Chapter 2

Theoretical Calculations

In an experiment, the measurements are validated by doing the comparison with the perturbative QCD (pQCD) theoretical calculations. The lowest order (LO) calculations describe well the shapes of the measured distributions but not the normalization due to the dependence on the unphysical renormalization (μ_r) and factorization (μ_f) scales. The next-to-leading order calculations (NLO) improves the precision by reducing the dependence on μ_r and μ_f scales and become an essential feature in the determination of fundamental parameters such as α_S and the parton density distributions. In this chapter, the next-to-leading order pQCD calculations are described in details. NLO pQCD calculations are corrected for the multiparton interactions (MPI) and hadronization effects by applying non-perturbative (NP) corrections and also corrected for the electroweak interactions (EW).

2.1 Fixed Order NLO Calculations

The predictions of the inclusive differential jet event cross section at NLO accuracy in pQCD are computed with the NLOJET++ program version 4.1.3 [18, 19]. The results are provided within the framework of fastNLO version 2.3 [20, 21] for use within fits. The parton distribution functions (PDFs) are accessed through the

LHAPDF6 library [22, 23]. The FASTNLO is preferred over the direct calculation with NLOJET++ as the calculations of the cross sections can be repeated several times with different PDFs as well as scale choices required for the calculating PDF and scale uncertainties. The factorization and renormalization scales are chosen equal to $H_{T,2}/2$, i.e. $\mu_f = \mu_r = H_{T,2}/2$.

In the current study, different PDF sets available for a series of different assumptions on $\alpha_s(M_Z)$ are used for NLO calculations. In Table 2.1 upper rows list the already existing PDF sets in LHC Run 1 whereas lower rows list 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 for different PDF sets. All sets uses a variable-flavour number scheme with at most five or six flavours apart from the ABM11 PDFs, which employ a fixed-flavour number scheme with $N_F = 5$. Mainly CT10 PDF set is considered for comparison between data and theory predictions as well as for calculating theoretical uncertainties. 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 [24–27].
- 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.

2.1.1 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 a NLO

Table 2.1: NLO PDF sets are available via LHAPDF6 with various assumptions on the value of $\alpha_s(M_Z)$. The upper rows list the already 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. A * behind the $\alpha_s(M_Z)$ values signifies that the parameter was fixed, not fitted.

Base set	N_F	M_t (GeV)	M_Z (GeV)	$\alpha_s(M_Z)$	$\alpha_s(M_Z)$ range
ABM11 [28]	5	180	91.174	0.1180	0.110–0.130
CT10 [29]	$\leq$5	172	91.188	0.1180*	0.112–0.127
MSTW2008 [30, 31]	$\leq$5	10^{10}	91.1876	0.1202	0.110–0.130
NNPDF2.3 [32]	$\leq$6	175	91.1876	0.1180*	0.114–0.124
CT14 [33]	$\leq$5	172	91.1876	0.1180*	0.113–0.123
HERAPDF2.0 [34]	$\leq$5	173	91.1876	0.1180*	0.110–0.130
MMHT2014 [35]	$\leq$5	10^{10}	91.1876	0.1180*	0.108–0.128
NNPDF3.0 [36]	$\leq$5	173	91.2	0.1180*	0.115–0.121

correction factor, k-factor, which is derived as the ratio of cross sections at NLO accuracy to that at LO i.e.

$$\text{k-factor} = \frac{\sigma_{\text{NLO}}}{\sigma_{\text{LO}}} \tag{2.1}$$

The size of k-factor determine the impact of the higher-order corrections. 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 2.1 shows the k-factors of the NLOJET++ calculations, for inclusive 2-jet and 3-jet events 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 events 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 $H_{\text{T},2}/2$. It is observed that for inclusive 3-jet events cross sections, k-factor jumps at lowest $H_{\text{T},2}/2$. 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 $H_{\text{T},2}/2$ (below 225 GeV) still suffer from these

kinematical effects, the minimum value of $H_{T,2}/2$ studied is 300 GeV.

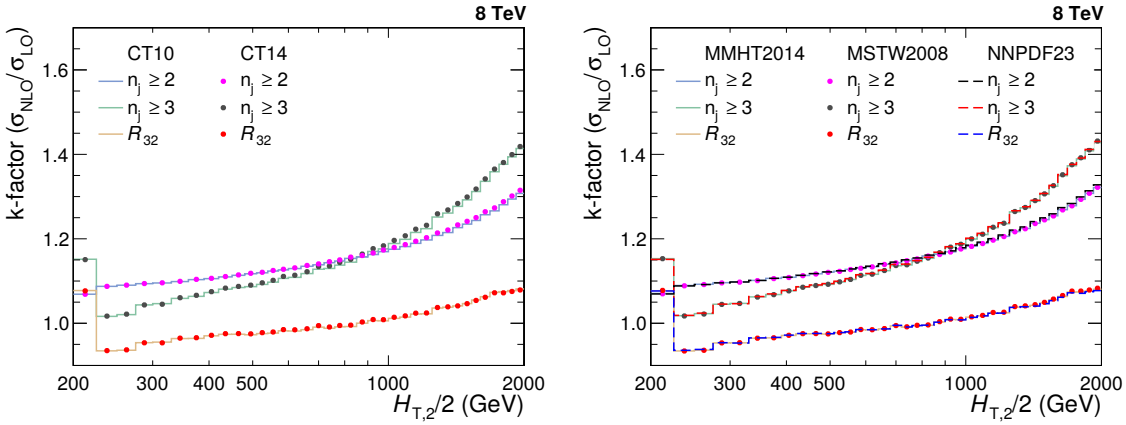

Figure 2.1: The k-factors of the NLOJET++ calculations, for inclusive 2-jet and 3-jet events cross sections and their ratio R_{32}, using five different PDF sets.

2.1.2 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 lower order perturbation theories and hence requires a parton shower (PS) correction. The scattering phenomena between partons within a colliding proton, other than the hard scattering, are known as multi-parton interactions (MPI). The partons of the hard scattering forms 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 unfolded data. So the corrections for non-perturbative effects (NP) should be taken into account in NLO calculations. The ratio of cross section predicted with a nominal event generation interfaced to the simulation of UE contributions and to the one without hadronization and MPI effects gives the correction factors which are defined as :

$$C^{\mathrm{NP}} = \frac{\sigma^{\mathrm{PS+HAD+MPI}}}{\sigma^{\mathrm{PS}}} \tag{2.2}$$

In this analysis, 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++ [37] with the default tune of version 2.3 and PYTHIA6 [2] with tune Z2*, and the NLO, POWHEG [38–40], MC event generators are considered. The matrix-element calculation is performed with POWHEG interfaced to PYTHIA8 with tune CUETS1 [41] for the UE simulation. The ratio, defined in Eq. 2.2, is obtained for each MC generator and is fitted by a power-law function defined in Eq. 2.3. 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 it is taken as the uncertainty on the NP correction factor.

$$f(H_{\mathrm{T},2}/2) = a \cdot \left(H_{\mathrm{T},2}/2\right)^{b} + c \tag{2.3}$$

The NP correction factors, $C^{\mathrm{NP}}_{\text{3-jet}}$ and $C^{\mathrm{NP}}_{\text{2-jet}}$ are calculated for $\mathrm{n_j} \geq 2$ and $\mathrm{n_j} \geq 3$ event cross sections respectively and then their ratio gives the correction factor for R_{32}. These are shown in Fig. 2.2 for the inclusive 2-jet (top left) and 3-jet event cross sections (top right), as well as for cross section ratio R_{32} (bottom). The NP corrections amount to $\sim$ 4-5% for inclusive 2-jet and 3-jet events cross section and $\sim$ 1% for ratio R_{32}, for $H_{\mathrm{T},2}/2 \sim 300$ GeV and decrease rapidly for increasing $H_{\mathrm{T},2}/2$. On comparing the NP correction factors of cross section ratio with that for individual cross sections, it has been observed that the non-perturbative effects get reduced in the cross section ratio.

2.1.3 Electroweak Corrections

In LHC, the centre-of-mass energy of proton-proton collisions is well beyond the electroweak (EW) scale $\sim$O(100 GeV). At such a high energy, the impact of higher

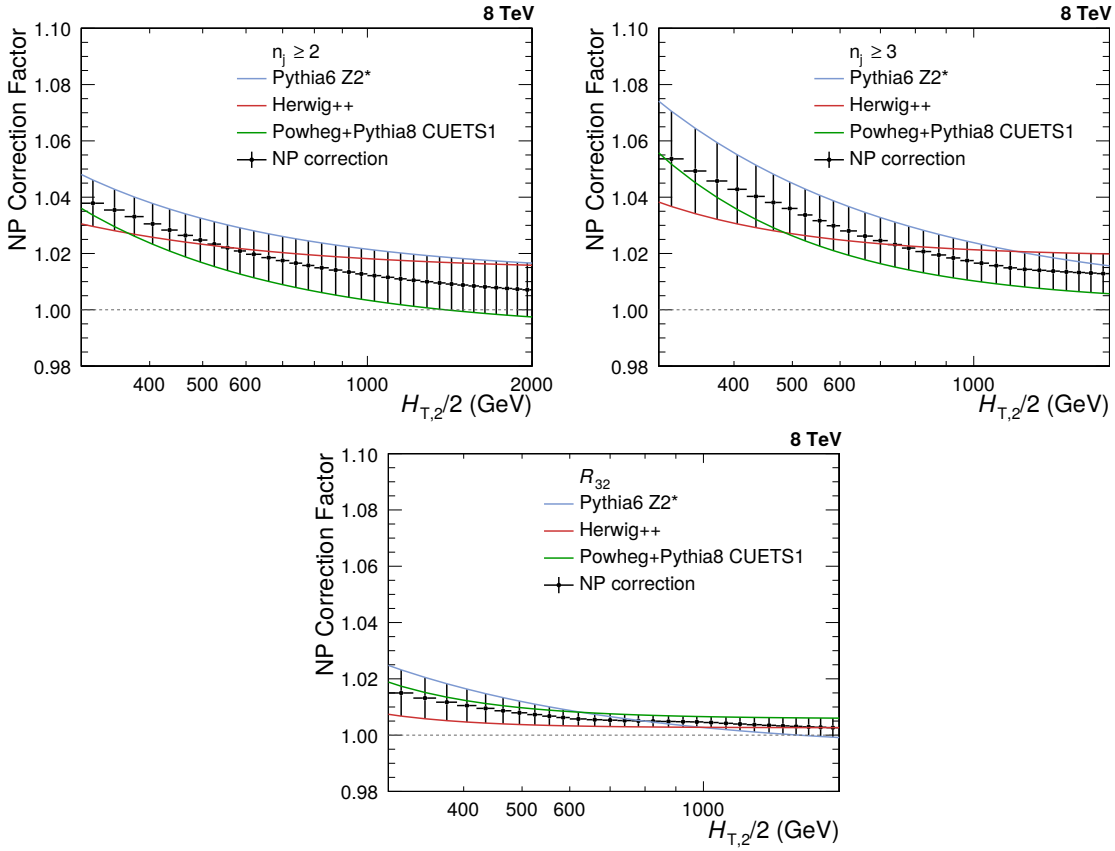

Figure 2.2: The nonperturbative (NP) corrections are presented as a function of $H_{T,2}/2$ for inclusive 2-jet (top left) and 3-jet (top right) event cross sections, as well as their ratio R_{32}. These corrections are calculated from the leading order HERWIG++ with the default tune of version 2.3 (red 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.

order EW corrections is much more with respect to QCD effects [42] and affect jet cross sections at large jet p_T. 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 corrections have been calculated for inclusive 1-jet and 2-jet case, in Ref. [43] but are not available yet for inclusive 3-jet production. The EW correction factors in the phase space of the measurement are shown as a function of $H_{T,2}/2$ in Fig. 2.3 for inclusive 2-jet

events cross sections. These correction factor increases up to 13% at high ends of $H_{T,2}/2$ and significantly improves the agreement between data and prediction. Since 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 the factor of 1. These corrections are applied as a bin-by-bin correction factor to the fixed-order calculation of NLOJET++ as well as to the Monte Carlo predictions .

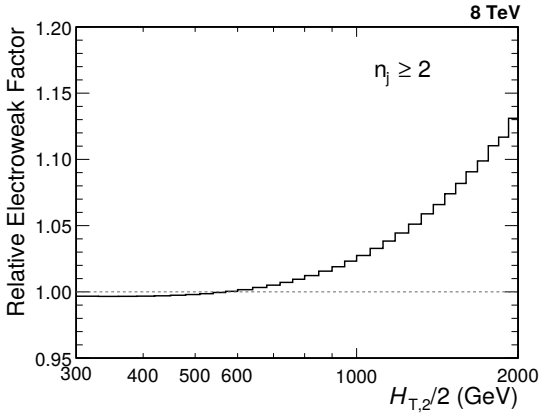

Figure 2.3: The electroweak (EW) corrections [43] in the phase space of the measurement are shown as a function of $H_{T,2}/2$ for inclusive 2-jet events cross sections. These corrections are applied as a bin-by-bin correction factor to the fixed-order calculation of NLOJET++ as well as the MC predictions of MADGRAPH5+PYTHIA6. The EW correction factor increases up to 13% at high ends of $H_{T,2}/2$ and significantly improves the agreement between data and prediction.

2.2 Theoretical Uncertainties

The measurements are not only sensitive to experimental uncertainties but also to the theoretical uncertainties. The various sources contributing to theoretical uncertainties are described below :

2.2.1 Scale Uncertainty

In perturbative QCD calculations of cross sections, one has to choose a renormalization (μ_r) and factorization (μ_f) scale. The dependence on scales is negligible

if these calculations are performed for all orders of the perturbative series. But the perturbative series is truncated at NLO, so there is a scale dependence of the measurement which is covered by systematic uncertainty known as scale uncertainty. The scale uncertainty is evaluated with the conventional recipe of varying the default scale $H_{T,2}/2$ chosen for μ_r and μ_f independently in the following six combinations: $(\mu_r/H_{T,2}/2,\ \mu_f/H_{T,2}/2)$ = (1/2,1/2), (1/2,1), (1,1/2), (1,2), (2,1) and (2,2). The maximal upwards and downwards deviations in cross section from the central prediction give the scale uncertainty. To calculate the scale uncertainty for cross section ratio R_{32}, first R_{32} is obtained for each above mentioned scale choice and then its difference from central R_{32} is taken. The scale uncertainty calculated using CT10-NLO PDF set ranges for inclusive 2-jet events from 5% to 13%, for inclusive 3-jet events from 11% to 17% and for R_{32} from 6% to 8%.

2.2.2 PDF Uncertainty

The calculation of jet cross sections in proton-proton collisions relies upon the knowledge of parton distribution functions (PDF). These PDF sets are determined 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 theory model, input parameters like the strong coupling constant α_S, the quark masses and the statistical and systematic uncertainty sources of the data included in the PDF fit. These sources contribute to PDF uncertainty which is evaluated according to the prescriptions given for each PDF set. The CT10 NLO PDF set [29,44] employ 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, ΔX^+ and ΔX^-, of a quantity X are given by Eq. 2.4 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.

$$\Delta X^{+} = \sqrt{\sum_{k=1}^{N_{\text{ev}}} \left[\max(X_k^{+} - X^0, X_k^{-} - X^0, 0)\right]^2}$$
$$\Delta X^{-} = \sqrt{\sum_{k=1}^{N_{\text{ev}}} \left[\min(X_k^{+} - X^0, X_k^{-} - X^0, 0)\right]^2} \tag{2.4}$$

The symmetric uncertainty ($\Delta X^{\pm}$) is given by half the difference of the upwards and downwards variations :

$$\Delta X^{\pm} = \sqrt{\sum_{k=1}^{N_{\text{ev}}} \left[\frac{X_k^{+} - X_k^{-}}{2}\right]^2} \tag{2.5}$$

The CT10 PDF set uncertainties are downscaled by a factor of 1.64 in order to have the uncertainties at the 68.3% confidence level CL(1σ) instead of 90% CL(2σ). The PDF uncertainty as derived with the CT10 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}.

2.2.3 Non-perturbative Uncertainty

As discussed in 2.1.2, 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 events cross sections respectively, and $< 1\%$ for cross section ratio R_{32}.

2.2.4 Total Theoretical Uncertainty

The total systematic theoretical uncertainties are obtained as the quadratic sum of the scale, PDF and NP uncertainties. Figure 2.4 presents the systematic theoretical uncertainties affecting the cross section measurement for inclusive 2-jet (top left) and 3-jet events (top right) and the cross section ratio R_{32} (bottom), using CT10 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 is dominated by PDF uncertainty. Table 2.2 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 $n_j \geq 3$ events cross sections and 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.

Table 2.2: Overview of all systematic theoretical uncertainties, obtained using CT10-NLO PDF set, affecting the measurement of cross sections for inclusive 2-jet (left) and inclusive 3-jet events (middle) and cross section ratio R_{32} (right).

Uncertainty Source	**Inclusive 2-jet**	**Inclusive 3-jet**	$\boldsymbol{R_{32}}$
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%

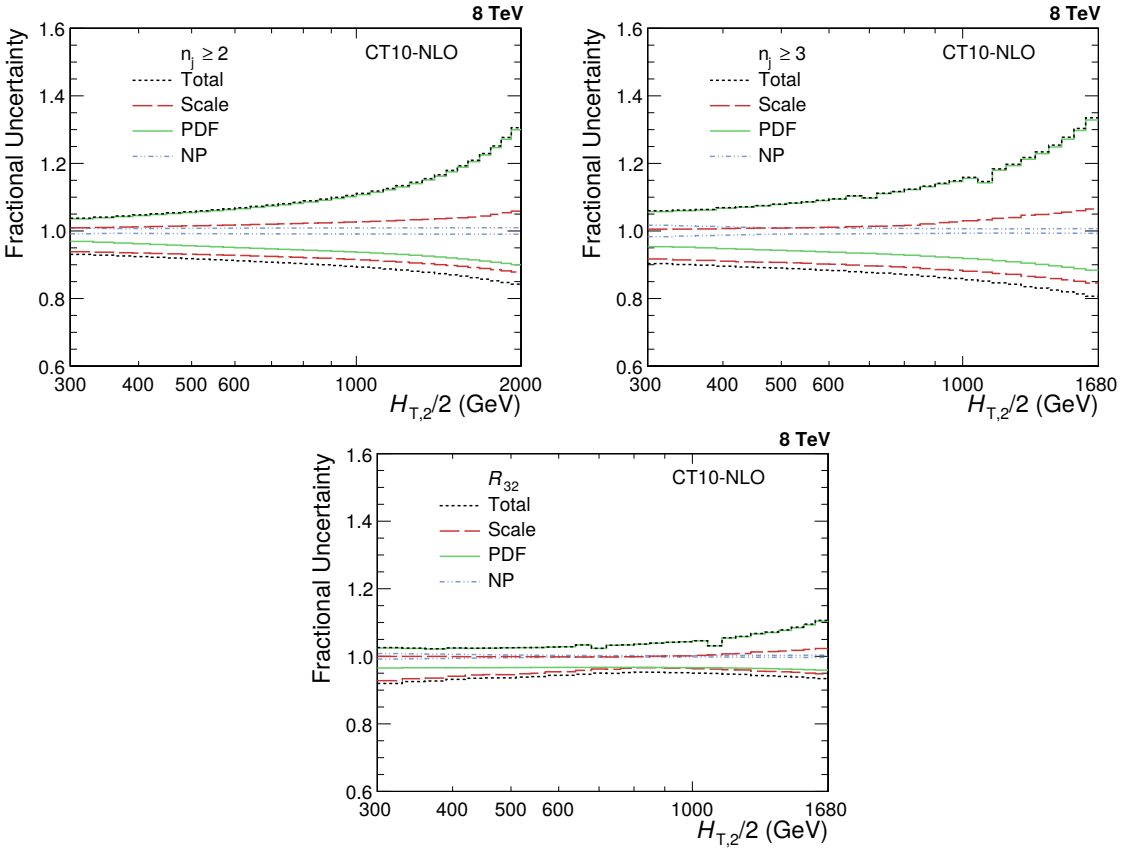

Figure 2.4: The systematic theoretical uncertainties affecting the cross section measurement for inclusive 2-jet (top left) and 3-jet events (top right) and their ratio R_{32} (bottom). The scale (red dashed line), PDF (green line) and NP (blue dashed line) uncertainties as well as total uncertainty (black dashed line) obtained using CT10-NLO PDF set are shown. The total theoretical uncertainty is asymmetric and is dominated by PDF uncertainty.

2.3 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 possible to compare the measured cross sections with the theory predictions. Figure 2.5 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 PDF set (line), corrected for NP effects and in addition for EW effects in the 2-jet case. On the right, the comparison is made to the predictions from MadGraph5+pythia6 (MG+P6)

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

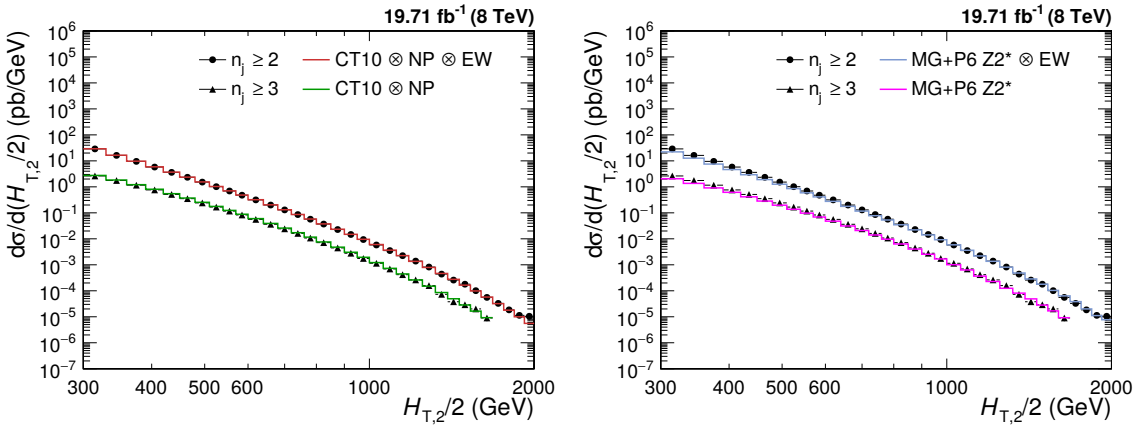

Figure 2.5: Comparison of the measured differential inclusive 2-jet and 3-jet event cross sections as a function of $H_{\mathrm{T},2}/2$ to theoretical predictions. On the left, the data (points) are shown together with NLOJET++ predictions (line) using the CT10 PDF set, corrected for non-perturbative (NP) and electroweak (EW) effects (2-jet) or only NP effects (3-jet). On the (right), the data (points) are compared to predictions from MADGRAPH5+PYTHIA6 (MG+P6) with tune Z2* (line), corrected for EW effects in the 2-jet case. The error bars give the total experimental uncertainty, gievn by the quadrature sum of the statistical and systematic uncertainties.

Figure 2.6 shows the cross section ratio R_{32} obtained from unfolded data 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 represents the total experimental uncertainty derived as quadratic sum from all uncertainty sources. The deviations of measured R_{32} from the theoretical predicted value can be explained by the electroweak effects which are not considered yet because of their unavailability for inclusive 3-jet events cross sections.

For better visibility, the ratios of data over the theory at NLO are also studied in details. In Fig. 2.7, the ratios of data over NLOJET++ predictions using the CT10 PDF set are shown for inclusive 2-jet (top left) and 3-jet event cross sections

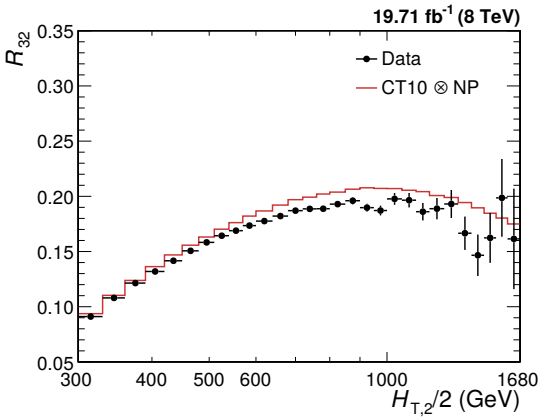

Figure 2.6: Cross section ratio R_{32} as a function of $H_{\mathrm{T},2}/2$ calculated from data (solid circles) in comparison to that from NLO pQCD predictions obtained using the CT10 NLO PDF set corrected with non-perturbative (NP) corrections (line). The error bars correspond to the total experimental uncertainty derived as quadratic sum from all uncertainty sources.

(top right) as well as their ratio R_{32} (bottom). The data are well described by the predictions within their uncertainty, which is dominated at large $H_{\mathrm{T},2}/2$ 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 $H_{\mathrm{T},2}/2$, is remedied by the inclusion of EWK corrections. In the 3-jet case the statistical precision of the data and the reach in $H_{\mathrm{T},2}/2$ is insufficient to observe any effect. The alternative PDF sets MSTW2008 and NNPDF2.3 exhibit a small underestimation of the cross sections at high $H_{\mathrm{T},2}/2$.

The POWHEG framework providing a NLO dijet calculation matched to the parton showers of PYTHIA8 employed with the CUETS1 and CUETM1 tunes [41] is also used for a comparison. The ratios of data over theory from POWHEG+PYTHIA8 with tune CUETS1 are shown for inclusive 2-jet (top left) and 3-jet event cross sections (top right) as well as their ratio R_{32} (bottom) in Fig. 2.8. 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 NLO prediction from POWHEG+PYTHIA8 with tune CUETM1 are shown as well. EW corrections have been accounted for in this comparison in the 2-jet case only. Significant dis-

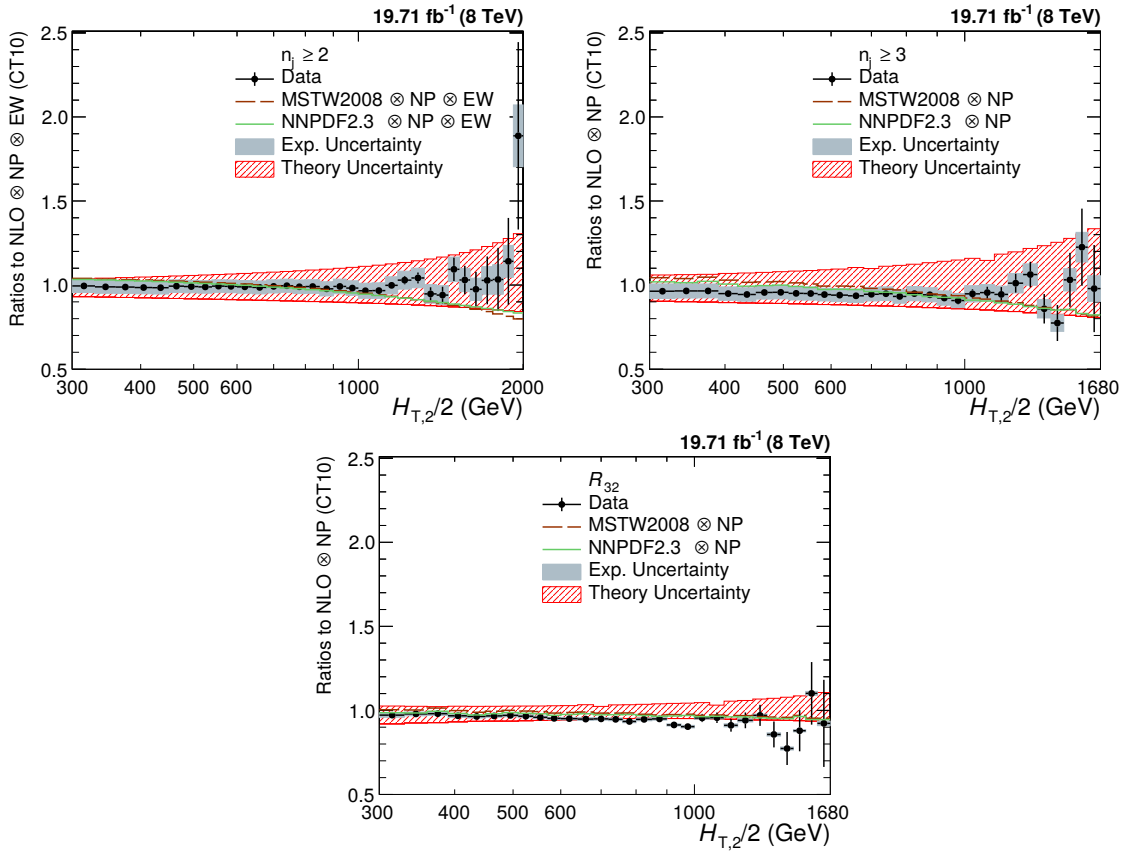

Figure 2.7: Ratio of data over theory using the CT10 PDF set for inclusive 2-jet (top left) and inclusive 3-jet event cross sections (top right) and their ratio R_{32} (bottom). For comparison predictions employing two other PDF sets, MSTW2008 and NNPDF2.3, are also shown. The error bars represents the statistical uncertainty of the data and the shaded rectangles to the total experimental systematic uncertainty. The shaded band around unity indicate the total uncertainty of the theory.

crepancies, 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 $H_{\mathrm{T},2}/2$. In contrast, the employed dijet MC powheg+pythia8 better describe the 2-jet event cross section, but fail for the 3-jet case.

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

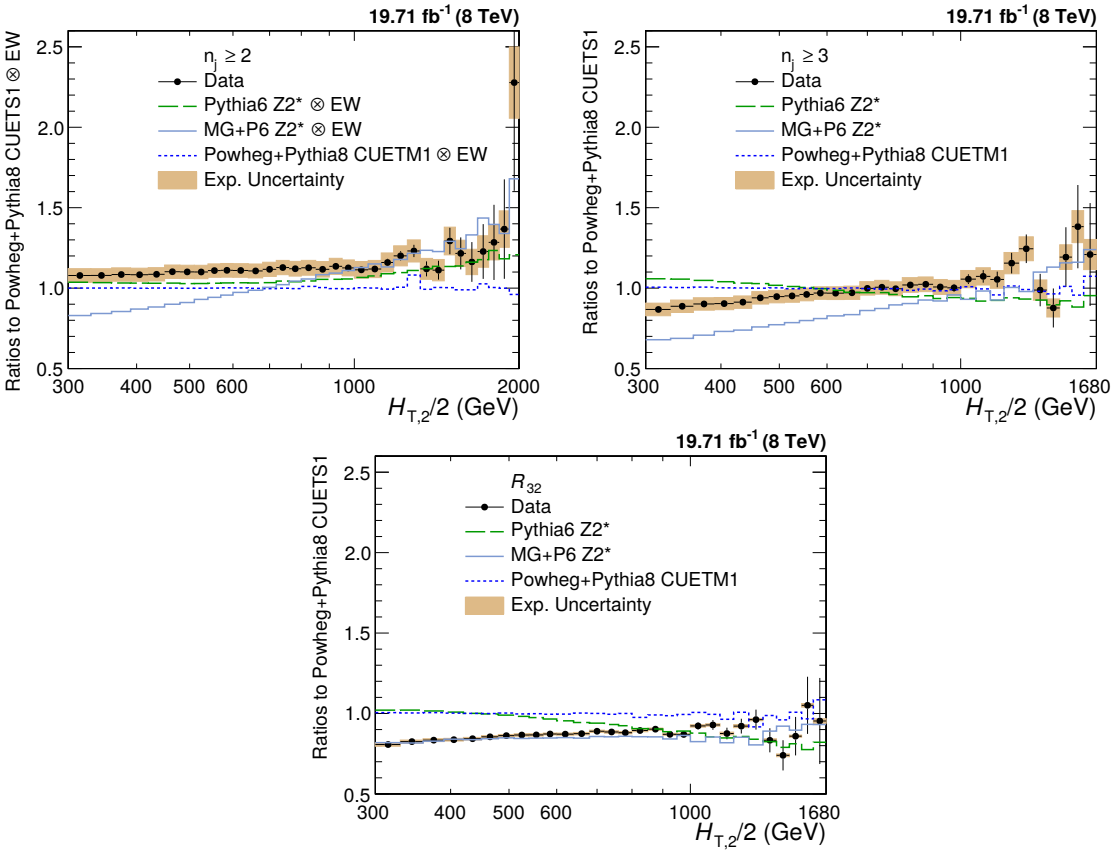

Figure 2.8: Ratio of data over the prediction from POWHEG+PYTHIA8 with tune CUETS1 are presented for inclusive 2-jet (top left) and 3-jet event cross sections (top right) as well as their ratio R_{32} (bottom). For comparison the alternative tune CUETM1 of POWHEG+PYTHIA8, the tree-level multi-leg improved prediction by MADGRAPH5+PYTHIA6 with tune Z2*, and the 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 in the 2-jet case only.