 0.5 % and all gain variations are within 3%. This illustrates that the gain determination behaves as expected and the operation of the SiPMs with a stable gain is possible. These results were presented at the CALOR 2014 conference [136] and are documented in Ref. [137].

## 9.2 MicroTCA

During the LHC upgrade at the time of LS1, the increase in LHC luminosity and center-of-mass energy increased the number of interactions per bunch crossing i.e. pileup. Hence, a large amount of the data became available which needed to be processed at a much faster rate than before. This required a very high speed DAQ system and an increase in the number of electronics readout channels to collect high quality data needed to perform physics analysis. Before the upgrade, the VERSAbus

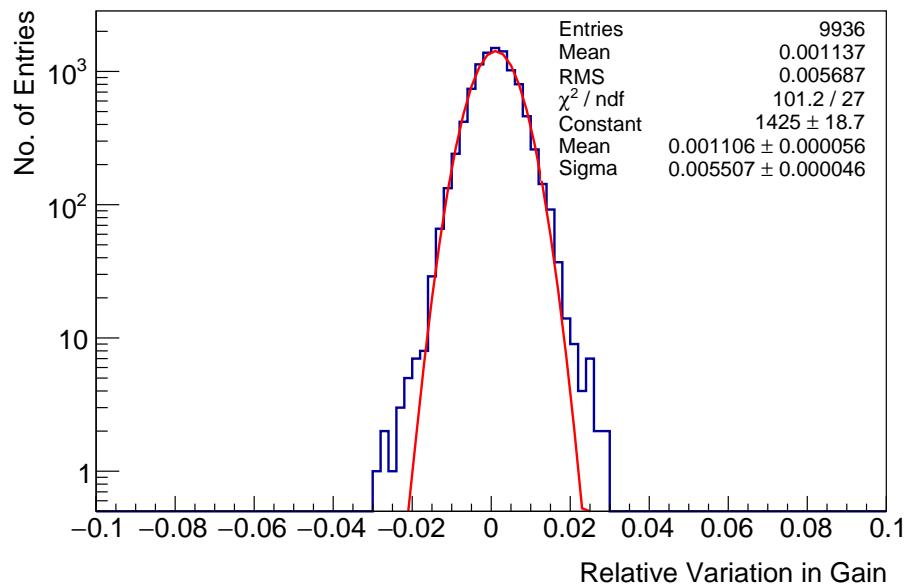


Figure 9.4: The distribution of the relative variations in gain for all the installed SiPMs is fitted with a Gaussian function. It has a width of only 0.5 % and all gain variations are within 3%.

Memory card (VME) based system was used but this could not support data transfer rate needed after LS1. So during the upgradation, the existing VME based system was replaced with  $\mu$ TCA (Micro Telecommunications Computing Architecture) standard system in HCAL back-end electronics [138].

The  $\mu$ TCA is an embedded, scalable architecture which offers flexibility to build robust systems. It was designed as a complimentary system to the Advanced Telecommunication Computing Architecture (ATCA), primarily for the core telecommunication networks. It is compact in size and less expensive than ATCA systems.  $\mu$ TCA is based on the Advanced Mezzanine Card (AMC) standard which was part of the ATCA. The ATCA standard specifies a crate which can host a large carrier of AMC cards, also known as  $\mu$ HTR (Micro HCAL Trigger & Readout) cards. In the simpler  $\mu$ TCA architecture, AMC cards are plugged directly into a backplane such that twelve standard AMC cards can be placed in a crate shown in Fig. 9.5. One or two special hub slots are also present in each crate where at-least

<sup>24</sup>Source: <https://www.vadatech.com>

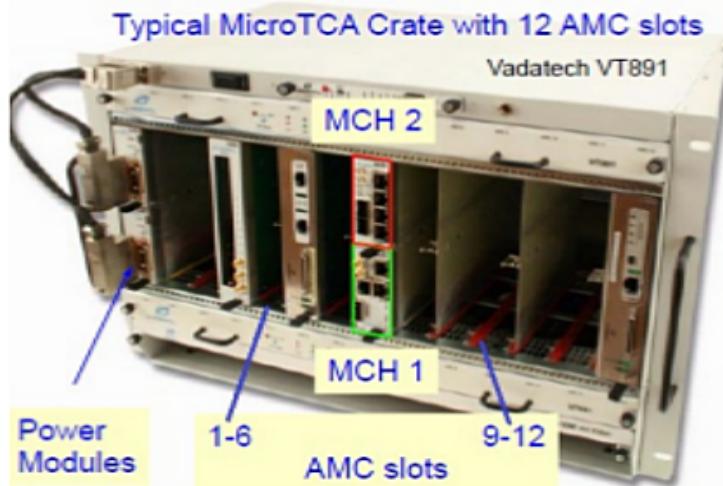


Figure 9.5:  $\mu$ TCA crate<sup>24</sup> where 12 AMC (Advanced Mezzanine Card) or  $\mu$ HTR (Micro HCAL Trigger & Readout) cards can be loaded along with one or two special hub slots occupied by MicroTCA Carrier Hub (MCH) cards. The MCH card is responsible for the control of the power to each slot.

one of these slots must be occupied by a MicroTCA Carrier Hub (MCH) card. The MCH card provides the control of the power to each slot and general house-keeping of the crate. The primary MCH site will hold the commercial MCH card responsible for crate management and the ethernet network. The secondary site is used for a CMS-common card, known as AMC13, which is responsible for distributing clock signals to the AMCs. The  $\mu$ HTR cards receive the data links from the front-ends, calculate and transmit trigger primitives. The Power Mezzanines/Auxiliary Power Mezzanines (PMs/APMs) mounted on  $\mu$ HTR cards will supply power to them.

The working of these mezzanines becomes very crucial as their failure may lead to loss of the data collection efficiency. Hence, a Power Mezzanine Testing program was designed to monitor or test  $\mu$ HTR PMs/APMs for the long term ( $\sim$ 39 hour) stability tests. To carry out these tests, a test-stand was designed as represented in Fig. 9.6. It includes a testing board with 5 slots for mounting 3 PMs and 2 APMs at a time, along with 2 Analog to Digital Converter (ADC) ICs to monitor the temperature, voltage and current. A SUB-20 module acted as a

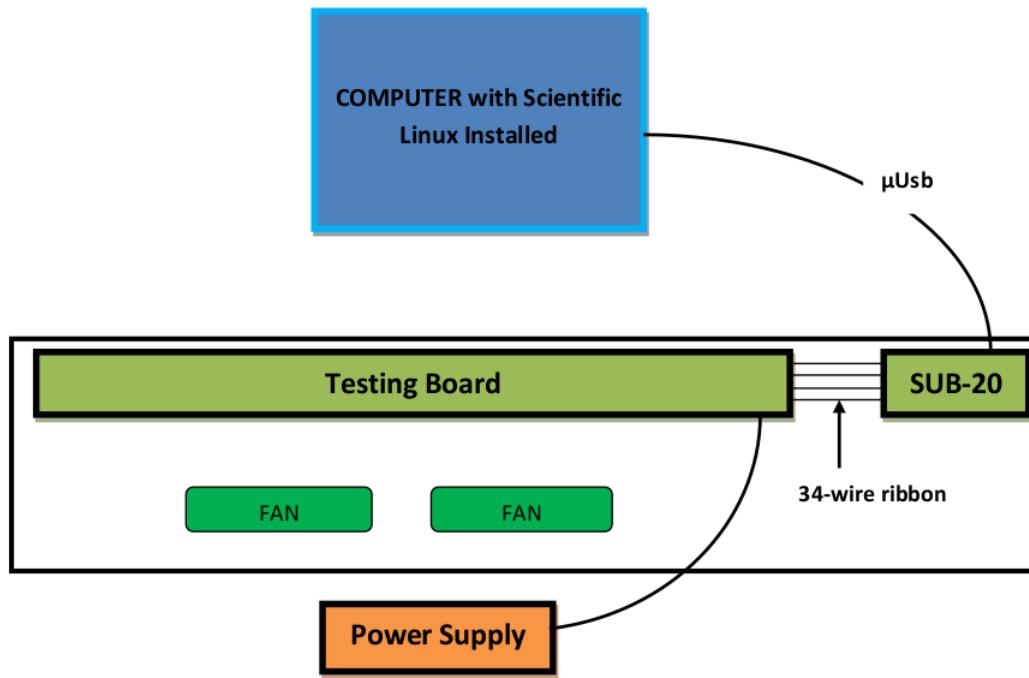


Figure 9.6: A test-stand designed to monitor the working of Power Mezzanines/Auxiliary Power Mezzanines (PMs/APMs) through stability tests. It included a testing board with appropriate slots for mounting PMs/APMs, a SUB-20 module acting as a bridge between testing board and a computer (PC), a 34-pin ribbon wire for a communication between module and PC, a  $\mu$ USB cable to supply power to module from PC, two fans for cooling of PMs/APMs and a power supply for powering the testing board and PMs/APMs.

communication bridge between the testing board and a computer (PC) with testing program installed. The SUB-20 module was connected to testing board through I2C (Inter Integrated circuit) and communicated via 34-pin ribbon. A  $\mu$ USB cable connected the PC and SUB-20 module and also supplied power to module from PC. Two fans were also mounted for cooling of PMs/APMs and resistors embedded into the testing board. A power supply providing voltage of 12 V and current of 10 A was used for powering the testing board and PMs/APMs mounted on it.

In the Power Mezzanine Testing program, two quick tests namely, Margin Up and Margin Down, were conducted by setting the output voltage 5% high and low, respectively. Each test was run for a duration of half an hour. These were then followed by long tests with nominal and high load nominal settings, each running

for 19 hours. During each test, the output voltage, current, power supplied and temperature of PMs/APMs were monitored and recorded after every 10 seconds. At the end of every test, the average and extremum values of every quantity were used to check the stability of PMs/APMs with time. As a part of the testing of  $\mu$ HTR cards, we participated in the testing of PMs/APMs for which a test-stand, as shown in Fig. 9.7, was installed at the Department of Physics, Panjab University, Chandigarh. We successfully tested three sets of PMs/APMs which were then sent to CERN to be used for  $\mu$ HTR cards. During the CERN visits, we participated in the testing of working of these  $\mu$ HTR cards at 904 building (Prevessin site in France). The Power Modules required to supply power to  $\mu$ TCA crates were also tested. The tested  $\mu$ HTR cards were then installed in  $\mu$ TCA crates at CMS P5 site.

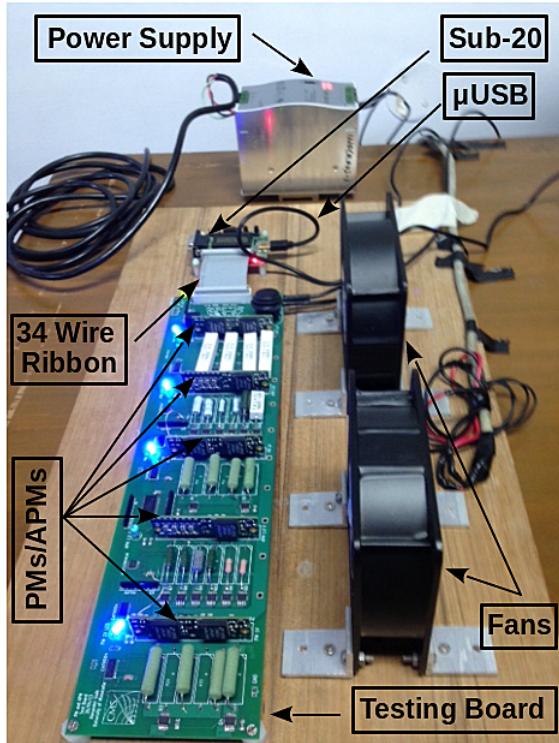


Figure 9.7: A test-stand installed at Department of Physics, Panjab University, Chandigarh to perform the stability tests for monitoring the working of Power Mezzanines/Auxiliary Power Mezzanines (PMs/APMs).

## 9.3 Other Activities

Along with performing physics analysis as well as participating in hardware and software activities, I was also involved in service work provided by CMS Collaboration. I worked in the Data Certification (DC) sub-group of the Data Quality Monitoring (DQM) group [139] of Physics Performance & Dataset (PPD) organization for performing the certification of 2016 CMS data. The DQM group is involved in many and major tasks related to CMS data. The DQM organization works at two different levels - online and offline. The online DQM organization takes care of centralization of the various online CMSSW monitoring modules provided by subsystems and Detector Performance Groups (DPGs), execution of the live monitoring applications and visualization tool called graphical user interface (DQM GUI) on the DQM cluster and organization of the central online DQM shifts. The online DQM spots problems in the CMS detector while it is running. The offline DQM and DC organization performs centralization of the offline CMSSW monitoring modules provided by DPGs and Data Certification Physics Object Groups (POGs), maintains DQM GUI, used for data certification and release validation and coordinates the certification and publication of the data suitable for physics analysis. I was part of the DC team which was involved in taking the inputs from certification experts, importing their results in CMS Web Based Monitoring's Run Registry (RR) and extracting the information to create the JSON files required for carrying out the physics analysis. In the data certification activity, a list of runs and lumi-sections (LS) is prepared which are good for physics analysis to be performed by the CMS collaboration. For this, one has to :

- Provide the list of physics runs (cosmics or collisions) to be certified by subsystems experts
- Keep the relevant information in the offline RR
- Provide help to DPG/POG certification experts

- Update the flags for run data quality depending on the feedback provided by experts
- Produce the JSON files which includes the final good runs and LS which are then used by physics analysts
- Announce the official JSON files through physics validation hypernews

More details of the data certification can be found in Ref. [140]. I also participated in on-going software development of a tool called Historic DQM (HDQM) which is beneficial to study and check stability of various sub-detectors with time.

# Appendix A

## A.1 Cross-section Ratio, $R_{32}$

Table A.1: Differential cross-sections ( $\times 10^{-3}(\text{pb}/\text{GeV})$ ) and the cross-section ratio  $R_{32}$  at detector level in each bin of  $H_{T,2}/2$ , along with statistical uncertainty (in %).

Bin	2-jet cross-section	Stat. unc.	3-jet cross-section	Stat. unc.	Ratio $R_{32}$	Stat. unc.
300 - 330	29772.726	0.211	2640.629	0.707	0.089	+0.665 -0.661
330 - 360	16792.917	0.231	1773.485	0.704	0.106	+0.523 -0.521
360 - 390	9889.326	0.182	1176.544	0.526	0.119	+0.485 -0.483
390 - 420	5976.777	0.179	778.034	0.492	0.130	+0.206 -0.206
420 - 450	3731.760	0.067	522.624	0.180	0.140	+0.167 -0.167
450 - 480	2398.741	0.084	357.622	0.217	0.149	+0.201 -0.200
480 - 510	1570.192	0.104	246.051	0.262	0.157	+0.241 -0.241
510 - 540	1048.665	0.127	171.080	0.314	0.163	+0.288 -0.287
540 - 570	713.042	0.154	119.566	0.376	0.168	+0.344 -0.343
570 - 600	490.776	0.186	84.798	0.447	0.173	+0.407 -0.406
600 - 640	325.046	0.198	57.463	0.470	0.177	+0.427 -0.426
640 - 680	205.727	0.248	37.282	0.583	0.181	+0.529 -0.527
680 - 720	133.674	0.308	24.859	0.714	0.186	+0.646 -0.643
720 - 760	87.911	0.380	16.560	0.875	0.188	+0.791 -0.786
760 - 800	58.657	0.465	11.056	1.071	0.188	+0.968 -0.961
800 - 850	38.106	0.516	7.318	1.178	0.192	+1.063 -1.054
850 - 900	23.587	0.656	4.600	1.485	0.195	+1.339 -1.326
900 - 950	15.130	0.819	2.896	1.872	0.191	+1.694 -1.672
950 - 1000	9.696	1.023	1.812	2.366	0.187	+2.151 -2.116
1000 - 1060	6.026	1.185	1.186	2.670	0.197	+2.414 -2.371
1060 - 1120	3.668	1.518	0.716	3.436	0.195	+3.118 -3.046
1120 - 1180	2.327	1.906	0.437	4.398	0.188	+4.024 -3.903
1180 - 1250	1.419	2.260	0.265	5.227	0.187	+4.798 -4.627
1250 - 1320	0.853	2.915	0.165	6.623	0.194	+6.080 -5.811
1320 - 1390	0.477	3.898	0.080	9.492	0.169	+8.951 -8.355
1390 - 1460	0.263	5.249	0.042	13.131	0.160	+12.619 -11.449
1460 - 1530	0.192	6.143	0.029	15.811	0.151	+15.437 -13.698
1530 - 1600	0.104	8.362	0.021	18.570	0.203	+17.571 -15.536
1600 - 1680	0.060	10.314	0.009	26.726	0.149	+27.132 -22.170

## A.2 Individual Sources of Jet Energy Correction Uncertainties

The sources of JEC considered in the current measurements are : AbsoluteStat, AbsoluteScale, AbsoluteFlavMap, AbsoluteMPFBias, Fragmentation, SinglePionECAL, SinglePionHCAL, FlavorQCD, RelativeJEREC1, RelativeJEREC2, RelativeJERHF, RelativePtBB, RelativePtEC1, RelativePtEC2, RelativePtHF, RelativeFSR, RelativeStatFSR, RelativeStatEC2, RelativeStatHF, PileUpDataMC, PileUpPtRef, PileUpPtBB, PileUpPtEC1, PileUpPtEC2 and PileUpPtHF. The AbsoluteFlavMap uncertainty is exactly zero for the 8 TeV and can be ignored. For the four sources : RelativeJERHF, RelativePtHF, RelativeStatHF, PileUpPtHF, the JEC uncertainty is exactly zero because of  $|y| < 2.5$  cut used in the analysis. So only 20 sources contribute to the total JEC uncertainty.

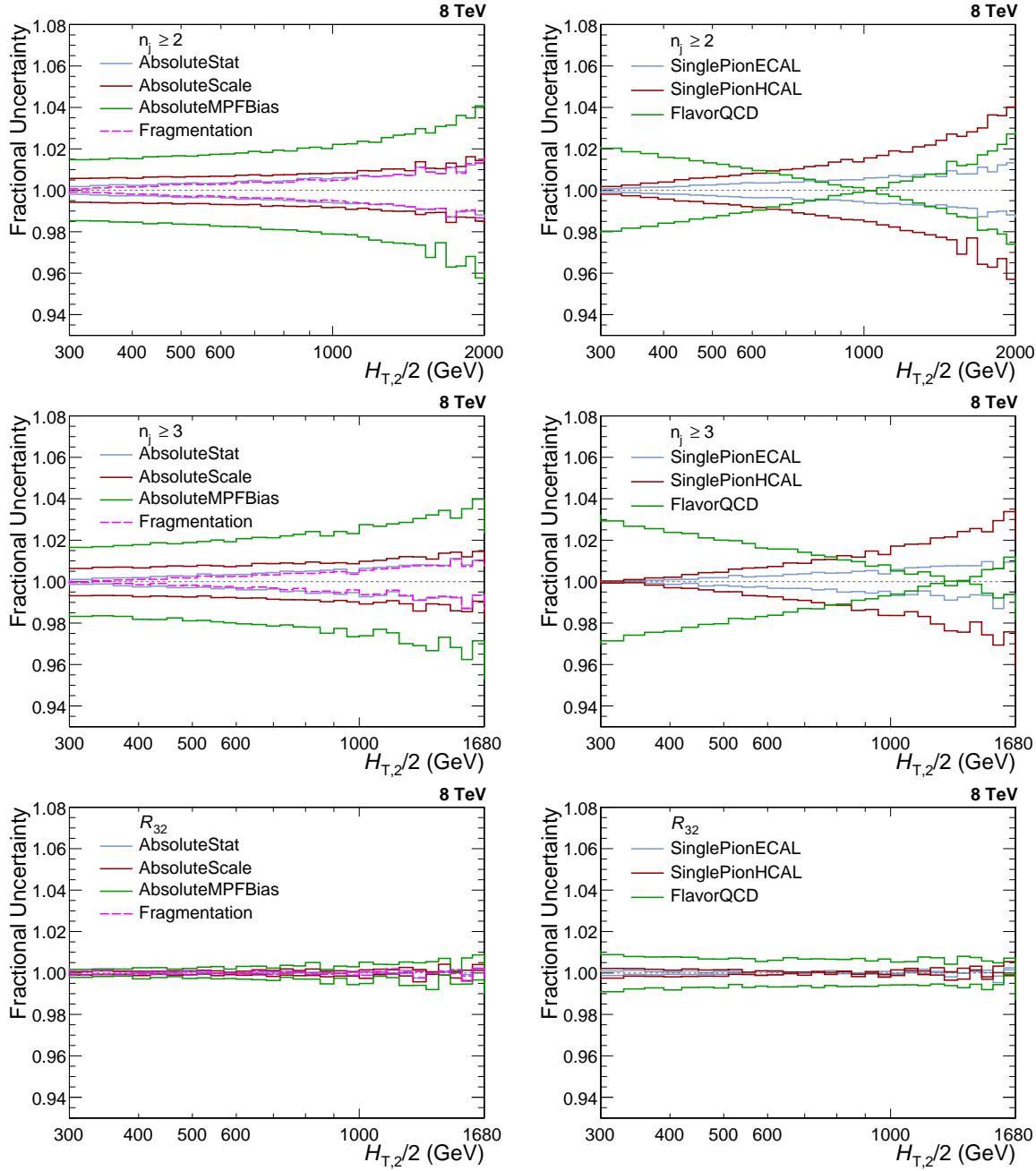


Figure A.1: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet (middle) events cross-sections and the cross-section ratio  $R_{32}$  (bottom). On left, JEC uncertainties are evaluated from AbsoluteStat (blue), AbsoluteScale (red), AbsoluteMPFBias (green) and Fragmentation (pink) sources whereas on right, these are evaluated from SinglePionECAL (blue), SinglePionHCAL (red) and FlavorQCD (green) sources.

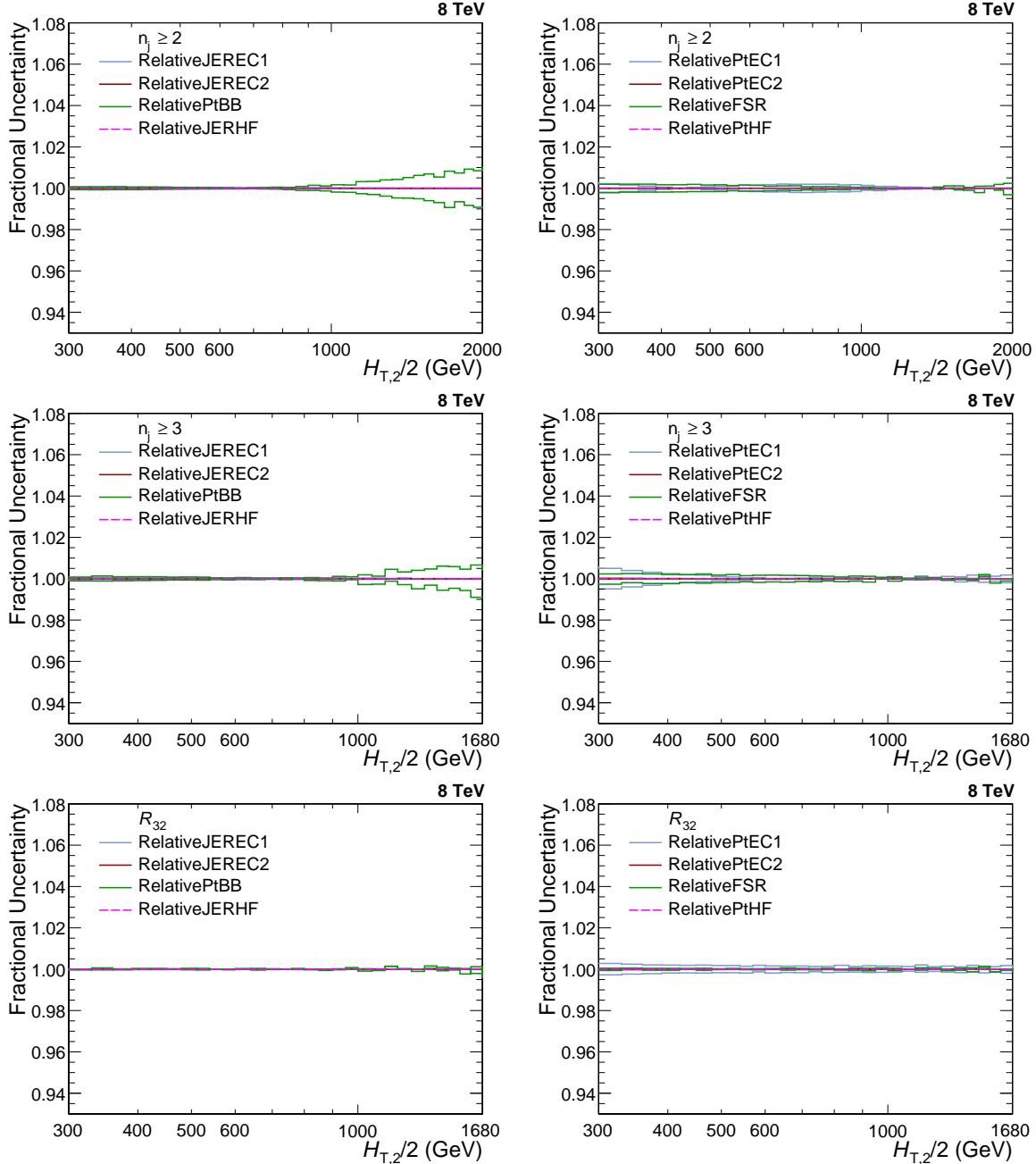


Figure A.2: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet (middle) events cross-sections and the cross-section ratio  $R_{32}$  (bottom). On left, JEC uncertainties are evaluated from RelativeJEREC1 (blue), RelativeJEREC2 (red), RelativePtBB (green) and RelativeJERHF (pink) sources whereas on right, these are evaluated from RelativePtEC1 (blue), RelativePtEC2 (red), RelativeFSR (green) and RelativePtHF (pink) sources.

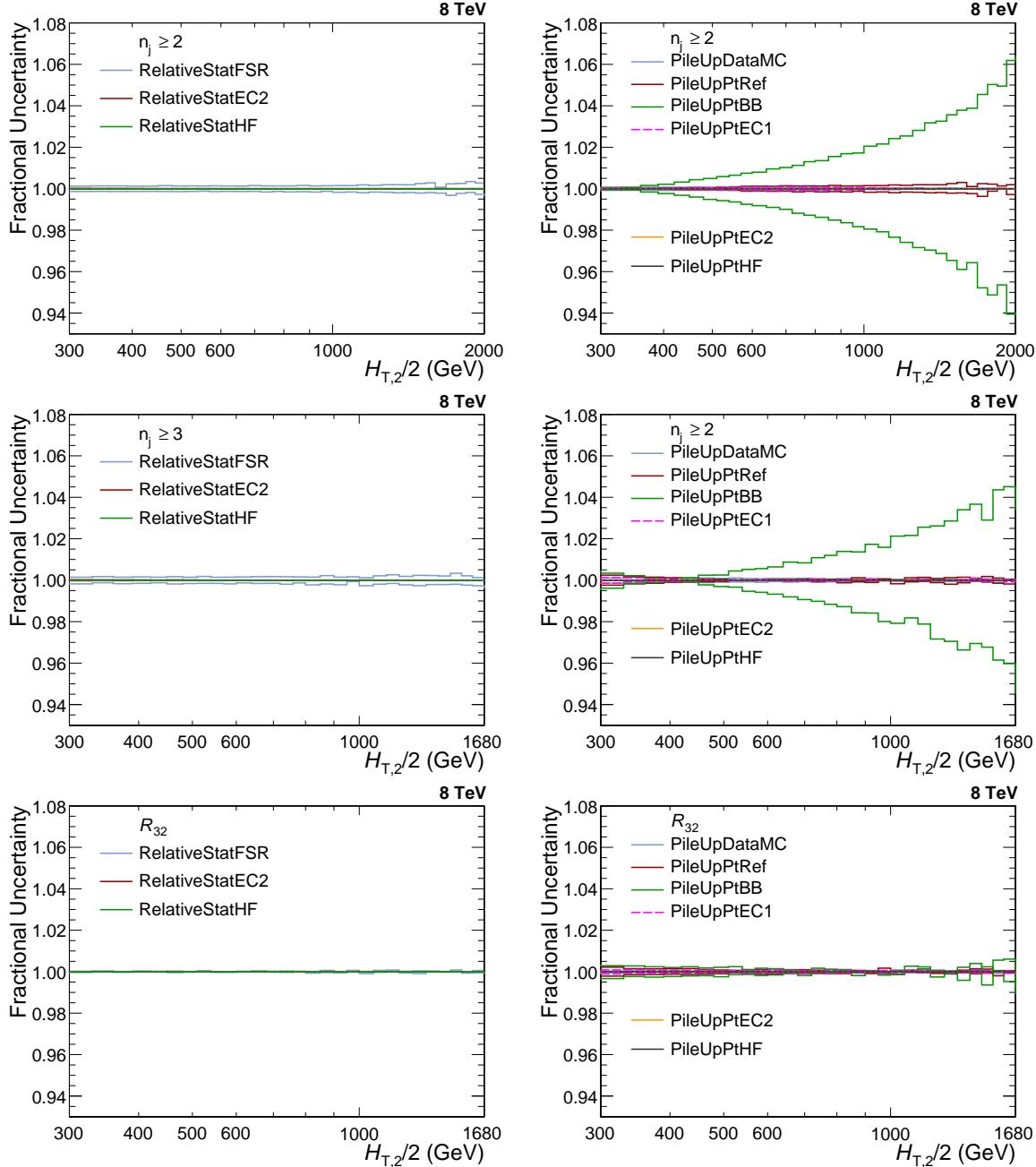


Figure A.3: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet (middle) events cross-sections and the cross-section ratio  $R_{32}$  (bottom). On left, JEC uncertainties are evaluated from RelativeStatFSR (blue), RelativeStatEC2 (red) and RelativeStatHF (green) sources whereas on right, these are evaluated from PileUpDataMC (blue), PileUpPtRef (red), PileUpPtBB (green), PileUpPtEC1 (pink), PileUpPtEC2 (orange) and PileUpPtHF (black) sources.

## A.3 Experimental Uncertainties

Table A.2: Experimental uncertainties (in %), from all sources as well as the total uncertainty, affecting the cross-section measurement in each bin of  $H_{\mathrm{T},2}/2$  for inclusive 2-jet events.

Bin	Statistical	JEC	Unfolding	Lumi	Residual	Total
300 - 330	0.242	+2.612 -2.565	+0.948 -0.928	2.6	1.0	+3.942 -3.906
330 - 360	0.258	+2.507 -2.473	+0.976 -0.969	2.6	1.0	+3.882 -3.858
360 - 390	0.202	+2.504 -2.465	+0.779 -0.783	2.6	1.0	+3.831 -3.807
390 - 420	0.193	+2.363 -2.381	+0.905 -0.904	2.6	1.0	+3.768 -3.780
420 - 450	0.084	+2.448 -2.422	+0.904 -0.895	2.6	1.0	+3.818 -3.799
450 - 480	0.096	+2.440 -2.352	+0.797 -0.795	2.6	1.0	+3.789 -3.733
480 - 510	0.107	+2.427 -2.406	+0.728 -0.715	2.6	1.0	+3.767 -3.751
510 - 540	0.128	+2.425 -2.395	+0.835 -0.862	2.6	1.0	+3.789 -3.775
540 - 570	0.154	+2.425 -2.376	+0.687 -0.674	2.6	1.0	+3.760 -3.726
570 - 600	0.180	+2.497 -2.474	+0.839 -0.827	2.6	1.0	+3.838 -3.820
600 - 640	0.209	+2.495 -2.491	+0.744 -0.743	2.6	1.0	+3.819 -3.816
640 - 680	0.264	+2.582 -2.545	+0.912 -0.912	2.6	1.0	+3.915 -3.891
680 - 720	0.320	+2.691 -2.574	+0.763 -0.756	2.6	1.0	+3.961 -3.880
720 - 760	0.387	+2.690 -2.755	+0.705 -0.712	2.6	1.0	+3.955 -4.001
760 - 800	0.465	+2.858 -2.846	+0.859 -0.846	2.6	1.0	+4.109 -4.098
800 - 850	0.548	+2.889 -2.913	+0.783 -0.787	2.6	1.0	+4.126 -4.143
850 - 900	0.698	+3.145 -3.102	+0.961 -0.958	2.6	1.0	+4.366 -4.334
900 - 950	0.847	+3.298 -3.233	+0.828 -0.829	2.6	1.0	+4.476 -4.429
950 - 1000	1.041	+3.291 -3.330	+0.895 -0.872	2.6	1.0	+4.525 -4.549
1000 - 1060	1.268	+3.598 -3.569	+0.945 -0.956	2.6	1.0	+4.817 -4.798
1060 - 1120	1.611	+3.759 -3.756	+0.970 -0.967	2.6	1.0	+5.043 -5.040
1120 - 1180	1.985	+4.154 -4.053	+1.089 -1.080	2.6	1.0	+5.490 -5.413
1180 - 1250	2.406	+4.251 -4.313	+1.062 -1.070	2.6	1.0	+5.722 -5.770
1250 - 1320	3.101	+4.696 -4.624	+1.151 -1.144	2.6	1.0	+6.384 -6.330
1320 - 1390	4.157	+4.934 -4.979	+1.343 -1.341	2.6	1.0	+7.155 -7.186
1390 - 1460	5.270	+5.148 -5.104	+1.185 -1.177	2.6	1.0	+7.965 -7.936
1460 - 1530	6.360	+5.890 -5.652	+1.405 -1.406	2.6	1.0	+9.213 -9.063
1530 - 1600	8.183	+5.924 -6.311	+1.598 -1.590	2.6	1.0	+10.601 -10.821
1600 - 1680	10.630	+5.969 -5.655	+1.607 -1.592	2.6	1.0	+12.608 -12.461
1680 - 1760	13.864	+7.245 -7.603	+1.821 -1.839	2.6	1.0	+15.993 -16.161
1760 - 1840	18.192	+7.781 -7.820	+1.902 -1.906	2.6	1.0	+20.071 -20.087
1840 - 1920	22.612	+7.647 -7.537	+1.588 -1.590	2.6	1.0	+24.085 -24.050
1920 - 2000	29.530	+9.199 -9.469	+1.511 -1.505	2.6	1.0	+31.092 -31.172

Table A.3: Experimental uncertainties (in %), from all sources as well as the total uncertainty, affecting the cross-section measurement in each bin of  $H_{\mathrm{T},2}/2$  for inclusive 3-jet events.

Bin	Statistical	JEC	Unfolding	Lumi	Residual	Total
300 - 330	0.796	+3.503 -3.475	+0.564 -0.552	2.6	1.0	+4.581 -4.558
330 - 360	0.781	+3.303 -3.186	+0.640 -0.633	2.6	1.0	+4.437 -4.350
360 - 390	0.583	+3.221 -3.094	+0.490 -0.496	2.6	1.0	+4.326 -4.233
390 - 420	0.531	+3.092 -3.149	+0.584 -0.584	2.6	1.0	+4.236 -4.278
420 - 450	0.224	+3.125 -2.996	+0.604 -0.592	2.6	1.0	+4.236 -4.140
450 - 480	0.248	+2.984 -2.890	+0.531 -0.528	2.6	1.0	+4.124 -4.056
480 - 510	0.269	+2.937 -2.963	+0.511 -0.512	2.6	1.0	+4.089 -4.108
510 - 540	0.318	+3.021 -2.797	+0.592 -0.612	2.6	1.0	+4.164 -4.007
540 - 570	0.375	+2.999 -2.935	+0.506 -0.500	2.6	1.0	+4.141 -4.094
570 - 600	0.434	+2.824 -2.906	+0.646 -0.620	2.6	1.0	+4.042 -4.096
600 - 640	0.497	+2.952 -2.956	+0.598 -0.604	2.6	1.0	+4.133 -4.136
640 - 680	0.617	+3.111 -3.001	+0.777 -0.786	2.6	1.0	+4.292 -4.215
680 - 720	0.739	+3.067 -2.984	+0.642 -0.611	2.6	1.0	+4.257 -4.194
720 - 760	0.895	+3.185 -3.111	+0.595 -0.607	2.6	1.0	+4.366 -4.313
760 - 800	1.068	+3.231 -3.166	+0.763 -0.774	2.6	1.0	+4.464 -4.419
800 - 850	1.250	+3.427 -3.295	+0.674 -0.687	2.6	1.0	+4.639 -4.544
850 - 900	1.578	+3.364 -3.540	+0.903 -0.898	2.6	1.0	+4.731 -4.857
900 - 950	1.961	+3.594 -3.524	+0.792 -0.793	2.6	1.0	+5.015 -4.965
950 - 1000	2.420	+3.603 -3.783	+0.846 -0.843	2.6	1.0	+5.226 -5.351
1000 - 1060	2.844	+4.164 -4.116	+0.916 -0.940	2.6	1.0	+5.834 -5.803
1060 - 1120	3.647	+4.038 -3.815	+0.963 -0.957	2.6	1.0	+6.188 -6.044
1120 - 1180	4.607	+4.278 -4.183	+1.084 -1.087	2.6	1.0	+6.961 -6.904
1180 - 1250	5.532	+4.894 -4.771	+1.074 -1.069	2.6	1.0	+7.967 -7.891
1250 - 1320	7.141	+5.144 -5.273	+1.222 -1.217	2.6	1.0	+9.312 -9.383
1320 - 1390	10.207	+5.542 -5.642	+1.414 -1.428	2.6	1.0	+12.027 -12.076
1390 - 1460	13.831	+5.630 -5.265	+1.257 -1.256	2.6	1.0	+15.242 -15.111
1460 - 1530	15.578	+5.576 -5.491	+1.546 -1.551	2.6	1.0	+16.850 -16.822
1530 - 1600	18.729	+6.409 -7.019	+1.718 -1.716	2.6	1.0	+20.063 -20.266
1600 - 1680	26.465	+7.017 -6.255	+1.775 -1.765	2.6	1.0	+27.578 -27.393

Table A.4: Experimental uncertainties (in %), from all sources as well as the total uncertainty, affecting the measurement of cross-section ratio  $R_{32}$ , in each bin of  $H_{T,2}/2$ .

<b>Bin</b>	<b>Statistical</b>	<b>JEC</b>	<b>Unfolding</b>	<b>Total</b>
300 - 330	0.741	+1.059 -1.097	+0.754 -0.751	+1.496 -1.522
330 - 360	0.587	+0.954 -0.923	+0.685 -0.689	+1.313 -1.292
360 - 390	0.519	+0.902 -0.855	+0.594 -0.593	+1.199 -1.163
390 - 420	0.236	+0.907 -0.952	+0.439 -0.438	+1.035 -1.074
420 - 450	0.192	+0.900 -0.835	+0.360 -0.361	+0.988 -0.930
450 - 480	0.209	+0.788 -0.802	+0.307 -0.308	+0.872 -0.884
480 - 510	0.245	+0.795 -0.867	+0.254 -0.235	+0.870 -0.931
510 - 540	0.287	+0.852 -0.682	+0.264 -0.268	+0.937 -0.787
540 - 570	0.326	+0.807 -0.803	+0.193 -0.189	+0.891 -0.887
570 - 600	0.397	+0.656 -0.774	+0.199 -0.219	+0.792 -0.898
600 - 640	0.447	+0.763 -0.797	+0.150 -0.154	+0.897 -0.926
640 - 680	0.573	+0.861 -0.781	+0.153 -0.140	+1.045 -0.979
680 - 720	0.663	+0.766 -0.787	+0.147 -0.164	+1.024 -1.042
720 - 760	0.774	+0.842 -0.769	+0.118 -0.118	+1.149 -1.097
760 - 800	0.970	+0.800 -0.729	+0.115 -0.096	+1.263 -1.218
800 - 850	1.116	+0.873 -0.775	+0.115 -0.104	+1.422 -1.363
850 - 900	1.436	+0.770 -0.896	+0.069 -0.069	+1.631 -1.694
900 - 950	1.716	+0.704 -0.752	+0.050 -0.051	+1.855 -1.874
950 - 1000	2.156	+0.824 -0.897	+0.089 -0.045	+2.310 -2.336
1000 - 1060	2.554	+0.812 -0.870	+0.045 -0.040	+2.680 -2.698
1060 - 1120	3.244	+0.792 -0.658	+0.018 -0.027	+3.339 -3.310
1120 - 1180	4.121	+0.985 -0.757	+0.025 -0.043	+4.237 -4.191
1180 - 1250	4.990	+1.031 -0.848	+0.023 -0.041	+5.095 -5.062
1250 - 1320	6.456	+0.750 -1.087	+0.079 -0.079	+6.500 -6.548
1320 - 1390	8.990	+1.112 -1.144	+0.080 -0.099	+9.059 -9.063
1390 - 1460	12.699	+1.157 -0.815	+0.076 -0.078	+12.751 -12.725
1460 - 1530	13.926	+0.768 -1.235	+0.143 -0.145	+13.948 -13.981
1530 - 1600	16.903	+1.050 -1.258	+0.120 -0.127	+16.936 -16.950
1600 - 1680	28.070	+1.471 -0.859	+0.178 -0.177	+28.109 -28.084

## A.4 Theoretical Uncertainties

Table A.5: Theoretical uncertainties (in %), calculated using CT10-NLO PDF set from all sources as well as the total uncertainty, affecting the cross-section measurement in each bin of  $H_{T,2}/2$  for inclusive 2-jet events.

Bin	Scale	PDF	NP	Total
300 - 330	+0.942 -6.149	+3.566 -3.090	0.825	+3.780 -6.931
330 - 360	+1.035 -6.289	+3.906 -3.342	0.736	+4.107 -7.159
360 - 390	+1.159 -6.438	+4.232 -3.573	0.696	+4.442 -7.396
390 - 420	+1.220 -6.536	+4.551 -3.794	0.723	+4.767 -7.592
420 - 450	+1.326 -6.660	+4.857 -3.997	0.745	+5.089 -7.802
450 - 480	+1.421 -6.776	+5.153 -4.186	0.765	+5.399 -8.001
480 - 510	+1.512 -6.888	+5.444 -4.365	0.782	+5.704 -8.192
510 - 540	+1.566 -6.967	+5.721 -4.527	0.797	+5.984 -8.347
540 - 570	+1.666 -7.082	+6.000 -4.682	0.810	+6.279 -8.528
570 - 600	+1.731 -7.172	+6.269 -4.825	0.822	+6.555 -8.683
600 - 640	+1.805 -7.271	+6.597 -4.979	0.833	+6.890 -8.852
640 - 680	+1.930 -7.416	+6.978 -5.143	0.845	+7.289 -9.064
680 - 720	+2.007 -7.527	+7.364 -5.295	0.856	+7.680 -9.243
720 - 760	+2.113 -7.663	+7.749 -5.437	0.865	+8.078 -9.436
760 - 800	+2.196 -7.781	+8.140 -5.569	0.873	+8.476 -9.609
800 - 850	+2.323 -7.945	+8.573 -5.706	0.881	+8.926 -9.822
850 - 900	+2.389 -8.062	+9.082 -5.863	0.889	+9.433 -10.008
900 - 950	+2.499 -8.227	+9.600 -6.018	0.896	+9.961 -10.232
950 - 1000	+2.631 -8.402	+10.134 -6.166	0.902	+10.509 -10.460
1000 - 1060	+2.738 -8.569	+10.747 -6.343	0.908	+11.127 -10.700
1060 - 1120	+2.853 -8.751	+11.431 -6.526	0.914	+11.817 -10.955
1120 - 1180	+2.992 -8.970	+12.183 -6.727	0.919	+12.579 -11.250
1180 - 1250	+3.135 -9.194	+13.019 -6.944	0.924	+13.423 -11.558
1250 - 1320	+3.324 -9.469	+14.004 -7.189	0.929	+14.423 -11.925
1320 - 1390	+3.434 -9.677	+15.080 -7.444	0.933	+15.494 -12.244
1390 - 1460	+3.629 -9.976	+16.223 -7.700	0.937	+16.650 -12.637
1460 - 1530	+3.760 -10.224	+17.505 -7.980	0.940	+17.929 -13.004
1530 - 1600	+3.894 -10.471	+18.891 -8.258	0.943	+19.311 -13.368
1600 - 1680	+4.107 -10.813	+20.496 -8.560	0.946	+20.925 -13.824
1680 - 1760	+4.421 -11.101	+22.481 -8.905	0.949	+22.931 -14.263
1760 - 1840	+4.921 -11.461	+24.654 -9.251	0.951	+25.158 -14.760
1840 - 1920	+5.404 -11.813	+27.143 -9.607	0.953	+27.692 -15.256
1920 - 2000	+5.867 -12.154	+29.986 -9.973	0.955	+30.570 -15.751

Table A.6: Theoretical uncertainties (in %), calculated using CT10-NLO PDF set from all sources as well as the total uncertainty, affecting the cross-section measurement in each bin of  $H_{T,2}/2$  for inclusive 3-jet events.

<b>Bin</b>	<b>Scale</b>	<b>PDF</b>	<b>NP</b>	<b>Total</b>
300 - 330	+0.539 -8.294	+5.716 -4.657	1.692	+5.986 -9.662
330 - 360	+0.550 -8.577	+5.977 -4.779	1.516	+6.191 -9.935
360 - 390	+0.599 -8.709	+6.187 -4.987	1.363	+6.363 -10.128
390 - 420	+0.719 -8.948	+6.751 -5.223	1.228	+6.900 -10.433
420 - 450	+0.799 -9.145	+7.031 -5.395	1.110	+7.162 -10.676
450 - 480	+0.847 -9.247	+7.404 -5.578	1.005	+7.520 -10.845
480 - 510	+0.847 -9.294	+7.837 -5.717	0.937	+7.938 -10.951
510 - 540	+0.922 -9.436	+8.198 -5.884	0.921	+8.301 -11.158
540 - 570	+0.974 -9.566	+8.529 -6.000	0.904	+8.632 -11.328
570 - 600	+1.086 -9.786	+8.970 -6.156	0.886	+9.079 -11.595
600 - 640	+1.107 -9.852	+9.402 -6.297	0.866	+9.506 -11.724
640 - 680	+1.278 -10.101	+10.310 -6.526	0.842	+10.423 -12.055
680 - 720	+1.384 -10.342	+9.682 -6.618	0.820	+9.815 -12.305
720 - 760	+1.415 -10.404	+11.051 -6.826	0.798	+11.170 -12.469
760 - 800	+1.547 -10.615	+11.565 -7.009	0.777	+11.694 -12.744
800 - 850	+1.679 -10.804	+12.242 -7.185	0.755	+12.379 -12.997
850 - 900	+2.085 -11.134	+13.097 -7.461	0.731	+13.282 -13.422
900 - 950	+2.475 -11.432	+13.889 -7.703	0.709	+14.125 -13.804
950 - 1000	+2.655 -11.608	+14.614 -7.915	0.688	+14.869 -14.066
1000 - 1060	+3.025 -11.926	+15.576 -8.173	0.667	+15.881 -14.473
1060 - 1120	+3.299 -12.189	+14.250 -8.441	0.645	+14.641 -14.840
1120 - 1180	+3.741 -12.584	+17.984 -8.787	0.625	+18.380 -15.361
1180 - 1250	+3.969 -12.843	+19.324 -9.127	0.625	+19.737 -15.768
1250 - 1320	+4.663 -13.452	+21.246 -9.517	0.642	+21.761 -16.490
1320 - 1390	+4.878 -13.702	+22.884 -9.899	0.657	+23.407 -16.916
1390 - 1460	+5.242 -14.095	+24.854 -10.332	0.670	+25.410 -17.489
1460 - 1530	+5.582 -14.464	+27.170 -10.733	0.682	+27.746 -18.024
1530 - 1600	+6.003 -14.907	+29.741 -11.165	0.692	+30.349 -18.637
1600 - 1680	+6.503 -15.418	+32.855 -11.617	0.702	+33.500 -19.317

Table A.7: Theoretical uncertainties (in %) calculated using CT10-NLO PDF set from all sources as well as the total uncertainty, affecting the measurement of cross-section ratio  $R_{32}$ , in each bin of  $H_{T,2}/2$ .

<b>Bin</b>	<b>Scale</b>	<b>PDF</b>	<b>NP</b>	<b>Total</b>
300 - 330	+0.038 -7.203	+2.458 -3.463	0.822	+2.592 -8.035
330 - 360	+0.027 -6.626	+2.317 -3.378	0.734	+2.431 -7.474
360 - 390	+0.024 -6.449	+2.149 -3.367	0.656	+2.247 -7.304
390 - 420	+0.084 -5.894	+2.411 -3.383	0.586	+2.482 -6.821
420 - 450	+0.113 -5.532	+2.345 -3.362	0.523	+2.405 -6.494
450 - 480	+0.109 -5.409	+2.390 -3.357	0.467	+2.438 -6.383
480 - 510	+0.073 -5.442	+2.506 -3.327	0.416	+2.541 -6.392
510 - 540	+0.107 -5.168	+2.559 -3.326	0.371	+2.588 -6.157
540 - 570	+0.112 -5.010	+2.586 -3.292	0.330	+2.609 -6.004
570 - 600	+0.163 -4.576	+2.729 -3.292	0.292	+2.750 -5.645
600 - 640	+0.146 -4.565	+2.824 -3.270	0.253	+2.839 -5.621
640 - 680	+0.198 -4.163	+3.368 -3.298	0.236	+3.382 -5.316
680 - 720	+0.155 -3.754	+2.352 -3.247	0.227	+2.368 -4.968
720 - 760	+0.196 -3.842	+3.267 -3.268	0.219	+3.280 -5.049
760 - 800	+0.126 -3.523	+3.366 -3.272	0.212	+3.375 -4.813
800 - 850	+0.110 -3.368	+3.596 -3.261	0.206	+3.604 -4.693
850 - 900	+0.048 -3.351	+3.909 -3.309	0.200	+3.915 -4.714
900 - 950	+0.116 -3.504	+4.148 -3.334	0.196	+4.154 -4.841
950 - 1000	+0.127 -3.511	+4.300 -3.335	0.192	+4.306 -4.846
1000 - 1060	+0.282 -3.683	+4.604 -3.357	0.204	+4.617 -4.988
1060 - 1120	+0.436 -3.779	+3.079 -3.375	0.224	+3.118 -5.071
1120 - 1180	+0.732 -3.982	+5.430 -3.452	0.241	+5.485 -5.276
1180 - 1250	+0.813 -4.031	+5.835 -3.511	0.258	+5.897 -5.352
1250 - 1320	+1.303 -4.414	+6.626 -3.591	0.275	+6.759 -5.697
1320 - 1390	+1.403 -4.471	+7.036 -3.659	0.290	+7.180 -5.785
1390 - 1460	+1.564 -4.590	+7.657 -3.778	0.304	+7.822 -5.953
1460 - 1530	+1.765 -4.738	+8.438 -3.853	0.316	+8.626 -6.115
1530 - 1600	+2.040 -4.972	+9.306 -3.962	0.328	+9.532 -6.366
1600 - 1680	+2.313 -5.179	+10.381 -4.075	0.339	+10.641 -6.599

## A.5 Definitions

**Clopper-Pearson Method :**

**Crystal Ball Function :** The Crystal Ball function, developed within the Crystal Ball Collaboration, is a probability density function which is often used as a fitting function in high energy physics. This function, described by Eq. A.1, consists of a Gaussian core with separate power-law low-end tails, below a certain threshold.

$$f = N \cdot \begin{cases} e^{-\frac{1}{2}\alpha_L^2} \cdot \left[ \left( \frac{\alpha_L}{n_L} \right) \left( \frac{n_L}{\alpha_L} - [\alpha_L + x] \right) \right]^{-n_L}, & x < -\alpha_L \\ e^{-\frac{1}{2}x^2}, & -\alpha_L \leq x \leq \alpha_H \\ e^{-\frac{1}{2}\alpha_H^2} \cdot \left[ \left( \frac{\alpha_H}{n_H} \right) \left( \frac{n_H}{\alpha_H} - [\alpha_H + x] \right) \right]^{-n_H}, & x > \alpha_H \end{cases} \quad (\text{A.1})$$

where  $N$  is a normalisation factor,  $\alpha_L$  and  $\alpha_H$  delimit the Gaussian core, which is replaced by a power-law behaviour proportional to  $1/n_L$  and  $1/n_H$  to the lower and higher side, respectively. The Crystal Ball function itself and its first derivative are continuous.



# Bibliography

- [1] CMS Collaboration, “Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at  $\sqrt{s} = 7$  TeV and first determination of the strong coupling constant in the TeV range”, Eur. Phys. J. C **73**, no. 10, 2604 (2013) doi:[10.1140/epjc/s10052-013-2604-6](https://doi.org/10.1140/epjc/s10052-013-2604-6) arXiv:[1304.7498](https://arxiv.org/abs/1304.7498) [hep-ex]
- [2] D. Griffiths, “Introduction to elementary particles; 2nd rev. version”, Wiley, New York (2008) <https://cds.cern.ch/record/111880>
- [3] D. H. Perkins, “Introduction to high energy physics”, Cambridge (1982) ISBN : 9780521621960
- [4] M. Herrero, “The Standard model”, NATO Sci. Ser. C **534**, 1 (1999) doi:[10.1007/978-94-011-4689-0\\_1](https://doi.org/10.1007/978-94-011-4689-0_1) arXiv:[hep-ph/9812242](https://arxiv.org/abs/hep-ph/9812242)
- [5] S. Weinberg, “A Model of Leptons”, Phys. Rev. Lett. **19**, 1264 (1967) doi:[10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264)
- [6] M. E. Peskin and D. V. Schroeder, “An Introduction to quantum field theory”, Addison-Wesley, USA (1995)
- [7] S. L. Glashow, “Towards a Unified Theory: Threads in a Tapestry”, Rev. Mod. Phys. **52**, 539 (1980) doi:[10.1103/RevModPhys.52.539](https://doi.org/10.1103/RevModPhys.52.539)

- [8] A. Salam, “Gauge Unification of Fundamental Forces”, Rev. Mod. Phys. **52**, 525 (1980) [Science **210**, 723 (1980)] doi:[10.1103/RevModPhys.52.525](https://doi.org/10.1103/RevModPhys.52.525)
- [9] H. Georgi and S. L. Glashow, “Unity of All Elementary Particle Forces”, Phys. Rev. Lett. **32**, 438 (1974) doi:[10.1103/PhysRevLett.32.438](https://doi.org/10.1103/PhysRevLett.32.438)
- [10] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, Phys. Rev. Lett. **13**, 508 (1964) doi:[10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508)
- [11] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, Phys. Rev. Lett. **13**, 321 (1964) doi:[10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321)
- [12] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, Phys. Lett. B **716**, 30 (2012) doi:[10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021) arXiv:[1207.7235](https://arxiv.org/abs/1207.7235) [hep-ex]
- [13] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, Phys. Lett. B **716**, 1 (2012) doi:[10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020) arXiv:[1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex]
- [14] R. K. Ellis, W. J. Stirling and B. R. Webber, “QCD and collider physics”, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8**, 1 (1996)
- [15] F. Halzen and A. D. Martin, “Quarks And Leptons: An Introductory Course In Modern Particle Physics”, New York, Usa: Wiley (1984) 396p
- [16] M. Gell-Mann, “Symmetries of baryons and mesons”, Phys. Rev. **125**, 1067 (1962) doi:[10.1103/PhysRev.125.1067](https://doi.org/10.1103/PhysRev.125.1067)
- [17] K. Rabbertz, “Jet Physics at the LHC : The Strong Force beyond the TeV Scale”, Springer Tracts Mod. Phys. **268**, pp.1 (2017) doi:[10.1007/978-3-319-42115-5](https://doi.org/10.1007/978-3-319-42115-5)

- [18] C. G. Callan, Jr., “Broken scale invariance in scalar field theory”, Phys. Rev. D **2**, 1541 (1970) doi:[10.1103/PhysRevD.2.1541](https://doi.org/10.1103/PhysRevD.2.1541)
- [19] G. ’t Hooft, “Dimensional regularization and the renormalization group”, Nucl. Phys. B **61**, 455 (1973) doi:[10.1016/0550-3213\(73\)90376-3](https://doi.org/10.1016/0550-3213(73)90376-3)
- [20] S. Weinberg, “New approach to the renormalization group”, Phys. Rev. D **8**, 3497 (1973) doi:[10.1103/PhysRevD.8.3497](https://doi.org/10.1103/PhysRevD.8.3497)
- [21] C. Patrignani *et al.* (Particle Data Group), “Review of Particle Physics”, Chin. Phys. C **40**, no. 10, 100001 (2016) doi:[10.1088/1674-1137/40/10/100001](https://doi.org/10.1088/1674-1137/40/10/100001)
- [22] J. C. Collins, D. E. Soper and G. F. Sterman, “Factorization of Hard Processes in QCD”, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1989) doi:[10.1142/9789814503266\\_0001](https://doi.org/10.1142/9789814503266_0001) arXiv:[hep-ph/0409313](https://arxiv.org/abs/hep-ph/0409313)
- [23] V. N. Gribov and L. N. Lipatov, “Deep inelastic e p scattering in perturbation theory”, Sov. J. Nucl. Phys. **15**, 438 (1972) [Yad. Fiz. **15**, 781 (1972)]
- [24] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.”, Sov. Phys. JETP **46**, 641 (1977) [Zh. Eksp. Teor. Fiz. **73**, 1216 (1977)]
- [25] G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language”, Nucl. Phys. B **126**, 298 (1977) doi:[10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4)
- [26] CMS Collaboration, “Measurement of the triple-differential dijet cross section in proton-proton collisions at  $\sqrt{s} = 8$  TeV and constraints on parton distribution functions”, Eur. Phys. J. C **77**, no. 11, 746 (2017) doi:[10.1140/epjc/s10052-017-5286-7](https://doi.org/10.1140/epjc/s10052-017-5286-7) arXiv:[1705.02628 \[hep-ex\]](https://arxiv.org/abs/1705.02628)

- [27] S. Dulat *et al.*, “New parton distribution functions from a global analysis of quantum chromodynamics”, Phys. Rev. D **93**, no. 3, 033006 (2016) doi:[10.1103/PhysRevD.93.033006](https://doi.org/10.1103/PhysRevD.93.033006) arXiv:[1506.07443](https://arxiv.org/abs/1506.07443) [hep-ph]]
- [28] L. A. Harland-Lang *et al.*, “Parton distributions in the LHC era: MMHT 2014 PDFs”, Eur. Phys. J. C **75**, no. 5, 204 (2015) doi:[10.1140/epjc/s10052-015-3397-6](https://doi.org/10.1140/epjc/s10052-015-3397-6) arXiv:[1412.3989](https://arxiv.org/abs/1412.3989) [hep-ph]
- [29] NNPDF Collaboration, “Parton distributions for the LHC Run II”, JHEP **1504**, 040 (2015) doi:[10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040) arXiv:[1410.8849](https://arxiv.org/abs/1410.8849) [hep-ph]
- [30] S. Alekhin, J. Blumlein and S. Moch, “Parton Distribution Functions and Benchmark Cross Sections at NNLO”, Phys. Rev. D **86**, 054009 (2012) doi:[10.1103/PhysRevD.86.054009](https://doi.org/10.1103/PhysRevD.86.054009) arXiv:[1202.2281](https://arxiv.org/abs/1202.2281) [hep-ph]
- [31] H1 and ZEUS Collaborations, “Combination of measurements of inclusive deep inelastic  $e^\pm p$  scattering cross sections and QCD analysis of HERA data”, Eur. Phys. J. C **75**, no. 12, 580 (2015) doi:[10.1140/epjc/s10052-015-3710-4](https://doi.org/10.1140/epjc/s10052-015-3710-4) arXiv:[1506.06042](https://arxiv.org/abs/1506.06042) [hep-ex]
- [32] B. Andersson, “The Lund model”, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **7**, 1 (1997) ISBN: 9780521420945
- [33] G. Marchesini and B. R. Webber, “Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation”, Nucl. Phys. B **310**, 461 (1988) doi:[10.1016/0550-3213\(88\)90089-2](https://doi.org/10.1016/0550-3213(88)90089-2)
- [34] B. R. Webber, “A QCD Model for Jet Fragmentation Including Soft Gluon Interference”, Nucl. Phys. B **238**, 492 (1984) doi:[10.1016/0550-3213\(84\)90333-X](https://doi.org/10.1016/0550-3213(84)90333-X)
- [35] D. Amati and G. Veneziano, “Preconfinement as a Property of Perturbative QCD”, Phys. Lett. **83B**, 87 (1979) doi:[10.1016/0370-2693\(79\)90896-7](https://doi.org/10.1016/0370-2693(79)90896-7)

- [36] G. F. Sterman and S. Weinberg, “Jets from Quantum Chromodynamics”, Phys. Rev. Lett. **39**, 1436 (1977) doi:[10.1103/PhysRevLett.39.1436](https://doi.org/10.1103/PhysRevLett.39.1436)
- [37] G. Hanson *et al.*, “Evidence for Jet Structure in Hadron Production by e+ e- Annihilation”, Phys. Rev. Lett. **35**, 1609 (1975) doi:[10.1103/PhysRevLett.35.1609](https://doi.org/10.1103/PhysRevLett.35.1609)
- [38] G. P. Salam, “Towards Jetography”, Eur. Phys. J. C **67**, 637 (2010) doi:[10.1140/epjc/s10052-010-1314-6](https://doi.org/10.1140/epjc/s10052-010-1314-6) arXiv:[0906.1833 \[hep-ph\]](https://arxiv.org/abs/0906.1833)
- [39] S. D. Ellis, Z. Kunszt and D. E. Soper, “The One Jet Inclusive Cross-Section at Order  $\alpha_s^3$ . 1. Gluons Only”, Phys. Rev. D **40**, 2188 (1989) doi:[10.1103/PhysRevD.40.2188](https://doi.org/10.1103/PhysRevD.40.2188)
- [40] G. C. Blazey *et al.*, “Run II jet physics”, arXiv:[hep-ex/0005012](https://arxiv.org/abs/hep-ex/0005012)
- [41] S. Weinzierl, “The SISCone jet algorithm optimised for low particle multiplicities”, Comput. Phys. Commun. **183**, 813 (2012) doi:[10.1016/j.cpc.2011.12.007](https://doi.org/10.1016/j.cpc.2011.12.007) arXiv:[1108.1934 \[hep-ph\]](https://arxiv.org/abs/1108.1934)
- [42] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, Phys. Rev. D **48**, 3160 (1993) doi:[10.1103/PhysRevD.48.3160](https://doi.org/10.1103/PhysRevD.48.3160) arXiv:[hep-ph/9305266](https://arxiv.org/abs/hep-ph/9305266)
- [43] S. Catani *et al.*, “Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions”, Nucl. Phys. B **406**, 187 (1993) doi:[10.1016/0550-3213\(93\)90166-M](https://doi.org/10.1016/0550-3213(93)90166-M)
- [44] . Catani, Y. L. Dokshitzer and B. R. Webber, “The  $K^-$  perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions”, Phys. Lett. B **285**, 291 (1992) doi:[10.1016/0370-2693\(92\)91467-N](https://doi.org/10.1016/0370-2693(92)91467-N)

- [45] Y. L. Dokshitzer *et al.*, “Better jet clustering algorithms”, JHEP **9708**, 001 (1997) doi:[10.1088/1126-6708/1997/08/001](https://doi.org/10.1088/1126-6708/1997/08/001) arXiv:[hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323)
- [46] M. Cacciari, G. P. Salam and G. Soyez, “The Anti-k(t) jet clustering algorithm”, JHEP **0804**, 063 (2008) doi:[10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063) arXiv:[0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189)
- [47] M. Cacciari, G. P. Salam and G. Soyez, “FastJet User Manual”, Eur. Phys. J. C **72**, 1896 (2012) doi:[10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2) arXiv:[1111.6097 \[hep-ph\]](https://arxiv.org/abs/1111.6097)
- [48] UA1 Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540\text{-GeV}$ ”, Phys. Lett. **122B**, 103 (1983) doi:[10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2)
- [49] UA2 Collaboration, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”, Phys. Lett. **122B**, 476 (1983) doi:[10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2)
- [50] L. Evans and P. Bryant, “LHC Machine”, JINST **3**, S08001 (2008) doi:[10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001)
- [51] S. Myers and E. Picasso, “The design, construction and commissioning of the CERN large Electron–Positron collider”, Contemporary Physics **31**, no. 6, 387 (1990) doi:[10.1080/00107519008213789](https://doi.org/10.1080/00107519008213789)
- [52] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, JINST **3**, S08002 (2008) doi:[10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002)
- [53] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, JINST **3**, S08003 (2008) doi:[10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003)
- [54] CMS Collaboration, “The CMS Experiment at the CERN LHC”, JINST **3**, S08004 (2008) doi:[10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004)

- [55] CMS Collaboration, “CMS Physics : Technical Design Report Volume 1: Detector Performance and Software”, CERN-LHCC-2006-001 (2006), CMS-TDR-8-1
- [56] CMS Collaboration, “CMS technical design report, volume II: Physics performance”, J. Phys. G **34**, no. 6, 995 (2007) doi:[10.1088/0954-3899/34/6/S01](https://doi.org/10.1088/0954-3899/34/6/S01)
- [57] LHCb Collaboration, “The LHCb Detector at the LHC”, JINST **3**, S08005 (2008) doi:[10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005)
- [58] LHCf Collaboration, “The LHCf detector at the CERN Large Hadron Collider”, JINST **3**, S08006 (2008) doi:[10.1088/1748-0221/3/08/S08006](https://doi.org/10.1088/1748-0221/3/08/S08006)
- [59] TOTEM Collaboration, “The TOTEM experiment at the CERN Large Hadron Collider”, JINST **3**, S08007 (2008) doi:[10.1088/1748-0221/3/08/S08007](https://doi.org/10.1088/1748-0221/3/08/S08007)
- [60] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, CMS Physics Analysis Summary **CMS-PAS-LUM-13-001** (2013) <http://cds.cern.ch/record/1598864>
- [61] S. van der Meer, “Calibration of the Effective Beam Height in the ISR”, Technical Report CERN-ISR-PO-68-31 (1968)
- [62] P. Adzic *et al.*, “Energy resolution of the barrel of the CMS electromagnetic calorimeter”, JINST **2**, P04004 (2007) doi:[10.1088/1748-0221/2/04/P04004](https://doi.org/10.1088/1748-0221/2/04/P04004)
- [63] CMS Collaboration, “Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data”, JINST **5**, T03012 (2010) doi:[10.1088/1748-0221/5/03/T03012](https://doi.org/10.1088/1748-0221/5/03/T03012) arXiv:[0911.4991 \[physics.ins-det\]](https://arxiv.org/abs/0911.4991)
- [64] I. Bird *et al.*, “LHC computing Grid. Technical design report”, CERN-LHCC-2005-024 <http://cds.cern.ch/record/840543>
- [65] CMS Collaboration, “CMS: The computing project. Technical design report”, CERN-LHCC-2005-023 (2005) <http://cds.cern.ch/record/838359>

- [66] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework”, Nucl. Instrum. Meth. A **389**, 81 (1997) doi:[10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X)
- [67] Z. Nagy, “Three jet cross-sections in hadron hadron collisions at next-to-leading order”, Phys. Rev. Lett. **88**, 122003 (2002) doi:[10.1103/PhysRevLett.88.122003](https://doi.org/10.1103/PhysRevLett.88.122003) arXiv:[hep-ph/0110315](https://arxiv.org/abs/hep-ph/0110315)
- [68] Z. Nagy, “Next-to-leading order calculation of three jet observables in hadron hadron collision”, Phys. Rev. D **68**, 094002 (2003) doi:[10.1103/PhysRevD.68.094002](https://doi.org/10.1103/PhysRevD.68.094002) arXiv:[hep-ph/0307268](https://arxiv.org/abs/hep-ph/0307268)
- [69] D. Britzger *et al.*, “New features in version 2 of the fastNLO project”, 217 (2012) doi:[10.3204/DESY-PROC-2012-02/165](https://doi.org/10.3204/DESY-PROC-2012-02/165) arXiv:[1208.3641 \[hep-ph\]](https://arxiv.org/abs/1208.3641)
- [70] T. Kluge, K. Rabbertz and M. Wobisch, “FastNLO: Fast pQCD calculations for PDF fits”, Proceedings DIS 2006, 14th International Workshop, 483 (2006) doi:[10.1142/9789812706706\\_0110](https://doi.org/10.1142/9789812706706_0110) arXiv:[hep-ph/0609285](https://arxiv.org/abs/hep-ph/0609285)
- [71] M. R. Whalley, D. Bourilkov and R. C. Group, “The Les Houches accord PDFs (LHAPDF) and LHAGLUE”, Proceedings : Workshop on the implications of HERA for LHC physics (2005) arXiv:[hep-ph/0508110](https://arxiv.org/abs/hep-ph/0508110)
- [72] A. Buckley, *et al.*, “LHAPDF6: parton density access in the LHC precision era”, Eur. Phys. J. C **75**, 132 (2015) doi:[10.1140/epjc/s10052-015-3318-8](https://doi.org/10.1140/epjc/s10052-015-3318-8) arXiv:[1412.7420 \[hep-ph\]](https://arxiv.org/abs/1412.7420)
- [73] N. Metropolis and S. Ulam, “The Monte Carlo Method”, Journal of the American Statistical Association **44**, no. 247, 335 (1949) doi:[10.1080/01621459.1949.10483310](https://doi.org/10.1080/01621459.1949.10483310)
- [74] B. Andersson *et al.*, “Parton fragmentation and string dynamics”, Physics Reports, **97**, no. 2, 31 (1983) doi:[10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7)

- [75] T. Sjostrand, S. Mrenna and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, JHEP **0605**, 026 (2006) doi:[10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026) arXiv:[hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175)
- [76] T. Sjostrand, S. Mrenna and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1”, Comput. Phys. Commun. **178**, 852 (2008) doi:[10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036) arXiv:[0710.3820 \[hep-ph\]](https://arxiv.org/abs/0710.3820)
- [77] R. Field, “Min-Bias and the Underlying Event at the LHC”, Acta Phys. Polon. B **42**, 2631 (2011) doi:[10.5506/APhysPolB.42.2631](https://doi.org/10.5506/APhysPolB.42.2631) arXiv:[1110.5530 \[hep-ph\]](https://arxiv.org/abs/1110.5530)
- [78] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, Eur. Phys. J. C **76**, no. 3, 155 (2016) doi:[10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x) arXiv:[1512.00815 \[hep-ex\]](https://arxiv.org/abs/1512.00815)
- [79] J. Alwall *et al.*, “MadGraph 5 : Going Beyond”, JHEP **1106**, 128 (2011) doi:[10.1007/JHEP06\(2011\)128](https://doi.org/10.1007/JHEP06(2011)128) arXiv:[1106.0522 \[hep-ph\]](https://arxiv.org/abs/1106.0522)
- [80] J. Alwall *et al.*, “A Standard format for Les Houches event files”, Comput. Phys. Commun. **176**, 300 (2007) doi:[10.1016/j.cpc.2006.11.010](https://doi.org/10.1016/j.cpc.2006.11.010) arXiv:[hep-ph/0609017](https://arxiv.org/abs/hep-ph/0609017)
- [81] G. Corcella *et al.*, “HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)”, JHEP **0101**, 010 (2001) doi:[10.1088/1126-6708/2001/01/010](https://doi.org/10.1088/1126-6708/2001/01/010) arXiv:[hep-ph/0011363](https://arxiv.org/abs/hep-ph/0011363)
- [82] M. Bahr *et al.*, “Herwig++ Physics and Manual”, Eur. Phys. J. C **58**, 639 (2008) doi:[10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9) arXiv:[0803.0883 \[hep-ph\]](https://arxiv.org/abs/0803.0883)
- [83] M. Bahr *et al.*, “Herwig++ 2.3 Release Note”, arXiv:[0812.0529 \[hep-ph\]](https://arxiv.org/abs/0812.0529)
- [84] S. Frixione, P. Nason and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, JHEP **0711**, 070 (2007) doi:[10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070) arXiv:[0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092)

- [85] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, JHEP **0411**, 040 (2004) doi:[10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040) arXiv:[hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146)
- [86] S. Alioli *et al.*, “Jet pair production in POWHEG”, JHEP **1104**, 081 (2011) doi:[10.1007/JHEP04\(2011\)081](https://doi.org/10.1007/JHEP04(2011)081) arXiv:[1012.3380 \[hep-ph\]](https://arxiv.org/abs/1012.3380)
- [87] C. Oleari, “The POWHEG-BOX”, Nucl. Phys. Proc. Suppl. **205-206**, 36 (2010) doi:[10.1016/j.nuclphysbps.2010.08.016](https://doi.org/10.1016/j.nuclphysbps.2010.08.016) arXiv:[1007.3893 \[hep-ph\]](https://arxiv.org/abs/1007.3893)
- [88] GEANT4 Collaboration, “GEANT4: A Simulation toolkit”, Nucl. Instrum. Meth. A **506**, 250 (2003) doi:[10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8)
- [89] CMS Collaboration, “The fast simulation of the CMS detector at LHC”, J. Phys. Conf. Ser. **331**, 032049 (2011) doi:<http://dx.doi.org/10.1088/1742-6596/331/3/032049>
- [90] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, CMS-PAS-PFT-09-001 (2009) <http://cds.cern.ch/record/1194487>
- [91] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector”, CMS-PAS-PFT-10-001 (2010) <http://cds.cern.ch/record/1247373>
- [92] W. Adam *et al.*, “Track reconstruction in the CMS tracker”, CERN-CMS-NOTE-2006-041 (2006) <http://cds.cern.ch/record/934067>
- [93] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting”, Nucl. Instrum. Meth. A **262**, 444 (1987) doi:[10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4)

- [94] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems”, Proceedings of the IEEE **86** (1998) doi:[10.1109/5.726788](https://doi.org/10.1109/5.726788)
- [95] R. Fruhwirth, W. Waltenberger and P. Vanlaer, “Adaptive vertex fitting”, J. Phys. G **34**, N343 (2007) doi:[10.1088/0954-3899/34/12/N01](https://doi.org/10.1088/0954-3899/34/12/N01)
- [96] T. Schörner-Sadenius, “The Large Hadron Collider : Harvest of Run 1”, Springer (2015) doi:[10.1007/978-3-319-15001-7](https://doi.org/10.1007/978-3-319-15001-7)
- [97] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, JINST **6**, P11002 (2011) doi:[10.1088/1748-0221/6/11/P11002](https://doi.org/10.1088/1748-0221/6/11/P11002) arXiv:[1107.4277 \[physics.ins-det\]](https://arxiv.org/abs/1107.4277)
- [98] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, JINST **12**, P02014 (2017) doi:[10.1088/1748-0221/12/02/P02014](https://doi.org/10.1088/1748-0221/12/02/P02014) arXiv:[1607.03663 \[hep-ex\]](https://arxiv.org/abs/1607.03663)
- [99] CMS Collaboration, “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at  $\sqrt{s} = 8$  TeV and cross section ratios to 2.76 and 7 TeV”, JHEP **1703**, 156 (2017) doi:[10.1007/JHEP03\(2017\)156](https://doi.org/10.1007/JHEP03(2017)156) arXiv:[1609.05331 \[hep-ex\]](https://arxiv.org/abs/1609.05331)
- [100] C. J. Clopper and E. S. Pearson, “The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial”, Biometrika **26**, 404 (1934) doi:[10.2307/2331986](https://doi.org/10.2307/2331986)
- [101] CMS Collaboration, “Jet Performance in pp Collisions at 7 TeV”, CMS Physics Analysis Summary **CMS-PAS-JME-10-003** (2010) <http://cds.cern.ch/record/1279362>
- [102] CMS Collaboration, “Jet Identification at 8 TeV”, URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID> (accessed on 2017-10-31)

- [103] CMS Collaboration, “Jet Energy Resolution at 8 TeV”, URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution> (accessed on 2016-01-22)
- [104] CMS Collaboration, “Jet Energy Resolution in CMS at  $\sqrt{s}=7$  TeV” CMS Physics Analysis Summary **CMS-PAS-JME-10-014** (2011) <http://cds.cern.ch/record/1339945>
- [105] Tim Adye, “Unfolding algorithms and tests using RooUnfold”, Proceedings of the PHYSTAT 2011 Workshop, CERN, Geneva, Switzerland 313 (January 2011) doi:[10.5170/CERN-2011-006.313](https://doi.org/10.5170/CERN-2011-006.313) arXiv:[1105.1160](https://arxiv.org/abs/1105.1160) [physics.data-an]
- [106] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem”, Nucl. Instrum. Meth. A **362**, 487 (1995) doi:[10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X)
- [107] G. D’Agostini, “Improved iterative Bayesian unfolding”, (2010) arXiv:[1010.0632](https://arxiv.org/abs/1010.0632) [physics.data-an]
- [108] CMS Collaboration, “Measurement of the Inclusive Jet Cross Section in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV”, Phys. Rev. Lett. **107**, 132001 (2011) doi:[10.1103/PhysRevLett.107.132001](https://doi.org/10.1103/PhysRevLett.107.132001) arXiv:[1106.0208](https://arxiv.org/abs/1106.0208) [hep-ex]
- [109] CMS Collaboration, “Jet energy scale uncertainty sources”, URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources> (accessed on 2017-10-31)
- [110] CMS Collaboration, “Measurements of differential jet cross sections in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the CMS detector”, Phys. Rev. D **87**, no. 11, 112002 (2013) Erratum: [Phys. Rev. D **87**, no. 11, 119902 (2013)] doi:[10.1103/PhysRevD.87.112002](https://doi.org/10.1103/PhysRevD.87.112002), [10.1103/PhysRevD.87.119902](https://doi.org/10.1103/PhysRevD.87.119902) arXiv:[1212.6660](https://arxiv.org/abs/1212.6660) [hep-ex]

- [111] ATLAS Collaboration, “Measurement of the inclusive jet cross section in pp collisions at  $\sqrt{s}=2.76$  TeV and comparison to the inclusive jet cross section at  $\sqrt{s}=7$  TeV using the ATLAS detector”, Eur. Phys. J. C **73**, no. 8, 2509 (2013) doi:[10.1140/epjc/s10052-013-2509-4](https://doi.org/10.1140/epjc/s10052-013-2509-4) arXiv:[1304.4739](https://arxiv.org/abs/1304.4739) [hep-ex]
- [112] ATLAS Collaboration, “Measurement of the inclusive jet cross-section in proton-proton collisions at  $\sqrt{s} = 7$  TeV using  $4.5 \text{ fb}^{-1}$  of data with the ATLAS detector”, JHEP **1502**, 153 (2015) Erratum: [JHEP **1509**, 141 (2015)] doi:[10.1007/JHEP02\(2015\)153](https://doi.org/10.1007/JHEP02(2015)153), [10.1007/JHEP09\(2015\)141](https://doi.org/10.1007/JHEP09(2015)141) arXiv:[1410.8857](https://arxiv.org/abs/1410.8857) [hep-ex]
- [113] CMS Collaboration, “Measurement of the inclusive 3-jet production differential cross section in proton–proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range”, Eur. Phys. J. C **75**, no. 5, 186 (2015) doi:[10.1140/epjc/s10052-015-3376-y](https://doi.org/10.1140/epjc/s10052-015-3376-y) arXiv:[1412.1633](https://arxiv.org/abs/1412.1633) [hep-ex]
- [114] CMS Collaboration, “Measurement of the inclusive jet cross section in pp collisions at  $\sqrt{s} = 2.76 \text{ TeV}$ ”, Eur. Phys. J. C **76**, no. 5, 265 (2016) doi:[10.1140/epjc/s10052-016-4083-z](https://doi.org/10.1140/epjc/s10052-016-4083-z) arXiv:[1512.06212](https://arxiv.org/abs/1512.06212) [hep-ex]]
- [115] H. L. Lai *et al.*, “New parton distributions for collider physics”, Phys. Rev. D **82**, 074024 (2010) doi:[10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024) arXiv:[1007.2241](https://arxiv.org/abs/1007.2241) [hep-ph]
- [116] A. D. Martin *et al.*, “Parton distributions for the LHC”, Eur. Phys. J. C **63**, 189 (2009) doi:[10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5) arXiv:[0901.0002](https://arxiv.org/abs/0901.0002) [hep-ph]
- [117] A. D. Martin *et al.*, “Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections”, Eur. Phys. J. C **64**, 653 (2009) doi:[10.1140/epjc/s10052-009-1164-2](https://doi.org/10.1140/epjc/s10052-009-1164-2) arXiv:[0905.3531](https://arxiv.org/abs/0905.3531) [hep-ph]
- [118] R. D. Ball *et al.*, “Parton distributions with LHC data”, Nucl. Phys. B **867**, 244 (2013) doi:[10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003) arXiv:[1207.1303](https://arxiv.org/abs/1207.1303) [hep-ph]

- [119] W. Hollik *et al.*, “Electroweak physics”, *Acta Phys. Polon. B* **35**, 2533 (2004) arXiv:[hep-ph/0501246](#)
- [120] S. Dittmaier, A. Huss and C. Speckner, “Weak radiative corrections to dijet production at hadron colliders”, *JHEP* **1211**, 095 (2012) doi:[10.1007/JHEP11\(2012\)095](#) arXiv:[1210.0438 \[hep-ph\]](#)
- [121] J. Pumplin *et al.*, “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **0207**, 012 (2002) doi:[10.1088/1126-6708/2002/07/012](#) arXiv:[hep-ph/0201195](#)
- [122] CMS Collaboration, “Determination of the top-quark pole mass and strong coupling constant from the t t-bar production cross section in pp collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Lett. B* **728**, 496 (2014) Erratum: [Phys. Lett. B **738**, 526 (2014)] doi:[10.1016/j.physletb.2014.08.040](#), [10.1016/j.physletb.2013.12.009](#) arXiv:[1307.1907 \[hep-ex\]](#)
- [123] CMS Collaboration, “Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at  $\sqrt{s} = 7$  TeV”, *Eur. Phys. J. C* **75**, no. 6, 288 (2015) doi:[10.1140/epjc/s10052-015-3499-1](#) arXiv:[1410.6765 \[hep-ex\]](#)
- [124] ATLAS Collaboration, “Measurement of transverse energy-energy correlations in multi-jet events in  $pp$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector and determination of the strong coupling constant  $\alpha_s(m_Z)$ ”, *Phys. Lett. B* **750**, 427 (2015) doi:[10.1016/j.physletb.2015.09.050](#) arXiv:[1508.01579 \[hep-ex\]](#)
- [125] D0 Collaboration, “Determination of the strong coupling constant from the inclusive jet cross section in  $p\bar{p}$  collisions at  $\text{sqrt}(s)=1.96$  TeV”, *Phys. Rev. D* **80**, 111107 (2009) doi:[10.1103/PhysRevD.80.111107](#) arXiv:[0911.2710 \[hep-ex\]](#)

- [126] D0 Collaboration, “Measurement of angular correlations of jets at  $\sqrt{s} = 1.96$  TeV and determination of the strong coupling at high momentum transfers”, Phys. Lett. B **718**, 56 (2012) doi:[10.1016/j.physletb.2012.10.003](https://doi.org/10.1016/j.physletb.2012.10.003) arXiv:[1207.4957](https://arxiv.org/abs/1207.4957) [hep-ex]
- [127] H1 Collaboration, “Measurement of multijet production in  $ep$  collisions at high  $Q^2$  and determination of the strong coupling  $\alpha_s$ ”, Eur. Phys. J. C **75**, no. 2, 65 (2015) doi:[10.1140/epjc/s10052-014-3223-6](https://doi.org/10.1140/epjc/s10052-014-3223-6) arXiv:[1406.4709](https://arxiv.org/abs/1406.4709) [hep-ex]
- [128] H1 Collaboration, “Measurement of Jet Production Cross Sections in Deep-inelastic ep Scattering at HERA”, Eur. Phys. J. C **77**, no. 4, 215 (2017) doi:[10.1140/epjc/s10052-017-4717-9](https://doi.org/10.1140/epjc/s10052-017-4717-9) arXiv:[1611.03421](https://arxiv.org/abs/1611.03421) [hep-ex]
- [129] ZEUS Collaboration, “Inclusive-jet photoproduction at HERA and determination of alphas”, Nucl. Phys. B **864**, 1 (2012) doi:[10.1016/j.nuclphysb.2012.06.006](https://doi.org/10.1016/j.nuclphysb.2012.06.006) arXiv:[1205.6153](https://arxiv.org/abs/1205.6153) [hep-ex]]
- [130] CDF Collaboration, “Measurement of the strong coupling constant from inclusive jet production at the Tevatron  $\bar{p}p$  collider”, Phys. Rev. Lett. **88**, 042001 (2002) doi:[10.1103/PhysRevLett.88.042001](https://doi.org/10.1103/PhysRevLett.88.042001) arXiv:[hep-ex/0108034](https://arxiv.org/abs/hep-ex/0108034)
- [131] R. D. Ball *et al.*, “A first unbiased global NLO determination of parton distributions and their uncertainties”, Nucl. Phys. B **838**, 136 (2010) doi:[10.1016/j.nuclphysb.2010.05.008](https://doi.org/10.1016/j.nuclphysb.2010.05.008) arXiv:[1002.4407](https://arxiv.org/abs/1002.4407) [hep-ph]
- [132] B. Schmidt and M. Steinhauser, “CRunDec: a C++ package for running and decoupling of the strong coupling and quark masses”, Comput. Phys. Commun. **183**, 1845 (2012) doi:[10.1016/j.cpc.2012.03.023](https://doi.org/10.1016/j.cpc.2012.03.023) arXiv:[1201.6149](https://arxiv.org/abs/1201.6149) [hep-ph]
- [133] K. G. Chetyrkin, J. H. Kuhn and M. Steinhauser, “RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses”,

- Comput. Phys. Commun. **133**, 43 (2000) doi:[10.1016/S0010-4655\(00\)00155-7](https://doi.org/10.1016/S0010-4655(00)00155-7)  
arXiv:[hep-ph/0004189](https://arxiv.org/abs/hep-ph/0004189)
- [134] CMS Collaboration, “Upgrade of the CMS Hadron Outer Calorimeter with SiPM sensors”, J. Phys. Conf. Ser. **404**, 012018 (2012) doi:[10.1088/1742-6596/404/1/012018](https://doi.org/10.1088/1742-6596/404/1/012018)
- [135] CMS Collaboration, “Upgrade of the CMS Hadron Outer Calorimeter with SIPMs”, Phys. Procedia **37**, 72 (2012) doi:[10.1016/j.phpro.2012.02.358](https://doi.org/10.1016/j.phpro.2012.02.358)
- [136] CMS Collaboration, “MPPC photon sensor operational experience in CMS”, Proceedings of 16th International Conference on Calorimetry in High Energy Physics (CALOR 2014) **587**, 12022 (2015) doi:[10.1088/1742-6596/587/1/012022](https://doi.org/10.1088/1742-6596/587/1/012022)
- [137] CMS Collaboration, “Commissioning and performance of the CMS Hadron Outer Calorimeter”, CMS Detector Note (Unpublished) **CMS DN-14-019** (2014)
- [138] CMS Collaboration, “CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter”, CERN-LHCC-2012-015, CMS-TDR-010, FERMILAB-DESIGN-2012-01 (2012) doi:[10.2172/1151651](https://doi.org/10.2172/1151651)
- [139] CMS Collaboration, “Data Quality Monitoring (DQM) and Data Certification (DC)”, URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/DQM> (accessed on 2018-2-10)
- [140] CMS Collaboration, “Data certification activity”, URL: <https://twiki.cern.ch/twiki/bin/view/CMS/DQMDataCertification> (accessed on 2018-2-10)

# List of Publications/Conferences

- CMS Public Documents

1. CMS Collaboration, “Determination of the strong coupling constant from the measurement of inclusive multijet event cross sections in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$ ”, **CMS-PAS-SMP-16-008** (2017).

- CMS Internal Documents

- Analysis Notes

1. A. Kaur *et al.*, “Search for light scalar resonances decaying to b quarks at  $\sqrt{s} = 13 \text{ TeV}$ ”, **CMS AN-2016/384** (2016).
2. A. Kaur *et al.*, “Measurement of inclusive multijet cross-sections in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$ ”, **CMS AN-2015/102** (2015).

- Detector Note

1. A. Kaur *et al.*, “Commissioning and performance of the CMS Hadron Outer Calorimeter”, **CMS DN-14-019** (2014).

- Papers in Refreed Journals

1. A. Kaur *et al.*, “Measurements of event properties and multi-differential jet cross sections and impact of CMS measurements on Proton Struc-

- ture and QCD parameters”, EPJ Web of Conferences **172** (2018) doi:[10.1051/epjconf/201817202001](https://doi.org/10.1051/epjconf/201817202001).
2. A. Kaur *et al.*, “Extraction of the strong coupling constant from the measurement of inclusive multijet event cross-sections in pp collisions at center of mass energy of 8 TeV”, Springer Proc. Phys. **203**, 02001 (2018) doi:10.1007/978-3-319-73171-1.
  3. M. Kaur and Anter P. Kaur, “Subjet Multiplicities at LHC Energies and the QCD Color Factor Ratio  $C_A/C_F$ ”, Advances in High Energy Physics **2013**, Article ID 585809 (2013) doi:10.1155/2013/585809.

- **Papers presented in Conferences, Workshops and Symposia**

1. A. Kaur *et al.*, “Measurements of event properties and multi-differential jet cross sections and impact of CMS measurements on Proton Structure and QCD parameters”, **XLVII International Symposium on Multiparticle Dynamics (ISMD2017)**, 11-15 September, 2017, Tlaxcala City, Mexico.
2. A. Kaur *et al.*, “Measurement of inclusive multijet cross sections in pp collisions using the CMS detector”, **11<sup>th</sup> Chandigarh Science Congress, CHASCON 2017**, 9-11 March, 2017, Panjab University, Chandigarh, India.
3. A. Kaur *et al.*, “Extraction of the strong coupling constant from the measurement of inclusive multijet event cross-sections in pp collisions at center of mass energy of 8 TeV”, **XXII DAE-BRNS High Energy Physics Symposium 2016**, 12-16 December, 2016, Delhi, India.
4. A. Kaur *et al.*, “Inclusive jets results from CMS”, **International Workshop on Frontiers in Electroweak Interactions of Leptons and Hadrons**, 2-6 November, 2016, Aligarh, India.
5. A. Kaur *et al.*, “HCAL Back-end Upgradation at CMS Detector”, **9<sup>th</sup> Chandigarh Science Congress, CHASCON 2015**, 25-27 February, 2015, Panjab

---

University, Chandigarh, India.

## • Conferences, Schools and Workshops attended

1. **Workshop on Upgrade of the CMS Calorimeter**, 23-24 November, 2015, Tata Institute of Fundamental Research, Mumbai, India.
2. **IX SERC SCHOOL ON EXPERIMENTAL HIGH ENERGY PHYSICS**, 2-21 December, 2013, Indian Institute of Technology Madras, Chennai, India.
3. **CMS Data Analysis School (CMSDAS)**, 7-11 November, 2013, Saha Institute of Nuclear Physics, Kolkata, India.
4. **5<sup>th</sup> Chandigarh Science Congress, CHASCON 2011**, 26-28 February, 2011, Panjab University, Chandigarh, India.

## • CMS Publications<sup>25</sup>

1. CMS Collaboration, “Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, Phys. Rev. Lett. **120**, no. 8, 081801 (2018).
2. CMS Collaboration, “Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, Phys. Lett. B **774**, 682 (2017).
3. CMS Collaboration, “Measurement of the differential cross sections for the associated production of a  $W$  boson and jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, Phys. Rev. D **96**, no. 7, 072005 (2017).
4. CMS Collaboration, “Measurements of jet charge with dijet events in pp collisions at  $\sqrt{s} = 8$  TeV”, JHEP **1710**, 131 (2017).

---

<sup>25</sup>I am a co-author of more than 300 “Journal Articles” with the CMS Collaboration and the full list of publications can be found at <http://inspirehep.net>.

5. CMS Collaboration, “Measurement of the triple-differential dijet cross section in proton-proton collisions at  $\sqrt{s} = 8$  TeV and constraints on parton distribution functions”, *Eur. Phys. J. C* **77**, no. 11, 746 (2017).
6. CMS Collaboration, “Measurements of the  $pp \rightarrow W\gamma\gamma$  and  $pp \rightarrow Z\gamma\gamma$  cross sections and limits on anomalous quartic gauge couplings at  $\sqrt{s} = 8$  TeV”, *JHEP* **1710**, 072 (2017).
7. CMS Collaboration, “Search for anomalous couplings in boosted WW/WZ  $\rightarrow \ell\nu q\bar{q}$  production in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Phys. Lett. B* **772**, 21 (2017).
8. CMS Collaboration, “Measurement of the cross section for electroweak production of  $Z\gamma$  in association with two jets and constraints on anomalous quartic gauge couplings in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Phys. Lett. B* **770**, 380 (2017).
9. CMS Collaboration, “Measurement of electroweak-induced production of  $W\gamma$  with two jets in pp collisions at  $\sqrt{s} = 8$  TeV and constraints on anomalous quartic gauge couplings”, *JHEP* **1706**, 106 (2017).
10. CMS Collaboration, “Measurements of the associated production of a Z boson and b jets in pp collisions at  $\sqrt{s} = 8$  TeV”, *Eur. Phys. J. C* **77**, no. 11, 751 (2017).
11. CMS Collaboration, “Measurements of differential production cross sections for a Z boson in association with jets in pp collisions at  $\sqrt{s} = 8$  TeV”, *JHEP* **1704**, 022 (2017).
12. CMS Collaboration, “Measurements of differential cross sections for associated production of a W boson and jets in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Phys. Rev. D* **95**, 052002 (2017).
13. CMS Collaboration, “Measurement of the WZ production cross section in pp collisions at  $\sqrt{s} = 7$  and 8 TeV and search for anomalous triple gauge couplings at  $\sqrt{s} = 8$  TeV”, *Eur. Phys. J. C* **77**, no. 4, 236 (2017).

14. CMS Collaboration, “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at  $\sqrt{s} = 8$  TeV and cross section ratios to 2.76 and 7 TeV”, JHEP **1703**, 156 (2017).
15. CMS Collaboration, “Measurement of the production cross section of a W boson in association with two b jets in pp collisions at  $\sqrt{s} = 8$  TeV”, Eur. Phys. J. C **77**, no. 2, 92 (2017).
16. CMS Collaboration, “Measurement of the ZZ production cross section and  $Z \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  branching fraction in pp collisions at  $\sqrt{s}=13$  TeV”, Phys. Lett. B **763**, 280 (2016), Erratum: [Phys. Lett. B **772**, 884 (2017)].
17. CMS Collaboration, “Measurement of the WZ production cross section in pp collisions at  $\sqrt{(s)} = 13$  TeV”, Phys. Lett. B **766**, 268 (2017)
18. CMS Collaboration, “Measurement of electroweak production of a W boson and two forward jets in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, JHEP **1611**, 147 (2016).
19. CMS Collaboration, “Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, JHEP **1702**, 096 (2017)
20. CMS Collaboration, “Measurement of the double-differential inclusive jet cross section in proton–proton collisions at  $\sqrt{s} = 13$  TeV”, Eur. Phys. J. C **76**, no. 8, 451 (2016).
21. CMS Collaboration, “Measurement of the differential cross section and charge asymmetry for inclusive  $pp \rightarrow W^\pm + X$  production at  $\sqrt{s} = 8$  TeV”, Eur. Phys. J. C **76**, no. 8, 469 (2016).
22. CMS Collaboration, “Measurement of the  $Z\gamma \rightarrow \nu\bar{\nu}\gamma$  production cross section in pp collisions at  $\sqrt{s} = 8$  TeV and limits on anomalous  $ZZ\gamma$  and  $Z\gamma\gamma$  trilinear gauge boson couplings”, Phys. Lett. B **760**, 448 (2016).
23. CMS Collaboration, “Measurement of dijet azimuthal decorrelation in pp col-

- lisions at  $\sqrt{s} = 8$  TeV”, Eur. Phys. J. C **76**, no. 10, 536 (2016).
24. CMS Collaboration, “Forward–backward asymmetry of Drell–Yan lepton pairs in pp collisions at  $\sqrt{s} = 8$  TeV”, Eur. Phys. J. C **76**, no. 6, 325 (2016).
  25. CMS Collaboration, “Measurement of the inclusive jet cross section in pp collisions at  $\sqrt{s} = 2.76$  TeV”, Eur. Phys. J. C **76**, no. 5, 265 (2016).
  26. CMS Collaboration, “Measurement of the  $W^+W^-$  cross section in pp collisions at  $\sqrt{s} = 8$  TeV and limits on anomalous gauge couplings”, Eur. Phys. J. C **76**, no. 7, 401 (2016).
  27. CMS Collaboration, “Comparison of the  $Z/\gamma^*$  + jets to  $\gamma$  + jets cross sections in pp collisions at  $\sqrt{s} = 8$  TeV”, JHEP **1510**, 128 (2015), Erratum: [JHEP **1604**, 010 (2016)].
  28. CMS Collaboration, “Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in  $\ell^+\ell^-b\bar{b}$  final states”, Phys. Lett. B **748**, 221 (2015).
  29. CMS Collaboration, “Search for a Higgs boson in the mass range from 145 to 1000 GeV decaying to a pair of W or Z bosons”, JHEP **1510**, 144 (2015).
  30. CMS Collaboration, “Search for Third-Generation Scalar Leptoquarks in the  $t\tau$  Channel in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV”, JHEP **1507**, 042 (2015), Erratum: [JHEP **1611**, 056 (2016)].

*Selected*

*Reprint*



# CMS Physics Analysis Summary

---

Contact: cms-pag-conveners-smp@cern.ch

2017/02/21

Determination of the strong coupling constant from the measurement of inclusive multijet event cross sections in pp collisions at  $\sqrt{s} = 8$  TeV

The CMS Collaboration

## Abstract

A measurement of inclusive multijet event cross sections is presented from proton-proton collisions recorded at  $\sqrt{s} = 8$  TeV with the CMS detector and corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$ . Jets are reconstructed with the anti- $k_t$  clustering algorithm for a jet size parameter  $R = 0.7$  in a phase space region ranging up to jet transverse momenta  $p_T$  of 2.0 TeV and an absolute rapidity of  $|y| = 2.5$ . The inclusive 2-jet and 3-jet event cross sections are measured as a function of the average  $p_T$  of the two leading jets. The data are well described by predictions at next-to-leading order in perturbative quantum chromodynamics and additionally are compared to several Monte Carlo event generators. The strong coupling constant at the scale of the  $Z$  boson mass is inferred from a fit of the ratio of the 3-jet over 2-jet event cross section giving  $\alpha_s(M_Z) = 0.1150 \pm 0.0010(\text{exp}) \pm 0.0013(\text{PDF}) \pm 0.0015(\text{NP})^{+0.0050}_{-0.0000}(\text{scale})$ .

## 1 Introduction

Inelastic collisions of protons are viewed as interactions between their constituent partons, the (anti-)quarks and gluons. Within the context of perturbative quantum chromodynamics (pQCD), the cross section of a high- $p_T$  scattering process can be expressed as a sum of terms with increasing powers of the strong coupling constant,  $\alpha_s$ , convoluted with the parton momentum distribution functions (PDFs) of the proton. The lowest-order  $\alpha_s^2$  term represents the production of two-parton final states. Terms of higher-order  $\alpha_s^3, \dots$  in the expansion signify the existence of multi-parton final states. The theoretical description of the transition from strongly interacting, colored partons to color-neutral hadrons, which are observable in detectors, relies on this nonperturbative phase or models implemented in Monte-Carlo (MC) event generators. To relate the collimated sprays of colorless hadrons to the initiating partons, jet algorithms are applied. These algorithms bundle together particles that are close in phase space and primarily move into the same direction. Hence, the constructed jets preserve energy and momentum of the initial partons so that the structure of the final jet system mirrors, to a large extent, the topology of the initial partonic system.

The inclusive jet cross section,  $pp \rightarrow \text{jet} + X$ , as a function of jet  $p_T$  and rapidity  $y$  is a fundamental observable providing essential information about the PDFs and the strong coupling constant. Corresponding measurements conducted by the experiments at the CERN LHC are reported in Refs. [1–11]. The investigation of inclusive multijet event cross sections  $\sigma_{\text{inel}}$  by  $pp \rightarrow j\text{jets} + X$ , as suggested here, permits more elaborate tests of QCD to be performed by subdividing the observed jet event sample into classes according to the presumed minimal power in  $\alpha_s$  necessary to describe theoretically such a topology. Moreover, the ratios of such cross sections,  $R_{mn} = \frac{\sigma_m}{\sigma_n}$ , with  $m > n$ , are proportional to  $\alpha_s^{m-n}$  while at the same time numerous theoretical and experimental uncertainties cancel. Thus, they provide an ideal tool to determine the strong coupling constant  $\alpha_s(M_Z)$ . A previous analysis of the ratio  $R_{32}$  as a function of the average transverse momentum,  $\langle p_{T,2} \rangle$ , of the two leading  $p_T$  jets in the event was performed at  $\sqrt{s} = 7\text{ TeV}$  by the CMS Collaboration and lead 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 [12].

In this analysis, a measurement of inclusive 2- and 3-jet event cross sections is presented using an event sample collected by the CMS experiment during 2012 at the LHC and corresponding to an integrated luminosity of  $19.7\text{ fb}^{-1}$  of  $pp$  collisions at a centre-of-mass energy of  $8\text{ TeV}$ . Jets are reconstructed using the infrared- and collinear-safe anti- $k_t$  clustering algorithm [13] with a jet size parameter  $R$  of  $0.7$ . All jets are required to satisfy  $p_T > 150\text{ GeV}$  and  $|y| < 5.0$ . The event sample is further reduced by requiring the two leading  $p_T$  jets to lie in the central detector region of  $|y| < 2.5$ .

The event scale is chosen as before to be the average transverse momentum of the two leading jets, but will be referred to as  $H_{T2}/2$  in this analysis. Fits of the strong coupling constant are performed for the 2-jet and 3-jet event cross sections separately and for their ratio  $R_{32}$ .

## 2 Event selection and reconstruction

The measurement uses data samples which were collected with six single-jet high-level triggers (HLT) [14]. They are seeded by Level 1 (L1) triggers based on calorimetric information and require at least one jet in the event with corrected jet  $p_T > 80, 140, 200, 260$ , and  $320\text{ GeV}$ . All except the highest-threshold trigger were prescaled during the 2012 run. The efficiency of each trigger is estimated using lower- $p_T$ -threshold triggers and it is found to be more than 99% in

Table 1: Trigger regions defined as ranges of the  $H_{T2}/2$  for every single-jet trigger used in the inclusive multijet cross section measurement along with the effective integrated luminosities.

HLT path	$H_{T2}/2$ range (GeV)	Integrated Luminosity ( $\text{pb}^{-1}$ )
PFJet80	120 – 188	2.12
PFJet40	188 – 263	$5.57 \times 10$
PFJet200	263 – 345	$2.61 \times 10^2$
PFJet260	345 – 406	$1.06 \times 10^3$
PFJet320	406 – 5000	$1.97 \times 10^4$

In the CMS experiment, all particles are reconstructed and identified using a particle-flow (PF) algorithm, which combines the information from the individual subdetectors [15, 16]. The four vectors of particle candidates, reconstructed by the above technique, are used as input to the anti- $k_t$  jet-clustering algorithm. The clustering is performed within the FASTJET package [17] using four-momentum summation.

The reconstructed jets require additional energy corrections to account for residual nonuniformities and nonlinearities in the detector response. These jet energy corrections [18] are derived using simulated events, generated by PYTHIA 6.4 [19] with tune Z2\* [20] and processed through the CMS detector simulation based on GEANT4 [21], and in situ measurements with dijet, photon+jet, and Z+jet events. The jet energy corrections, which depend on the  $\eta$  (pseudorapidity) and  $p_T$  of the jet, are applied to the jet four-momentum vector as a multiplicative factor [18]. For a jet with a  $p_T$  of  $100\text{ GeV}$ , the typical correction is about 10%, and decreases with increasing  $p_T$ . An additional offset correction is applied to take into account the extra energy clustered into jets from additional proton-proton interactions within the same or neighbouring bunch crossings (in-time and out-of-time pileup) [18]. Pileup effects are important only for jets with low  $p_T$  and become negligible for jets with  $p_T > 200\text{ GeV}$ . The current measurement is therefore largely insensitive to pileup effects.

Each selected event is required to have at least one offline-reconstructed vertex [22] along the beam line that is within  $24\text{ cm}$  of the nominal interaction point. To suppress nonphysical jets, i.e. jets resulting from noise in the electromagnetic calorimeter (ECAL) and/or brass/scintillator hadron calorimeter (HCAL), tight identification criteria [23] are applied: each jet should contain at least two particles, one of which is a charged hadron, and the jet energy fraction carried by neutral hadrons and photons should be less than 90%. These criteria have an efficiency greater than 99% for physical jets. Jets not satisfying the tight identification requirements are discarded.

A sample of multijet events is selected which has two or more jets with transverse momentum greater than  $150\text{ GeV}$  and  $|y| < 5.0$  in the event. Events, in which the two leading  $p_T$  jets have  $|y| < 2.5$  are selected. Further jets are counted only if they lie within the same central rapidity range of  $|y| < 2.5$ .

In QCD pure jet events are balanced in  $p_T$  and thus exhibit a low level of missing transverse energy, which predominantly is caused by jet calibration and resolution effects of the detector. Therefore, the ratio of missing transverse energy to the total transverse energy  $\frac{E_T^{\text{miss}}}{\sum E_T}$ , both derived from the reconstructed particle-flow objects, is required to be less than 0.3 to select well measured jet events.

### 3 Measurement of the inclusive 2-jet and 3-jet event cross sections

The inclusive differential multijet cross sections are measured 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.

For inclusive 2-jet events sufficient data are available up to  $H_{\mathrm{T}2}/2 = 2 \text{ TeV}$ , while for inclusive 3-jet events (and the ratio  $R_{32}$ ) the accessible range in  $H_{\mathrm{T}2}/2$  is limited to  $H_{\mathrm{T}2}/2 < 1.68 \text{ TeV}$ . In the following, results for the inclusive 2-jet and 3-jet event selections will be labelled as  $n_j \geq 2$  and  $n_j \geq 3$ , respectively.

The inclusive differential jet event cross section is defined as :

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

where  $\epsilon$  is the product of the trigger and jet selection efficiencies, which are greater than 99%, in an  $H_{\mathrm{T}2}/2$  bin, and  $\Delta(H_{\mathrm{T}2}/2)$  are the bin widths. The measurements are reported in units of  $(\text{pb}/\text{GeV})$ .

To compare the measured cross sections with theoretical predictions at particle level, an unfolding procedure based on the iterative D'Agostini method [24] as implemented in the ROOTJUNFOLD software package [25] is applied. This correction for detector resolution effects is regularized by an early stopping after four iterations similarly as in a previously published 3-jet measurement [26], which prevents the buildup of large-scale correlations. The response matrix describes the mapping between the particle-level  $H_{\mathrm{T}2}/2$  spectrum and the reconstructed  $H_{\mathrm{T}2}/2$  spectrum. To construct the response matrix, the particle-level  $H_{\mathrm{T}2}/2$  spectrum is taken from a fit to the theoretically predicted  $H_{\mathrm{T}2}/2$  spectrum. The reconstructed  $H_{\mathrm{T}2}/2$  spectrum is obtained by smearing this particle-level prediction. The resolution in  $H_{\mathrm{T}2}/2$  is evaluated from CMS detector simulation based on the MADGRAPH5 + PYTHIA6 MC event generator using a jet-based smearing according to the jet energy resolution (JER). The JER from simulation is corrected (increased) for residual differences between data and simulation following Ref. [18].

Figure 1 shows the response matrices derived using a Toy MC procedure for inclusive 2-jet (left) and 3-jet events (right). The matrices are normalized to the number of generated events in each column and are mostly diagonal with small off-diagonal elements describing migrations between close-by  $H_{\mathrm{T}2}/2$  bins.

Through the unfolding procedure the final statistical uncertainties become correlated among bins. The size of these correlations varies typically between 10 and 20%. As a consequence, statistical fluctuations present in data before the unfolding might affect neighbouring bins after the unfolding. The fluctuations observed in the unfolded spectrum are compatible with the the statistical uncertainties of the underlying distributions and the statistical uncertainty after unfolding increases with respect to the original one of the measured data.

The unfolding procedure is affected by uncertainties of the JER. Alternative response matrices, which were built by varying the JER one standard deviation up and down [18], are used to unfold the measured spectra for comparison and introduce a corresponding uncertainty on the cross sections. In addition, to account for a model dependence of the theoretical  $H_{\mathrm{T}2}/2$  spectrum, two different functions are assumed when fitting the theoretically predicted  $H_{\mathrm{T}2}/2$  spectra. Finally, a supplementary uncertainty is attributed by comparison to an unfolding with a 30% reduced resolution as compared to the one extracted from simulation. This accounts for shortcomings in the detector simulation of the theory spectra leading to small nonclosures

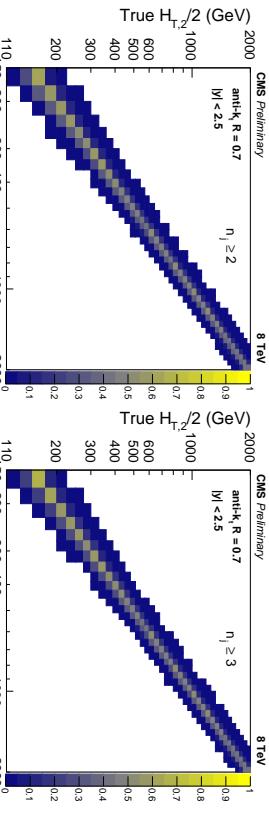


Figure 1: Response matrices derived using a Toy MC procedure for the inclusive 2-jet (left) and 3-jet event samples (right).

observed in the unfolding. All three uncertainties are added quadratically to give the unfolding uncertainty, which increases from about 1% at low  $H_{\mathrm{T}2}/2$  up to 2% at the high  $H_{\mathrm{T}2}/2$  end of the cross sections.

The dominant detector-related contribution to the experimental systematic uncertainty of the measured cross sections is caused by the jet energy corrections (JEC) [18]. The JEC uncertainty ranges for inclusive 2-jet events from 3% to 10% and for inclusive 3-jet events from 3% to 8%, respectively.

The uncertainty on the integrated luminosity, which propagates directly to the cross sections, is 2.6% [27] and at low  $H_{\mathrm{T}2}/2$  is of a similar size as the one from the JEC. To account for residual effects of small inefficiencies from triggering and jet identification, an uncorrelated uncertainty of 1% is assumed across all  $H_{\mathrm{T}2}/2$  bins, similar as in previous CMS jet cross-section measurements [8].

The total experimental systematic uncertainty on the measured cross section is obtained by summing in quadrature the single contributions. The upper panels of Fig. 2 give an overview of all experimental uncertainties affecting the cross section measurement for inclusive 2-jet (top left) and 3-jet events (top right). The error bars indicate the statistical uncertainty after unfolding. The colored lines represent the systematic uncertainties resulting from JEC, the luminosity, residual effects, and the unfolding including JER effects. The total experimental uncertainty, indicated by dashed black lines, is calculated by adding in quadrature all the sources of uncertainty. Beyond about 1.4 (1.2) TeV the statistical uncertainty is dominating for the 2-jet (3-jet) event cross sections, respectively.

The cross section ratio  $R_{32}$  as a function of  $H_{\mathrm{T}2}/2$  is extracted from the ratio of unfolded differential cross sections for each bin in  $H_{\mathrm{T}2}/2$ . The systematic experimental uncertainties are propagated from the cross sections to the ratio taking into account correlations. The uncertainties due to luminosity and residual effects cancel completely in this ratio. The statistical uncertainty including bin-by-bin correlations and statistical correlations between the 3-jet and 2-jet event cross sections is derived by directly unfolding the measured ratio  $R_{32}$ . Figure 2 "bottom panel" presents an overview of all experimental uncertainties affecting the cross section ratio  $R_{32}$ . The JEC and unfolding uncertainties for  $R_{32}$  amount to about 1.2% and  $\approx 1\%$ , respectively. The total uncertainty calculated by adding in quadrature all individual sources of uncertainty, is dominated by statistical effects beyond about 0.8 TeV in  $H_{\mathrm{T}2}/2$ .

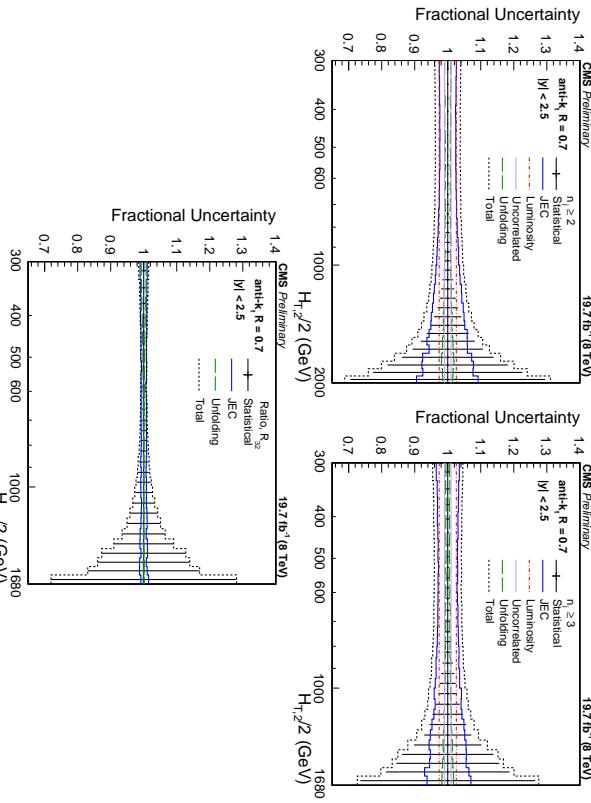


Figure 2: Overview of all experimental uncertainties affecting the inclusive 2-jet (top left) and 3-jet event cross sections (top right) and their ratio  $R_{32}$  (bottom). The error bars indicate the statistical uncertainty after unfolding. The colored lines represent the systematic uncertainties resulting from JEC, the luminosity, residual effects, and the unfolding including JER effects. Uncertainties due to luminosity and residual effects are cancelled completely in the ratio. The total experimental uncertainty, indicated by dashed black lines, is calculated by adding in quadrature all the sources of uncertainty.

Predictions at NLO accuracy in pQCD are computed with the NLOJET++ program version 4.1.3 [28, 29]. The results are provided within the framework of FASTNLO version 2.3 [30] for use within fits. The renormalization and factorization scales  $\mu_r$  and  $\mu_f$  are chosen equal to  $H_{T,2}/2$ . PDF sets at NLO available for a series of different assumptions on  $\alpha_s(M_Z)$  via the LHAPDF6 package [31] are listed in Table 2. All sets employ a variable-flavour number scheme with at most five or six flavours apart from the ABM11 PDFs, which use a fixed-flavour number scheme with  $N_F = 5$ .

Out of these eight PDF sets the following three will not be considered further:

- At NLO, predictions based on ABM11 do not describe LHC jet data at small jet rapidity, cf. Refs. [4, 5, 26, 32].
- The HERAPDf2.0 set exclusively fits HERA DIS data with only weak constraints on the gluon PDF.
- The range in values available for  $\alpha_s(M_Z)$  is too limited for the NNPDF3.0 set.

PDF uncertainties are evaluated according to the prescriptions given for each PDF set. Uncertainties on  $\alpha_s(M_Z)$  are not considered, since this value is later on determined from a fit to the data. The PDF uncertainty as derived with the CT10 PDF set ranges from 2% to 30% for inclusive 2- and 3-jet cross sections and from 2% to 7% for  $R_{32}$ .

Table 2: NLO PDF sets available via LHAPDF6 for comparisons to data with various assumptions on the value of  $\alpha_s(M_Z)$ . Sets existing already in LHC Run 1 (upper rows) and newer sets for Run 2 (lower rows) are listed together 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. A \* behind the  $\alpha_s(M_Z)$  values signifies that the parameter was fixed, not fitted.

Base set	Refs.	$N_F$	$M_t$ (GeV)	$M_Z$ (GeV)	$\alpha_s(M_Z)$	$\alpha_s(M_Z)$ range
ABM11	[33]	5	180	91.174	0.1180	0.110–0.130
CT10	[34]	≤5	172	91.188	0.1180*	0.112–0.127
MSTW2008	[35, 36]	≤5	10 <sup>10</sup>	91.1876	0.1202	0.110–0.130
NNPDF2.3	[37]	≤6	175	91.1876	0.1180*	0.114–0.124
CT14	[38]	≤5	172	91.1876	0.1180*	0.113–0.123
HERAPDf2.0	[39]	≤5	173	91.1876	0.1180*	0.110–0.130
MMHT2014	[40]	≤5	10 <sup>10</sup>	91.1876	0.1180*	0.108–0.128
NNPDF3.0	[41]	≤5	173	91.12	0.1180*	0.115–0.121

The uncertainty related to unknown higher orders of the perturbative series is evaluated with the conventional recipe of varying the default scale  $H_{T,2}/2$  chosen for  $\mu_r$  and  $\mu_f$  independently in the following six combinations:  $(\mu_r/H_{T,2}, \mu_f/H_{T,2}) = (1/2, 1/2), (1/2, 1), (1, 1/2), (1, 2), (2, 1)$  and  $(2, 2)$ . The maximal upwards and downwards deviations in cross section from the central prediction are taken as scale uncertainty. This uncertainty ranges for inclusive 2-jet events from 5% to 13%, for inclusive 3-jet events from 11% to 17% and for their ratio  $R_{32}$  from 6% to 8%.

The computation of the NLO predictions with NLOJET++ is also subject to statistical fluctuations from the numerical integrations. For the inclusive 2-jet event cross sections this uncer-

tainty is smaller than about one per mille, while for the inclusive 3-jet event cross section it amounts to 1–9 per mille.

Higher order effects of electroweak origin affect jet cross sections at large jet  $p_T$ . These electroweak (EWK) corrections have been calculated for the inclusive 1-jet and 2-jet case, cf. Ref. [42], but are not yet known for 3-jet production. Therefore, they are considered for the 2-jet events, while for the 3-jet event cross section and for the ratio they have been neglected.

The impact of nonperturbative (NP) effects, i.e. from multiple-parton interactions (MPI) and hadronization, are evaluated by using samples obtained from different MC event generators with a simulation of parton-shower and underlying-event (UE) contributions. The leading order (LO) MC event generators **HERWIG++** [43], with the default tune of version 2.3 and **PYTHIA6** [19] with tune  $Z2^*$ , are considered, and the dijet NLO prediction from **POWHEG** [44–46] interfaced to **PYTHIA8** with tune **CUEFTS1** [47] for full event generation. The cross section ratios between a nominal event generation and a sample without hadronization and MPI effects are taken as correction separately for inclusive 2-, and 3-jet events, and as their ratio for  $R_{32}$ . This ratio is fitted by a power-law function. The differences in the correction factors obtained from the various MC event generators are assigned as an uncertainty. The central correction factors  $C_{NP}$  are determined by the centre of the envelope which covers all predictions and half of the spread is taken as the uncertainty.

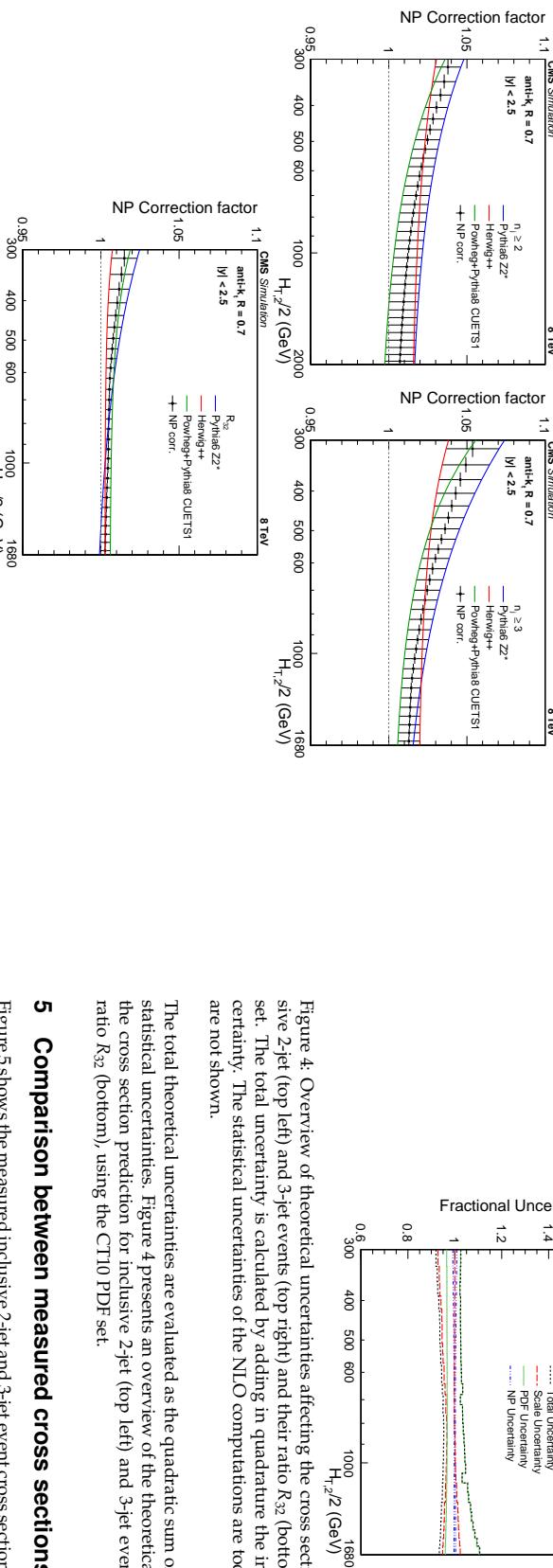


Figure 3: Fits to the nonperturbative corrections obtained for inclusive 2-jet (top left) and 3-jet (top right) event cross sections and their ratio  $R_{32}$  (bottom) as a function of  $H_{T2}/2$  within  $|y| < 2.5$  for the three investigated MC event generators.

The NP corrections are shown in Fig. 3 for the inclusive 2-jet (top left) and 3-jet event cross

sections (top right) as well for  $R_{32}$  (bottom). They amount to  $\approx 4\text{--}5\%$  for inclusive 2-jet and 3-jet events and  $\approx 1\%$  for  $R_{32}$  at  $H_{T2}/2 \approx 300\text{ GeV}$  and decrease for increasing  $H_{T2}/2$ . The uncertainty assigned to the NP corrections is of the order of 1–2%. The non-perturbative effects are reduced in the cross section ratio.

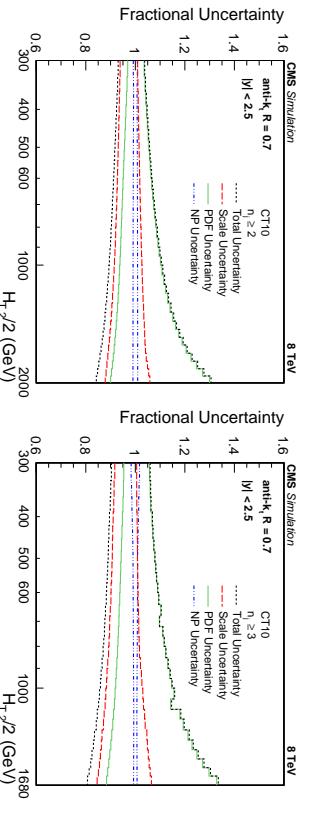


Figure 4: Overview of theoretical uncertainties affecting the cross section prediction for inclusive 2-jet (top left) and 3-jet events (top right) and their ratio  $R_{32}$  (bottom), using the CT10 PDF set. The total uncertainty is calculated by adding in quadrature the individual sources of uncertainty. The statistical uncertainties of the NLO computations are too small to be visible and are not shown.

The total theoretical uncertainties are evaluated as the quadratic sum of the scale, PDF, NP, and statistical uncertainties. Figure 4 presents an overview of the theoretical uncertainties affecting the cross section prediction for inclusive 2-jet (top left) and 3-jet events (top right) and their ratio  $R_{32}$  (bottom), using the CT10 PDF set.

## 5 Comparison between measured cross sections and theory

Figure 5 shows the measured inclusive 2-jet and 3-jet event cross sections as a function of  $H_{T2}/2$  after unfolding for detector effects. On the left, the measurements are compared to the NLO-JET++ predictions using the CT10 PDF set, corrected for NP effects and in addition for EWK effects in the 2-jet case. On the right, the comparison is made to the predictions from MADGRAPH5 + PYTHIA6 with tune  $Z2^*$  ( $MG+Pythia6 Z2^*$ ), corrected for EWK effects in the 2-jet case. On a logarithmic scale, the data are in agreement with the NLO predictions over the whole

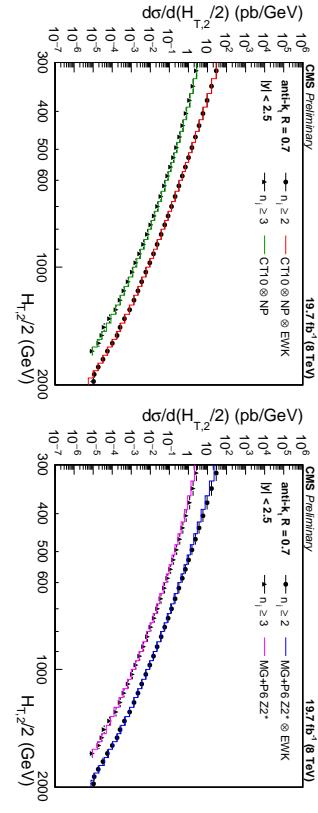


Figure 5: Comparison of the inclusive 2-jet and 3-jet event cross sections as a function of  $H_{T2}/2$  to theoretical predictions. On the (left), the data (points) are shown together with NLO JET++ predictions (line) using the CT10 PDF set, corrected for NP and EWK (2-jet) or only NP effects (3-jet). On the (right), the data (points) are compared to predictions from MADGRAPH5+PYTHIA6 with tune ZZZ\* (line), corrected for EWK effects in the 2-jet case. The error bars correspond to the total uncertainty, for which the statistical and systematic uncertainties are added in quadrature.

range of  $H_{T2}/2$  from 300 GeV up to 2.0 (2-jet) and 1.68 TeV (3-jet) respectively.

For better visibility the ratios of data over the NLO JET++ predictions using the CT10 PDF set are shown in Fig. 6. The data are well described by the predictions within their uncertainty, which is dominated at large  $H_{T2}/2$  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  $H_{T2}/2$ , is remedied by the inclusion of EWK corrections. In the 3-jet case the statistical precision of the data and the reach in  $H_{T2}/2$  is insufficient to observe any effect. The alternative PDF sets MSTW2008 and NNPDF2.3 exhibit a small underestimation of the cross sections at high  $H_{T2}/2$ .

As for the NP corrections, the POWHEG framework providing a NLO dijet calculation matched to the parton showers of PYTHIA8 is used for a comparison. Here, POWHEG + PYTHIA8 are employed with the CUET5I and CUETM1 tunes. The ratios of data over theory from POWHEG + PYTHIA8 with tune CUET5I are shown in Fig. 7. For comparison the LO prediction from PYTHIA6 with tune ZZZ\* , the tree-level multi-leg improved prediction by MADGRAPH5 + PYTHIA6 with tune ZZZ\*, and the matched NLO prediction from POWHEG + PYTHIA8 with tune CUETM1 are shown as well. 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 ZZZ\*, in particular for small  $H_{T2}/2$ . In contrast, the employed dijetMC PYTHIA8 and POWHEG + PYTHIA8 better describe the 2-jet event cross section, but fail for the 3-jet case.

The cross section ratio  $R_{32}$  as a function of  $H_{T2}/2$  is extracted from the data by dividing the differential cross sections for each bin in  $H_{T2}/2$ . Figure 8 presents this ratio as obtained from unfolded data in comparison to that from NLO pQCD. The error bars correspond to the total experimental uncertainty.

For a better comparison of the behaviour of the 2- and 3-jet event cross sections and their ratio Figs. 9–11 present the respective ratios with respect to theory for varying assumptions on PDFs and  $\alpha_s(M_Z)$ . A small slope increasing with  $H_{T2}/2$  is visible for most PDFs in both cross

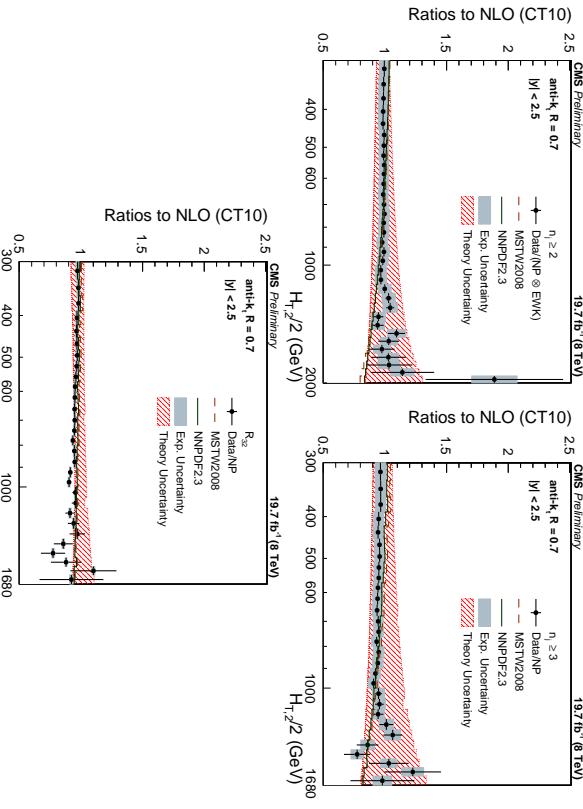


Figure 6: Ratio of data over theory using the CT10 PDF set for inclusive 2-jet (top left) and inclusive 3-jet event cross sections (top right) and their ratio  $R_{32}$  (bottom). For comparison predictions employing two other PDF sets are also shown. The error bars correspond to the statistical uncertainty of the data and the shaded rectangles to the total experimental systematic uncertainty. The shaded band around unity represents the total uncertainty of the theory.

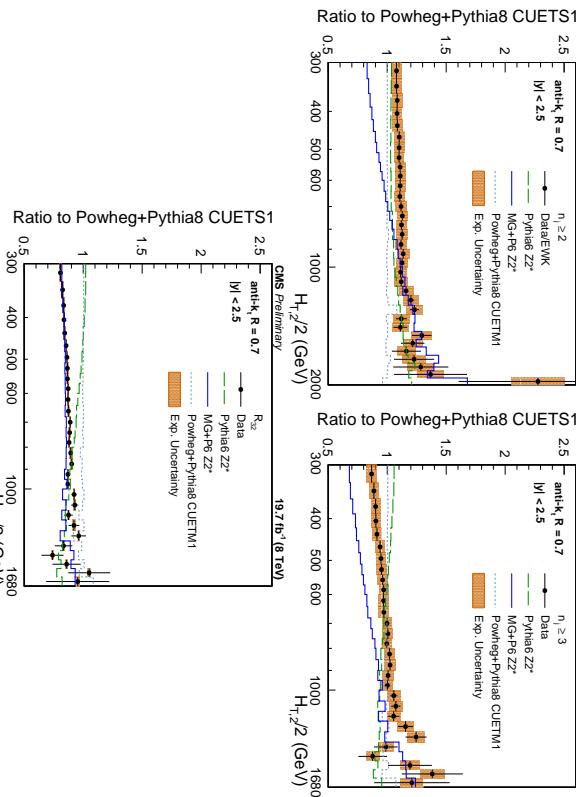


Figure 7: Ratio of data over the prediction from POWHEG + PYTHIA8 with tune CUETs1. For comparison the alternative tune CUETM1 of powHEG + PYTHIA8, the tree-level multi-leg improved prediction by MADGRAPH5 + PYTHIA8 with tune ZZ\*, and the LOMC predictions from PYTHIA6 tune ZZ\* are shown as well. The error bars correspond to the statistical uncertainty of the data and the shaded rectangles to the total experimental systematic uncertainty. EWK corrections have been accounted for in this comparison in the 2-jet case.

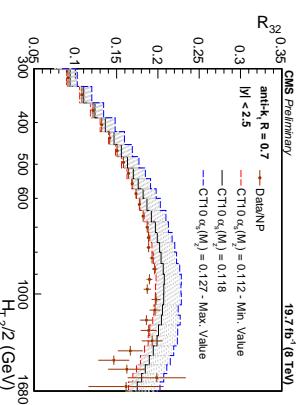


Figure 8: Cross section ratio  $R_{32}$  as a function of  $H_{T2}/2$  calculated from data (solid circles) in comparison to that from NLO pQCD (lines). The error bars correspond to the total experimental uncertainty derived as quadratic sum from all uncertainty sources. The NLO predictions using the CT10 NLO PDF set corrected with NP corrections are shown for a series of values assumed for  $\alpha_s(M_z)$  (dashed lines) together with the central prediction (solid line) where  $\alpha_s(M_z) = 0.118$ . The assumption on  $\alpha_s(M_z)$  is varied in steps of 0.001 in the range of 0.112–0.127. For brevity the relative factor of NP between data and theory has been indicated as “Data/NP” in the legend.

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  $H_{T2}/2$  up to 1.68 TeV. Moreover, the different sensitivity to  $\alpha_s(M_z)$  caused by the leading power in  $\alpha_s$  in the expansion of the 2-jet inclusive ( $\propto \alpha_s^2$ ) and the 3-jet inclusive cross section ( $\propto \alpha_s^3$ ), 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 when passing through Figs. 9–11. This also demonstrates the potential of ratios  $R_{mn}$  with  $m - n > 1$ .

## 6 Fits of the strong coupling constant

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  $\alpha_s(M_z)$ . The value of  $\alpha_s(M_z)$  is determined by minimizing the  $\chi^2$  between the experimental measurement and the theoretical predictions. The fit procedure here follows closely the one previously used in Refs. [12] and [48]. The  $\chi^2$  is defined as:

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

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

$$M^j = D^j - T^j \quad (3)$$

and  $C$  is the covariance matrix including all experimental uncertainties as described in Section 3 and some theoretical uncertainties. More precisely,  $C = C_{\text{Exp}} + C_{\text{Theo}}$  is defined as the sum of

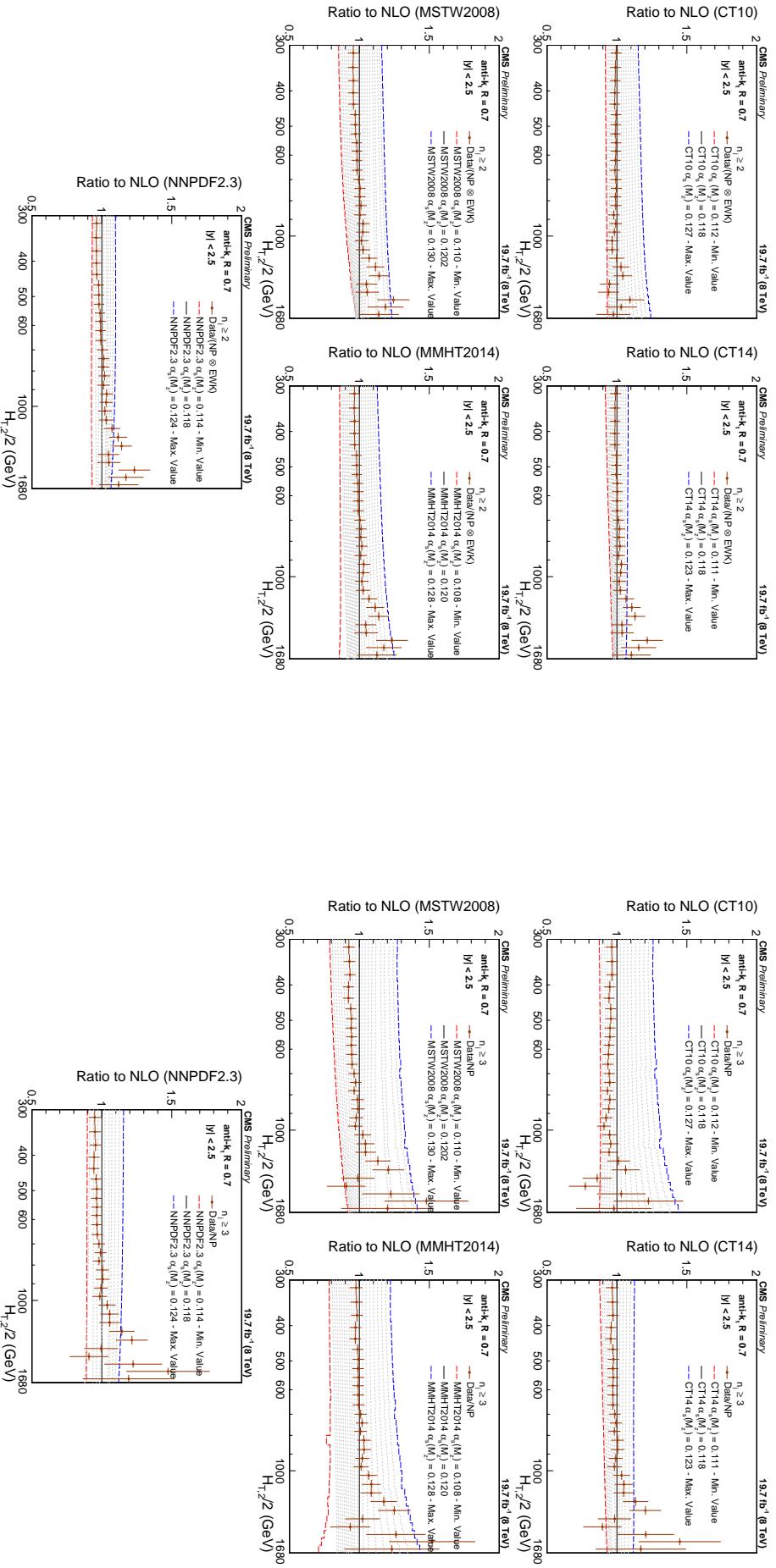


Figure 9: Ratio of measured 2-jet inclusive event cross section (data points) over NLO theory times NP corrections for various PDF sets at their respective default value for  $\alpha_s(M_Z)$  (black solid line at unity). The error bars correspond to the total experimental uncertainty. The NLO predictions have been derived with the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and the NNPDF2.3 PDF sets (bottom) for the series of assumptions on  $\alpha_s(M_Z)$  available for the respective PDF set as specified in Table 2. For brevity, the relative factor of NP between data and theory has been indicated as “Data/NP” in the legend.

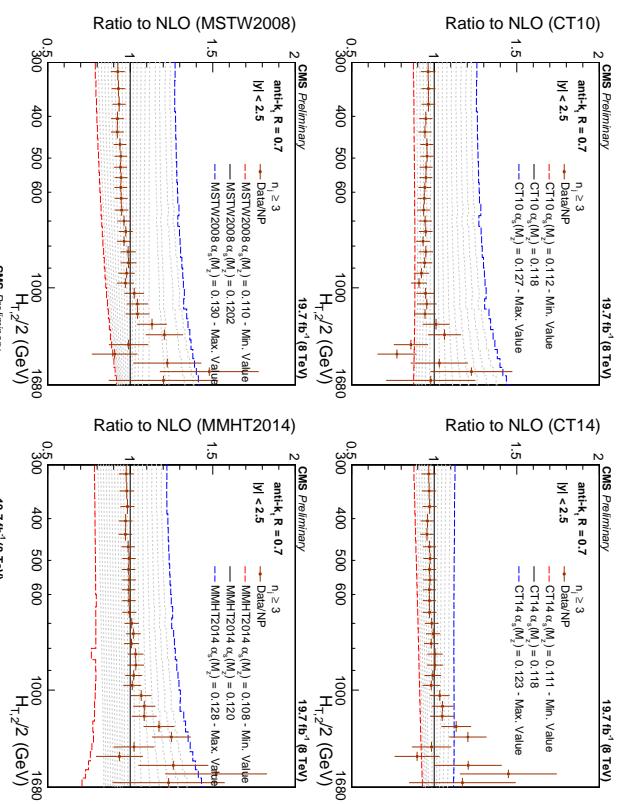


Figure 10: Ratio of measured 3-jet inclusive event cross section (data points) over NLO theory times NP corrections for various PDF sets at their respective default value for  $\alpha_s(M_Z)$  (black solid line at unity). The error bars correspond to the total experimental uncertainty. The NLO predictions have been derived with the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and the NNPDF2.3 PDF sets (bottom) for the series of assumptions on  $\alpha_s(M_Z)$  available for the respective PDF set as specified in Table 2. For brevity, the relative factor of NP between data and theory has been indicated as “Data/NP” in the legend.

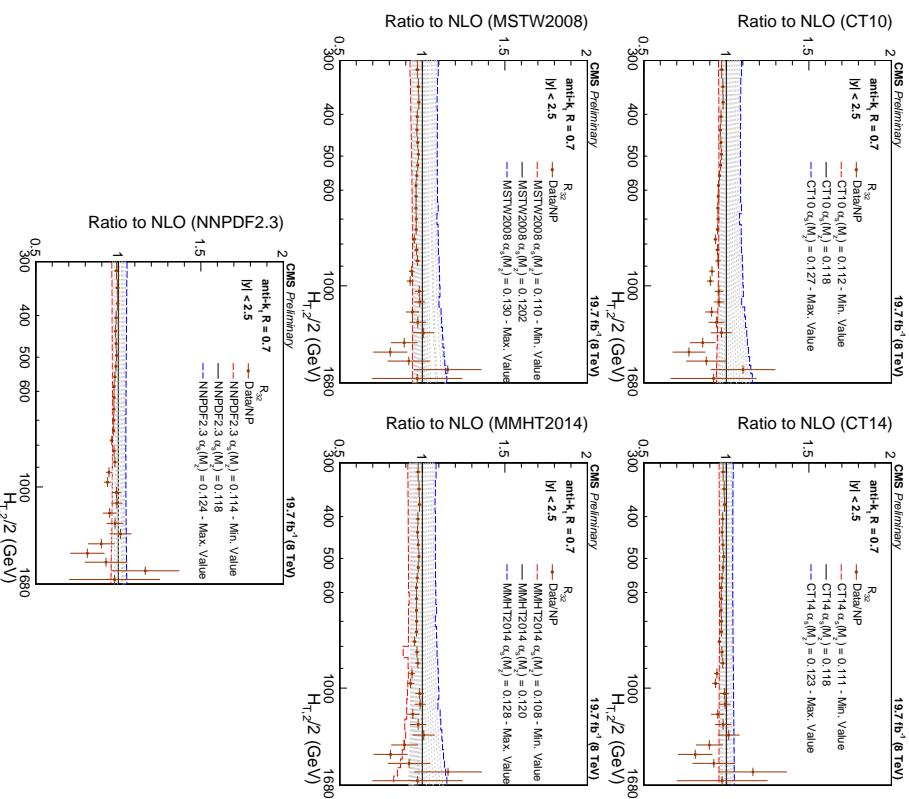


Figure 11: Ratio of measured  $R_{32}$  ratio (data points) over NLO theory times NP corrections for various PDF sets at their respective default value for  $\alpha_s(M_Z)$  (black solid line at unity). The error bars correspond to the total experimental uncertainty. The NLO predictions have been derived with the CT10 (top left), the CT14 (top right), the MSTW2008 (middle left), the MMHT2014 (middle right) and the NNPDF2.3 PDF sets (bottom) for the series of assumptions on  $\alpha_s(M_Z)$  available for the respective PDF set as specified in Table 2. For brevity, the relative factor of NP between data and theory has been indicated as “Data/NP” in the legend.

covariances of experimental and theoretical sources of uncertainty as follows

$$C_{\text{exp}} = \text{Cov}^{\text{ExpStat}} + \sum \text{Cov}^{\text{JEC}} + \text{Cov}^{\text{Unfolding}} + \text{Cov}^{\text{Lumi}} + \text{Cov}^{\text{Uncor}}, \quad (4)$$

$$C_{\text{theo}} = \text{Cov}^{\text{TheoStat}} + \text{Cov}^{\text{NP}} + \text{Cov}^{\text{PDF}}, \quad (5)$$

where the labelled covariance matrices account for the following effects:

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

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 JEC, unfolding, and luminosity uncertainties are treated as multiplicative to avoid the statistical bias that arises when estimating uncertainties from data.

The derivation of PDF uncertainties depends on each PDF set. The CT10 PDF set consists of  $N_{\text{ev}} = 26$  eigenvectors with two PDF members per eigenvector  $k$ , which lead to the predictions  $S_k^\pm$  that follow from PDF variations with respect to the plus and minus directions of eigenvector  $k$ . Symmetric uncertainties as required by the use of covariance matrices are then computed by [49]:

$$(\Delta X)^2 = \frac{1}{4} \sum_{k=1}^{N_{\text{ev}}} [X(S_k^+) - X(S_k^-)]^2, \quad (6)$$

where  $\Delta X$  is the uncertainty of the cross section and  $X(S_k^\pm)$  is the predicted cross section for each eigenvector orientation,  $+$  or  $-$ .

Scale uncertainties of the pQCD predictions are taken into account employing the offset method, i.e. by performing separate fits with varying scale factors as described in the previous section. 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. 4. 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  $\alpha_s(M_Z)$ .

First, fits to the cross sections are performed, where the range in  $H_T/2$  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  $p_T$ , which are available for

the dijet case only. The results are reported in Table 3 for the 2-jet and 3-jet event cross sections. For comparison, a simultaneous fit to both cross sections ignoring any correlations, and a fit to their ratio fully accounting for correlations are given in Table 4. Also, EWK 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; for the ratio  $R_{32}$  somewhat smaller values are obtained. A common issue, except for the ratio fits, is the rather small  $\chi^2/n_{\text{dof}}$ . 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  $\chi^2/n_{\text{dof}}$  values lie around unity while the  $\alpha_s(M_Z)$  values are still compatible with the previous results but with slightly reduced uncertainties.

Table 3: 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  $H_{\text{T2}}/2$  bins in the restricted range of  $0.3 < H_{\text{T2}}/2 < 1.0 \text{ TeV}$ .

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

Table 4: 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 (19)  $H_{\text{T2}}/2$  bins in the restricted range of  $0.3 < H_{\text{T2}}/2 < 1.0 \text{ TeV}$ . 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	2- & 3-jets		$R_{32}$		$\chi^2/n_{\text{dof}}$
	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	$\alpha_s(M_Z)$	$\pm \Delta \alpha_s(M_Z)$	
CT10	0.1170	0.0026	8.2/37	0.1141	0.0028
CT14	0.1161	0.0029	9.1/37	0.1139	0.0032
MSTW2008	0.1161	0.0021	11./37	0.1150	0.0023
MMHT2014	0.1168	0.0025	11./37	0.1142	0.0022
NNPDF2.3	0.1188	0.0019	15./37	0.1184	0.0021
					12./18

To investigate how the EWK corrections affect the fit results for  $\alpha_s(M_Z)$ , the range in  $H_{\text{T2}}/2$  is extended to  $0.3 < H_{\text{T2}}/2 < 1.68 \text{ TeV}$ . Table 5 reports the values obtained for  $\alpha_s(M_Z)$  from fits to the 2-jet event cross section in this range with or without EWK correction factors. The largest impact is a reduction in  $\chi^2/n_{\text{dof}}$ , which indicates a better agreement when EWK effects are included. In addition, a tendency to slightly smaller  $\alpha_s(M_Z)$  values is observed without the EWK corrections. For the ratio  $R_{32}$  it is expected that these effects are much reduced. From Fig. 11 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

Table 5: Determination of  $\alpha_s(M_Z)$  from the inclusive 2-jet event cross section using five PDF sets at NLO with (right) and without (left) EWK corrections. Only total uncertainties without scale variations are quoted. The results are obtained from a simultaneous fit to all 29  $H_{\text{T2}}/2$  bins in the range of  $0.3 < H_{\text{T2}}/2 < 1.68 \text{ TeV}$ .

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

$Q = H_{\text{T2}}/2$ . Tables 6–8 give the complete results for MSTW2008 and MMHT2014 for the full range in  $H_{\text{T2}}/2$  of 300 GeV up to 1.68 TeV, for scale variations in this range, and for subranges in  $H_{\text{T2}}/2$ . Using the MSTW2008 PDF set, which dates from before the LHC start, the strong coupling constant finally is determined to

$$\alpha_s(M_Z) = 0.1150 \pm 0.0010(\text{exp}) \pm 0.0013(\text{PDF}) \pm 0.0015(\text{NP})^{+0.0050}_{-0.0000}(\text{scale})^{+0.0030}_{-0.0000}(\text{scale}).$$

The MMHT2014 PDF set, although using LHC jet data to determine the PDF parameters, leads to a very similar result of

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

In contrast to fits at NLO using cross sections, where the scale uncertainty recipe usually leads to a very asymmetric behaviour with the larger uncertainty towards smaller values of  $\alpha_s(M_Z)$ , this is inverted for the fits to the cross section ratio.

Table 9 provides in addition to the extracted  $\alpha_s(M_Z)$  value for each range in  $H_{\text{T2}}/2$  the  $\alpha_s(Q)$  values with total uncertainty as evolved to the respective cross-section averaged scale  $\langle Q \rangle$  in that range. The evolution is performed for five flavours at 2-loop order with the RUNDEC program [50, 51]. The obtained  $\alpha_s(Q)$  points are illustrated in Fig. 12 together with the world average [52] and results from other measurements of the CMS [11, 12, 26, 48, 53], ATLAS [54], D0 [55, 56], H1 [57, 58], and ZEUS [59] experiments.

## 7 Summary

A measurement of the inclusive 2-jet (3-jet) event cross sections has been presented in a range of  $0.3 < H_{\text{T2}}/2 < 1.68 \text{ TeV}$  ( $0.3 < H_{\text{T2}}/2 < 1.68 \text{ TeV}$ ) for the average  $p_T$  of the two leading jets at central rapidity of  $|y| < 2.5$ . The data sample has been collected from proton-proton collisions at 8 TeV centre-of-mass energy and corresponds to an integrated luminosity of  $19.7 \text{ fb}^{-1}$ . The data are found to be well described and corresponds to calculations at NLO in pQCD completed with NP corrections that are important at low  $H_{\text{T2}}/2$ . The upwards trend seen in the 2- and 3-jet data at high  $H_{\text{T2}}/2$  in comparison to the prediction at NLO QCD, is explained by the onset of EWK

Table 6: Determination of  $\alpha_s(M_Z)$  from the ratio  $R_{32}$  using the two most compatible PDF sets MSTW2008 and MMHT2014 at NLO. The results are obtained from a simultaneous fit to all 29  $H_{T2}/2$  bins in the full range of  $0.3 < H_{T2}/2 < 1.68 \text{ TeV}$ .

PDF set	$R_{32}: \Delta\alpha_s(M_Z) \times 1000$						
	$\alpha_s(M_Z)$	exp	PDF	NP	all exc. scale	scale	$\chi^2/n_{\text{dof}}$
MSTW2008	0.1150	$\pm 10$	$\pm 13$	$\pm 15$	$\pm 23$	$^{+50}_{-49}$	26./28
MMHT2014	0.1142	$\pm 10$	$\pm 13$	$\pm 14$	$\pm 22$	$^{+50}_{-49}$	24./28

Table 7: Fitted values of  $\alpha_s(M_Z)$  using  $R_{32}$  in the  $H_{T2}/2$  range from 0.3 up to 1.68 TeV at the central scale and for the six scale factor combinations for the two PDF sets MSTW2008 and MMHT2014.

$\mu_f/H_{T2}/2$	$\mu_f/H_{T2}/2$	MSTW2008		MMHT2014	
		$\alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$	$\alpha_s(M_Z)$	$\chi^2/n_{\text{dof}}$
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 8: Uncertainty composition for  $\alpha_s(M_Z)$  from the determination of  $\alpha_s$  from the jet event rate  $R_{32}$  in bins of  $H_{T2}/2$ . 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  $H_{T2}/2$ , are assumed to cancel between the numerator and denominator.

$H_{T2}/2$	MSTW2008: $\Delta\alpha_s(M_Z) \times 1000$				MMHT2014: $\Delta\alpha_s(M_Z) \times 1000$				
	$\alpha_s(M_Z)$	exp	PDF	NP	$\alpha_s(M_Z)$	exp	PDF	NP	
(GeV)									
300–420	0.1157	$\pm 15$	$\pm 14$	$\pm 19$	$^{+53}_{-50}$	0.1158	$\pm 14$	$\pm 10$	$^{+52}_{-50}$
420–600	0.1153	$\pm 11$	$\pm 14$	$\pm 18$	$^{+50}_{-50}$	0.1154	$\pm 11$	$\pm 12$	$^{+56}_{-50}$
600–1000	0.1134	$\pm 13$	$\pm 16$	$\pm 19$	$^{+53}_{-50}$	0.1140	$\pm 12$	$\pm 18$	$^{+45}_{-50}$
1000–1680	0.1147	$\pm 29$	$\pm 17$	$\pm 18$	$^{+63}_{-11}$	0.1154	$\pm 25$	$\pm 14$	$^{+50}_{-11}$
300–1680	0.1150	$\pm 10$	$\pm 13$	$\pm 15$	$^{+50}_{-6}$	0.1142	$\pm 10$	$\pm 13$	$^{+49}_{-6}$

Table 9: Evolution of the strong coupling constant between the scale of the  $Z$  boson mass and the cross-section averaged  $H_{T2}/2$  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 [50, 51].

$H_{T2}/2$	$\langle Q \rangle$	$\alpha_s(M_Z)$	$\alpha_s(Q)$	No. of data points	$\chi^2/n_{\text{dof}}$
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

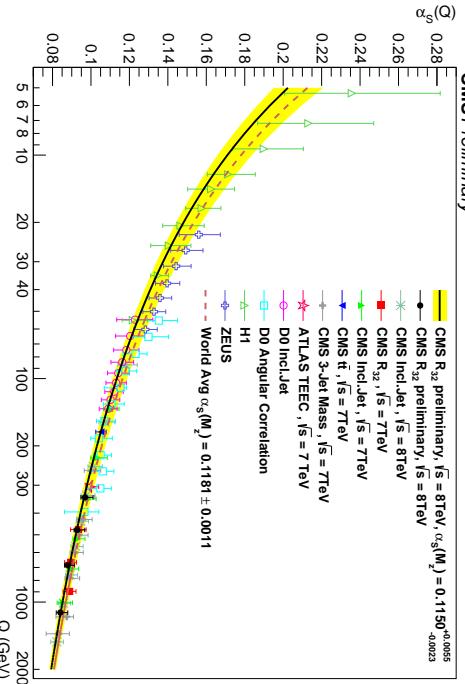


Figure 12: The running  $\alpha_s(Q)$  as a function of the scale  $Q$  is shown as obtained by using the MSTW2008 NLO PDF set. The solid line and the uncertainty band are drawn by evolving the extracted  $\alpha_s(M_Z)$  values using the 2-loop 5-flavour renormalization group equations as implemented in RUNDEC [50, 51]. The dashed line represents the evolution of the world average [52] and the black circles correspond to the  $\alpha_s(Q)$  determinations presented in Table 9. Results from other measurements of CMS [11, 12, 26, 48, 53], ATLAS [54], D0 [55, 56], H1 [57, 58], and ZEUS [59] are superimposed.

corrections in the 2-jet case. For the 3-jet event cross section these correction have not yet been computed.

In the 3-jet to 2-jet cross section ratio the EWK corrections are assumed to cancel. In fact, NLO QCD provides an adequate description of  $R_{32}$  in the accessible range of  $H_{12}/2$ . In contrast, LO tree-level MC predictions exhibit significant deviations.

Based on the observed agreement, the strong coupling constant is determined in a fit to the  $R_{32}$  measurement to

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

using the MSTW2008 PDF set. Employing the MMHT2014 PDF set instead leads to very similar results. Equally compatible determinations of  $\alpha_s(M_Z)$  are achieved with separate fits to the inclusive 2-jet and 3-jet event cross sections employing various PDF sets provided the range in  $H_{12}/2$  is restricted to  $0.3 < H_{12}/2 < 1.0 \text{ TeV}$ . The result for  $\alpha_s(M_Z)$  is in agreement with previous determinations obtained by the ATLAS and CMS collaborations [11, 12, 26, 48, 53, 54] and with the world average value of  $\alpha_s(M_Z) = 0.1181 \pm 0.0011$  derived in Ref. [52].

- 
- ## References
- [1] ALICE Collaboration, "Measurement of the inclusive differential jet cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 2.76 \text{ TeV}$ ", *Phys. Lett. B* **722** (2013) 262, doi:10.1016/j.physletb.2013.04.026, arXiv:1301.3475.
  - [2] ATLAS Collaboration, "Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector", *Eur. Phys. J. C* **71** (2011) 1512, doi:10.1140/epjc/s10052-010-1512-2, arXiv:1009.5908.
  - [3] ATLAS Collaboration, "Measurement of inclusive jet and dijet production in  $p\bar{p}$  collisions at  $\sqrt{s} = 7 \text{ TeV}$  using the ATLAS detector", *Phys. Rev. D* **86** (2012) 014022, doi:10.1103/PhysRevD.86.014022, arXiv:1112.6297.
  - [4] ATLAS Collaboration, "Measurement of the inclusive jet cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 2.76 \text{ TeV}$  and comparison to the inclusive jet cross section at  $\sqrt{s} = 7 \text{ TeV}$  using the ATLAS detector", *Eur. Phys. J. C* **73** (2013) 2509, doi:10.1140/epjc/s10052-013-2509-4, arXiv:1304.4739.
  - [5] ATLAS Collaboration, "Measurement of the inclusive jet cross-section in proton-proton collisions at  $\sqrt{s} = 7 \text{ TeV}$  using 4.5  $\text{fb}^{-1}$  of data with the ATLAS detector", *JHEP* **02** (2015) 153, doi:10.1007/JHEP02(2015)153, arXiv:1410.8857.
  - [6] CMS Collaboration, "Measurement of the Inclusive Jet Cross Section in  $p\bar{p}$  Collisions at  $\sqrt{s} = 7 \text{ TeV}"$ , *Phys. Rev. Lett.* **107** (2011) 132001, doi:10.1103/PhysRevLett.107.132001, arXiv:1106.0208.
  - [7] CMS Collaboration, "Measurement of the inclusive production cross sections for forward jets and for dijet events with one forward and one central jet in  $p\bar{p}$  collisions at  $\sqrt{s} = 7 \text{ TeV}"$ , *JHEP* **06** (2012) 036, doi:10.1007/JHEP06(2012)036, arXiv:1202.0704.
  - [8] CMS Collaboration, "Measurements of differential jet cross sections in proton-proton collisions at  $\sqrt{s} = 7 \text{ TeV}$  with the CMS detector", *Phys. Rev. D* **87** (2013) 112002, doi:10.1103/PhysRevD.87.112002, arXiv:1212.6660.
  - [9] CMS Collaboration, "Measurement of the ratio of inclusive jet cross sections using the anti- $k_T$  algorithm with radius parameters  $R = 0.5$  and  $0.7$  in  $p\bar{p}$  collisions at  $\sqrt{s} = 7 \text{ TeV}"$ , *Phys. Rev. D* **90** (2014) 072006, doi:10.1103/PhysRevD.90.072006, arXiv:1406.0324.
  - [10] CMS Collaboration, "Measurement of the double-differential inclusive jet cross section in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}"$ , *Eur. Phys. J. C* **76** (2016), no. 8, 451, doi:10.1140/epjc/s10052-016-4286-3, arXiv:1605.04436.
  - [11] CMS Collaboration, "Measurement and QCD analysis of double-differential inclusive jet cross-sections in  $p\bar{p}$  collisions at  $\sqrt{s} = 8 \text{ TeV}$  and ratios to 2.76 and 7 TeV", (2016), arXiv:1609.05331. Submitted to *JHEP*.
  - [12] CMS Collaboration, "Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 7 \text{ TeV}$  and first determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **73** (2013) 2604, doi:10.1140/epjc/s10052-013-2604-6, arXiv:1304.7498.

- [13] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [14] CMS Trigger and Data Acquisition Group Collaboration, “The CMS high level trigger”, *Eur. Phys. J. C* **46** (2006) 605, doi:10.1140/epjc/s2006-02495-8, arXiv:hep-ex/0512077.
- [15] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, technical report, CERN, 2009.
- [16] CMS Collaboration, “Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV”, technical report, CERN, 2010.
- [17] M. Cacciari, G. P. Salam, and G. Soyez, “Fastjet User Manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [18] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, (2016), arXiv:1607.03663. Submitted to *JINST*.
- [19] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [20] CMS Collaboration, “Charged particle multiplicities in  $p\bar{p}$  interactions at  $\sqrt{s} = 0.9, 2.36$ , and 7 TeV”, *JHEP* **01** (2011) 079, doi:10.1007/JHEP01(2011)079, arXiv:1011.5531.
- [21] S. Agostinelli et al., “GEANT4: A Simulation toolkit”, *Nuclear Instruments & Methods in Physics Research A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [22] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014) no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [23] CMS Collaboration, “Jet Performance in pp Collisions at  $\sqrt{s} = 7$  TeV”, technical report, CERN, 2010.
- [24] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem”, *Nucl. Instrum. Meth. A* **362** (1995) 487, doi:10.1016/0168-9002(95)00274-X.
- [25] T. Adye, “Unfolding algorithms and tests using RooUnfold”, in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, p. 313. Geneva, Switzerland, January 17-20, 2011. arXiv:1105.1160. doi:10.5170/CERN-2011-006.
- [26] CMS Collaboration, “Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range”, *Eur. Phys. J. C* **75** (2015) 186, doi:10.1140/epjc/s10052-015-3376-y, arXiv:1412.1633.
- [27] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, technical report, CERN, 2013.
- [28] Z. Nagy, “Three jet cross-sections in hadron hadron collisions at next-to-leading order”, *Phys. Rev. Lett.* **88** (2002) 122003, doi:10.1103/PhysRevLett.88.122003, arXiv:hep-ph/0110315.
- [29] Z. Nagy, “Next-to-leading order calculation of three-jet observables in hadron-hadron collisions”, *Phys. Rev. D* **68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268.
- [30] D. Britzger, K. Rabbertz, F. Stöber, and M. Wobisch, “New features in version 2 of the fastNLO project”, in *Proceedings, XX. International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012)*, p. 217. Bonn, Germany, March 26-30, 2012. arXiv:1208.3641. doi:10.3204/DESY-PROC-2012-02/165.
- [31] A. Buckley et al., “LHAPDF6: parton density access in the LHC precision era”, *Eur. Phys. J.* **C75** (2015) 132, doi:10.1140/epjc/s10052-015-3318-8, arXiv:1412.7420.
- [32] CMS Collaboration, “Measurement of the inclusive jet cross section in pp collisions at  $\sqrt{s} = 2.76$  TeV”, *Eur. Phys. J. C* **76** (2016) 265, doi:10.1140/epjc/s10052-016-4083-z, arXiv:1512.06212.
- [33] S. Alekhin, J. Blümlein, and S. Moch, “Parton Distribution Functions and Benchmark Cross Sections at NNLO”, *Phys. Rev. D* **86** (2012) 054009, doi:10.1103/PhysRevD.86.054009, arXiv:1202.2281.
- [34] H.-L. Lai et al., “New parton distributions for collider physics”, *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.
- [35] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC”, *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.
- [36] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Uncertainties on  $\alpha_S$  in global PDF analyses and implications for predicted hadronic cross sections”, *Eur. Phys. J. C* **64** (2009) 653, doi:10.1140/epjc/s10052-009-1164-2, arXiv:0905.3531.
- [37] R. D. Ball et al., “Parton distributions with LHC data”, *Nucl. Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.
- [38] S. Dulat et al., “New parton distribution functions from a global analysis of quantum chromodynamics”, *Phys. Rev. D* **93** (2016) 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.
- [39] H1 ZEUS Collaboration, “Combination of measurements of inclusive deep inelastic  $e^{\pm}p$  scattering cross sections and QCD analysis of HERA data”, *Eur. Phys. J. C* **75** (2015) 580, doi:10.1140/epjc/s10052-015-3710-4, arXiv:1506.06042.
- [40] L. Harland-Lang, A. Martin, P. Motylinski, and R. Thorne, “Parton distributions in the LHC era: MMHT 2014 PDFs”, *Eur. Phys. J. C* **75** (2015) no. 5, 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.
- [41] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [42] S. Dittmaier, A. Huss, and C. Speckner, “Weak radiative corrections to dijet production at hadron colliders”, *JHEP* **11** (2012) 095, doi:10.1007/JHEP11(2012)095, arXiv:1210.0438.
- [43] M. Bär et al., “Herwig++ Physics and Manual”, *Eur. Phys. J. C* **58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

- [44] P. Nason, "A New method for combining NLO QCD with shower Monte-Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [45] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [46] S. Alioli et al., "Jet pair production in POWHEG", *JHEP* **04** (2011) 081, doi:10.1007/JHEP04(2011)081, arXiv:1012.3380.
- [47] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2015) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [48] CMS Collaboration, "Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at  $\sqrt{s} = 7$  TeV", *Eur. Phys. J. C* **75** (2015) 288, doi:10.1140/epjc/s10052-015-3499-1, arXiv:1410.6765.
- [49] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.
- [50] K. G. Chetyrkin, J. H. Kuhn, and M. Steinhauser, "RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses", *Comput. Phys. Commun.* **133** (2000) 43, doi:10.1016/S0010-4655(00)00155-7, arXiv:hep-ph/0004189.
- [51] B. Schmidt and M. Steinhauser, "CRunDec: a C++ package for running and decoupling of the strong coupling and quark masses", *Comput. Phys. Commun.* **183** (2012) 1845, doi:10.1016/j.cpc.2012.03.023, arXiv:1201.6149.
- [52] C. Patrignani and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **40** (2016) 100001, doi:10.1088/1674-1137/40/10/100001.
- [53] CMS Collaboration, "Determination of the top-quark pole mass and strong coupling constant from the  $t\bar{t}$  production cross section in pp collisions at  $\sqrt{s} = 7$  TeV", *Phys. Lett. B* **728** (2014) 496, doi:10.1016/j.physletb.2013.12.009, arXiv:1307.1907.
- [54] ATLAS Collaboration, "Measurement of transverse energy-energy correlations in multi-jet events in  $p\bar{p}$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector and determination of the strong coupling constant  $a_s(m_Z)$ ", *Phys. Lett. B* **750** (2015) 427, doi:10.1016/j.physletb.2015.09.050, arXiv:1508.01579.
- [55] D0 Collaboration, "Determination of the strong coupling constant from the inclusive jet cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV", *Phys. Rev. D* **80** (2009) 111107, doi:10.1103/PhysRevD.80.111107, arXiv:0911.2710.
- [56] D0 Collaboration, "Measurement of angular correlations of jets at  $\sqrt{s} = 1.96$  TeV and determination of the strong coupling at high momentum transfers", *Phys. Lett. B* **718** (2012) 56, doi:10.1016/j.physletb.2012.10.003, arXiv:1207.4957.
- [57] H1 Collaboration, "Measurement of multijet production in  $e p$  collisions at high  $Q^2$  and determination of the strong coupling  $a_s$ ", *Eur. Phys. J. C* **75** (2015) 65, doi:10.1140/epjc/s10052-014-3223-6, arXiv:1406.4709.
- [58] H1 Collaboration, "Measurement of jet Production Cross Sections in Deep-inelastic  $e p$  Scattering at HERA", (2016) arXiv:1611.03421. Submitted to *Eur. Phys. J. C*.
- [59] ZEUS Collaboration, "Inclusive-jet photoproduction at HERA and determination of  $a_s$ ", *Nucl. Phys. B* **864** (2012) 1, doi:10.1016/j.nuclphysb.2012.06.006, arXiv:1205.6153.