

Chapter 1

¹² Measurement of the Differential Inclusive Multijet Cross Sections and their Ratio

The inclusive n -jet event samples include the events with number of jets $\geq n$, where $n = 2$ and 3 in the current study. The differential inclusive multijet cross sections 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, and are defined by :

$$\frac{d\sigma}{d(H_{T,2}/2)} = \frac{1}{\epsilon \mathcal{L}_{int,eff}} \frac{N_{event}}{\Delta(H_{T,2}/2)} \quad (1.1)$$

where N_{event} is the number of inclusive 2- or 3-jet events counted in an $H_{T,2}/2$ bin, ϵ is the product of the trigger and jet selection efficiencies, which are greater than 99%, $\mathcal{L}_{int,eff}$ is the effective integrated luminosity, and $\Delta(H_{T,2}/2)$ are the bin widths. The measurements are reported in units of (pb/GeV).

The cross section ratio R_{32} , defined in Eq. 1.2 is obtained by dividing the differential cross sections of inclusive 3-jet events to that of inclusive 2-jet one, for

⁴⁷
each bin in $H_{T,2}/2$.

$$R_{32} = \frac{\frac{d\sigma_{3-jet}}{d(H_{T,2}/2)}}{\frac{d\sigma_{2-jet}}{d(H_{T,2}/2)}} \quad (1.2)$$

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

1.1 Data Samples

³⁰
This measurement uses the data which was collected at the center of mass energy of 8 TeV by 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 run B-D, the JetMon stream datasets contain prescaled low trigger threshold paths (HLTPFJet40, 80, 140, 200 and 260) while the JetHT stream datasets 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 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 file¹ is based on the final event reconstruction of the 2012 CMS data sets. The datasets used in the current study are mentioned in the Table 1.1 along with the luminosity of each dataset.

The data sets have the LHC luminosity increasing with period, full data sample of 2012 corresponds to an integrated luminosity of 19.71 fb^{-1} .

³⁹
¹Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12.JSON

Table 1.1: Run range and luminosity of the proton-proton collisions data collected²⁴ at the center of mass energy of 8 TeV by CMS experiment in the year of 2012 in four different run periods A, B, C and D.

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

1.1.1 Monte Carlo Samples

To have a comparison of data results with the simulated events, the MADGRAPH5 [1]

Monte-Carlo (MC) event generator has been used. The MADGRAPH5 generates matrix elements for High Energy Physics processes, such as decays and $2 \rightarrow n$ scatterings. It has been interfaced to PYTHIA6 [2] by the LHE event record [3], to generate the rest of the higher-order effects using the Parton Showering (PS) model, with tune Z2* to model the underlying event. Matching algorithms make sure that there is no double-counting between the tree-level and the PS-model-generated partons.

The MC samples are processed through the complete CMS detector simulation to allow studies of the detector response and compare 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 increasing 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 are tabulated in Table 1.2 along with their cross sections and number of events generated.

Table 1.2: 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 motorization of the events. The cross section and number of events generated are mentioned for each sample.

Generator	Sample	Events	Cross Section pb
MADGRAPH5 + PYTHIA 6	/QCD_HT-100To250_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	50129518	1.036×10^7
	/QCD_HT-250To500_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	27062078	2.760×10^5
	/QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	30599292	8.426×10^3
	/QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	13843863	2.040×10^2

1.2 Event Selection

The events are selected according to several quality criteria which ensures 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 fake calorimeter energy deposits.

1.2.1 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 data. It consists of one L1 trigger seed and multiple HLT filters. The L1 jet trigger uses transverse energy sums computed 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) used for the current study are tabulated in Table 1.3. 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 1.3: The single jet trigger paths used in the analysis are listed here. 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 divided by the total integrated luminosity of 19.71 fb^{-1} , gives the effective prescale applied on a trigger over the whole run period.

Trigger Path	L1 threshold GeV	HLT threshold GeV	$H_{T,2}/2$, 99% GeV	Eff. Lumi fb^{-1}
HLT_PFJet80	36	80	120.0	0.21×10^{-2}
HLT_PFJet140	68	140	187.5	0.56×10^{-1}
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. 1.3 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.

$$\epsilon_{\text{HLT_PFJetY}} = \frac{H_{T,2}/2 \left(\text{HLT_PFJetX} + (\text{L1Object_}p_T > Z) + (\text{HLTOBJECT_}p_T > Y) \right)}{H_{T,2}/2(\text{HLT_PFJetX})} \quad (1.3)$$

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 ³ 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 ¹ which the trigger efficiency is larger than 99% is determined by fitting the turn-on ³⁸ distribution with a sigmoid function described in Eq. 1.4. The trigger turn-on curves ¹⁰ as a function of $H_{T,2}/2$ can be seen in Fig. 1.1 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 [4].

$$f_{fit}(x) = \frac{1}{2} \left(1 + erf\left(\frac{x - \mu}{\sqrt{2}\sigma}\right) \right) \quad (1.4)$$

1.2.2 Primary Vertex Selection

¹⁵ The reconstructed tracks, selected based on the transverse impact parameter significance with respect to the beam line, number of strip and pixel hits and the normalized track χ^2 , identify ³ 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 ¹ 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 be not be greater than 2 cm. The number of degrees of freedom ⁴⁶ in fitting for the position of each vertex using its associated tracks should be at-least four in number.

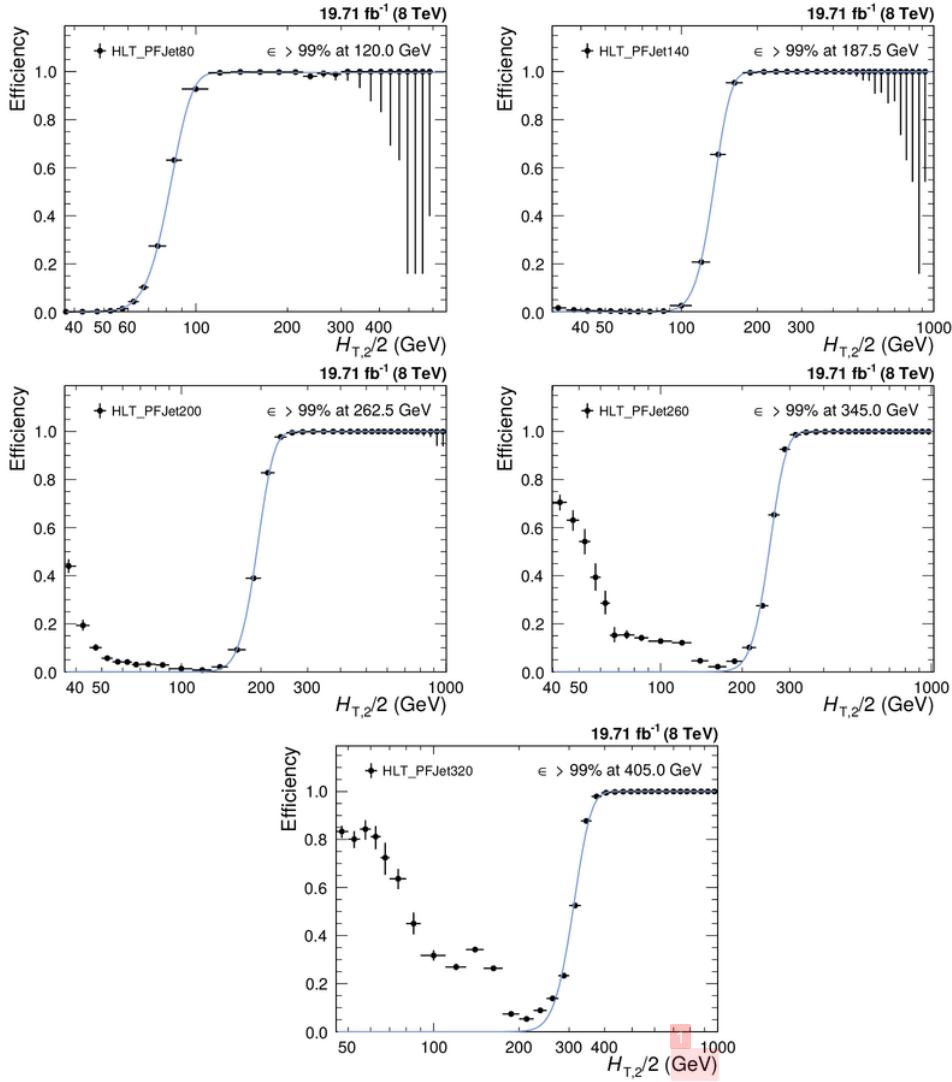


Figure 1.1: Trigger efficiencies turn-on curves for the single jet trigger paths are fitted with a sigmoid function (blue line) to obtain the 99% efficiency threshold. The error bars give the uncertainty on the efficiency which is calculated using Clopper-Pearson confidence intervals [4].

1.2.3 Missing Transverse Energy

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

The ratio of missing transverse energy to the total transverse energy
 $E_T^{\text{miss}}/\sum E_T$, shown in Fig. 1.2 for $n_j \geq 2$ (left) and $n_j \geq 3$ events (right), shows a
discrepancy between 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 data. Such
events are absent in QCD simulated events in MC. Hence $E_T^{\text{miss}}/\sum E_T$ is required
to be less than 0.3 to reject events with high E_T^{miss} .

1.2.4 Jet Identification

In order to suppress fake jets, arising from detector noise or misreconstructed particles, jet identification criteria (ID) has been applied. Instead of applying it event-wise, it is applied it on each jet. The algorithm works on reconstructed jets using information of the clustered particle candidates. The official tight jet ID [5], recommended by JETMET group [6] 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 selec-

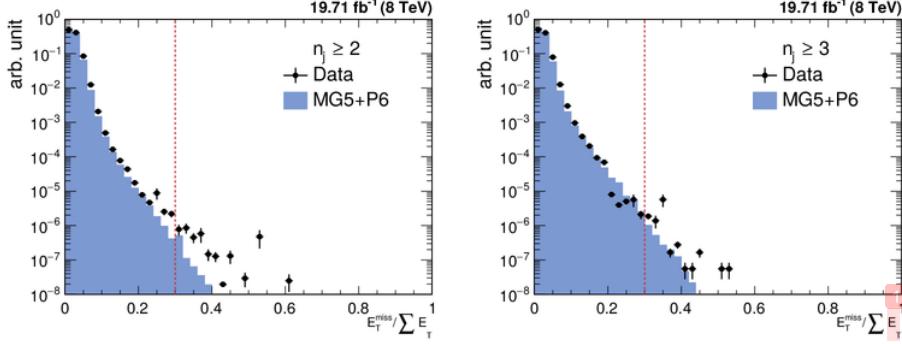


Figure 1.2: Missing transverse energy fraction of the total transverse energy per event in data (black solid circles) and simulated Monte Carlo events (blue histogram) in inclusive 2-jet (left) and 3-jet events (right). To remove background and noise, events with a fraction exceeding a certain threshold, here indicated with the red dashed line, are rejected.

tion criteria are implemented as selection cut on jet fractions. Table 1.4 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. 1.3 and 1.4 show the distributions of the jet constituents observed in data (black solid circles) and simulated MC events (blue histogram) for $n_j \geq 2$ and $n_j \geq 3$, respectively.

1.2.4.1 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

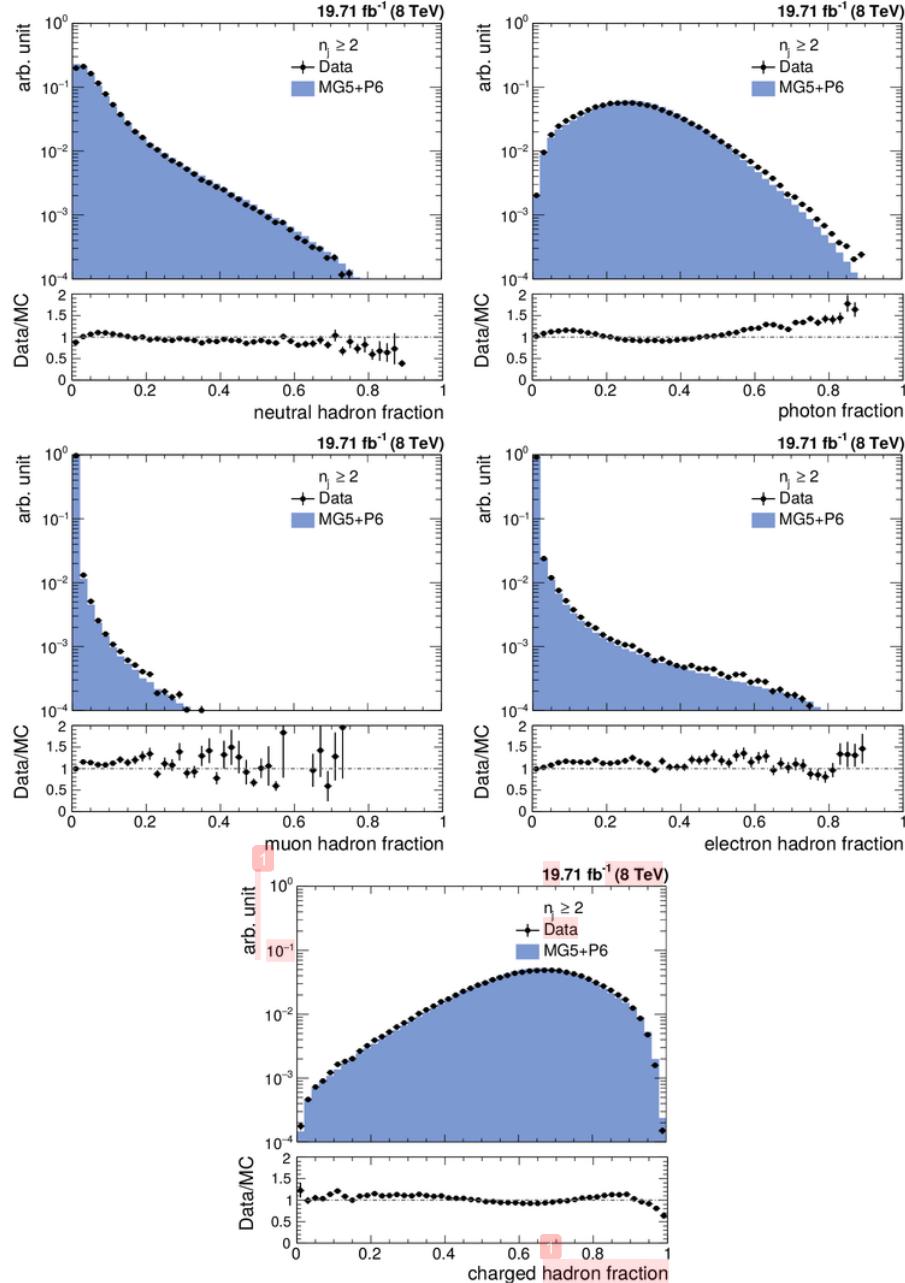


Figure 1.3: The fractions of jet constituents as observed in data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 2-jet events. Data and simulations are normalized to the same number of events. The distributions are shown after the application of the jet ID.

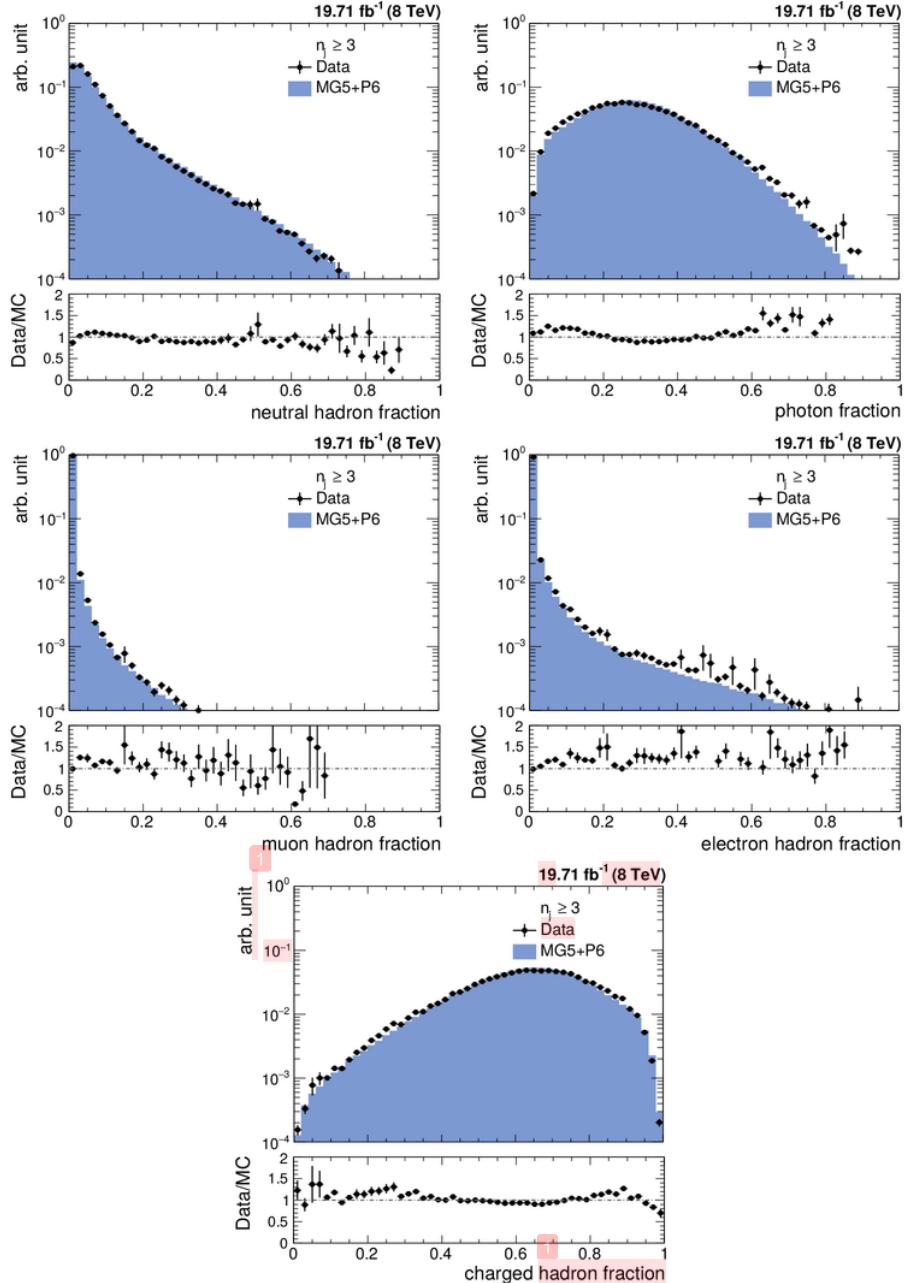


Figure 1.4: The fractions of jet constituents as observed in data (black solid circles) and simulated Monte Carlo events (blue histogram) for different types of PF candidates for inclusive 3-jet events. Data and simulations are normalized to the same number of events. The distributions are shown after the application of the jet ID.

Table 1.4: The jet ID removes noise and fake jets based on the properties of the reconstructed jets and the clustered particle candidates. All the selection cuts which are recommended by the JETMET group are applied [6].

	Property	Loose ID	Tight ID
Whole η region	neutral hadron fraction	< 0.99	< 0.90
	neutral EM fraction	< 0.99	< 0.90
	number of constituents	> 1	> 1
	muon fraction	< 0.80	< 0.80
only $ \eta < 2.4$	charged hadron fraction	> 0	> 0
	charged multiplicity	> 0	> 0
	charged EM fraction	< 0.99	< 0.90

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

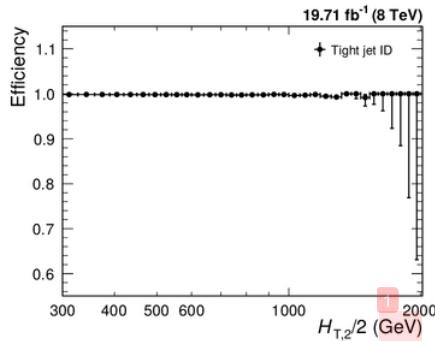


Figure 1.5: The jet ID efficiency is studied as a function of $H_{T,2}/2$ with tag-and-probe technique using dijet event topologies and it always exceeds 99%.

1.2.5 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. 17. These corrections are applied to jets in both data² as well those in simulated events³. The jet selection, based on phase space cuts on transverse momentum and rapidity of jets in an event, is as follows :

- All jets having $p_T > 150$ 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 $|y| < 2.5$.

These cuts assure high detector acceptance and exactly same selection is applied in the measurement, simulated events as well in theoretical calculations for a consistent comparison.

1.3 Comparison with Simulation

1.3.1 Pile-up 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 pile-up events implemented in the simulation $N_{\text{MC}}(N_{\text{PU,truth}})$, does not match exactly with the one measured actually in data $N_{\text{data}}(N_{\text{PU,est.}})$. To match the pile-up distributions in data, a reweighting factor w_{PU} , as given by Eq. 1.5 is applied

²Winter14_V8 jet energy corrections

³START53_V27 jet energy corrections

to the simulated events. In Fig. 1.6 the number of reconstructed vertices are shown before (left) and after pile-up reweighting (right). It is observed that before pile-up reweighting there was a significant mismatch of the pile-up distributions in data (black solid circles) and simulated MC events (blue histogram), which completely vanishes after reweighting.

$$w_{\text{PU}} = \frac{N_{\text{data}}(N_{\text{PU},\text{est.}}) / \sum N_{\text{data}}}{N_{\text{MC}}(N_{\text{PU},\text{truth}}) / \sum N_{\text{MC}}} \quad (1.5)$$

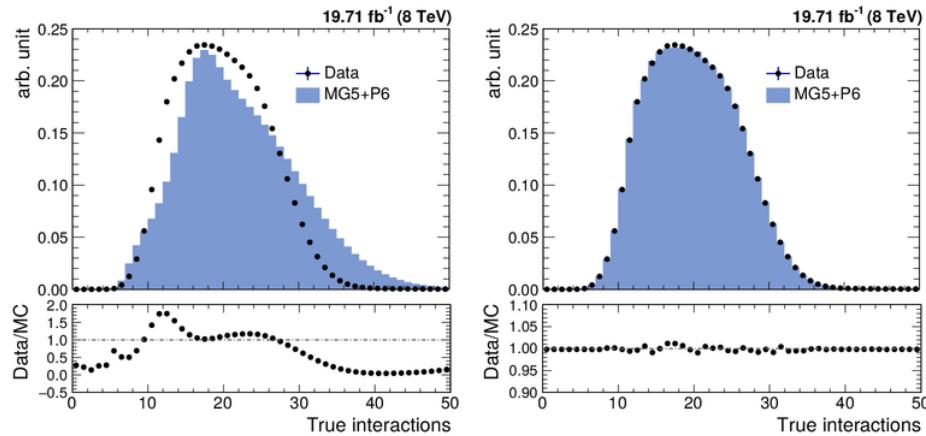


Figure 1.6: Number of reconstructed vertices in data (black solid circles) and simulated Monte Carlo events (blue histogram) before (left) and after (right) the pile-up reweighting.

1.3.2 Comparison of Cross Sections and their Ratio

The measured data distribution of differential cross section at 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 1.7 shows the comparison of differential cross section as a function of $H_{\text{T},2}/2$ for $n_j \geq 2$ (left) and $n_j \geq 3$ events (right), for 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 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 describes 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. As mentioned in Sec. 1 that the minimum kinematical cut on $H_{T,2}/2$ is 300 GeV. Hence the differential cross sections are studied in the range $300 \text{ GeV} \leq H_{T,2}/2 < 2000 \text{ GeV}$ for $n_j \geq 2$ and $300 \text{ GeV} \leq H_{T,2}/2 < 1680 \text{ GeV}$ for $n_j \geq 3$ events.

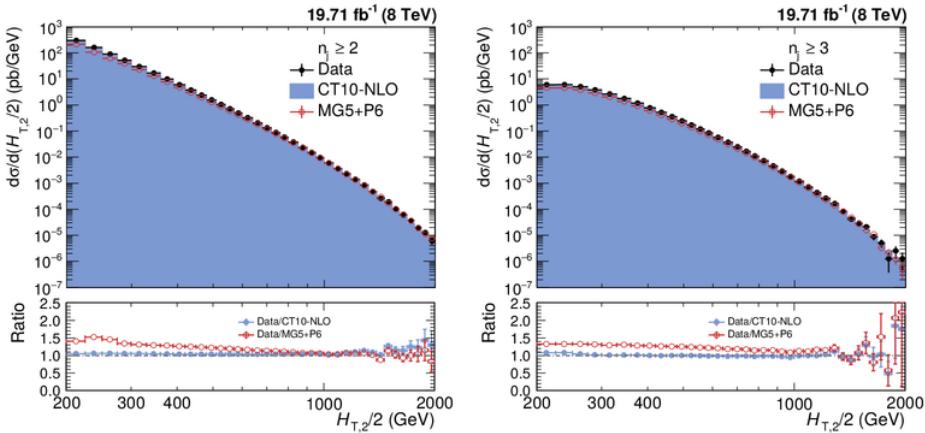


Figure 1.7: The differential cross sections are compared for data (black solid circles) and LO MADGRAPH5 + PYTHIA6 (MG5+P6) Monte Carlo (red empty circles), at reconstructed level with CT10-NLO theory predictions (blue histogram), as a function of $H_{T,2}/2$ for inclusive 2-jet (left) and 3-jet events (right). Ratios of 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.

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

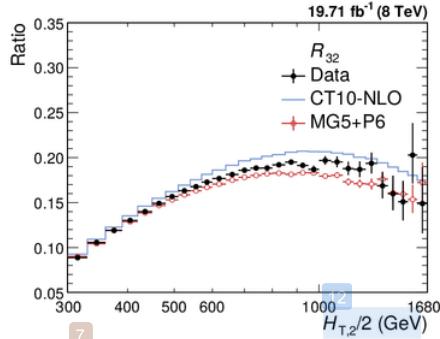


Figure 1.8: Comparison of the cross section ratio R_{32} as a function of $H_{T,2}/2$, for data (black solid circles) and LO MADGRAPH5 + PYTHIA6 (MG5+P6) Monte Carlo (red empty circles), at reconstructed level with CT10-NLO theory predictions (blue line). The error bars give the asymmetrical statistical uncertainty, calculated by the Wilson score interval method which takes into the account the correlation between the numerator and denominator.

statistical uncertainty 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 give asymmetric errors. Figure 1.8 shows the comparison of the cross section ratio R_{32} as a function of $H_{T,2}/2$, for 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 $n_j \geq 3$ events, the enough statistics for differential cross section is available upto 1680 GeV of $H_{T,2}/2$ only, R_{32} is also studied in the range $300 \text{ GeV} \leq H_{T,2}/2 < 1680 \text{ GeV}$. 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.

1.4 Jet Energy Resolution (JER)

In an ideal experiment, the value of a physical quantity would be determined exactly with an infinite precision. For e.g. whenever a particle with energy E passes 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 ³ might differ from the value E. This difference of ³ 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 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 which is centered around the true value of the measured quantity ³ and its width is generally interpreted as detector resolution. Hence the importance of the detector resolution ⁴² lies in the fact that it indicates how much the measured value ³ 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 the more efficient is the detector.

Due to finite resolution of the CMS detector, the measured transverse momentum of jets gets smeared. Since the observable in this study i.e. $H_{T,2}/2$ is the average sum of transverse momentum of leading and sub-leading jets, the resolution of the detector has to be studied in terms of the observable. 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 ²⁶ that the jet energy resolution in data is actually worse than in simulation [7]. So the reconstructed jet transverse momentum needs to be smeared additionally ¹ to match the resolution in data. Table 1.5 shows the scaling factors (c) which need to be applied on the transverse momentum of simulated reconstructed jets. The scaling factors depend on the absolute η of the jet. The uncertainty on these measured scaling factors ($c_{central}$) ² 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 1.5: JETMET working group at CMS has shown that the jet energy resolution in data is actually worse than in simulation [7]. To match the resolution in data, the reconstructed jet transverse momentum in simulated events need to be smeared by applying the scale factors. The uncertainty on the resolution is given by an upwards and downwards variation c_{up} and c_{down} of the measured scaling factor $c_{central}$.

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

The reconstructed jet p_T is smeared randomly using a Gaussian width widened by the scaling factor ($c_{central}$)

$$p_T \rightarrow \text{Gauss} \left(\mu = p_T, \sigma = \sqrt{c_{central}^2 - 1} \cdot \text{JER}(p_T) \right) \quad (1.6)$$

where $\text{JER}(p_T)$ is the resolution determined as a function of jet p_T using MG5+P6 MC simulated events. After smearing transverse momentum of each reco 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. 1.7.

$$R = \frac{\text{Reco } H_{T,2}/2}{\text{Gen } H_{T,2}/2} \quad (1.7)$$

²⁵ 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 the σ of a Gaussian fit of the response distribution. A double-sided Crystal-Ball function takes into account the non-Gaussian tails of the jet response distribution. The resolution as a function of $H_{T,2}/2$ is calculated separately for both $n_j \geq 2$ and $n_j \geq 3$ events. A

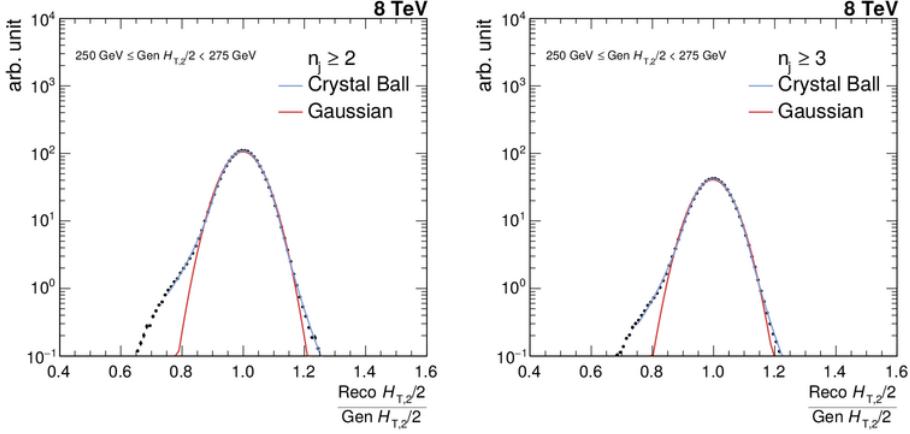


Figure 1.9: Fitting of the resolution distribution, obtained using LO MADGRAPH5 + PYTHIA6 (MG5+P6) Monte Carlo simulated events, as a function of $H_{T,2}/2$ for inclusive 2-jet (left) and inclusive 3-jet events (right). The blue line shows the double-sided Crystal Ball function fit of $\frac{\text{Reco } H_{T,2}/2}{\text{Gen } H_{T,2}/2}$ in each Gen $H_{T,2}/2$ bin, overlaid by Gaussian fitting the core of the resolution (red line).

fit example for one Gen $H_{T,2}/2$ bin is shown in Fig. 1.9 for $n_j \geq 2$ (left) and inclusive 3-jet events (right). 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. 1.10 that the Crystal Ball function (blue solid circles) describes the measured distributions better as compared to Gaussian function fit (red solid circles), especially in the low- $H_{T,2}/2$ region where the non-Gaussian tails are more pronounced. Hence JER is determined using Crystal Ball function fit. Figure 1.11 shows the final relative jet energy resolution (JER) which is described by a modified version of the NSC formula (blue solid line) [8], as mentioned in Equation 1.8. To consider the migration to lower as well higher bins and to obtain the resolution with reasonable statistics over the full range of Gen $H_{T,2}/2$, the fit function is extrapolated to 80 GeV and upto 2000 GeV which is

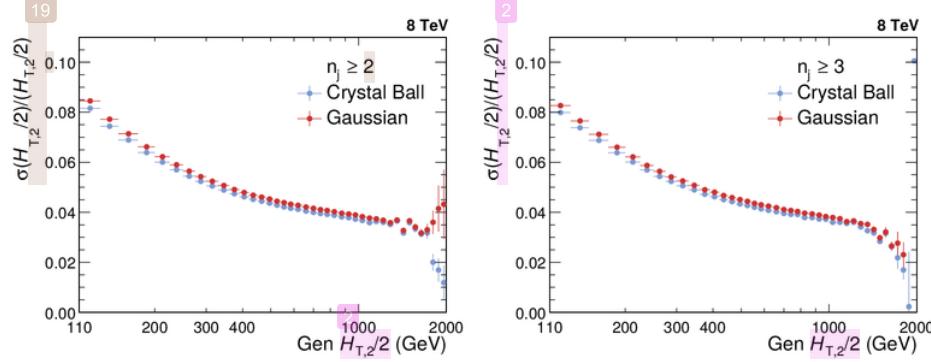


Figure 1.10: Comparison of jet energy resolution calculated using Crystal-Ball fit function (blue solid circles) and Gaussian fit function (red solid circles) for inclusive 2-jet (left) and inclusive 3-jet events (right).

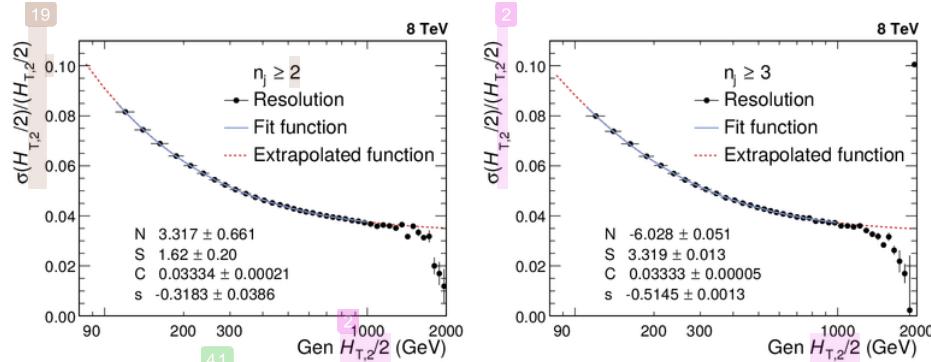


Figure 1.11: Jet energy resolution (JER) is shown as a function of Gen $H_{T,2}/2$ for inclusive 2-jet (left) and inclusive 3-jet events (right). JER (black solid circles) is fitted by using the modified NSC-formula (blue solid line) which is extrapolated to 80 GeV and upto 2000 GeV (red dashed line) to consider the migration into lower as well as higher bins.

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 and is independent of $H_{T,2}/2$; a stochastic component S due to sampling fluctuation and EM fraction fluctuation per hadrons; and a constant term C because of dead material, magnetic field and calorimeter cell to cell fluctuation. 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 ob-

tain slightly better fits. The parameters obtained after fitting the relative resolution using the above mentioned NSC formula are tabulated in Table 1.6 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 details in Sec. 1.5.1. Since JER in $n_j \geq 2$ events is similar to that one in $n_j \geq 3$ events, so N, S and C fit parameters obtained for $n_j \geq 3$ events are used for unfolding R_{32} .

$$\frac{\sigma(x)}{x} = \sqrt{\operatorname{sgn}(N) \cdot \frac{N^2}{x^2} + S^2 \cdot x^{s-1} + C^2} \quad (1.8)$$

⁴
Table 1.6: 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 inclusive 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. 1.12 for $n_j \geq 2$ (left) and $n_j \geq 3$ events (right). Some shortcomings in the detector simulation of the theory spectra leads 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 is presented upto 1680 GeV only.

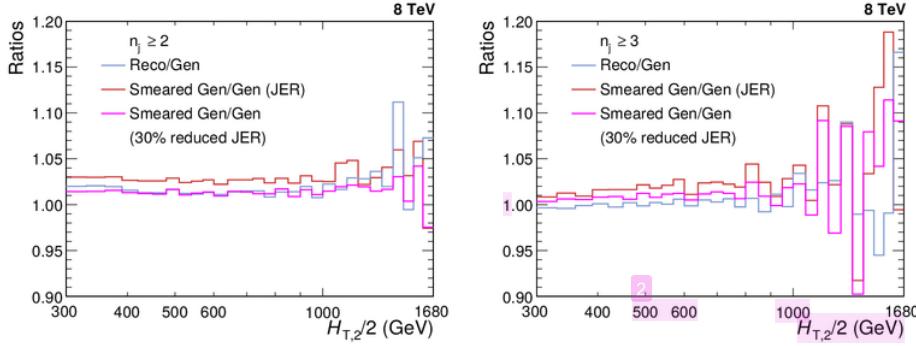


Figure 1.12: MADGRAPH5 + PYTHIA6 (MG5+P6) Gen smeared using extracted jet energy resolution (JER) shows a discrepancy from simulated Reco as Smeared Gen/Gen ratio (red line) does not match with Reco/Gen ratio (blue line), for both inclusive 2-jet (left) and inclusive 3-jet events (right). Smeared Gen/Gen 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.

1.5 Unfolding

One of the main goals in an experimental measurement is to do the comparison of 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 [9, 10] as implemented in RooUnfold software package [11]. In this algorithm, the number of iterations regularize the unfolding process. The obtained distribution in one iteration is taken as the input in the next one. χ^2 between two successive iterations is given by Eq. 1.9. The number of iterations stop when χ^2/N_{bins} is < 1 . A reduced χ^2 is obtained by a higher number of iterations but this will also increase the uncertainty and there

are larger bin-by-bin fluctuations and correlations. So the optimization of number of iterations is very important. In the current analysis, unfolding done with “four” iterations gives the best results with low χ^2 and low bin-by-bin correlations.

$$\chi^2 = \sum_{i=1}^{N_{bins}} \left(\frac{n_i^{j+1} - n_i^j}{\sqrt{n_i^j}} \right)^2 \quad (1.9)$$

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 are explained in the next section.

1.5.1 Response Matrices

The response matrix is a two dimensional mapping between the true and measured distributions. It is usually derived from simulated Monte Carlo (MC) samples, which takes the true distribution from MC as an input and smears it by taking into account the detector resolution. Then this response matrix is used to unfold the measured

data spectrum. 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.

However, 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 is constructed from these two distributions is used for the unfolding procedure.

11 1.5.1.1 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. 1.10 and 1.13. These functions describes the shape as well as normalization of the distribution.

- **Function I :**

$$f(H_{T,2}/2) = N[x_T]^{-a}[1 - x_T]^b \times \exp[-c/x_T] \quad (1.10)$$

where N is normalization factor and a, b, c are fit parameters.

This function is derived from the below function [12] :

$$f(p_T; \alpha, \beta, \gamma) = N_0[p_T]^{-\alpha} \left[1 - \frac{2 p_T \cosh(y_{min})}{\sqrt{s}} \right]^\beta \times \exp[-\gamma/p_T] \quad (1.11)$$

using

$$\alpha = \frac{40}{a}, \quad \beta = b, \quad \gamma = c * \sqrt{s}/2, \quad x_T = \frac{2 * H_{T,2}/2 * \cosh(y_{min})}{\sqrt{s}} = \frac{2 * H_{T,2}/2}{\sqrt{s}}$$
(1.12)

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 :**

$$f(H_{T,2}/2) = A_0 \left(1 - \frac{H_{T,2}/2}{A_6}\right)^{A_7} \times 10^{F(H_{T,2}/2)}, \text{ where } F(x) = \sum_{i=1}^5 A_i \left(\log\left(\frac{x}{A_6}\right)\right)^i$$
(1.13)

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

Figure 1.13 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 inclusive 3-jet events (right). 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. 1.6.3. To include the migration 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. 1.13) provides weights to the flat spectrum. A total of ten million events are generated randomly (in $H_{T,2}/2$ range 80-2000). These generated values are then smeared with a Gaussian function, where σ of the Gaussian is determined from the relative resolution

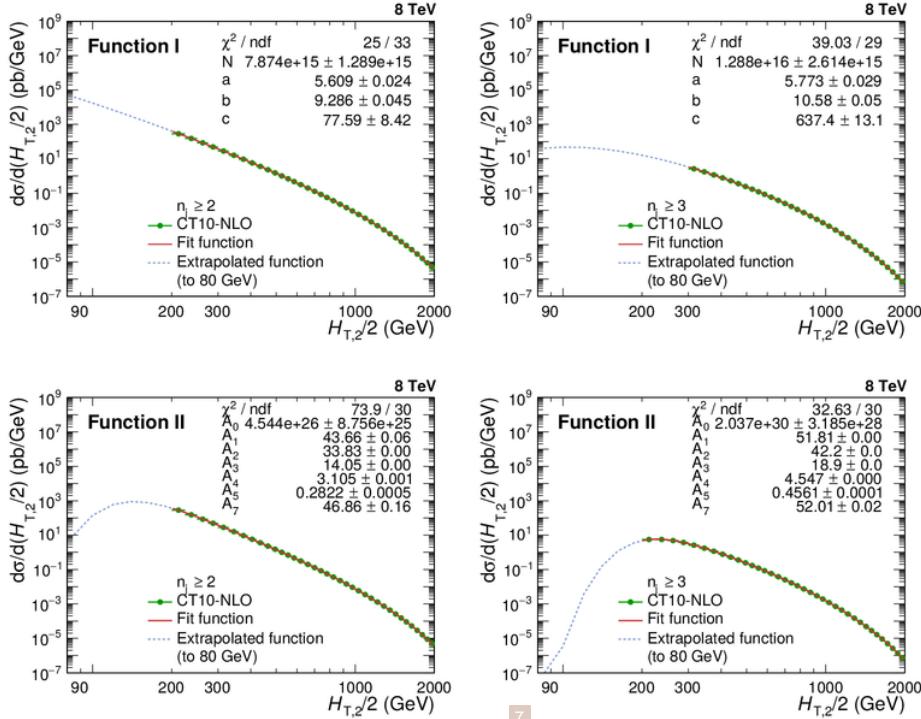


Figure 1.13: Fitted CT10-NLO spectrum of differential cross section as a function of $H_{T,2}/2$ (green solid circles) using Function I (top) defined in Eq. 1.10 and using Function II (bottom) given by Eq. 1.13, for inclusive 2-jet (left) and inclusive 3-jet events (right). To consider the migration to lower $H_{T,2}/2$ bins, the fit functions described by red lines are extrapolated to 80 GeV (blue dashed lines).

parametrization as a function of $H_{T,2}/2$ calculated from NSC formula mentioned in equation 1.8. The parameters N, S, C used for smearing are taken from Table 1.6. These randomly generated (Gen_{Toy}) and smeared ($\text{Measured}_{\text{Toy}}$) values are used to fill the response matrices. Figure 1.14 shows the response matrices derived using the Toy MC for $n_j \geq 2$ (left) and $n_j \geq 3$ events (right). 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.

1

3

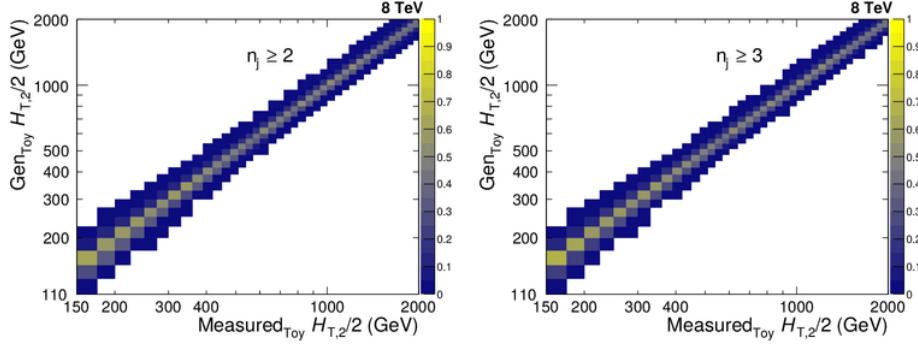


Figure 1.14: The response matrices are derived using the Toy Monte Carlo and forward smearing method, for inclusive 2-jet (left) and inclusive 3-jet events (right). The matrices are normalized to the number of events in each column and are diagonal with small off-diagonal migrations between close-by $H_{T,2}/2$ bins.

1.5.1.2 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. 1.5.1.1 for differential cross sections. To obtain the true spectrum for R_{32} , the ratio of cross section spectrum described by Eq. 1.10 for inclusive 3-jet to that of 2-jet events is taken. This ratio is shown by green solid circles in Fig. 1.15 (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. 1.15 (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.

1.5.2 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 1.14. It is expected that the same Gen_{Toy} spectrum should be re-obtained after unfolding. Figure 1.16 confirms that the unfolded $\text{Measured}_{\text{Toy}}$

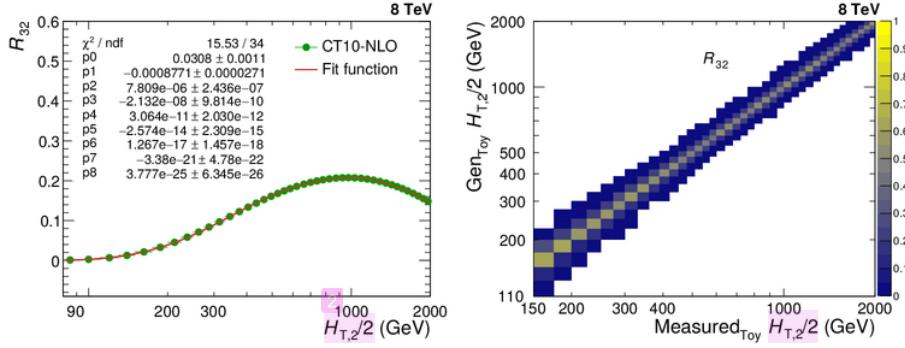


Figure 1.15: Left : The ratio of cross sections described by Eq. 1.10 for inclusive 3-jet to that of 2-jet events is shown as a function of $H_{T,2}/2$ (green solid circles). It is fit using a polynomial function of degree 8 (red line). Right : The response matrix is derived using the Toy Monte Carlo and forward smearing method, for the cross section ratio R_{32} . The matrix is normalized to the number of events in each column and is diagonal with small off-diagonal migrations between close-by $H_{T,2}/2$ bins.

spectrum matches exactly with Gen_{Toy} spectrum as the ratio of these distributions is perfectly flat at one for both $n_j \geq 2$ (top left) and $n_j \geq 3$ events (top right) 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 well closure is not obtained. This is represented by blue line in Fig. 1.17 for $n_j \geq 2$ (top left) and $n_j \geq 3$ events (top right). As observed in Fig. 1.12 in Sec. 1.4, 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. 1.17. 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).

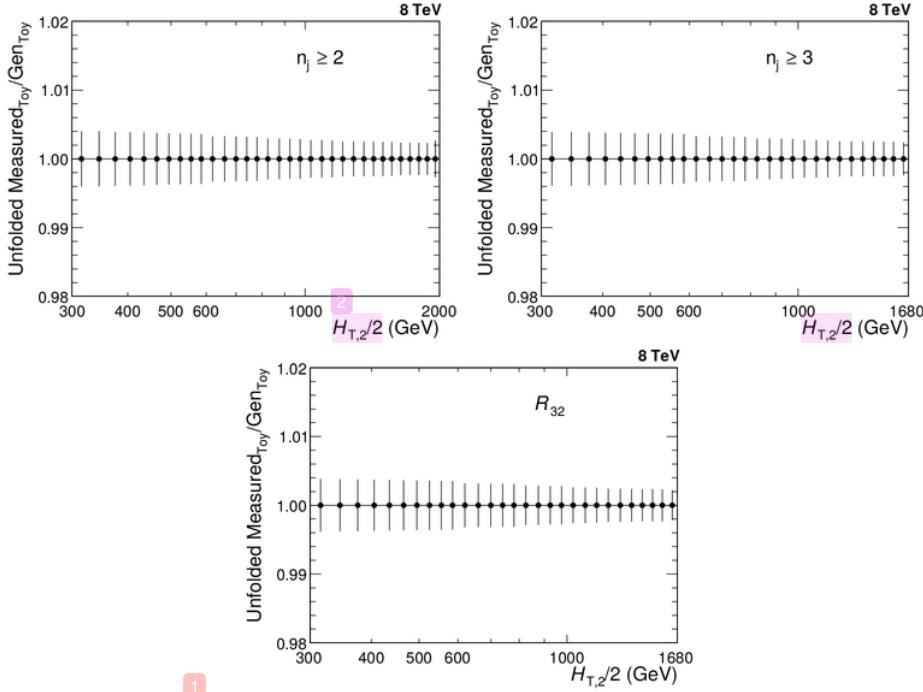


Figure 1.16: Closure test of the unfolding technique where the smeared spectrum obtained from Toy Monte Carlo method (Measured_{Toy}), is unfolded using the constructed response matrices (obtained by forward smearing the randomly generated spectrum (Gen_{Toy}) using extracted jet energy resolution (JER)). As expected, the unfolded measured_{Toy} spectrum matches exactly with Gen_{Toy} spectrum as the ratio of these distributions is perfectly flat at one for both inclusive 2-jet (top left) and inclusive 3-jet events (top right) cross sections as well as the cross section ratio R_{32} (bottom).

1.5.3 Unfolding of the Measurement

After validity 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. 1.18 for $n_j \geq 2$ (top left) and $n_j \geq 3$ events (top right) 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 data is done with response matrices using JER (blue solid

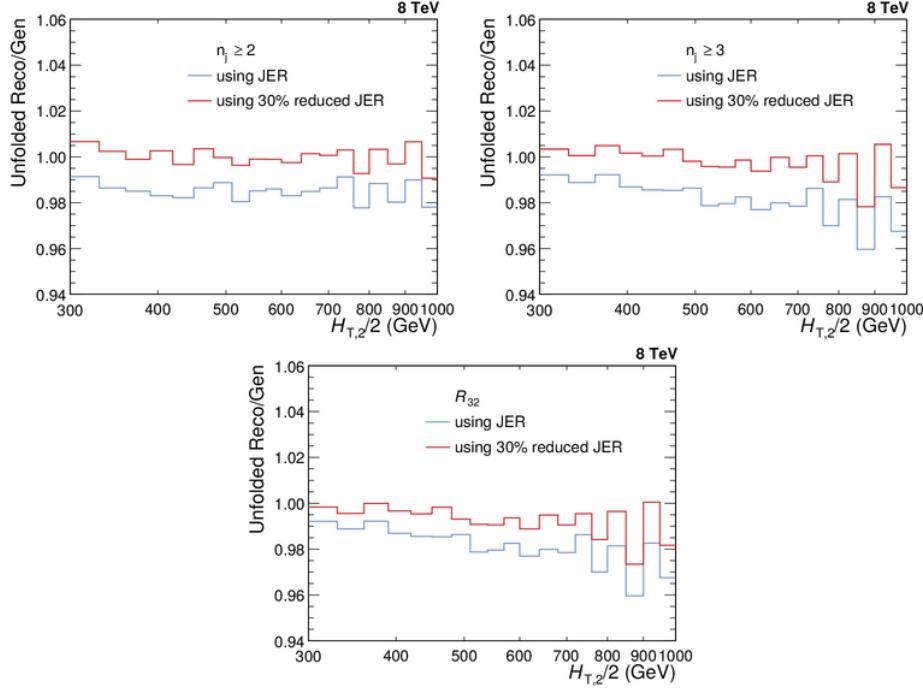
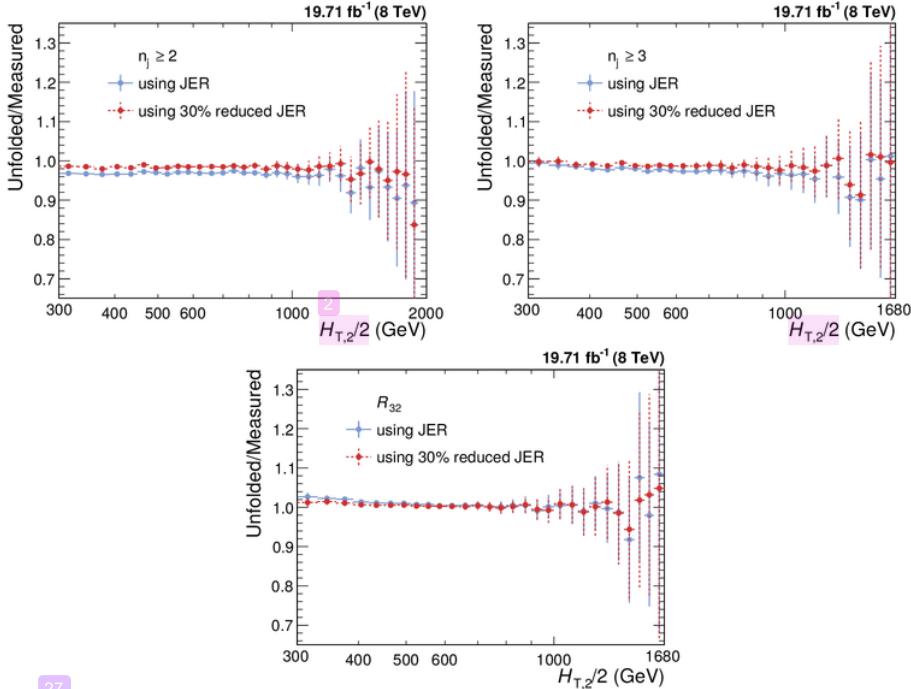


Figure 1.17: Reco MADGRAPH5 + PYTHIA6 Monte Carlo (MG5+P6 MC) differential cross section distributions unfolded with the response matrices (obtained by forward smearing the randomly generated spectrum (Gen) using extracted jet energy resolution (JER)), does not give a good closure with Gen MG5+P6 MC (blue line), for inclusive 2-jet (top left) and inclusive 3-jet events (top right). 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).

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.
14

1.6 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



27 Figure 1.18: 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 inclusive 3-jet events cross sections (right) 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.

11 uncertainties arise due to random fluctuations depending on the number of events.

The more the number of events, less 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. 32 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.

1.6.1 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. Such smeared spectrums are produced million in number and the unfolding is performed multiple times for each smeared spectra. The differences between the unfolded spectrums and the measured one give the statistical uncertainty. The unfolding process introduces more statistical fluctuations which can be observed in Fig. 1.19. 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 $n_j \geq 3$ events cross sections (top right) 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. 1.20, the correlations of the statistical uncertainty after the unfolding can be seen for $n_j \geq 2$ (top left) and $n_j \geq 3$ events cross sections (top right) and for the cross section ratio R_{32} (bottom). These correlations must be considered while performing the fits to extract the value of the strong coupling constant, α_S .

1.6.2 Jet Energy Corrections Uncertainties

As explained in Sec. 1, the measured jet energy is corrected for a variety of detector effects by applying jet energy corrections (JEC) [13]. There are 25 mutually independent sources which contribute to JEC. Each source presents a 1σ shift and is

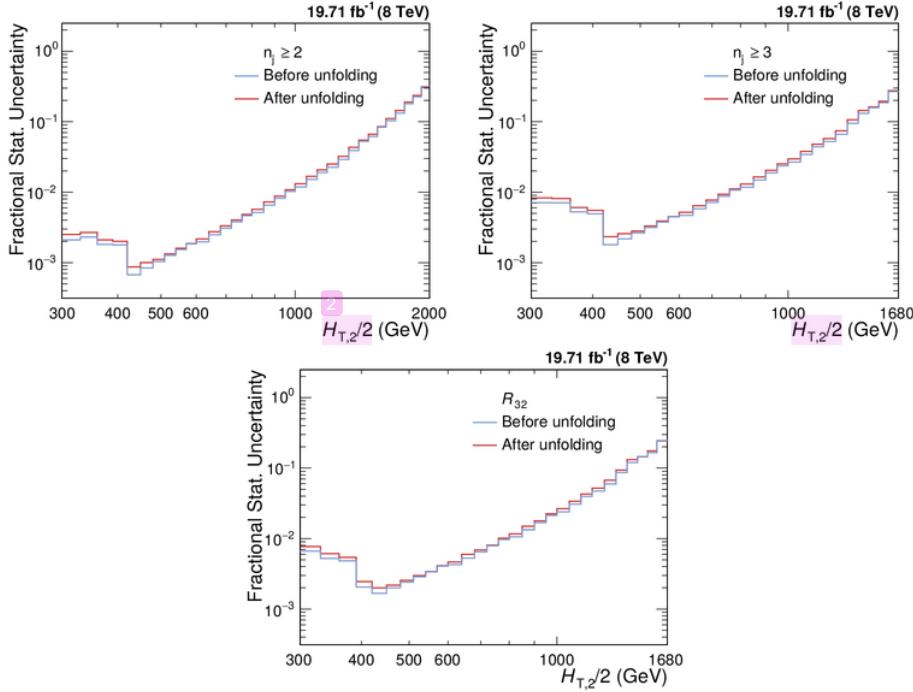


Figure 1.19: The fractional statistical uncertainties of the unfolded data (red line) are compared with those of the measured one (blue line) for inclusive 2-jet (top left) and inclusive 3-jet events cross sections (top right) as well as for the cross section ratio R_{32} (bottom). After unfolding, the statistical uncertainty increases slightly.

1 fully correlated in p_T and η but uncorrelated to all other sources. The observable is
 2 studied with the nominal values of the jet energy which gives nominal distributions
 3 as well as by varying up and down the energy of all jets by the uncertainty. The
 4 differences between the nominal distributions and the ones obtained by varying the
 5 jet energy gives the uncertainties from each source. The JEC uncertainties can be
 6 asymmetric in nature which leads to separate treatment of upwards and downwards
 7 variation of each source. The sum in quadrature of uncertainties from all sources
 8 gives the total JEC uncertainty. In the current analysis, JEC uncertainties are a
 9 dominant source of experimental uncertainty at low $H_{T,2}/2$. The JEC uncertainty
 10 ranges from 3% to 10% for $n_j \geq 2$ and from 3% to 8% for $n_j \geq 3$ events cross sections

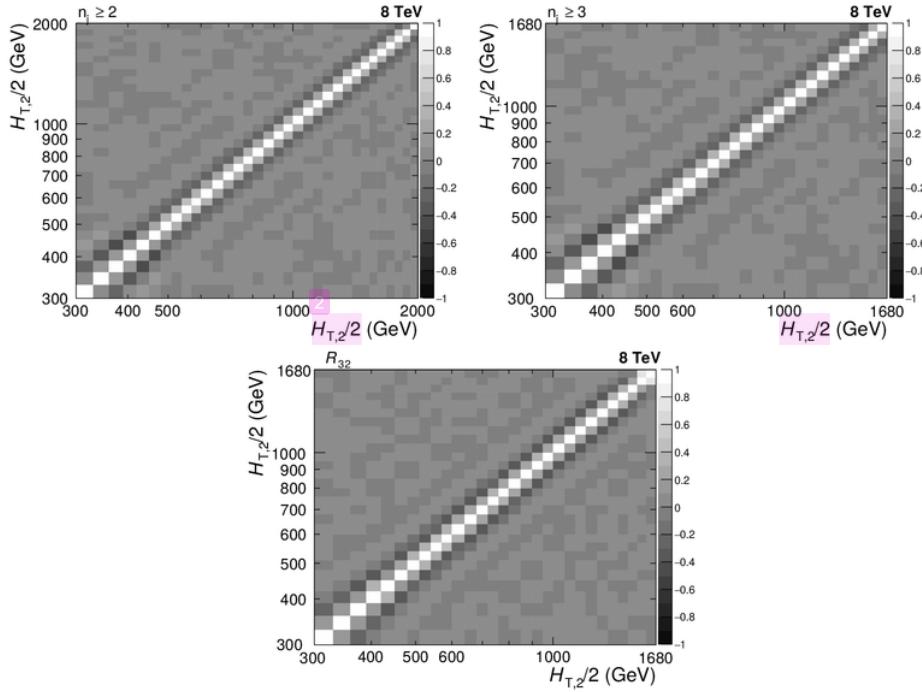


Figure 1.20: The unfolding procedure introduces the correlations of the statistical uncertainty through bin migrations which are shown here for inclusive 2-jet (top left) and and 3-jet events cross sections (top right) as well as for the cross section ratio R_{32} (bottom). The correlation (anti-) is more significant between neighbouring bins than far-ff ones.

measurement. 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 1-2% over all $H_{T,2}/2$ bins.

The sources of JEC considered in this analysis are : AbsoluteStat, AbsoluteScale, AbsoluteFlavMap, AbsoluteMPFBias, Fragmentation, SinglePionECAL, SinglePionHCAL, FlavorQCD, RelativeJEREC1, RelativeJEREC2, Rel-

ativeJERHF, 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. The Figs. A.1-A.3 show the JEC uncertainty from each source separately for ² inclusive 2-jet (top) and 3-jet events cross sections (middle) as for cross section R_{32} (bottom). Depending on the origin of sources, 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 pile-up inside and outside of jets), in simulated events. This uncertainty is derived from Z/γ +jet, dijet and multijet data using fit procedure to estimate the residual pileup uncertainty after the calibration.
2. **Relative** : The forward jets are calibrated by the relative η -dependent corrections using dijet events. The main contribution to the uncertainty comes from jet energy resolution (JER), derived by varying JER scale factors up and down by quoted uncertainties and the initial and final state radiation bias corrections.
3. **Absolute** : A global fit to Z/γ +jet and multi-jet events gives the absolute calibration of the jet energy scale. The uncertainties are related to the lepton momentum scale for muons in $Z \rightarrow \mu\mu$ +jet and the single pion response in the HCAL.
4. **Flavor** : Flavor response differences are studied from simulation by cross-checking the results with quark- and gluon-tagged γ +jet and Z +jet events.

These uncertainties are based on PYTHIA6.4 and HERWIG++2.3 differences propagated through the data-based calibration method.

The details of the jet energy corrections and uncertainties can be found in [14].

1.6.3 Unfolding Uncertainties

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. 1.4. As mentioned before, the measured jet transverse momentum (p_T) in simulated MC events needs to be smeared additionally to match the resolution in data. This smearing is done by using measured scale factors ($c_{central}$) mentioned in Table 1.5. 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 spectrums to the nominal ones accounts for a systematic JER uncertainty.
2. **Model Dependence :** It is explained in Sec. 1.5.1 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. 1.10. Using the alternative function, Function II given by Eq. 1.13, 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 gives the model dependence uncertainty.

3. Additional Uncertainty : Small nonclosures observed in Fig. 1.12 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} .

1.6.4 Luminosity Measurement Uncertainty

The measurement of the luminosity delivered to CMS detector by LHC in the proton-proton collisions in the year of 2012 is done by using the silicon pixel cluster counting method [15]. 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} .

1.6.5 Residual Uncertainty

The small trigger and jet identification inefficiencies account for smaller than 1% uncertainties on the cross section measurements [16, 17]. 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.

1.6.6 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. The values of uncertainties (in %) from each source as well as total uncertainty, for each $H_{T,2}/2$ bin, are tabulated in Tables A.2, A.3 and A.4 for $n_j \geq 2$ and $n_j \geq 3$ events cross sections and cross section ratio R_{32} , respectively. Figure 1.21 shows the experimental uncertainties, from different sources as well as the total uncertainty, affecting the measurement of $n_j \geq 2$ (top left) and $n_j \geq 3$ events cross sections (top right) and cross section ratio R_{32} (bottom). The error bars represent the statistical uncertainty obtained after unfolding. The systematic uncertainties due to jet energy corrections (JEC by blue line), luminosity (red dashed line), unfolding (green dashed line) and residual effects (light purple line) are also presented. The uncertainties due to luminosity and residual effects cancel completely in R_{32} . The total uncertainty (black dashed line) on the measurements is asymmetric in nature and dominated by the uncertainty due to the jet energy corrections (JEC) at lower $H_{T,2}/2$ values and by statistical uncertainty at higher $H_{T,2}/2$ values. The experimental uncertainties from each source as well as total uncertainty are also quoted in Table 1.7.

The complete data analysis of the differential inclusive 2-jet and inclusive 3-jet events cross sections as well as their ratio R_{32} has been presented as a function of

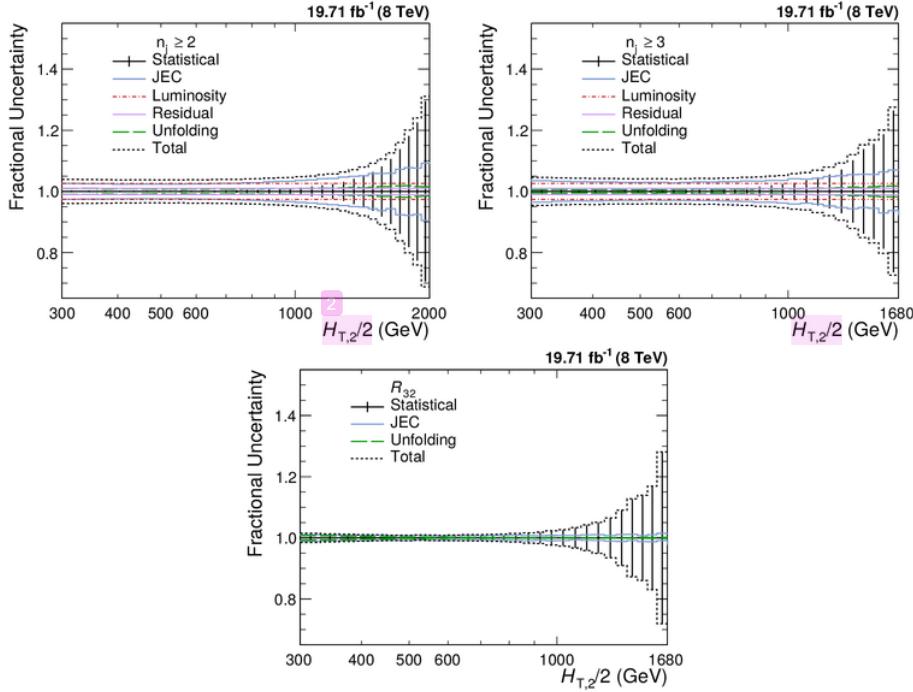


Figure 1.21: Experimental uncertainties from different sources affecting the measurement of cross sections for inclusive 2-jet (top left) and inclusive 3-jet events (top right) and cross section ratio R_{32} (bottom). The error bars represent the statistical uncertainty after unfolding. The systematic 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 calculated by adding in quadrature the individual sources of uncertainty.

$H_{T,2}/2$. The measured spectrums after correcting for detector effects through the unfolding procedure, are compared with the next-to-leading order (NLO) pQCD calculations in the next chapter.

Table 1.7: Overview of all experimental uncertainties affecting the measurement of cross sections for inclusive 2-jet (left) and inclusive 3-jet events (middle) and cross section ratio R_{32} (right). The uncertainties due to luminosity and residual effects cancel completely in R_{32} . The total uncertainty is calculated by adding in quadrature 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	1-2%	1-2%	< 1%
Luminosity	2.6%	2.6%	cancels
Residual	1%	1%	cancels
Total	4 to 32%	4 to 28%	1 to 28%

Chapter 2

Theoretical Calculations

Appendix A

A.1 Cross section ratio, R_{32}

Table A.1: Differential cross sections and cross section ratio at detector level in each bin of $H_{T,2}/2$, along with statistical uncertainty (in %).

Bin	2-jet x-section (x 10^{-3} (pb/GeV))	Stat. unc.	3-jet x-section (x 10^{-3} (pb/GeV))	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

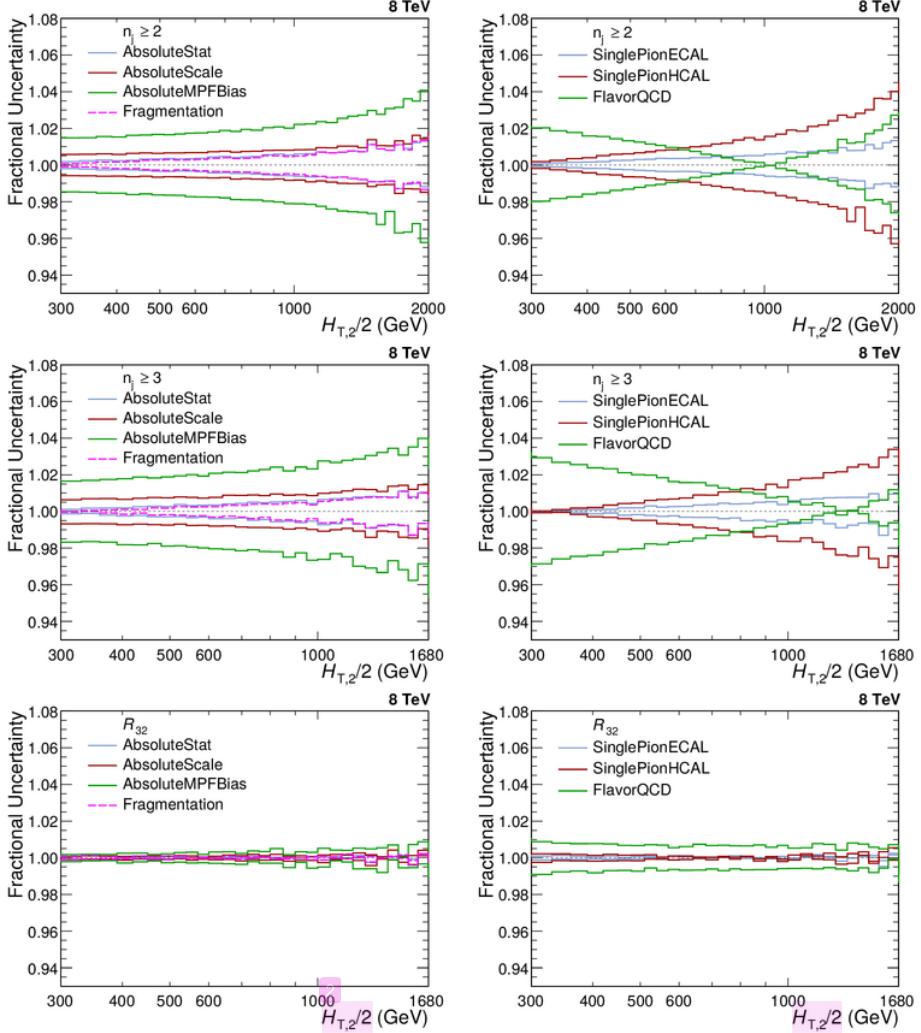


Figure A.1: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet events cross sections (middle); and cross section ratio R_{32} (bottom). On left, JEC uncertainties are evaluated from **AbsoluteStat**, **AbsoluteScale**, **AbsoluteMPFBias**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

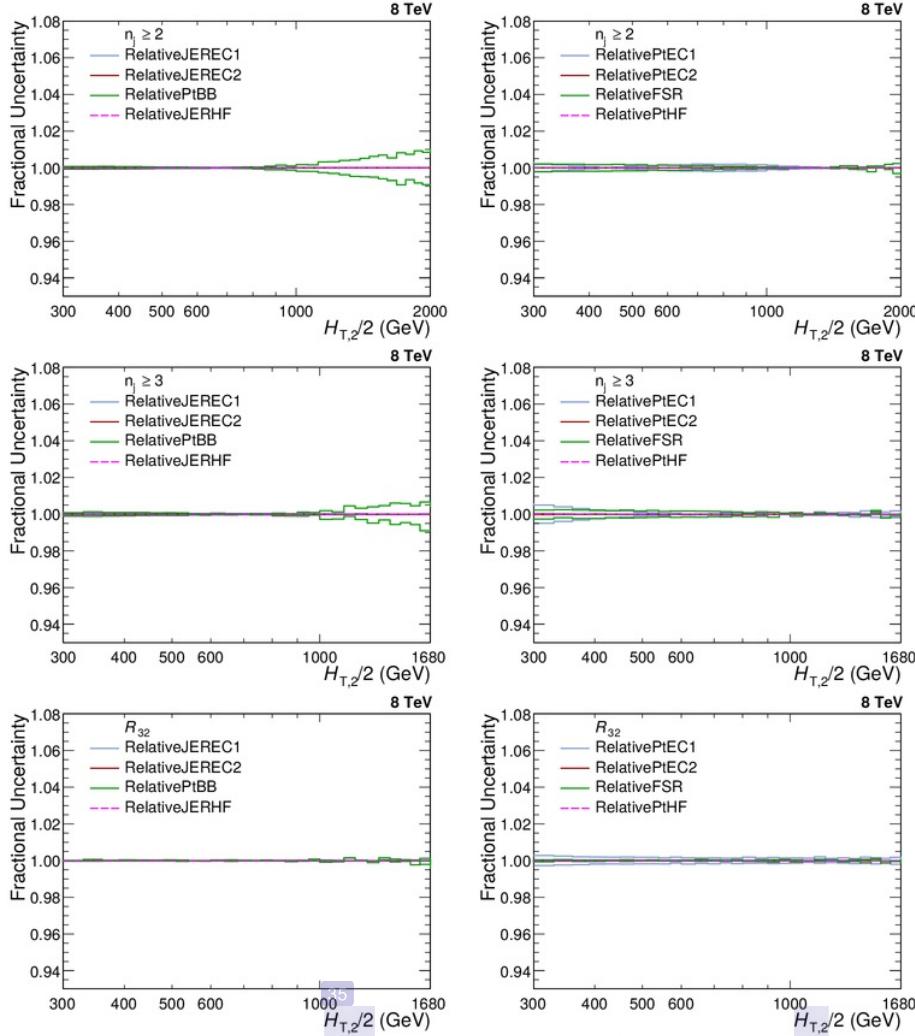


Figure A.2: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet events (middle); and cross section ratio R_{32} (bottom). On left, JEC uncertainties are evaluated from [RelativeJEREC1](#), [RelativeJEREC2](#), [RelativePtBB](#)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

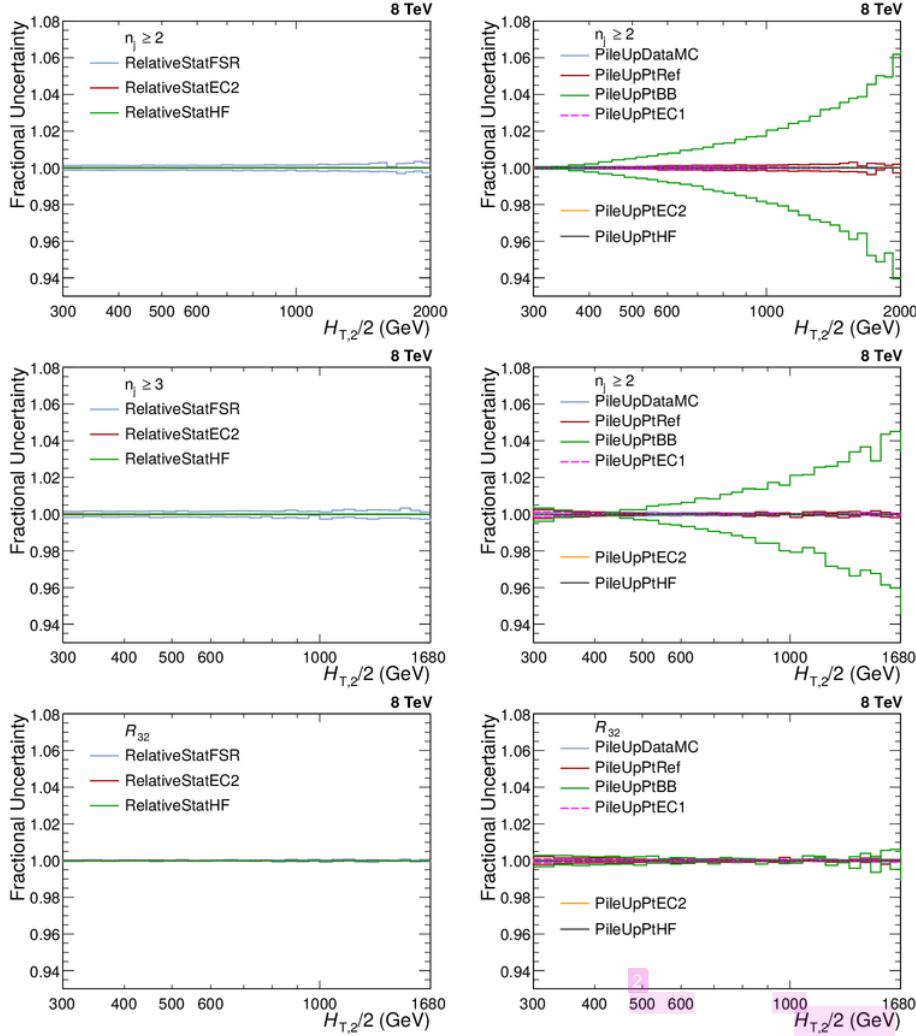


Figure A.3: The fractional jet energy correction (JEC) uncertainties from individual sources are shown for inclusive 2-jet (top) and 3-jet events cross sections (middle); and cross section ratio R_{32} (bottom). On left, JEC uncertainties are evaluated from [RelativeStatFSR](#), [RelativeStatEC2](#) and [RelativeStatHF](#) sources whereas on right, these are evaluated from [PileUpDataMC](#), [PileUpPtRef](#), [PileUpPtBB](#), [PileUpPtEC1](#), [PileUpPtEC2](#) and [PileUpPtHF](#) 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 $H_{T,2}/2$ 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 $H_{T,2}/2$.

Bin	Statistical	JEC	Unfolding	Total
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