

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2013/05/09

Head Id:

Archive Id: 185341:185359M

Archive Date: 2013/03/12

Archive Tag: trunk

Measurement Of Inclusive Jet Cross Section at $\sqrt{s} = 8$ TeV TeV at CMS

S. Ganguly³, M. Guchait³, M. Kaya¹, O. Kaya¹, K. Kousouris², K. Rabbertz⁴, and M. Voutilainen⁵

¹ Bogazici University, Istanbul, Turkey

² Cern, Geneva, Switzerland

³ Tata Inst. of Fundamental Research, Mumbai, India

⁴ Karlsruhe Institute of Technology

⁵ University of Helsinki

Abstract

A measurement of inclusive jet production cross section is presented. Data from LHC proton-proton collisions at $\sqrt{s} = 8$ TeV, corresponding to $10.71 fb^{-1}$ of integrated luminosity, have been collected with the CMS detector. Jets are reconstructed with the anti- k_T clustering algorithm of size parameter $R = 0.7$, extending to rapidity $|y| = 3$, and transverse momentum $p_T = 2.5$ TeV. The measured cross section is corrected for detector effects and compared with $NLO \times NP$ theory predictions.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: S. Ganguly, M. Guchait, M. Kaya, O. Kaya, K. Kousouris, K. Rabbertz
PDFTitle: Measurement of the Inclusive Jet Cross Section at s=8TeV
PDFSubject: CMS
PDFKeywords: CMS, physics, software, computing

Please also verify that the abstract does not use any user defined symbols

Contents

1	1	Introduction	1
2	2	Datasets and Event Selection	2
3	2.1	Data Samples	2
4	2.2	Jet Reconstruction	2
5	2.3	Trigger Studies	3
6	2.4	Event selection	5
7	2.5	Jet Selection and properties	5
8	2.6	Jet ID efficiency	6
9	2.7	Jet energy resolution	6
10	3	Experimental Measurement	14
11	3.1	Unfolding	14
12	3.2	Experimental Uncertainties	16
13	4	Theory Comparison	32
14	4.1	NLO Calculation	32
15	4.2	Non-Perturbative Effects	32
16	4.3	Theory Uncertainties	33
17	5	Results	34
18	6	Spectrum Fitting	51
19	7	Comparison With 7TeV	51
20	8	PileUp Stability	52
21	9	Summary	52
22			

1 Introduction

The theory of Quantum Chromodynamics(QCD) is one of the fundamental underlying theory to describe interaction among quarks and gluons. In QCD partons(quarks and gluons) are produced in hadron-hadron collision with large cross sections. Experimentally partons are detected as jets. The inclusive jet cross section measurements is one of the process by which QCD can be tested in hadron colliders. In the note, we present our study in the measurements of double differential cross sections for the production of inclusive jets ($p + p \rightarrow \text{jet} + X$) in terms of transverse momentum and rapidity of jets at the center of mass energy $\sqrt{s} = 8 \text{ TeV}$. The similar analysis for $\sqrt{s} = 7 \text{ TeV}$ can be found in [1]. The data analyzed, is collected by Compact Muon Solenoid (CMS) detector at the CERN large Hadron Collider(LHC) during the 2012 run accumulating luminosity of 10.71 fb^{-1} . The measured cross section is corrected for detector effects and compared to the QCD predictions. The momentum fraction of the incoming proton x , carried by the parton, probed in this measurement cover the range $0.019 < x < 0.625$.

Datasets	Run Range	Luminosity(pb ⁻¹)	GT
/Jet/Run2012A	190456-193686	611	FT_53_V10_AN3
/JetMon/Run2012B	193752-197044	4140	FT_53_V6C_AN3
/JetMon/Run2012C	198934-203755	5966	FT_P_V42C_AN3
Pythia6	-	-	START53_V7G

Table 1: Datasets used along with the corresponding run numbers and luminosity.

2 Datasets and Event Selection

2.1 Data Samples

In this current study data collected by CMS at the center of mass energy $\sqrt{S} = 8$ TeV are used for the inclusive jet cross section measurement, corresponding to total luminosity 10.71 fb^{-1} .

For the analysis the reconstructed data is prepared as a ROOT ntuple format, with calorimeter cleaning and basic data quality cuts (JSON file) already performed at the preselection stage. The jet transverse momentum spectra are filled independently for each trigger and E_T^{miss} are applied to clean the samples of residual noise.

The following dataset is used for the inclusive jet cross section:

```
/Jet/Run2012A-13Jul2012-v1/AOD
/JetMon/Run2012B-13Jul2012-v1/AOD
/JetHT/Run2012B-13Jul2012-v1/AOD
/JetMon/Run2012C-PromptReco-v2/AOD
/JetHT/Run2012C-PromptReco-v2/AOD
```

In order to use data of very good quality the, official json file used is the following:

```
Cert_190456-208686_8TeV_PromptReco_Collisions12_JSON.txt
```

The corresponding integrated luminosities calculated by LumiCalc2.py are also presented in Table 1,

2.2 Jet Reconstruction

Jets are reconstructed using particle flow technique [2] using the anti k_T clustering algorithm [3] with size parameter $R=0.7$. In particle flow technique [2] energy and momentum of each particles are reconstructed taking inputs from all sub detector components. Then the cluster is formed taking inputs of each particle momenta which come within the chosen size parameter and then by four momentum summation the total momentum of jets are formed. The rapidity and transverse momentum of jets are calculated using $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ and $p_T = \sqrt{(p_x^2 + p_y^2)}$ respectively. Because of non uniform and as well as non linear response of the CMS calorimeters, the reconstructed jets require additional energy correction, particularly the neutral hadron component of the jet, as the charged hadron component and photons are measured precisely by tracker and ECAL respectively.

The JEC is used based on events generated by PYTHIA6.4.22 (PYTHIA6) [4] and then simulated using GEANT4 [5], and *in situ* measurements with dijet and photon+jet events [6]. In addition an offset corection is also taken into account to subtract out the extra energy deposited in the jets due to pile up. The level of correction obviously depend on the η and p_T of the jet and it is

obtained as a multiplicative factor to the jet four momentum vector. The multiplicative factor is in general smaller than 1.2. For a jet $p_T = 100$ GeV the typical factor is 1.1, decreasing towards 1.0 with increasing p_T .

2.3 Trigger Studies

The CMS trigger system is designed to control event rates consistent with available bandwidth. It consists of two parts, viz. the Level-1 trigger(L1) and High-Level trigger(HLT), where the former one is mainly a hardware based trigger, where as the later one is a software based trigger. In this analysis the trigger paths which were used are single jet triggers viz, L1SingleJet and HLTJet which combinedly forms the HLTJet trigger path. It is to be noted that HLT jets used in the trigger paths are corrected AK5 particle flow jets.

HLTPath	PFJet40	PFJet80	PFJet140	PFJet200	PFJet260	PFJet320
p_T threshold(GeV)	74 - 133	133 - 220	220 - 300	300 - 395	395 - 507	507 - 2500
2012A	60000	3000	230	80	10	1
	0.010	0.195	2.782	14.08	60.57	611.02
2012B	250000	7000	270	70	15	1
	0.017	0.630	16.30	64.08	266.286	4140
2012C	220500	10000	400	80	20	1
	0.020	0.533	15.767	79.292	317.158	5966
Total effective luminosity(pb^{-1})	0.047	1.358	34.85	157.45	644.014	10717.02

Table 2: Prescale factors and effective luminosities(in pb^{-1}) for each trigger path and run era

The current 2012 data samples contain several jet triggers at the level of HLT which are presented in Table 3 along with the corresponding L1 triggers.

L1SingleJet16	L1SingleJet36	L1SingleJet68	L1SingleJet92	L1SingleJet128	L1SingleJet128
HLT_PFJet40	HLT_PFJet80	HLT_PFJet140	HLT_PFJet200	HLT_PFJet260	HLT_PFJet320

Table 3: HLT jet triggers and corresponding L1SingleJet trigger for the 2012 data set.

Because of the high rates of QCD events and existing limited bandwidths (~ 400 Hz) most of these trigger paths are subject to prescale factors. In addition, due to the high instantaneous luminosity in 2012 datasets, pile up may have sizable effects which in turn may affect the trigger efficiencies. In order to obtain the lower threshold with maximum efficiency of the jet triggers we produce trigger turn on curves for each HLT trigger paths. The trigger efficiency, for HLT_PFJetY is defined as:

$$\text{HLT_PFJet}_{\text{eff}}Y = \frac{\text{InclusiveRecoJet_}p_T(\text{HLT_PFJetX} + \text{L1Object_}p_T > Z + \text{HLTObject_}p_T > Y)}{\text{InclusiveRecoJet_}p_T(\text{HLT_PFJetX})} \quad (1)$$

Here the denominator is the number of events for which the trigger path HLT_PFJetX has fired. Here 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. The numerator is the number of events for which HLT_PFJetX has fired and the p_T of L1Object corresponding to the trigger path HLT_JetX is $\geq Z$ (where Z is the L1 seed value corresponding to the trigger path HLT_PFJetY) and the p_T of HLTObject corresponding to the trigger path HLT_JetX is $\geq Y$. This efficiency is plotted against the corrected inclusive RecoJet p_T .

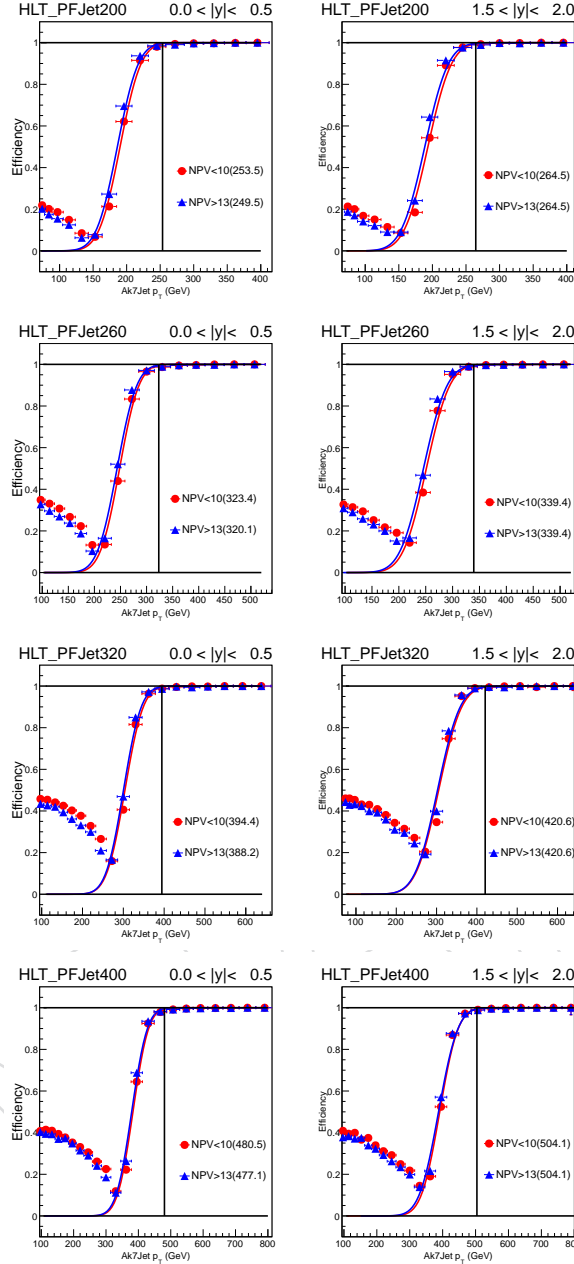


Figure 1: Trigger efficiencies in two different rapidity regions.

For example, in order to obtain turn on curve for HLT_PFJet260, the immediate HLT path of lower threshold HLT_PFJet200 is chosen, the p_T cut on L1Object corresponding to the trigger path HLT_PFJet200 is 92 GeV and p_T cut on HLTObject corresponding to the trigger path HLT_Jet200 is 260 GeV. In Fig. 1, we display the trigger turn on curves for HLT_PFJet140 to HLT_PFJet400 trigger paths for two rapidity regions, for barrel ($0 < |y| < 0.5$) and endcap ($1.5 < |y| < 2.0$) with respect to inclusive jet p_T for PF jets.

The pile up effects are studied by investigating the trigger turn on curve for lower and higher number of primary vertex(NPV). As for example, we study for each case the distribution when number of primary vertex, $NPV < 10$ and $NPV > 13$. The HLT turn on curves are unaffected by pile up(PU) as during the reconstruction of HLTjet the effect of pile-up has been subtracted.

2.4 Event selection

In the analysis to select good events certain event selection cuts are applied. The purpose of these cuts to reduce beam background and some noise. The event selection cuts for each event are following:

- At least one good reconstructed primary vertex with $z < 24$ cm.
- In QCD events has low level of missing energy, so a cut $\frac{E_T^{miss}}{\sum E_T} < 0.3$ helps to select clean QCD events only. In Fig. 2 we see how the Met/SumET behaves for data and MC samples for two extreme triggers.

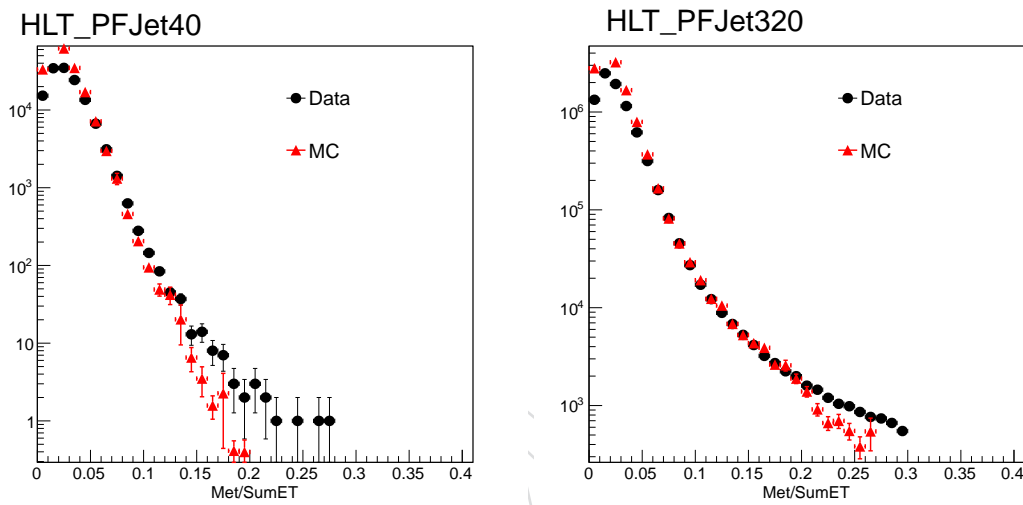


Figure 2: Met/SumET distribution for two extreme Trigger Path

2.5 Jet Selection and properties

In order to obtain good quality of jets, certain identification criteria are imposed, which are as follows:

- at least one PF particle.
- at least one charged hadron and one neutral hadron particle, if $\eta \leq 2.4$
- neutral hadron energy fraction is to be less than 0.9.
- photon energy fraction is to be less than 0.9.
- muon energy fraction is to be less than 0.9
- electron energy fraction is to be less than 0.9

In order to study the Jet compositions and carry out a Data-MontCarlo comparison first we carried out PU reweighting of the MC samples. The PU reweighting has been carried out for each individual trigger path. In Fig. 3 the comparison between data and MC shown after PU reweighting, which shows that after PU reweighting the agreement between data and MC distribution are moderate confirming that PU reweighting is working at a significant level.

Also a p_T threshold on the jets for each individual trigger path are used to carry out the data-MC comparison to ensure that in both the cases we are accessing the same region in phase space. In table. 4 we list the p_T threshold used for individual trigger paths.

$p_T(\text{GeV})$	74	133	220	300	395	507
HLT Path	HLT_PFJet40	HLT_PFJet80	HLT_PFJet140	HLT_PFJet200	HLT_PFJet260	HLT_PFJet320

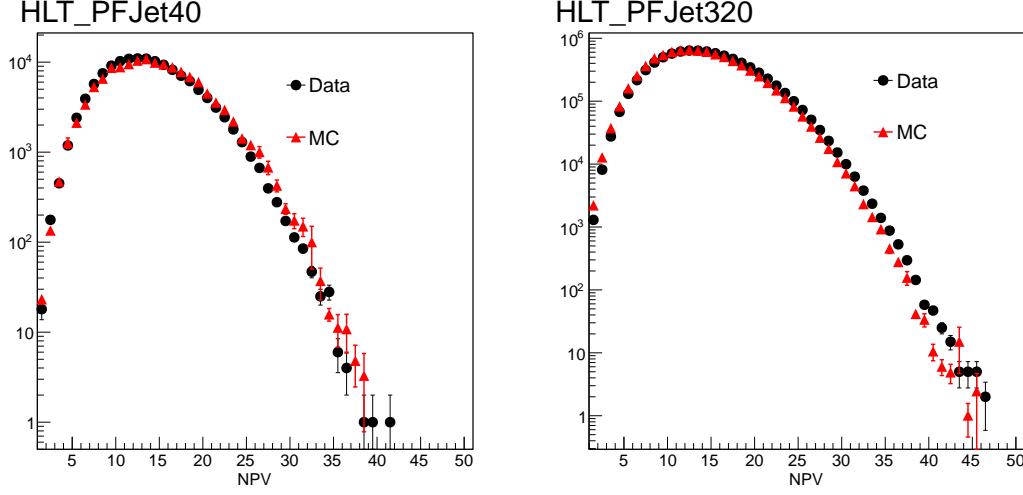
Table 4: HLT jet triggers and corresponding p_T cut.

Figure 3: Primary Vertex Distribution for two extreme Trigger paths

In the following Figures we present the compositions of jets for various p_T threshold and for different η regions. These are also compared with the predictions from MC.

2.6 Jet ID efficiency

In order to estimate the jet ID efficiency, a "tag and probe" method is employed, using dijet events. The two leading jets are required to be back-to-back in the azimuthal plane $\Delta\Phi > 2.7$. One jet is selected randomly as a "tag" jet, required to satisfy the ID criteria, while the other jet is called "probe" jet with no further cuts applied. 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. Figure 10 shows the ID efficiency for different rapidity bins. As expected, the jet ID efficiency is close to 100%.

2.7 Jet energy resolution

In the ongoing analysis we have studied the Jet Energy Resolution from monte carlo samples. We have made the distribution of $\frac{p_T(\text{Reco})}{p_T(\text{Gen})}$, where the chosen Gen Jet has the least $\Delta R(\sqrt{\Delta\eta^2 + \Delta\phi^2})$ with the corresponding Reco jet, and extracted the Gaussian fitted resolution from above distribution. This resolution is plotted against the Gen jet p_T . The distribution is fitted to obtain the average behaviour of jet energy resolution as a function of Gen jet.

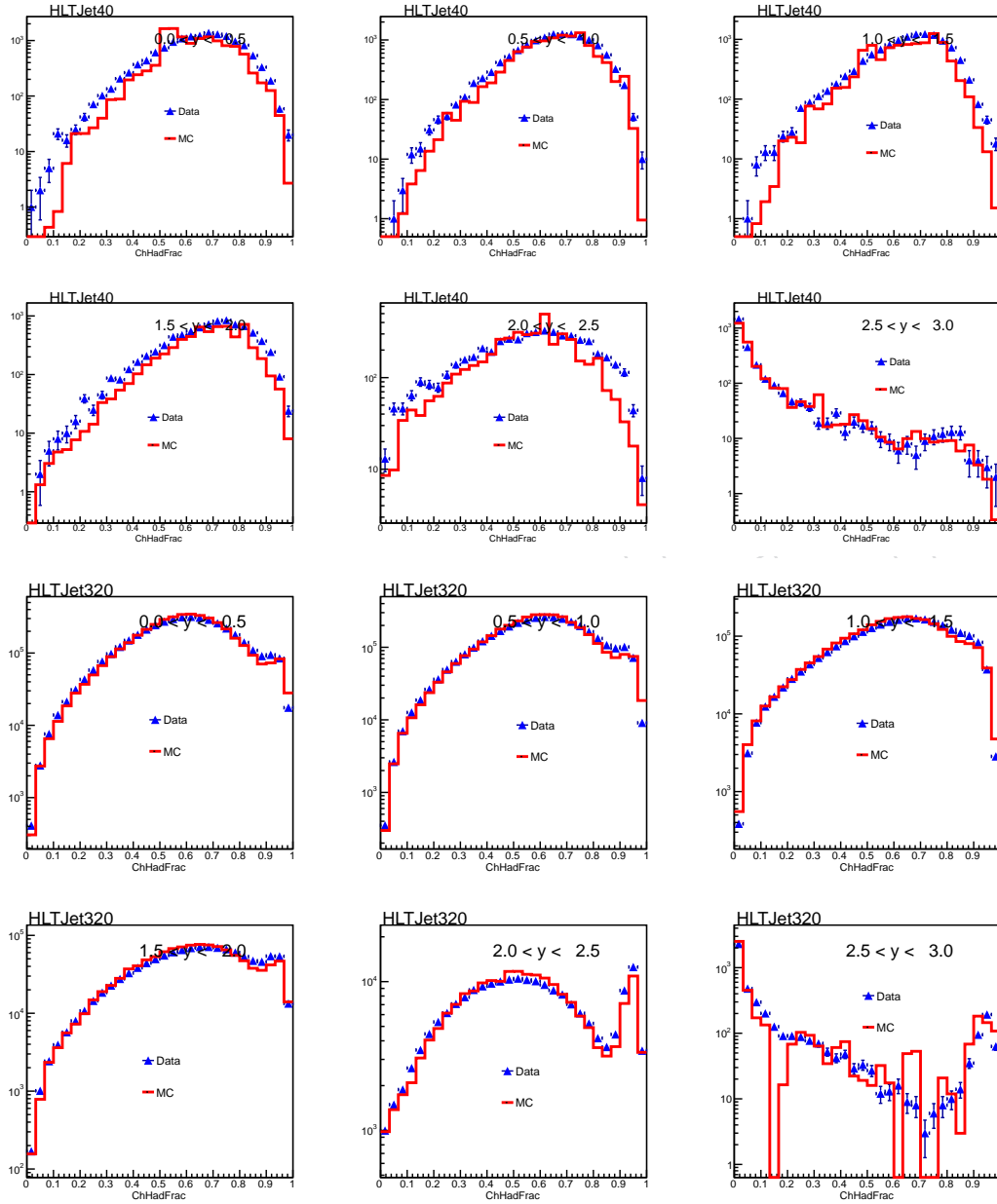


Figure 4: Charged hadron energy fraction in the jets for two extreme trigger paths for six rapidity bins.

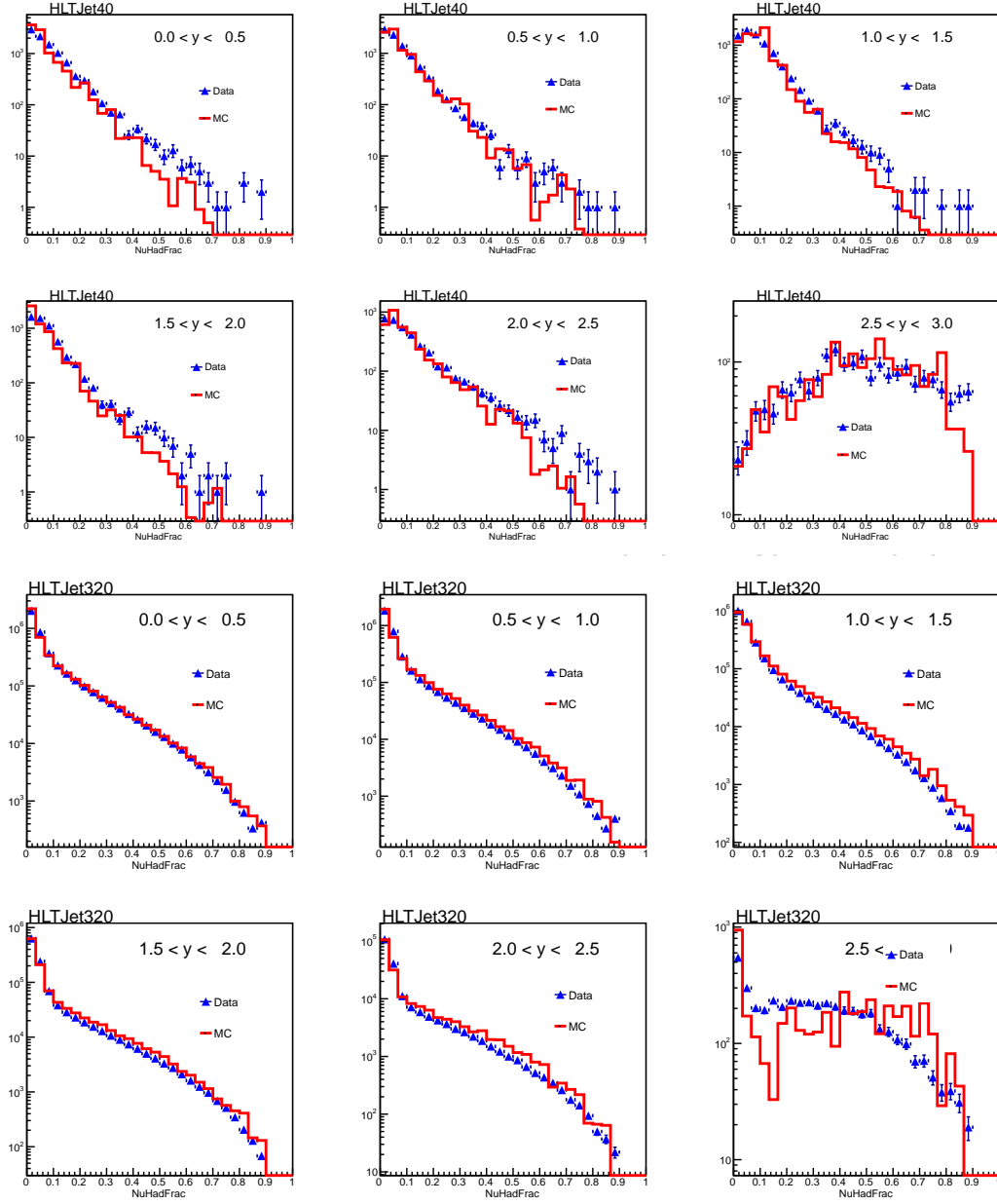


Figure 5: Neutral hadron energy fraction in the jets for two extreme trigger paths for six rapidity bins.

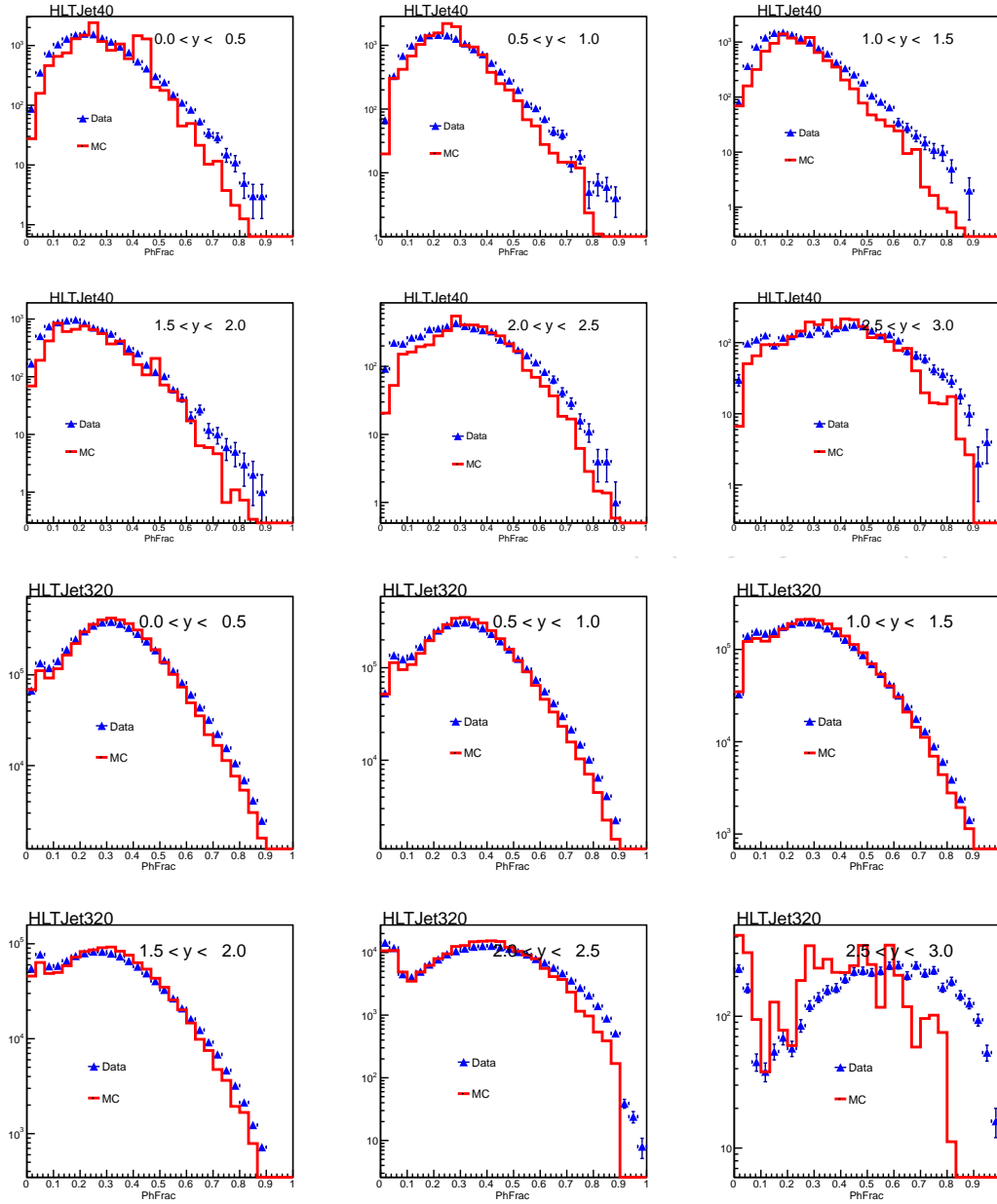


Figure 6: Photon energy fraction in the jets for two extreme trigger paths for six rapidity bins.

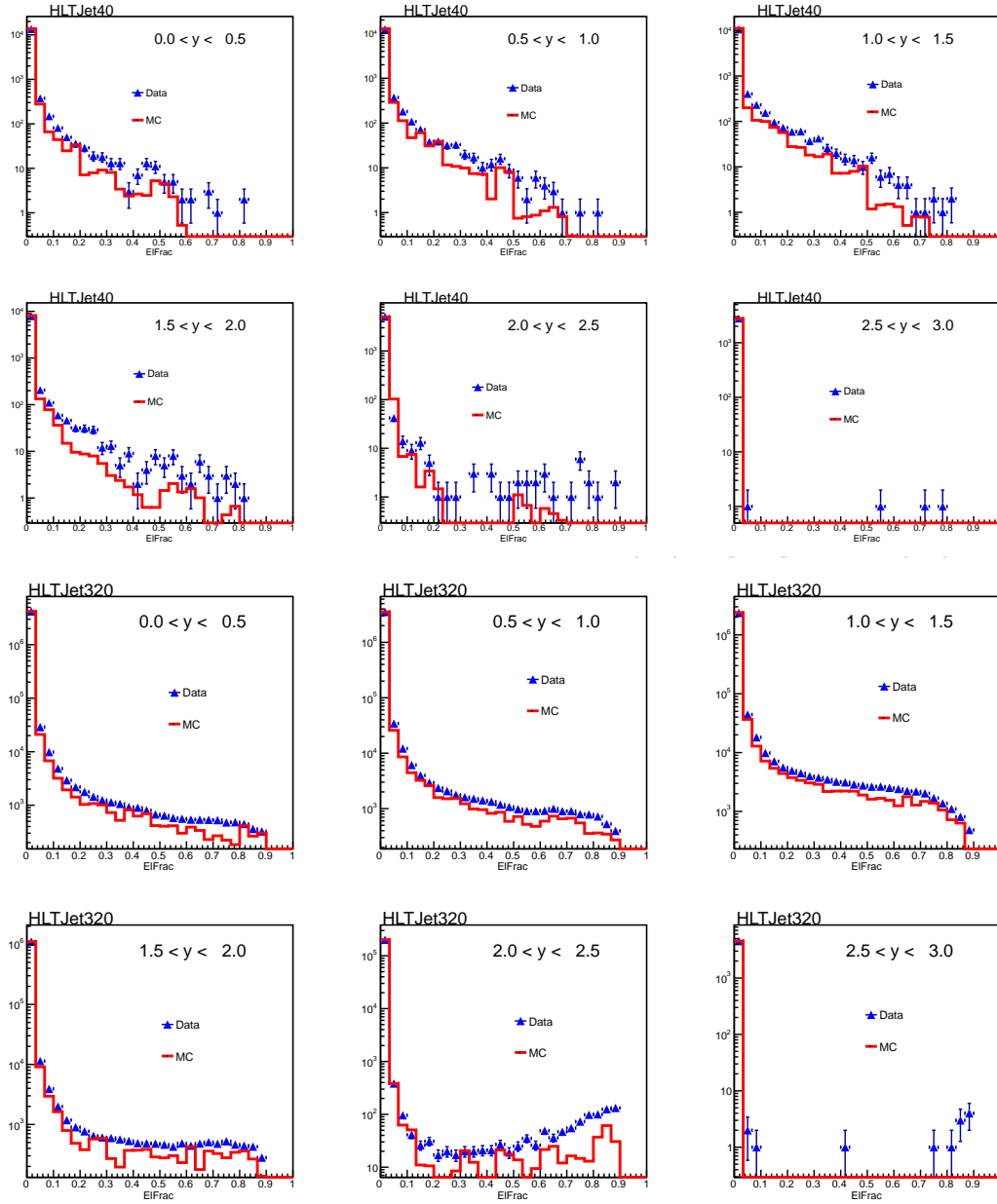


Figure 7: Electron energy fraction in the jets for two extreme trigger paths for six rapidity bins.

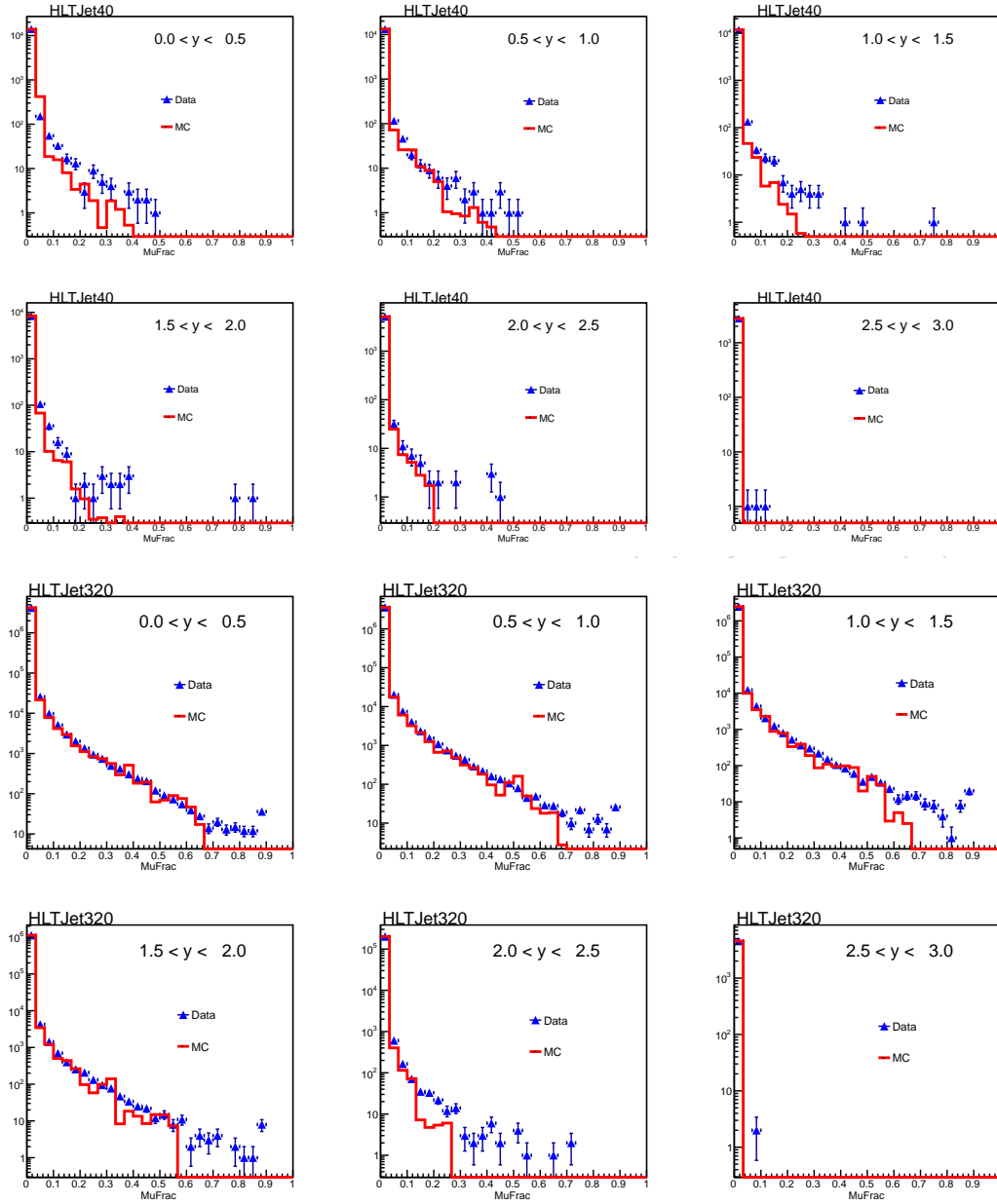


Figure 8: Muon energy fraction in the jets for two extreme trigger paths for six rapidity bins.

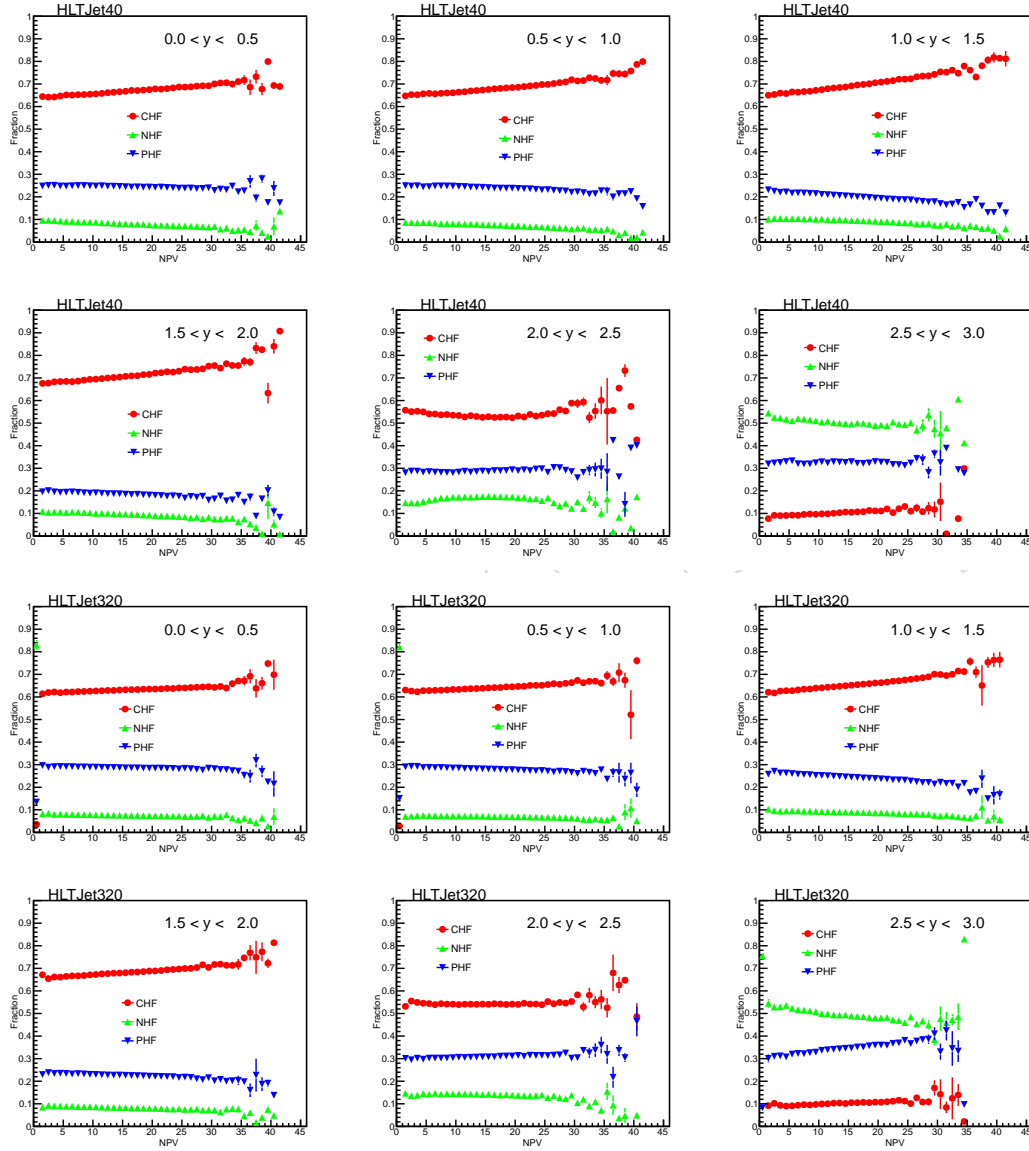


Figure 9: Average energy fractions vs PileUp

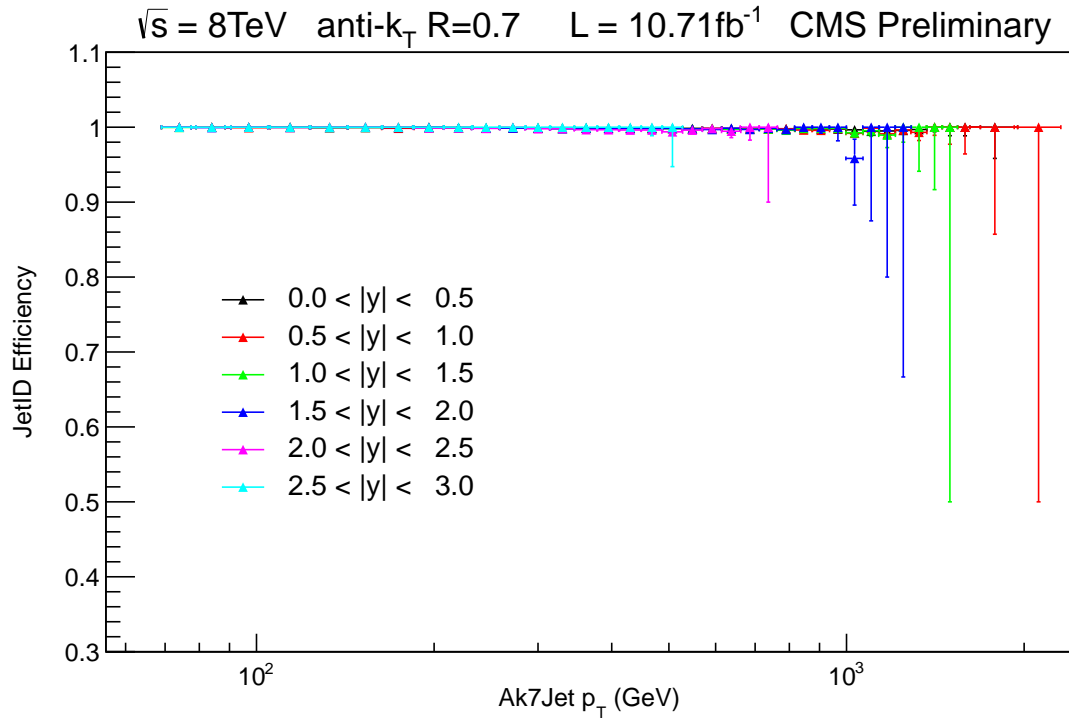


Figure 10: Jet ID efficiency determined by Tag and Probe method for individual rapidity bins.

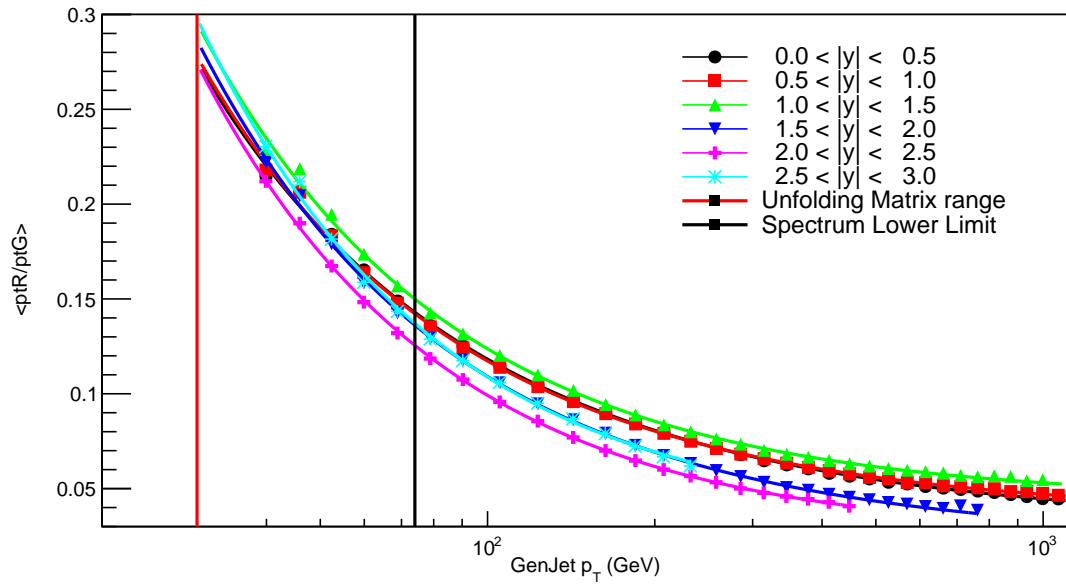


Figure 11: Fitted Jet Energy Resolution determined from monte Carlo for individual rapidity bins.

3 Experimental Measurement

The spectrum construction is achieved by combining the p_T spectra from individual trigger paths. In order to avoid double counting of jets, in each jet p_T bin only the lowset prescale, fully efficient trigger contributes.

The measured yields are transformed to double-differential cross sections as:

$$\frac{d^2\sigma}{dp_T d|y|} = \frac{1}{\epsilon \mathcal{L}_{\text{eff}}} \frac{N}{\Delta p_T \Delta |y|}, \quad (2)$$

where N is the number of jets in the bin, \mathcal{L}_{eff} is the effective integrated luminosity, taking into account the trigger prescales, of the data sample from which the events are taken. Here ϵ is the product of the trigger and jet selection efficiencies (greater than 99%), and Δp_T and $\Delta |y|$ are the transverse momentum and rapidity bin widths, respectively. The width of the p_T bins is progressively increased, proportional to the p_T resolution. Figure 12 shows the inclusive jet cross sections (yet uncorrected for resolution smearing effects).

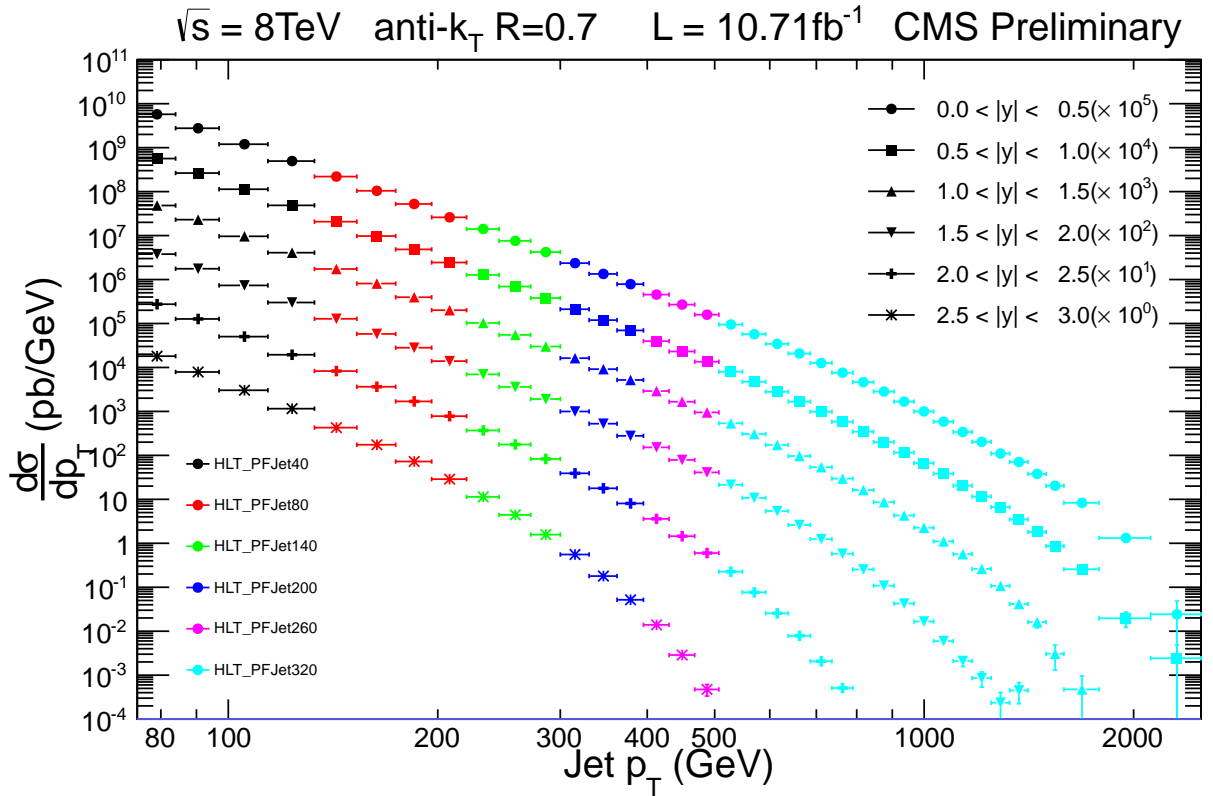


Figure 12: Inclusive jet cross sections, uncorrected for resolution smearing effects.

In table 5 we tabulate the jet yield for each rapidity bins for all the p_T bins. The effect of trigger prescale factors are prominent from the variation of jet yield along the p_T bins.

3.1 Unfolding

The finite detector resolution along with the steeply falling jet p_T spectrum distorts the measured cross section with respect to the particle level cross-section. Each p_T bin content is a composition of original events which are also in the same bin at the particle level and that of events

which migrated from neighbouring bins. In order to allow a direct comparison of experimental measurements with corresponding results from other experiments, an unfolding process of the data for detector effects should be followed. The response matrix R in the unfolding procedure is usually derived using simulated Monte Carlo (MC) samples. The CMS detector response of the simulation is based on GEANT4 [5]. There are several drawbacks of constructing response matrix from the MC. First of all, MC is not perfectly calibrated. Also, number of events used in the MC simulation is usually a limiting factor for estimating the transition rates across the p_T bins, especially in outer rapidities. Dealing with the systematic uncertainties is practically impossible with this technique. However, there is an indirect way of constructing the response matrix which uses a custom Toy Monte Carlo by utilizing MC simulation. Figure 13 shows the Response matrix generated from toy montecarlo. Here the true p_T binning goes much lower than the actual measurement to get least effect at the boundries. Jet p_T values at the particle level are generated randomly, following the spectrum predicted by NLOJET++ and FastNLO, and are subsequently smeared with a Gaussian function centered at the generated response curve. The σ of the Gaussian is determined from the relative resolution parametrization. Jet energy resolution figure shows the fitted jet energy resolution(JER). The JER values obtained from this fit are further modified by the results obtained from real data. The modified values have been used to smear the generated toy MC spectrum. These generated and smeared values are used to fill the response matrix object by using the RooUnfold package [7].

First we unfold the smeared toy montecarlo spectrum using the constructed response matrices and check how good the unfolding procedure is working. Figure 14 shows that after unfolding, the smeared toy mc matches exactly with the randomly generated mc spectrum.

In figures 18, 19, we show the the closure test done for two different theory spectrum generated by $NLO \times NP$ CT10 spectrum and LO flat Pythia6(tune Z2star) spectrum. In both the cases the smeared spectrum is unfolded using the previously constructed response matrix and a ratio of unfolded to gen spectrum is studied to check howmuch the unfolding closure test works across varying theory spectra. From these two figures we can see that within statistical fluctuations, the unfolding produces almost the generator level spectrum.

Then, Bayesian Unfolding method is applied to unfold the spectrum constructed from the data.

Finally, the technique is used to unfold the final measured spectra. In figure 15 and 16, it can be seen that the unfolding corrections to the measured spectra is always greater than unity as it is expected. The phenomena holds true for both Bayesian and BinByBin unfolding methods. In this figure we also see that in the p_T region of interest, the Bayesian unfolded over measured spectrum distribution is very close to the distribution of true over smeared distribution obtained from toy monte carlo. The shape of the curves is closely related to the p_T dependence of the resolution and the spectrum slope. At lower p_T values the resolution is relatively worse than it is at higher p_T values and the effect of smearing is larger. At higher p_T , even if the resolution improves, the spectrum becomes steeper which leads again to a larger smearing effect.

Figure 17 shows the comparison of relative error before and after unfolding. We clearly see fom this figure that the unfolding procedure always introducess some excess amount of statistical fluctuation.

We extract the JER by fitting the core of the jet response (p_{T_Reco}/p_{T_Gen}) by Gaussian distribution. It is also cross checked that whether the tail of the response function is effecting the JER and hence the unfolding significantly or not. To do this, we fit the response distribution by double Gaussian function, so that the right part of the response tail is properly covered. Then JER is derived from the width of the fitted double Gaussian function. Figure 21 shows an example plot of the fitting of the response by double gaussian function. Unfolding is repeated

by the new set of JER parameters. A ratio of the two sets of unfolded spectrum, obtained with two different JER parameters, is obtained as a function of jet p_T . Figure 22 shows that within statistical fluctuation, the non-gaussian tail parts of the response distribution doesn't produce significant changes to the final unfolded spectrum.

The jet energy resolution (JER) is taken from montecarlo simulation and further corrected by the factors derived from data. In general the JER in simulated MC is 10% different than what is measured from data. The jet energy resolution $\sigma(p_T)$ is parametrized as a function of jet p_T as [6]

$$\frac{\sigma(p_T)}{p_T} = \sqrt{\frac{N^2}{p_T^2} + \frac{S^2}{p_T} + C^2} \quad (3)$$

The parameters N, S, C used in the above equation are tabulated in table 6.

The final statistical errors include the correlations between the bins.

In the Bayesian unfolding process[8, 9], the number of estimated events in the i -th bin of the unfolded distribution ('estimated causes') $\hat{n}(C_i)$, as the result of applying the unfolding matrix M_{ij} on the j -th bin of raw distribution ('effects') $n(E_j)$, is given by:

$$\hat{n}(C_i) = \sum_{j=1}^{n_E} M_{ij} n(E_j)$$

where

$$M_{ij} = \frac{P(E_j|C_i)n_0(C_i)}{\epsilon_i \sum_{l=1}^{n_C} P(E_j|C_l)n_0(C_l)}$$

Here $P(E_j|C_i)$ is the $n_E \times n_C$ response matrix, where n_E and n_C are the number of bins in raw and unfolded distributions respectively. This response matrix causes correlation among different bins in the unfolded distribution. Here $\epsilon_i = \sum_{j=1}^{n_E} P(E_j|C_i)$ are efficiencies for each bin and $n_0(C_l)$ is the number of events in the l -th bin of the prior distribution. Output of each iteration of unfolding, goes as prior distribution to the next iteration.

The order of correlation between different bins vary between 10% to 20% for this measurement.

To construct the response matrix, the theory spectrum is fitted with a function

$$f(x) = A_0 \left(\frac{x}{A_3} \right)^{-A_1} \left(1 - \frac{x}{A_3} \right)^{A_2} \quad (4)$$

We build the response matrices using $NLO \times NP$ NNPDF spectrum. Figure 23 shows the distribution of relative difference between the theory and fitted distribution. The pull distribution for the same theory distribution and fitted functions are shown in figure 24.

3.2 Experimental Uncertainties

There is an asymmetric systematic uncertainty on the measured cross section due to Jet Energy Scale(JES), which is almost same for different rapidity bins. The JES varies between 2% and 4% and it gives rise to 10% to 15% uncertainty on the measured cross section. This is estimated by shifting the jet p_T according to the JES uncertainty. The estimation of uncertainty on integrated luminosity is 5% which directly goes to the measurement of cross section. The jet-energy resolution uncertainty of 10% affects the unfolding, and introduces a 1%-7% uncertainty on cross section measurement.

The individual uncertainty sources, which are uncorrelated among each other are the following: Absolute, HighPtExtra, SinglePion[ECAL][HCAL], Flavor, Time, RelativeJER[EC1][EC2][HF], RelativePt[EC1][EC2][HF], RelativeStat[EC2][HF], RelativeSample, PileUpDataMC, PileUpPt[BB][EC][HF], PileUpBias .

Other sources of experimental uncertainty, such as the jet angular resolution and the theoretical p_T spectrum that is used to calculate the response matrix, introduce negligible uncertainties on the measured cross section.

In Figure 25 we can see several individual components of JES and the total jet energy Scale Uncertainty as a function of Jet p_T for different rapidity bins.

DRAFT

pT (GeV)	Bin	YBin1	YBin2	YBin3	YBin4	YBin5	YBin6
74-84	1	26142	25511	21820	17289	12733	8520
84-97	2	16369	15621	13582	10427	7634	4885
97-114	3	9303	8715	7490	5749	3927	2457
114-133	4	4268	4149	3536	2683	1725	1039
133-153	5	55847	52487	43968	32715	21455	11408
153-174	6	27774	25886	21489	15552	10001	4880
174-196	7	14638	13507	11042	7897	4788	2116
196-220	8	7941	7384	6167	4261	2466	930
220-245	9	115077	103714	83979	57233	30548	9611
245-272	10	66962	61256	48594	31857	15849	4049
272-300	11	38474	34680	27328	17517	7717	1502
300-330	12	104201	93158	71795	44085	17404	2505
330-362	13	63154	56148	43023	24868	8461	863
362-395	14	38089	33808	25160	13562	3955	258
395-430	15	95734	83357	61129	32115	7633	304
430-468	16	61603	52833	37990	18068	3338	67
468-507	17	37381	31718	22222	9692	1419	12
507-548	18	415878	351136	235840	94180	9944	43
548-592	19	268878	223076	143787	50849	3618	
592-638	20	169226	137889	85055	26587	1260	
638-686	21	106712	85867	49747	13412	405	
686-737	22	68958	54045	29417	6807	113	
737-790	23	43139	33161	16843	3288	29	
790-846	24	27733	20773	9749	1523	5	
846-905	25	17902	12609	5439	686	1	
905-967	26	11113	7673	2858	283		
967-1032	27	6966	4638	1575	116		
1032-1101	28	4332	2836	818	44		
1101-1172	29	2593	1561	431	16		
1172-1248	30	1656	943	213	7		
1248-1327	31	933	563	91	2		
1327-1410	32	630	310	37	4		
1410-1497	33	354	169	15			
1497-1588	34	199	84	3			
1588-1784	35	175	54	1			
1784-2116	36	47	7	0			
2116-2500	37	1	1				

Table 5: Jet Yield for each p_T bin for different rapidity binsTable 6: The values of the JER parameters N , S and C used in equation 3

$ y $ Bin	0.0 – 0.5	0.5 – 1.0	1.0 – 1.5	1.5 – 2.0	2.0 – 2.5	2.5 – 3.0
N	0.032	0.038	0.043	0.017	0.016	0.022
S	1.136	1.053	1.112	1.034	0.878	0.945
C	5.886	6.577	6.743	8.147	8.334	7.534

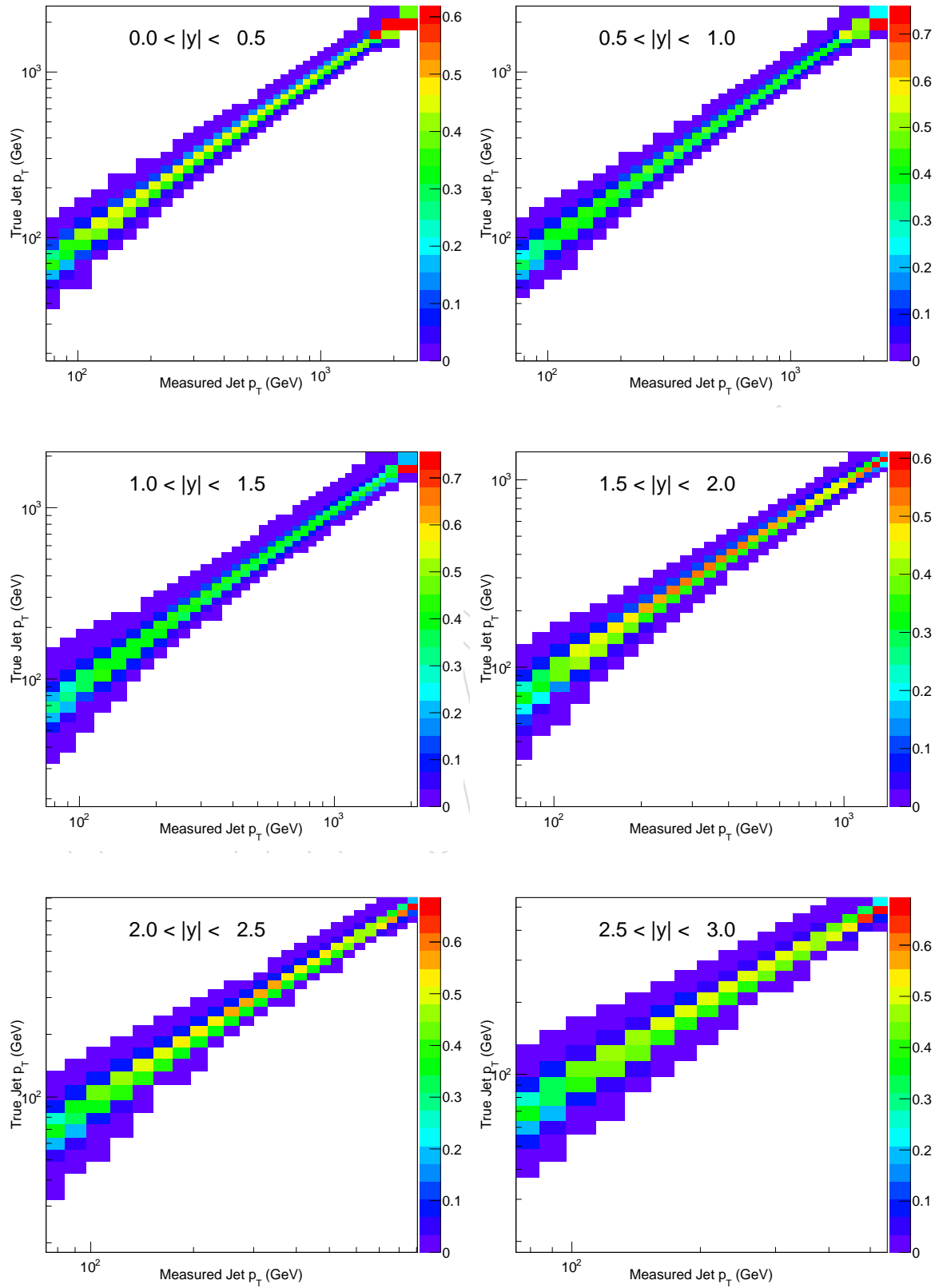


Figure 13: Response Matrix for each Rapidity Bin constructed from Toy MC.

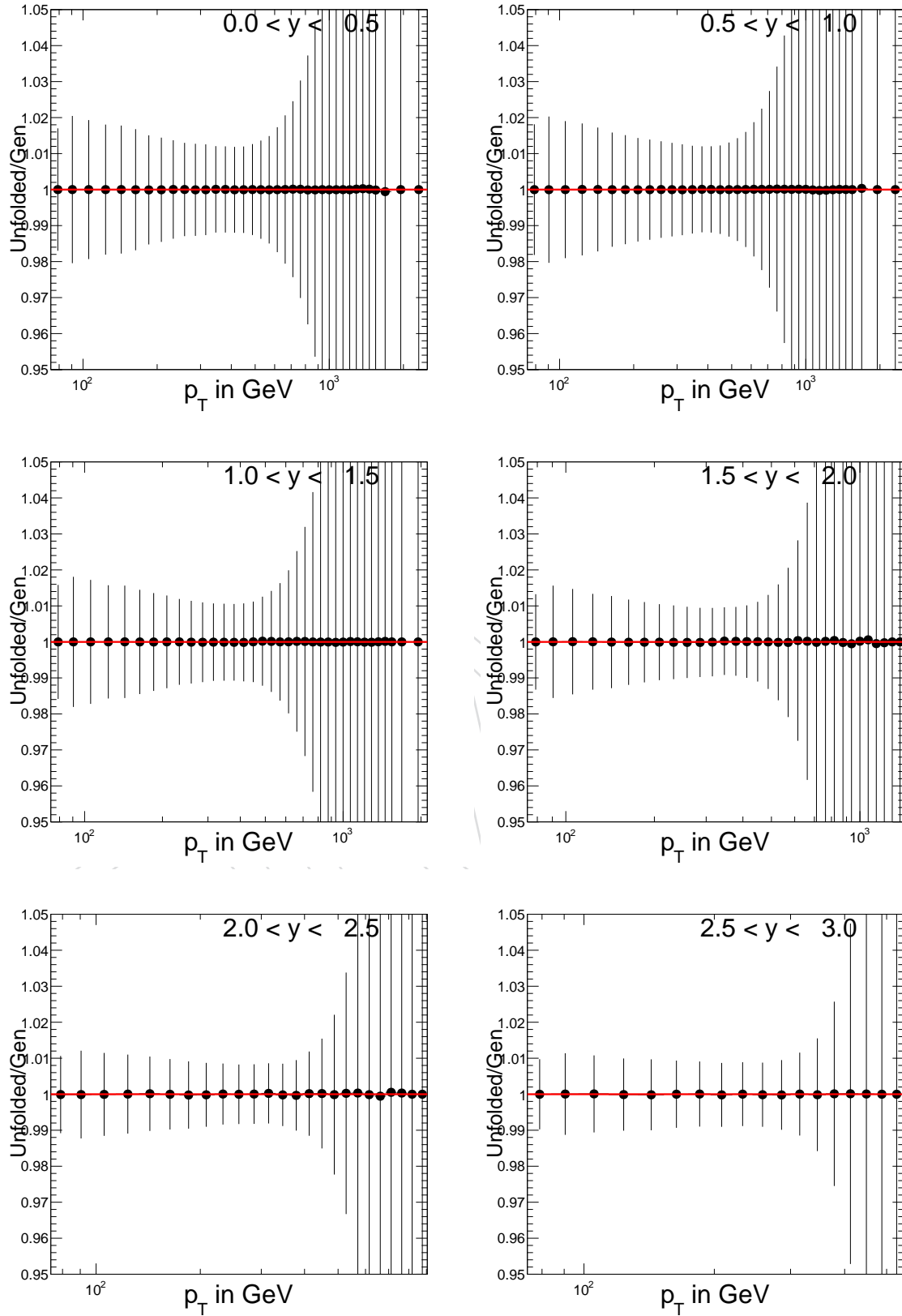


Figure 14: Closure test for unfolding from Toy-MC shows that the unfolding procedure is working correctly for each rapidity bins across all the p_T bins.

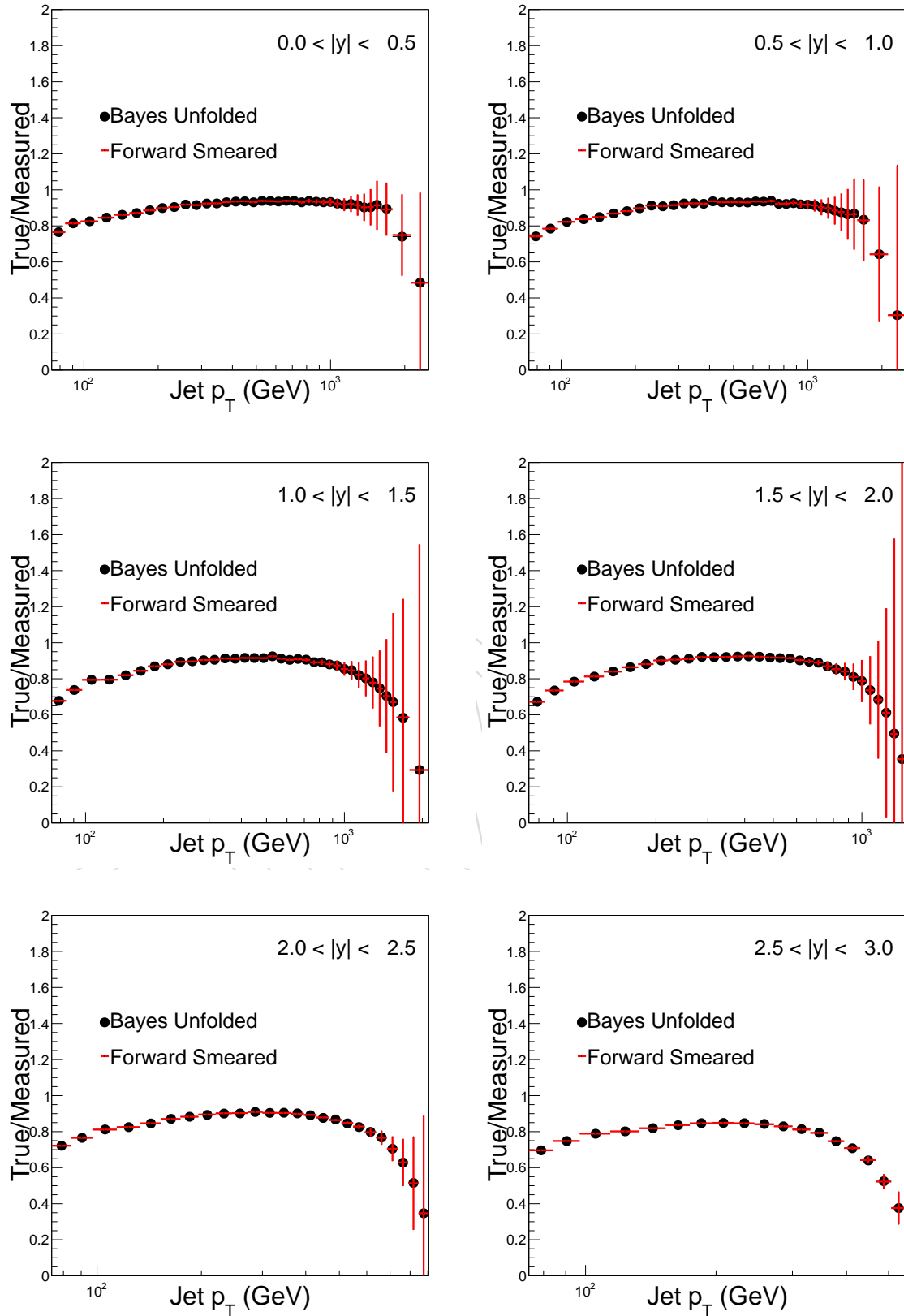


Figure 15: Comparison between Bayesian and BinByBin unfolding method for Toy MC for each Rapidity Bin

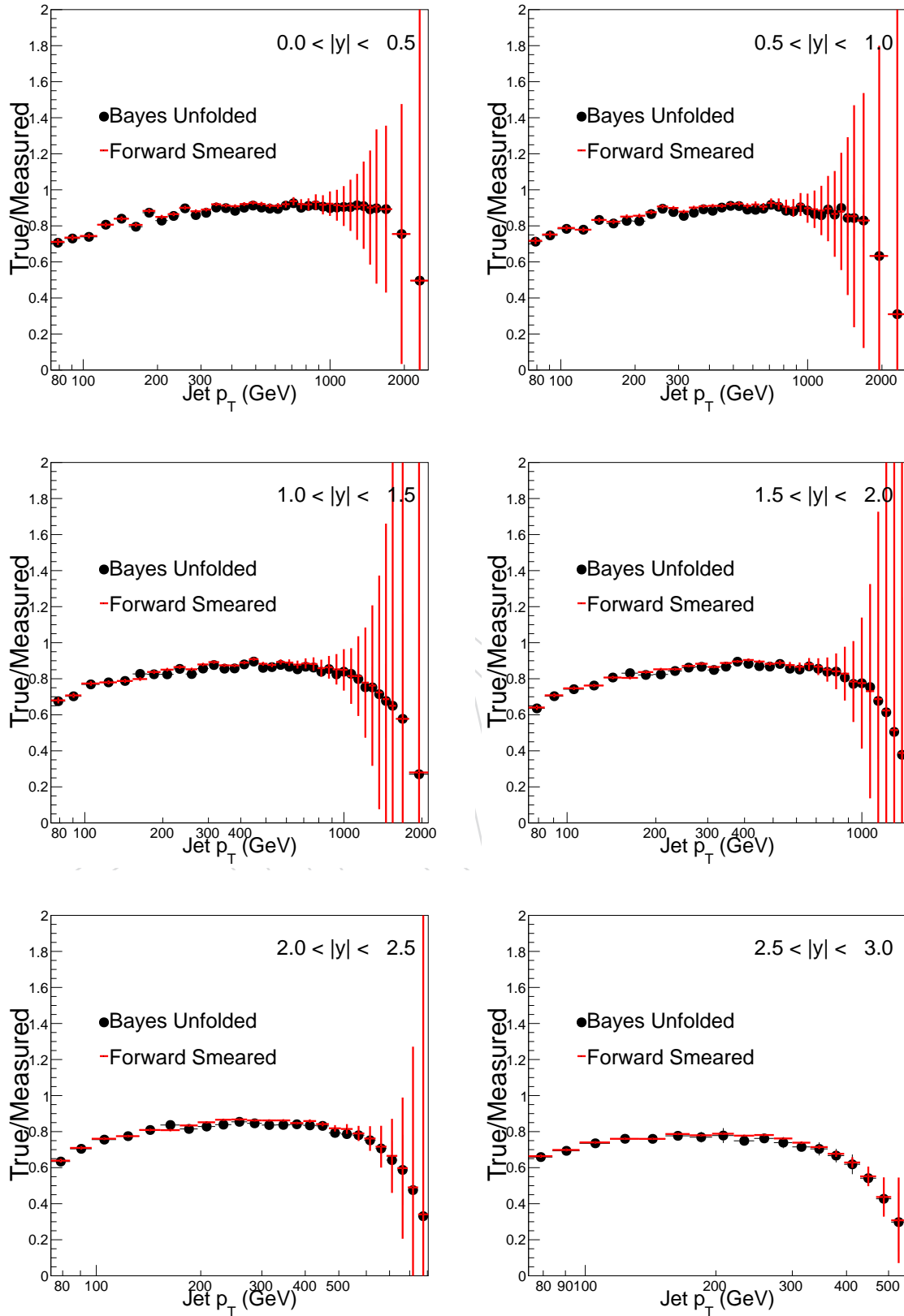


Figure 16: Comparison between Bayesian and BinByBin unfolding method for the measured spectrum for each Rapidity Bin

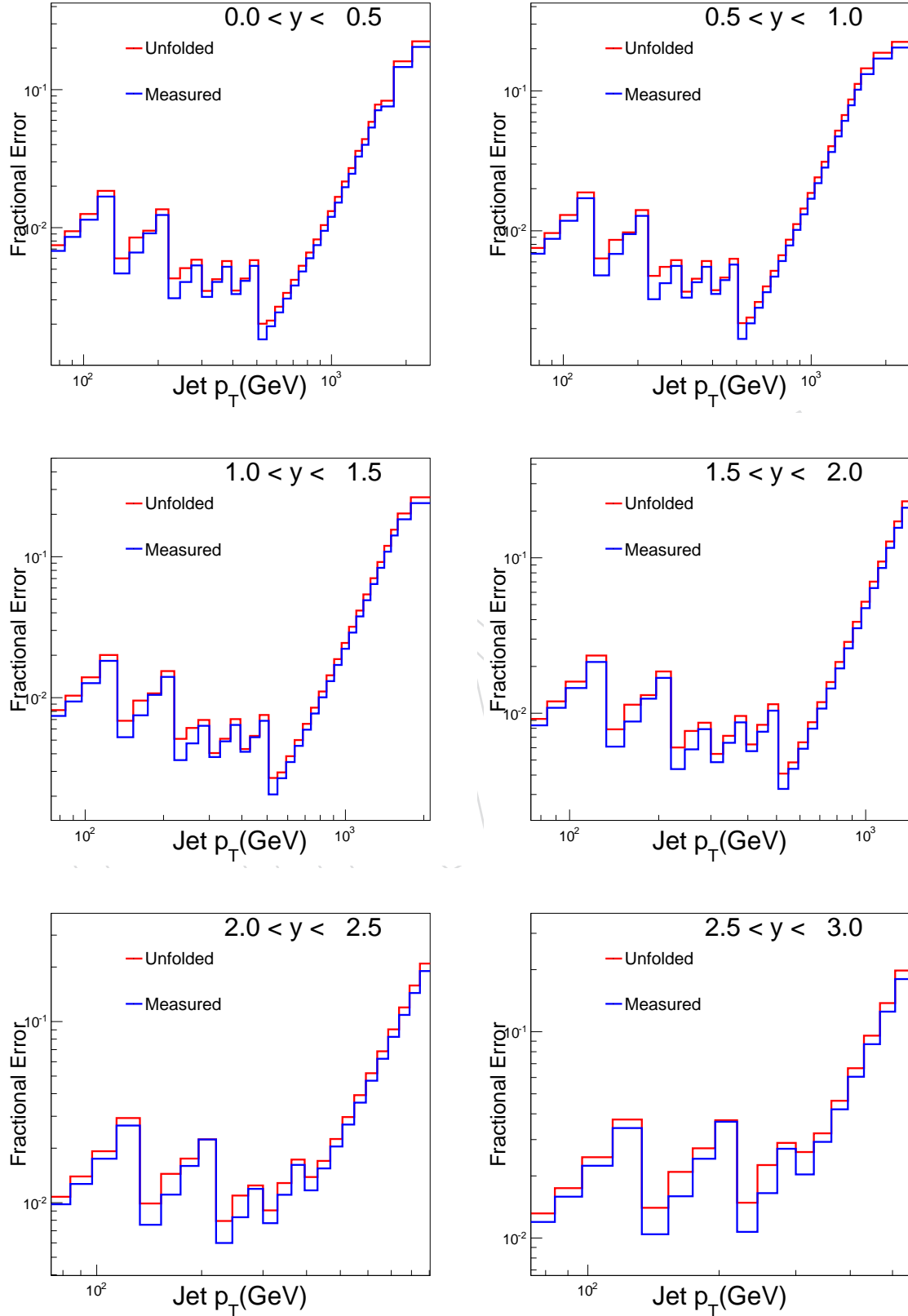


Figure 17: Relative Fractional Error for each Rapidity Bin before and after unfolding

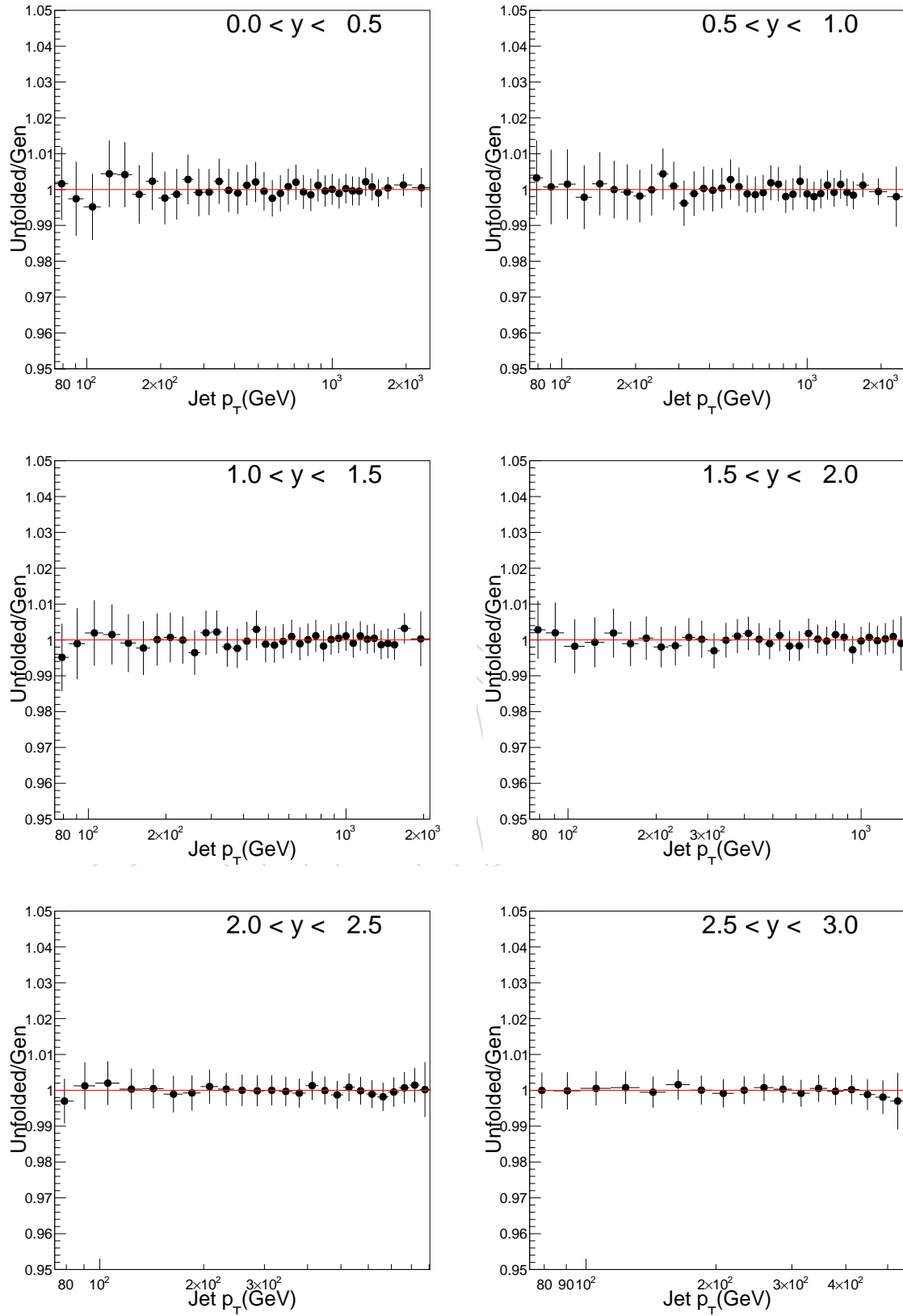


Figure 18: Ratio of unfolded CT10 reco spectrum to CT10 gen spectrum. CT10 spectrum unfolded using NNPDF response matrix. Within statistical error bars the closure is well established

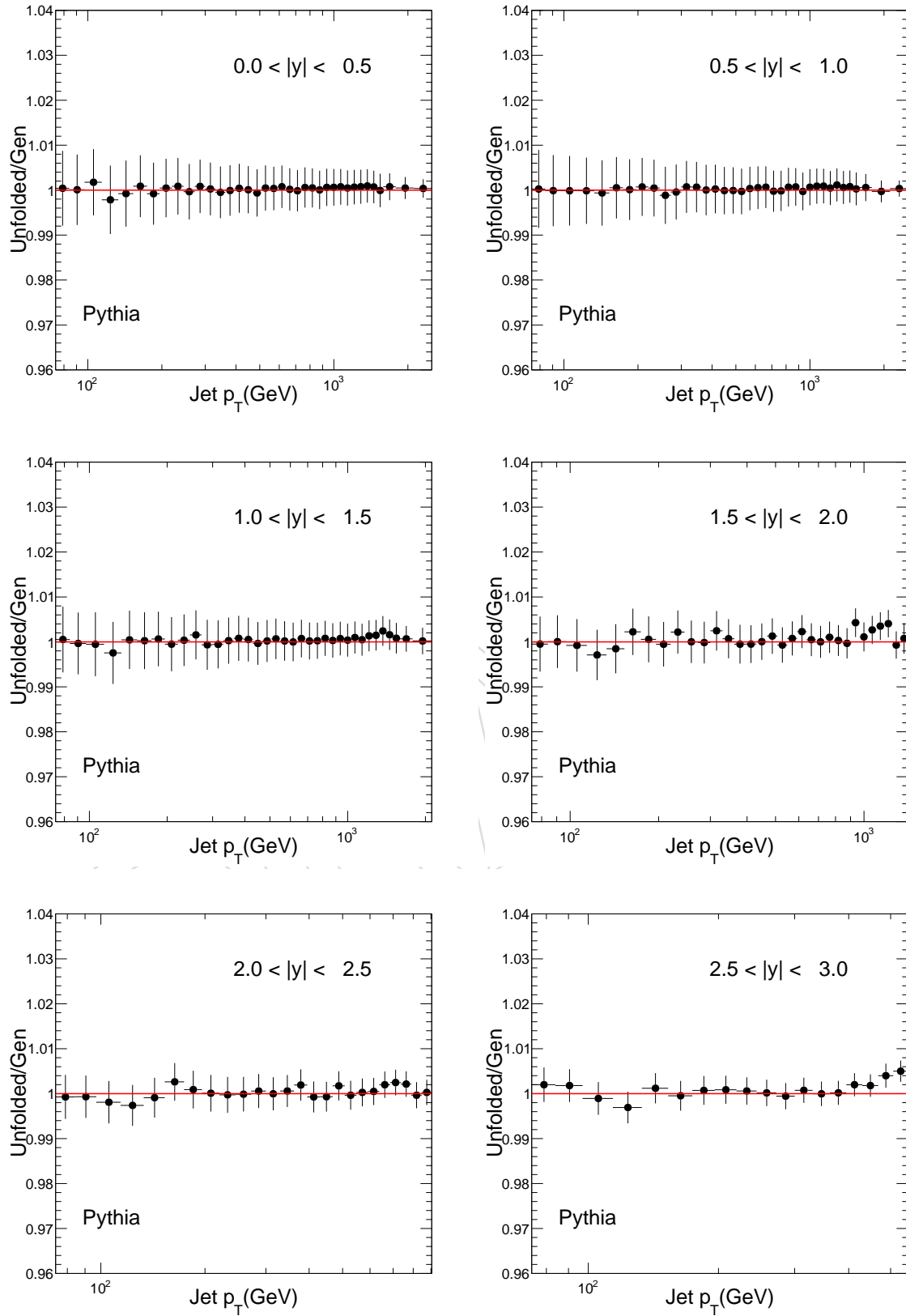


Figure 19: Ratio of unfolded Pythia reco spectrum to Pythia gen spectrum. Pythia spectrum unfolded using NNPDF response matrix. Within statistical error bars the closure is well established

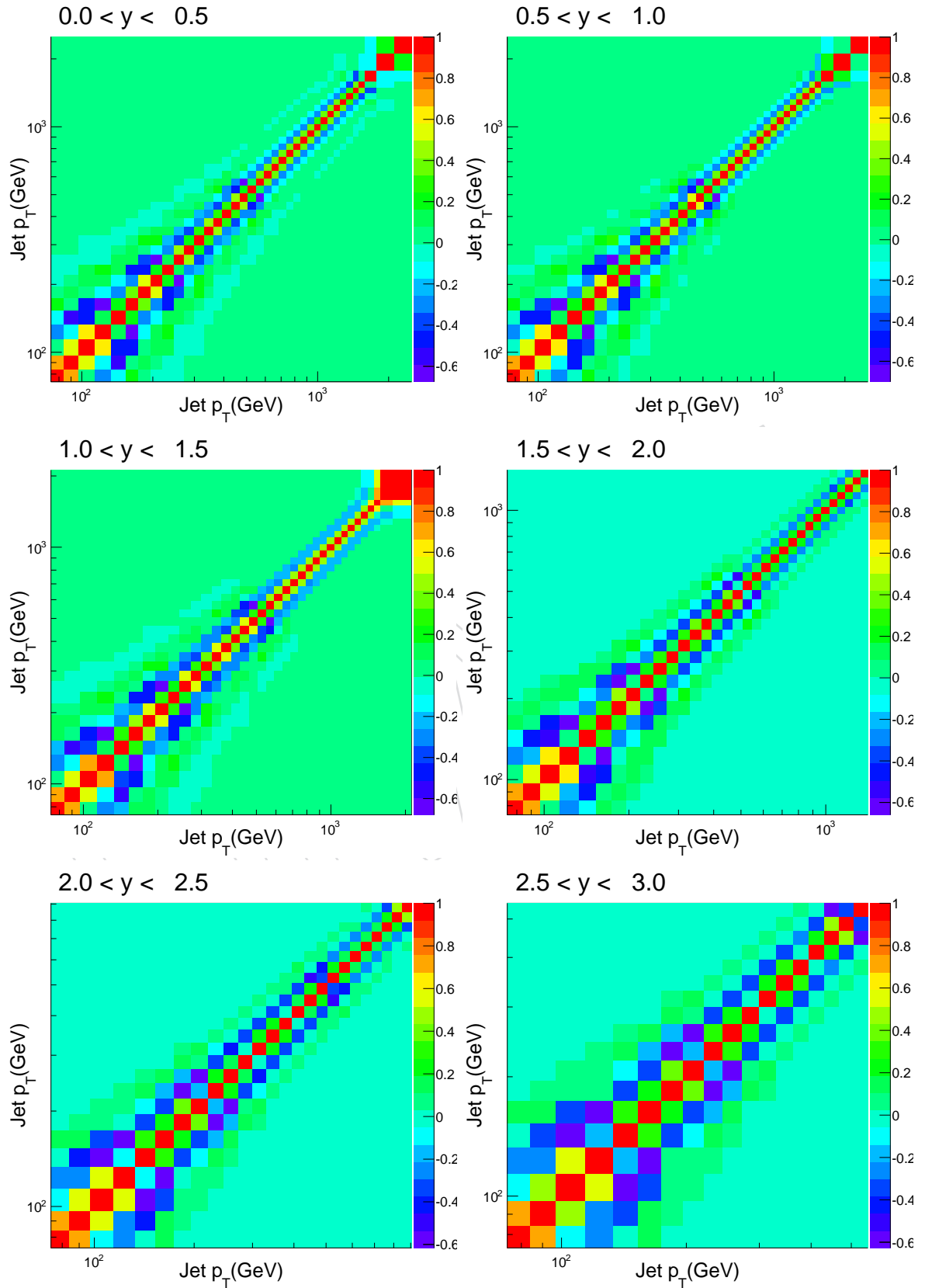


Figure 20: Correlation matrix obtained from unfolding for each rapidity bin.

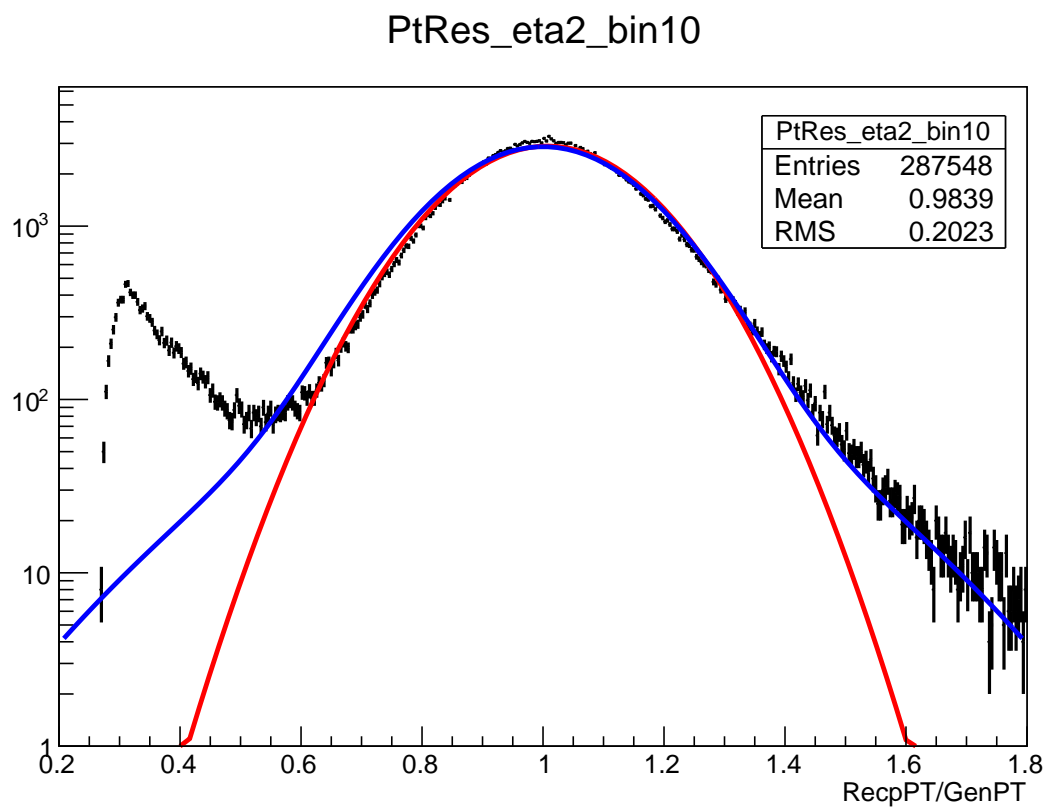


Figure 21: Fitting of the response distribution. The red curve shows the gaussian function which fits the core of the response distribution. The blue curve shows the fitted double gaussian curve covering non-gaussian tails of the response distribution.

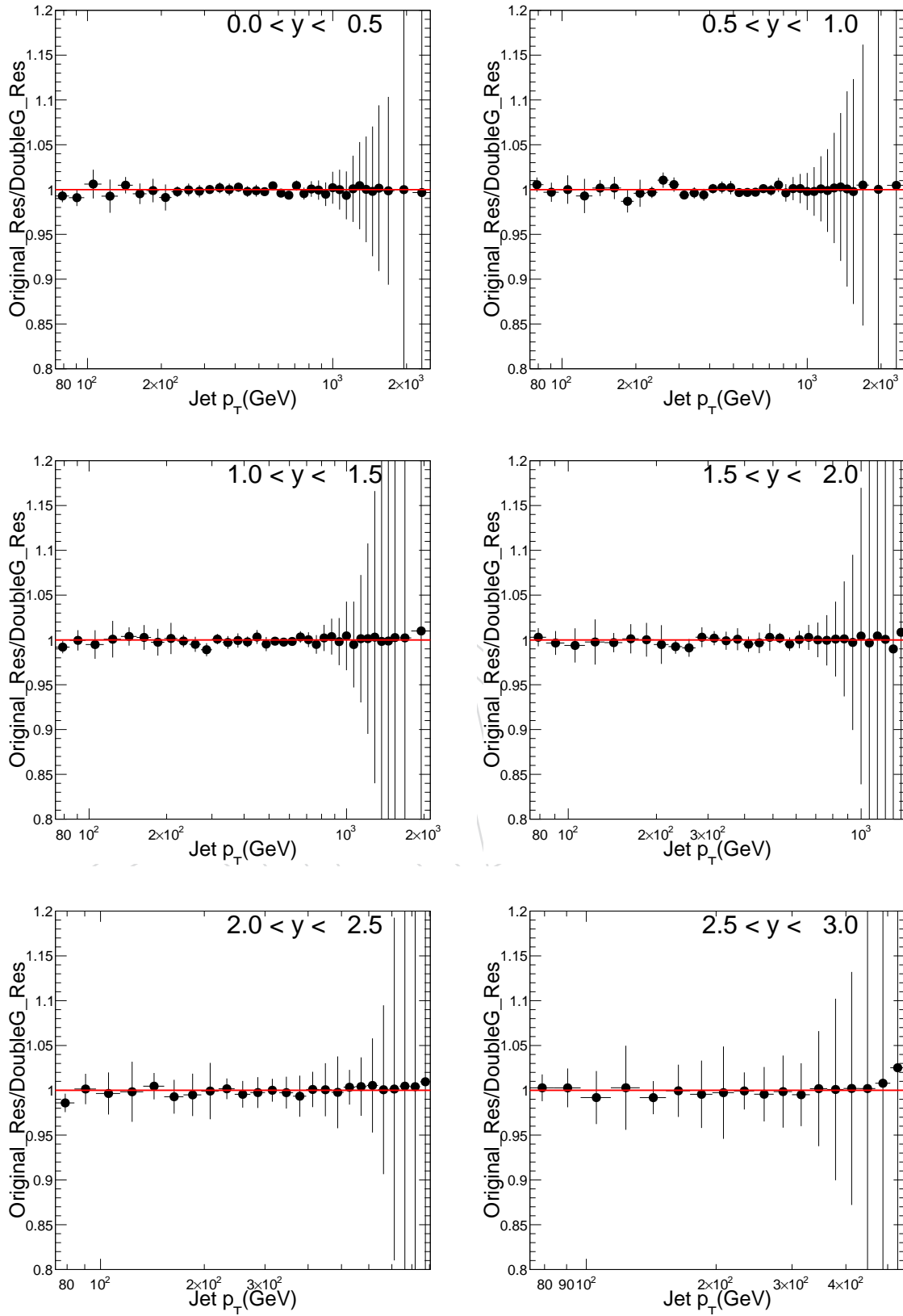


Figure 22: Ratio of original unfolded spectrum to the new unfolded spectrum, where in the later case the JER parameters are derived from double Gaussian fitting of the response function. Within statistical fluctuation, we see that the non-gaussian tails of the response function don't effect the unfolding significantly.

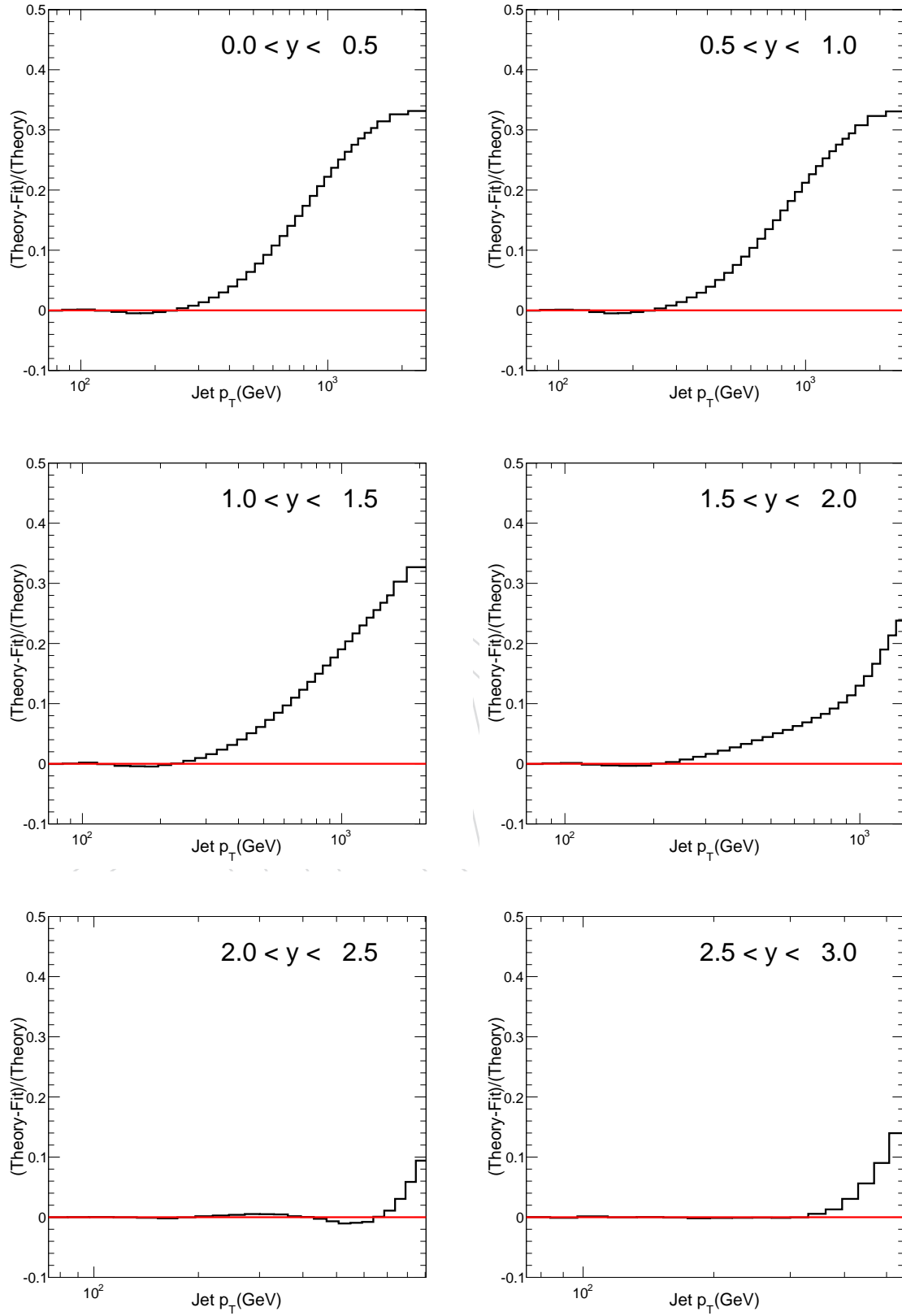


Figure 23: Relative difference between theory spectrum and the fitted function. The theory spectrum is derived using NNPDF PDF.

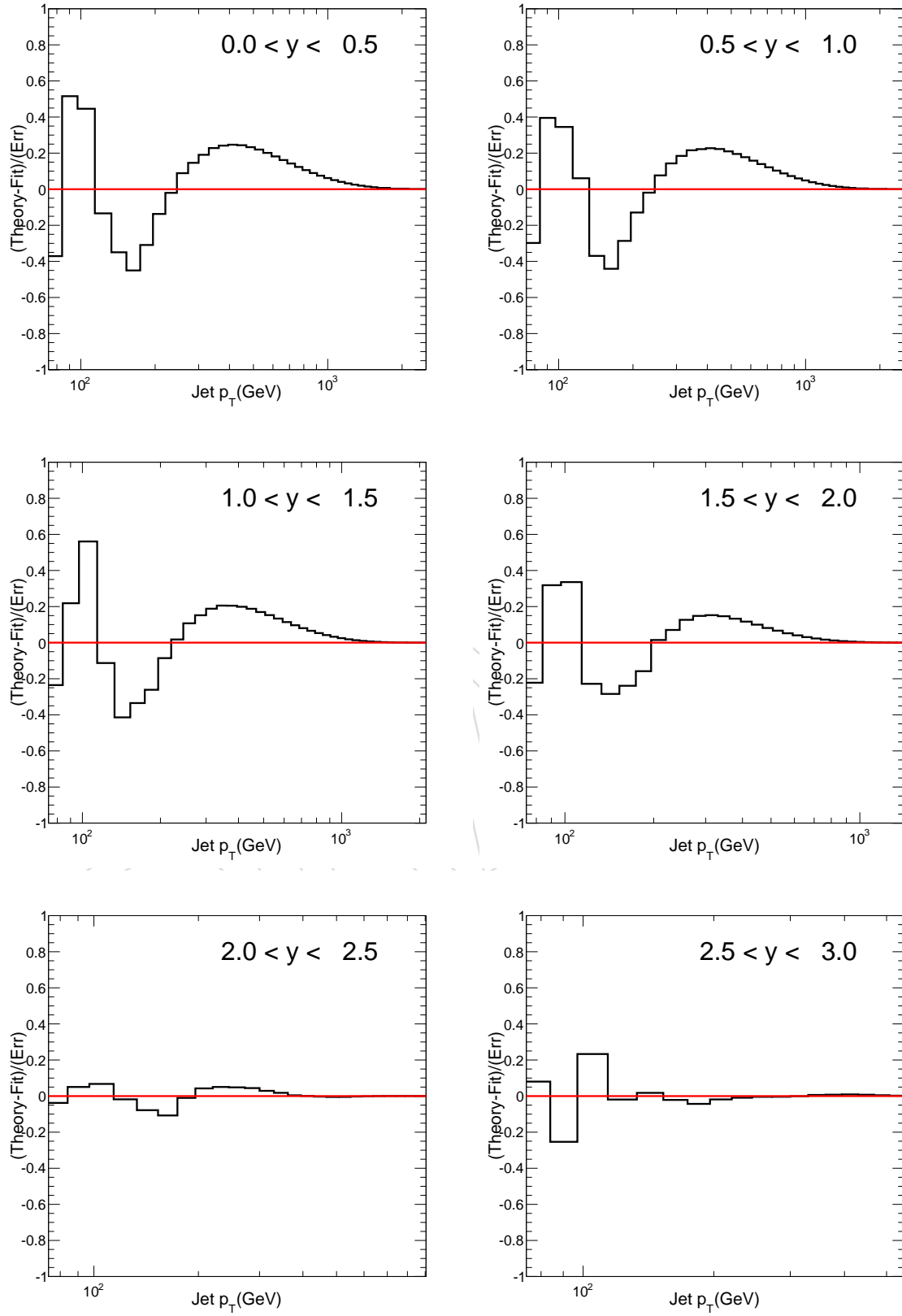


Figure 24: The PULL distribution showing the fluctuation between the fitted spectrum and the theory prediction compared to the standard deviation in each bin. The theory spectrum is derived using NNPDF PDF.

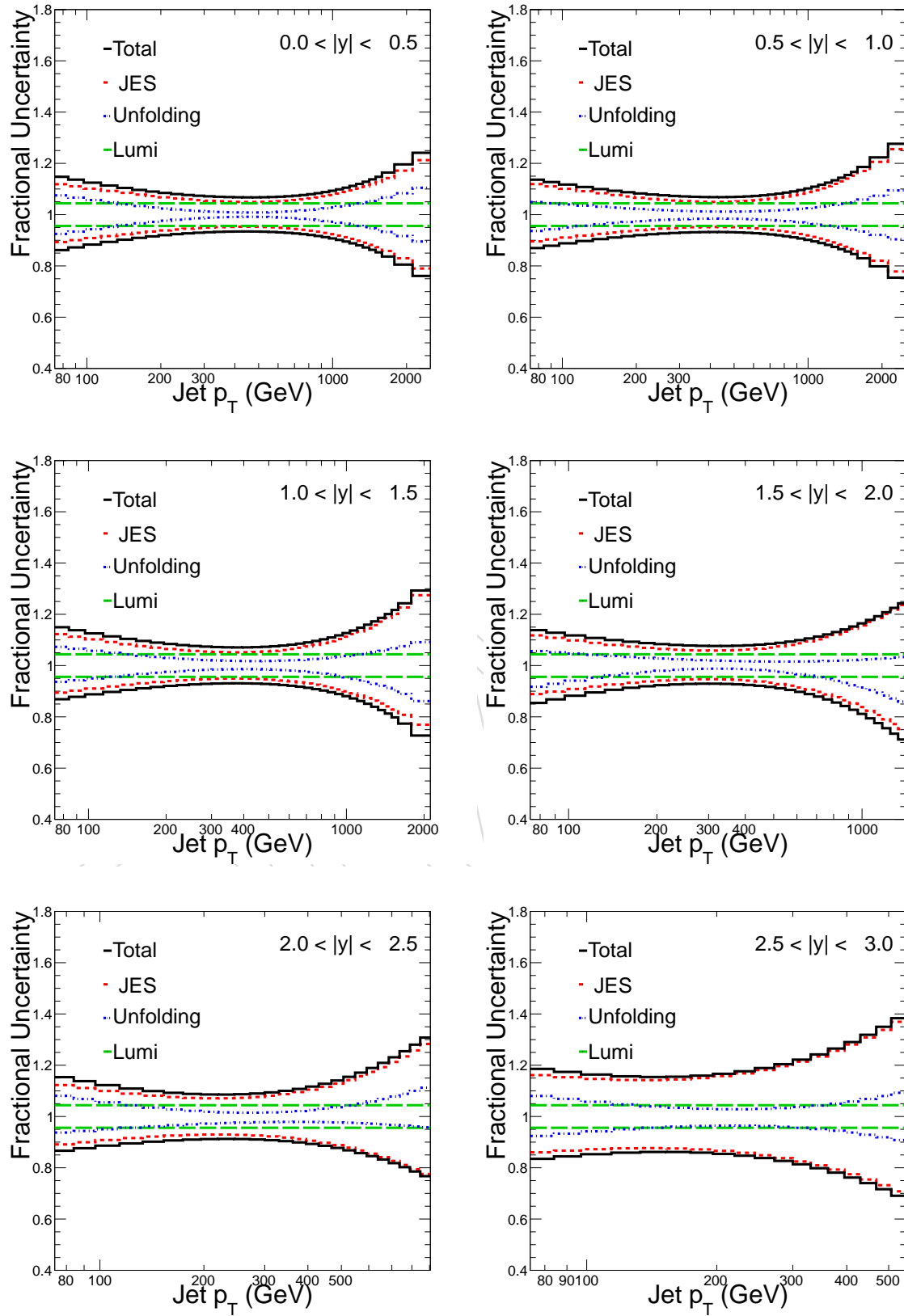


Figure 25: Relative experimental uncertainty on measured cross section due to different JES Sources

4 Theory Comparison

The theoretical predictions for the jet cross sections consist of a next-to-leading-order QCD calculation and a nonperturbative correction to account for the multiparton interactions (MPI) and hadronisation effects.

4.1 NLO Calculation

The NLO calculations are done using the NLOJet++ program (v2.0.1) [10] within the framework of the fastNLO package (v1.4) [11]. The renormalization and factorization scales (μ_R and μ_F) for the inclusive jet measurement are equal to the jet p_T . The NLO calculation is performed using five different PDF sets: CT10 [12], MSTW2008NLO [13], NNPDF2.1 [14], HERAPDF1.0 [15], and ABM11 [16] at the corresponding default values of the strong coupling constant $\alpha_s(M_Z) = 0.1180, 0.120, 0.119, 0.1176$, and 0.1134 , respectively.

4.2 Non-Perturbative Effects

We know that perturbative QCD (pQCD) predictions include initial-state radiation which form parton scattering from hadron-hadron interactions. At the experimental high energy physics, to compare these predictions with experimental data is important. For this purpose, experimental data are corrected for detector effects and collider effects which name is particle level, while next-to-leading order (NLO) predictions of pQCD generated by MC events are corrected for non-perturbative (NP) effects which include hadronization and multiple parton interaction (MPI).

4.2.1 Hadronization

pQCD is valid at short distances. At the long distances, pQCD breaks down. In this case, the colored quarks transform into colorless objects, and this process is called hadronization. We are not able to observe partons from hard scattering at the detector level but observe hadrons which pass on hadronization process. After quark-anti-quark collide at hard scattering, the outgoing partons carry color charge and create strong color fields between themselves and the rest of the proton. The strong color fields between a charge and an anti charge increase linearly with a separation of charges until getting enough energy to create additional quark-anti quark pairs. Due to this reason, original parton lost some of the own energy and momentum every time until havent enough to energy to create new a quark-anti quark pairs and color charge has become neutral. Due to this reason, every time original parton loses some of the own energy and momentum until having not enough energy to create a new quark-anti quark pairs and thus color charge becomes neutral. End of the hadronization, the original parton transforms into hadrons which known as a particle jets. Thus the energy of hadrons collimate into calorimeter cluster which known as calorimeter jets and observed by detector.

4.2.2 Multiple Parton Interaction (MPI)

Multiple Parton Interactions are defined as a scatters that dont take a role at the hard scattering directly. In the proton beam, the initiator partons produce an initial-state radiation that transforms from parton level to detector level, and leaves behind pair of parton remnant called spectators. These spectators that are separated from original color charge produce color fields and then also it may produce parton showers as well. Thus, The spectators interact with outgoing partons and also have soft interaction with each other. Moreover, spectators introduce degrees of freedom by sharing some of the energy of original parton. This energy causes more

energy than we have to get the energy from particle jets and calorimeter jets. All of these parton interaction which are not calculated from hard scattering, can be grouped as an Underlying Event.

4.2.3 Non-Perturbative Corrections

At the hadron-hadron interactions, passing from initial-radiation state to final-radiation state we need a non-perturbative corrections(NP) which known as hadronization and multiple parton interactions. In order to derive NP correction factor, the following MC event generator is used:

PYTHIA6(Tune Z2) and HERWIG++.

For the event generator, *Nominal* and *noMPI – noHAD* settings are applied. Nominal setting refers to cross section obtained from MC generations which includes hadronization and multiple hadron interaction, while *noMPI-noHAD* setting refers to cross section obtained from MC generations with NP effects off. NP correction factor is defined as the ratio of Nominal setting to (*noMPI,noHAD*) setting,

$$C = \frac{\sigma(Nominal)}{\sigma(noMPI,noHAD)}$$

The correction factor calculated for each rapidity region as a function of PF Jet p_T and fitted with same fit function, $[0]+[1]/\text{pow}(x,[2])$ for PYTHIA Z2. Thus only one correction factor, which are produced from both MC generators, is used for NLO of pQCD predictions.

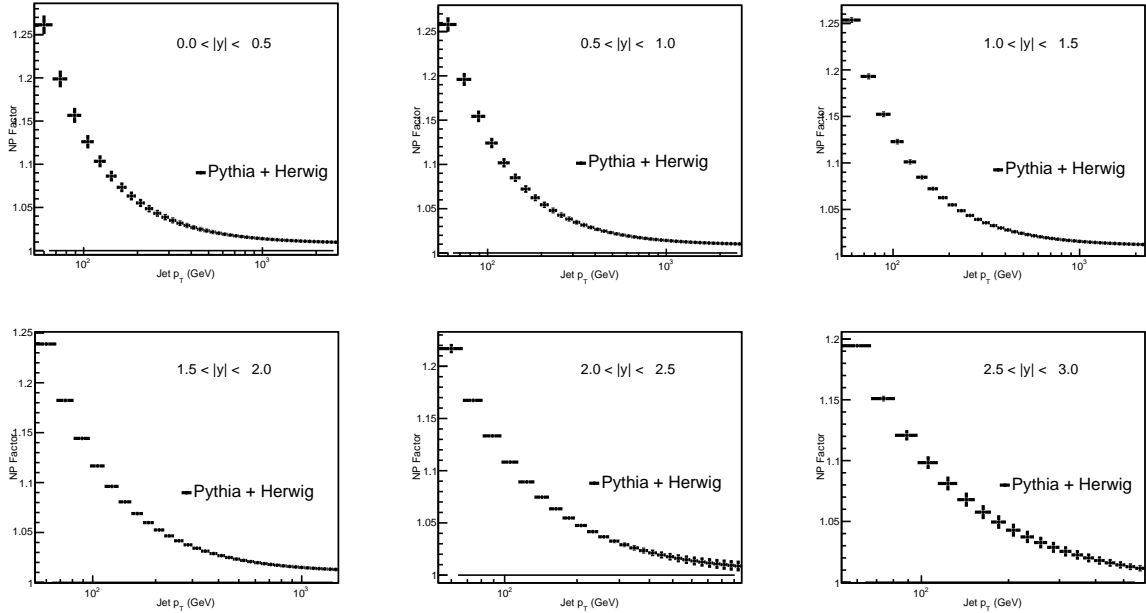


Figure 26: NP correction factor from PythiaZ2 for different rapidity region

In the Figure 26 the bin by bin values are used for the final non perturbative corrections.

4.3 Theory Uncertainties

The renormalization and factorization scale uncertainty is estimated as the maximum deviation at the six points $(\mu_F/\mu, \mu_R/\mu) = (0.5, 0.5), (2, 2), (1, 0.5), (1, 2), (0.5, 1), (2, 1)$, where $\mu = p_T$ (inclusive). NP factor ranges from 1% to 25% in the p_T region of interest. The PDF variation introduces uncertainties on the theoretical prediction up to 30% within $|y| \leq 1.5$, where

as beyond this region the pdf uncertainty crosses 50% for very high p_T bins. Moreover the variation of $\alpha_S(M_Z)$ by 0.001 introduces an additional 1–2% uncertainty. The renormalization and factorization scale uncertainty is estimated as the maximum deviation at the six points $(\mu_F/\mu, \mu_R/\mu) = (0.5, 0.5), (2, 2), (1, 0.5), (1, 2), (0.5, 1), (2, 1)$, where $\mu = p_T$ (inclusive). An additional uncertainty of at most 10% is caused by the non-perturbative correction at lower jet p_T around 100 GeV and decreases to less than 1% at higher p_T . Scale uncertainty ranges from 5% to 10% for $|y| < 1.5$ but increases to 40% for the outer $|y|$ bins and for high jet p_T s. Overall, the PDF uncertainty is dominant. For HERA and CT10 pdf sets, the PDF uncertainty for higher p_T bin reaches upto 100%.

5 Results

Figure 32 to Figure 36 shows the unfolded double-differential cross sections as a function of jet p_T compared to the QCD prediction. In order to reveal the details of the agreement between the data and the theory prediction, the ratio of the two is taken. Figure 42 to Figure 46 show the comparison of the measurement with the central value of all PDF sets, accompanied by the total theoretical uncertainties. The theoretical uncertainties are considerably different among the various PDF sets, and in particular in the high p_T .

It is noticeable, that in all cases the fluctuation from unity is within the total theoretical and experimental uncertainty bands, except for the Figure 42 corresponding to PDF set ABM11, and for the range of $0 < |y| < 2.0$ where the deviation upto 40% of this ratio is observed for $p_T \geq 300$ GeV onwards. Otherwise, for all other PDF sets a very nice agreement within 5-10% is observed. In Fig. 37 - 41 a comparison of the ratio of double differential cross sections for both data over theory prediction corresponding to each PDF sets are presented. In addition, the same ratio corresponding to that given PDF sets with all other PDF sets are also presented in the same plane. We find the same pattern of agreement as before, except for the case ABM11 PDF set, Figure 37, as observed in Fig. 42. Whereas in the Figure. 40, 41 we observe that the ratio between different NLO theory predictions have similar level of agreement to the data over theory predictions.

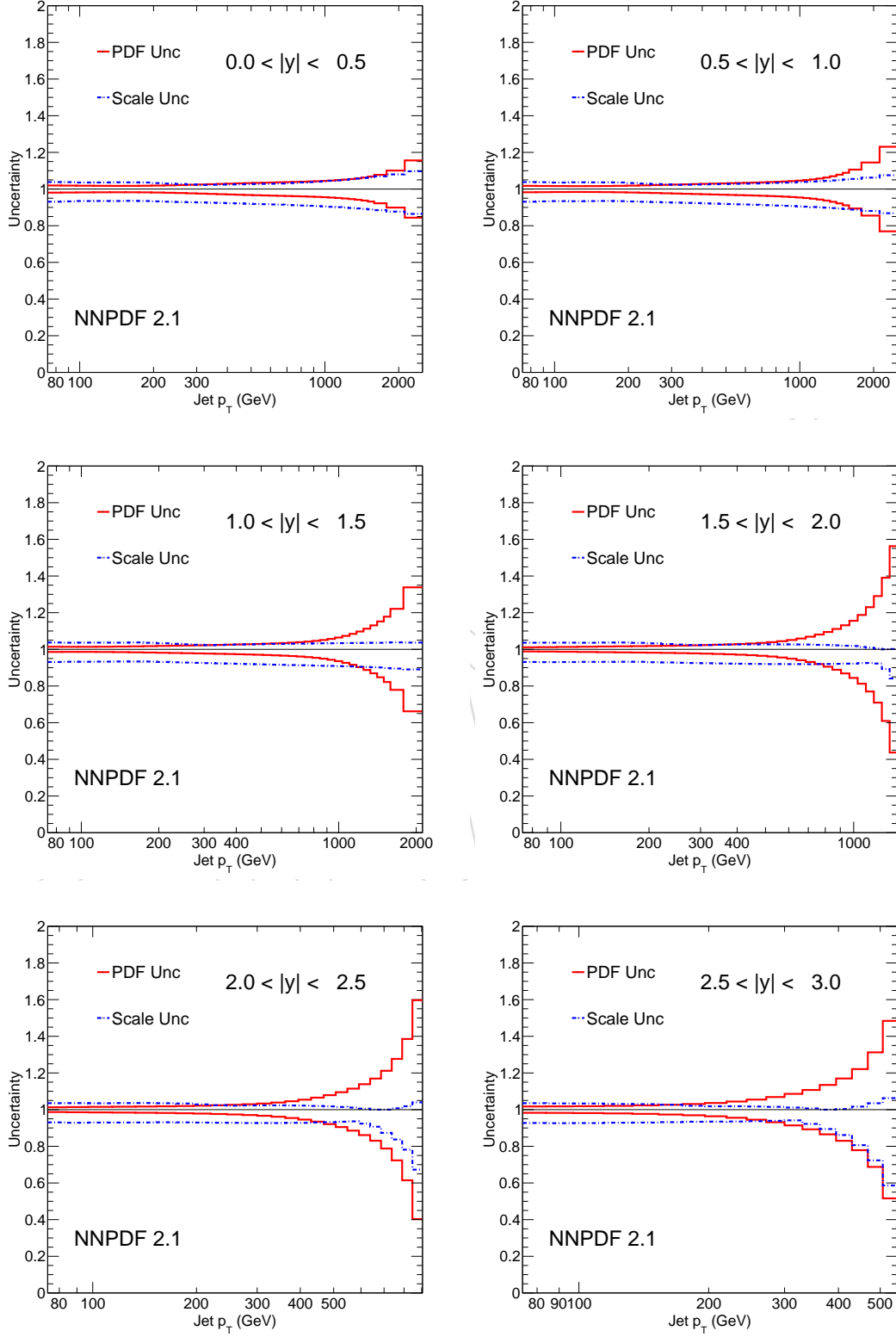


Figure 27: Theory Uncertainty for NNPDF2.1

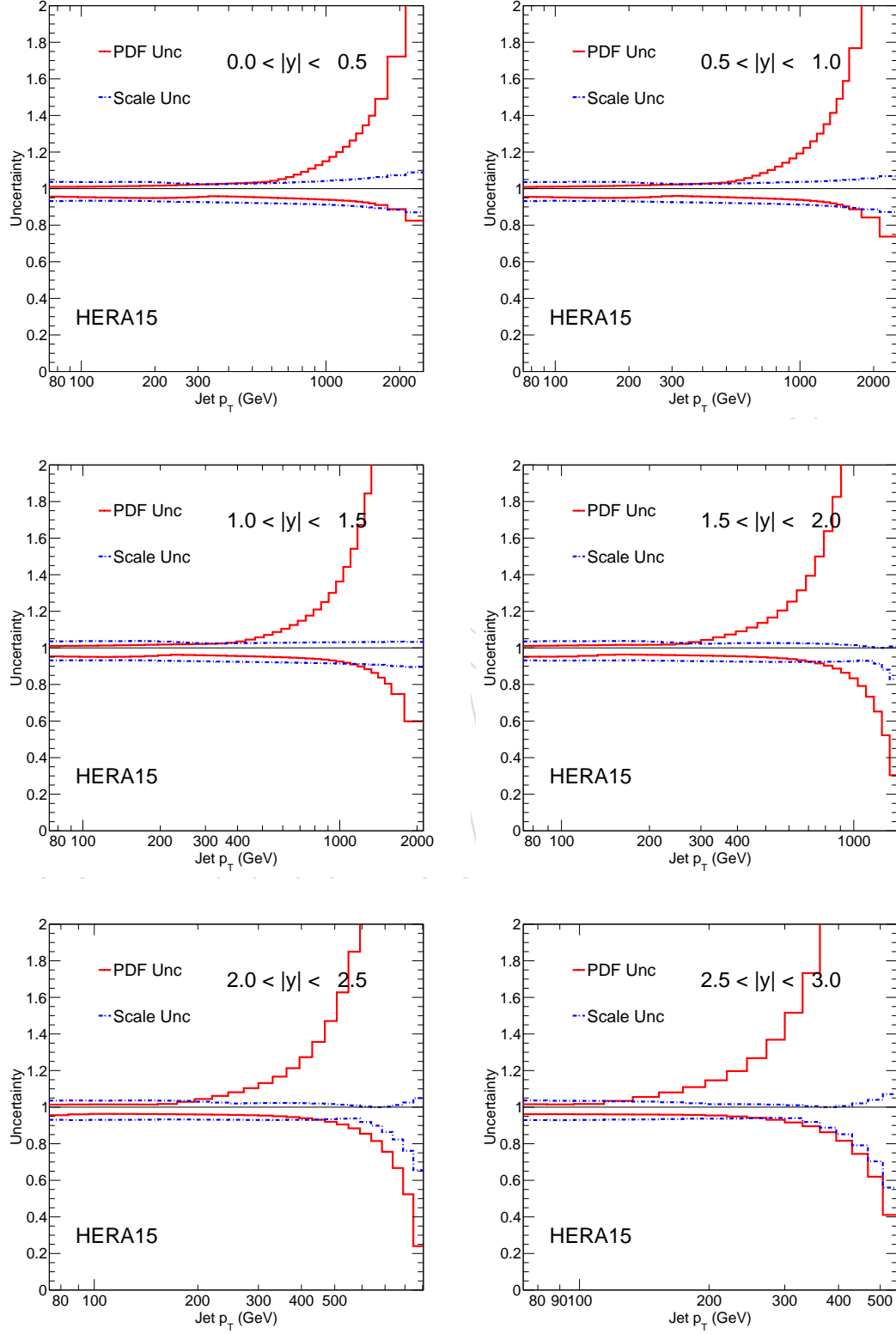


Figure 28: Theory Uncertainty for HERA

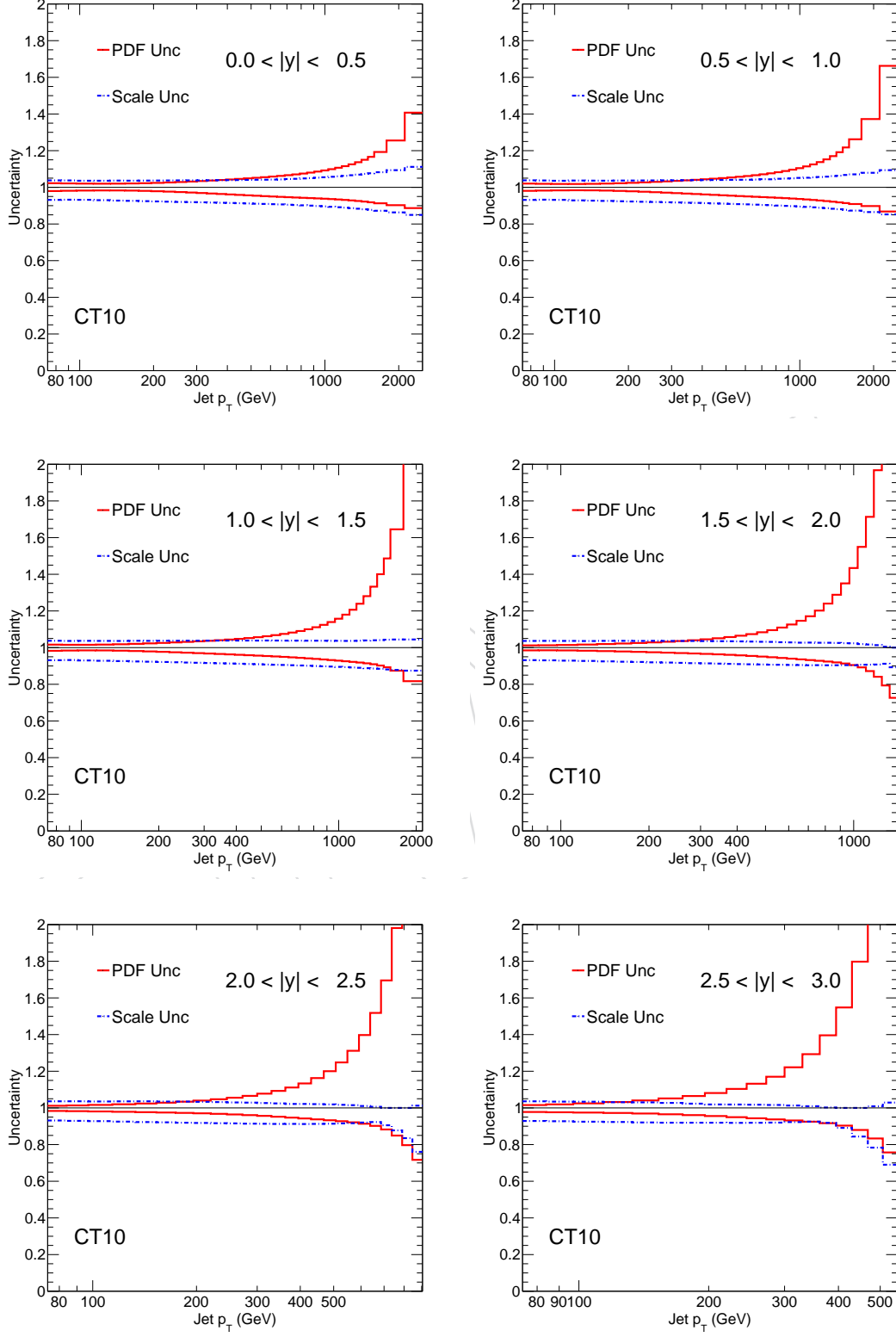


Figure 29: Theory Uncertainty for CT10

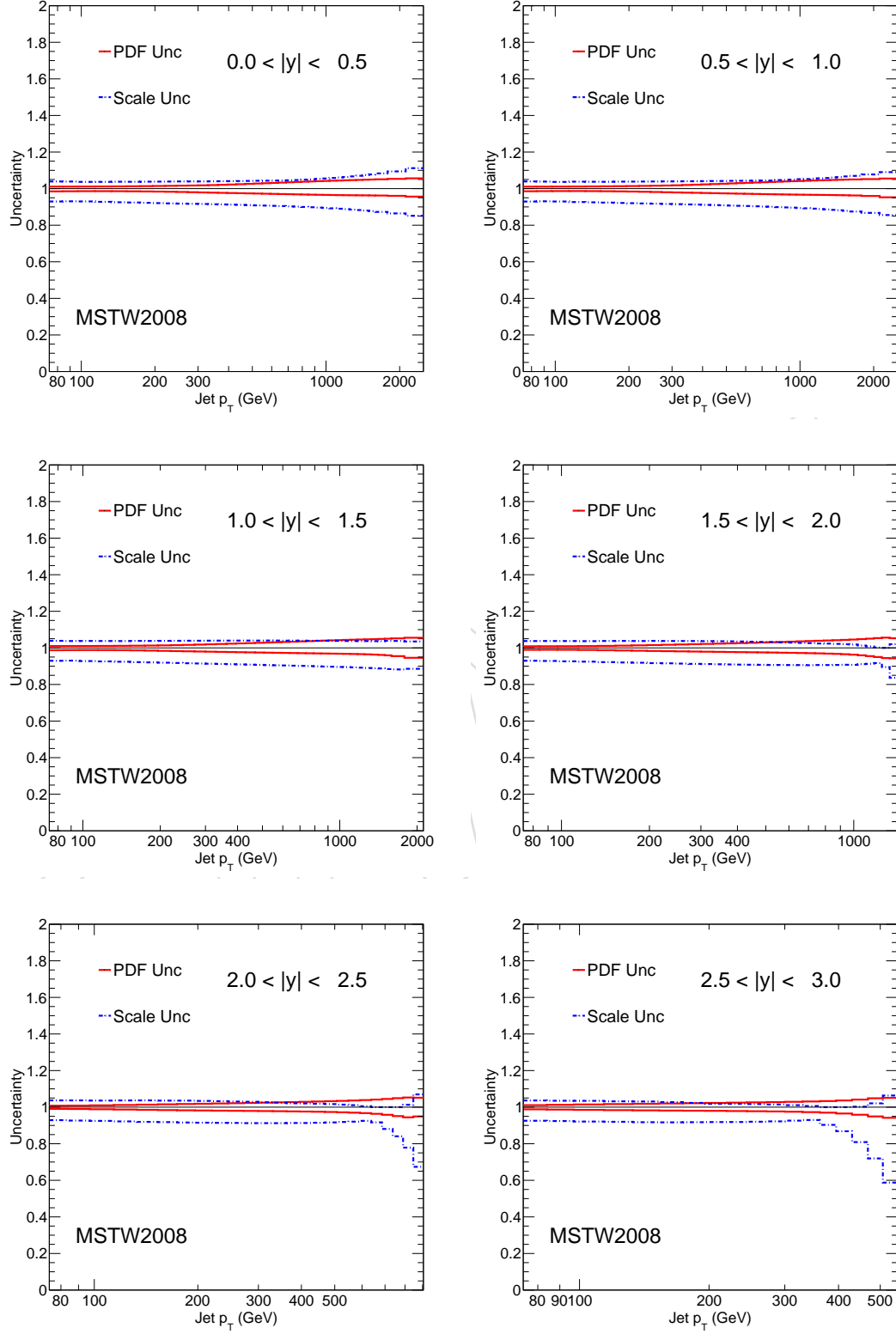


Figure 30: Theory Uncertainty for MSTW

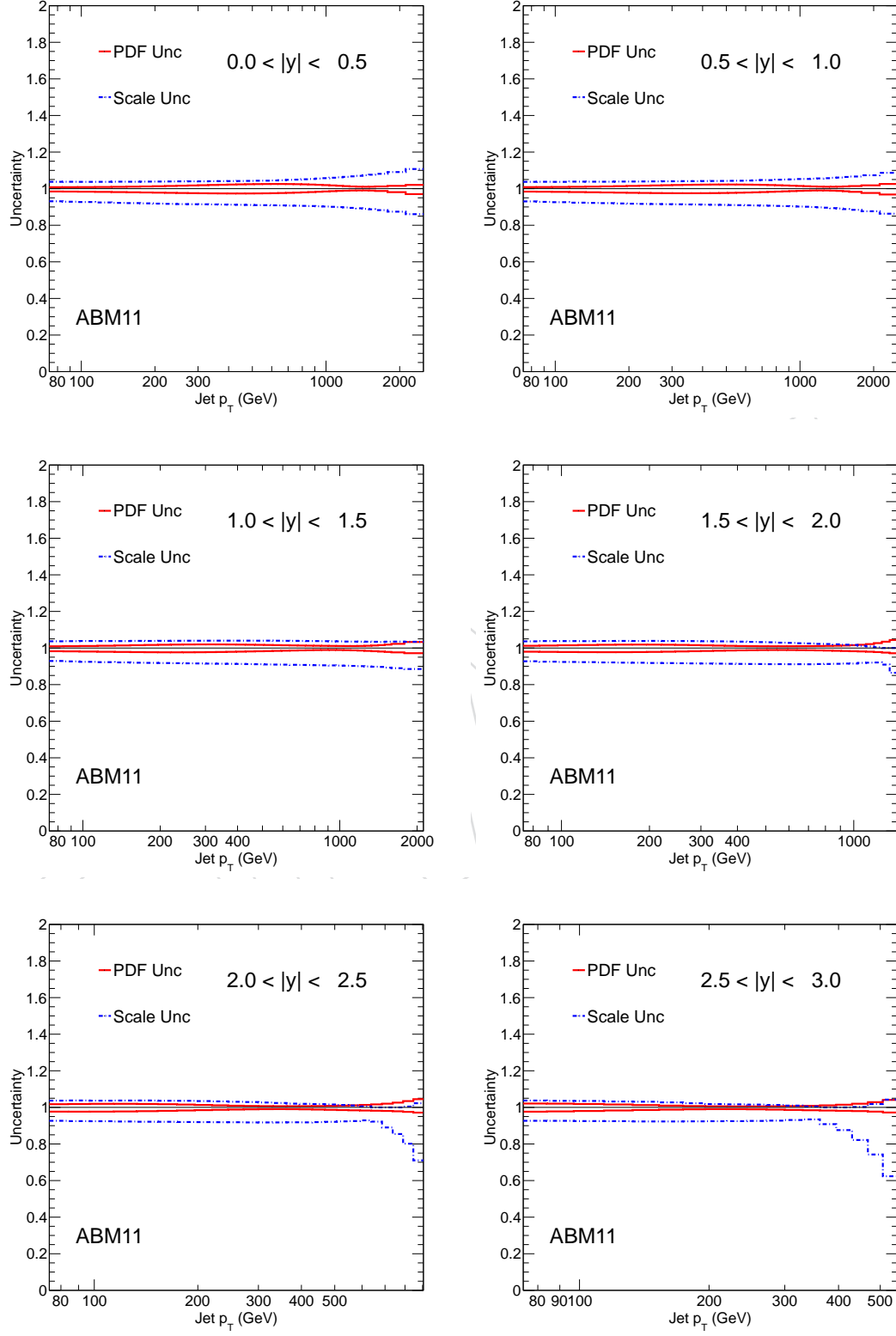


Figure 31: Theory Uncertainty for ABM11

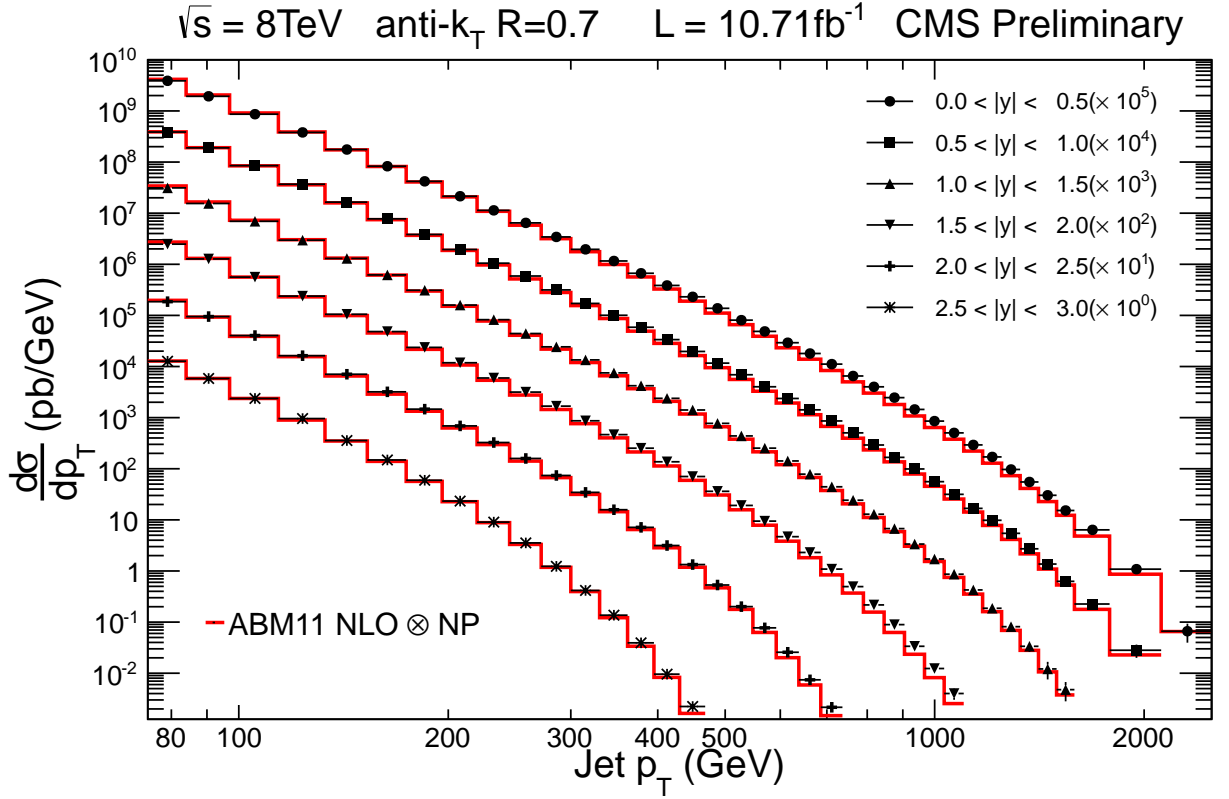


Figure 32: Unfolded Spectrum Compared to ABM Theory Prediction

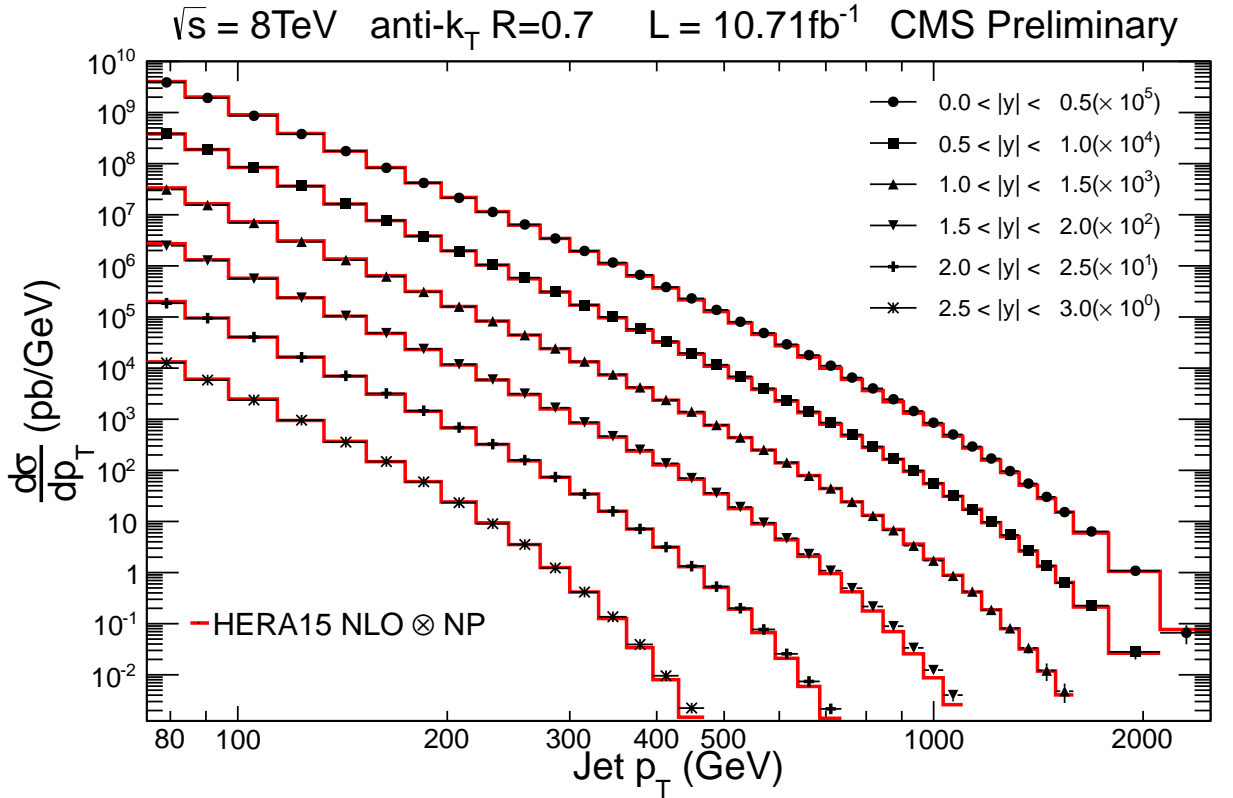


Figure 33: Unfolded Spectrum Compared to HERA Theory Prediction

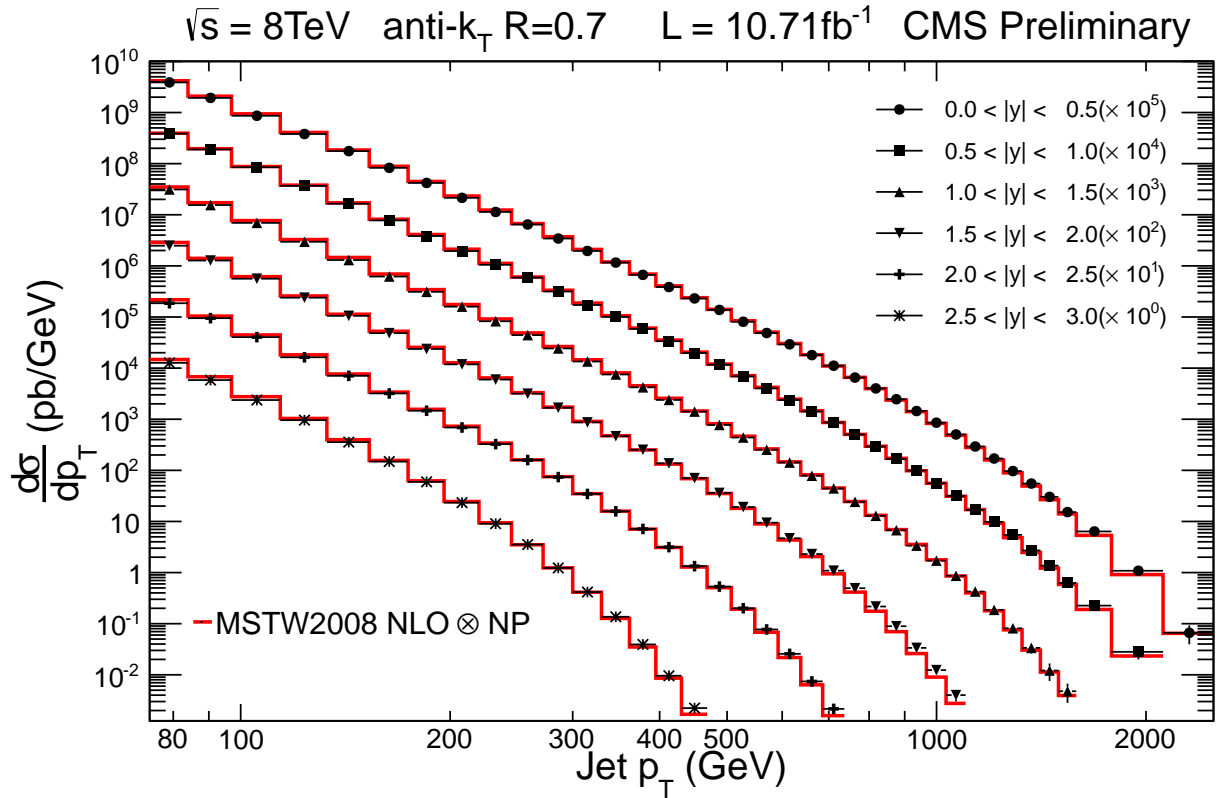


Figure 34: Unfolded Spectrum Compared to MSTW Theory Prediction

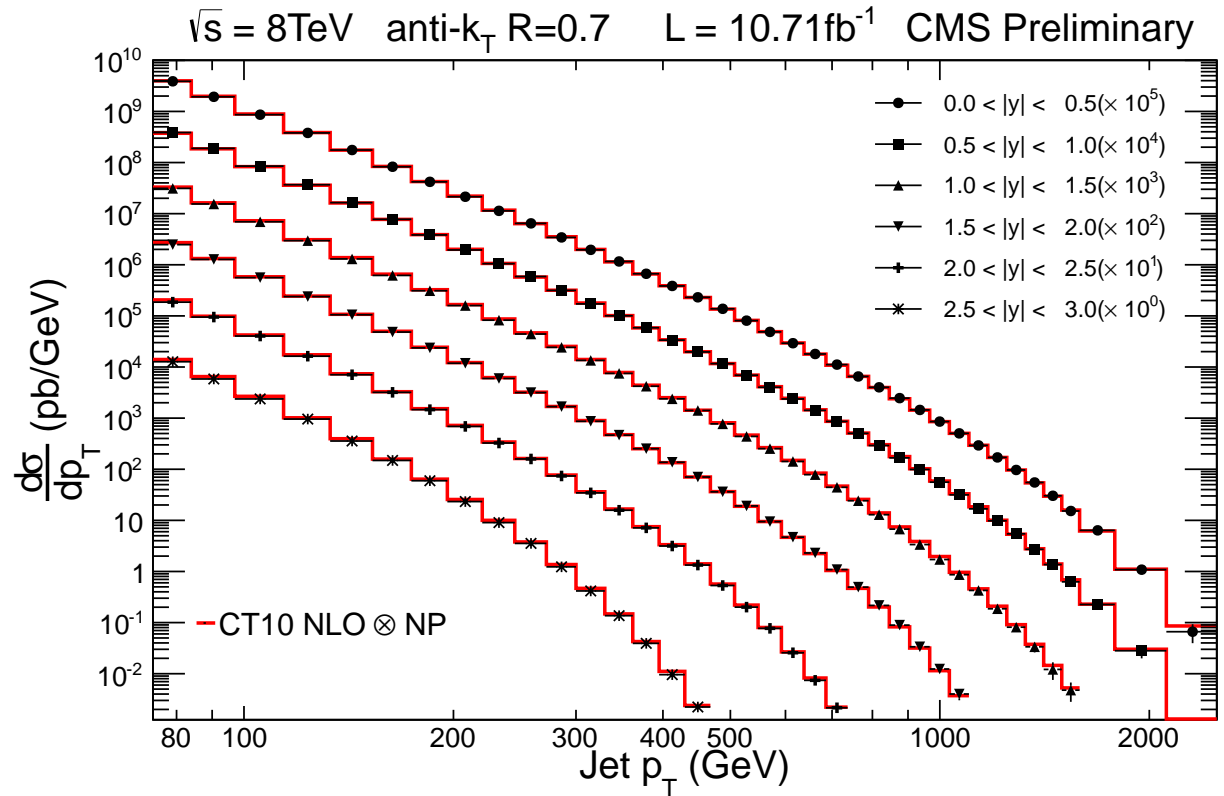


Figure 35: Unfolded Spectrum Compared to CT10 Theory Prediction

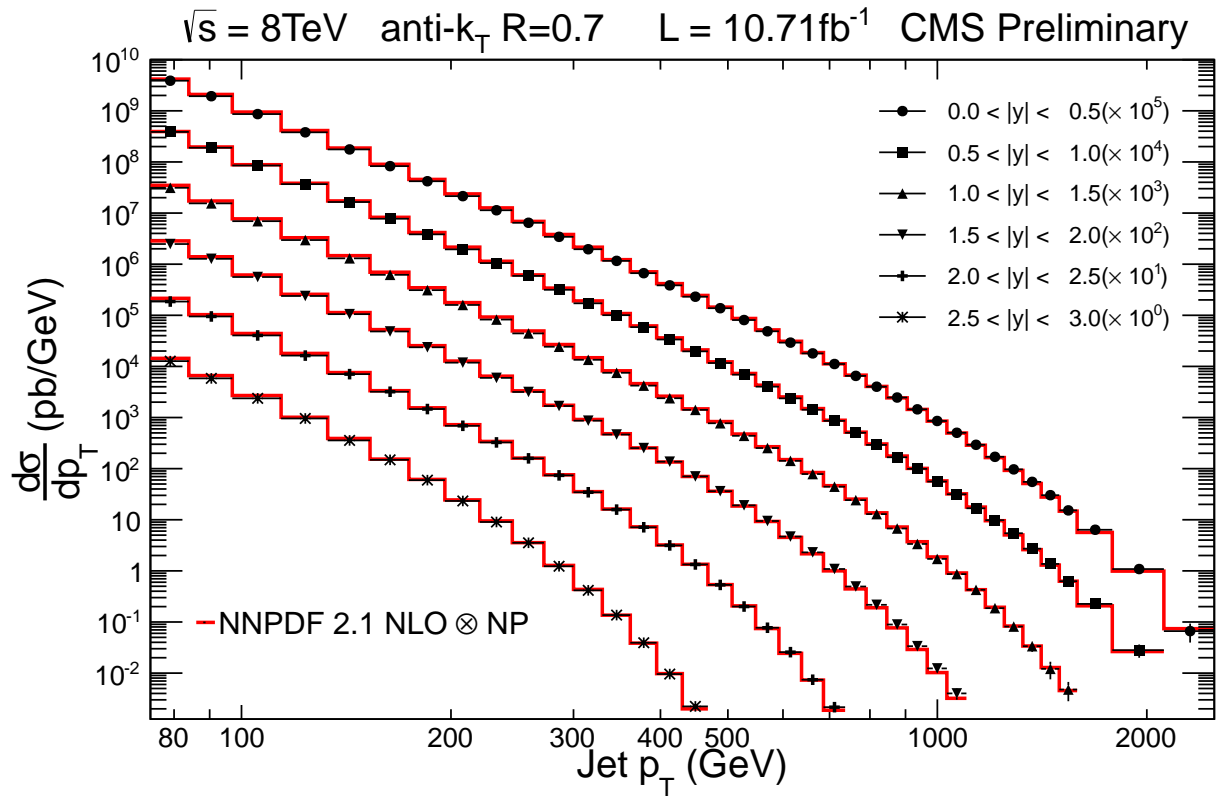


Figure 36: Unfolded Spectrum Compared to NNPDF Theory Prediction

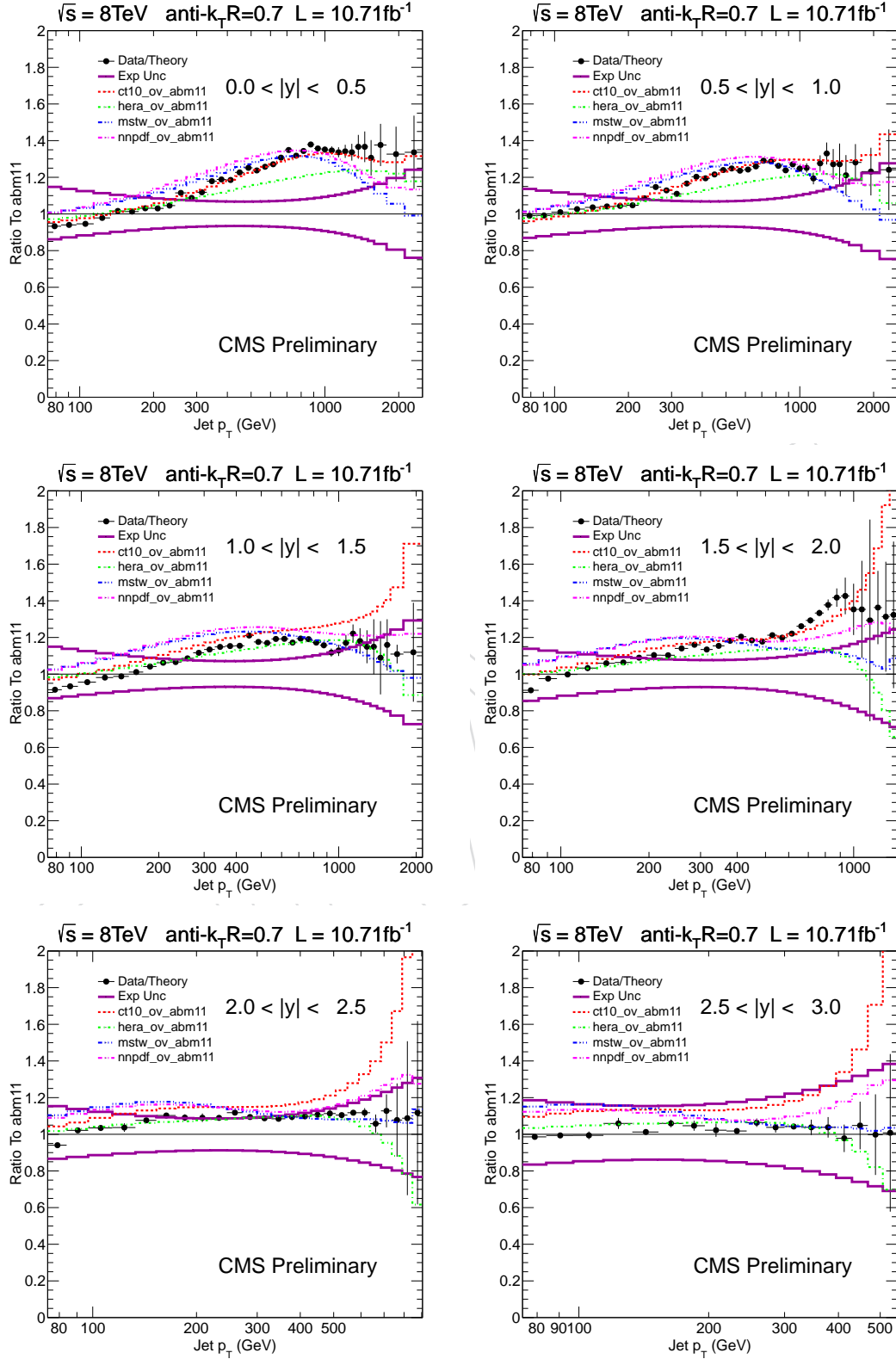


Figure 37: Ratio to ABM11

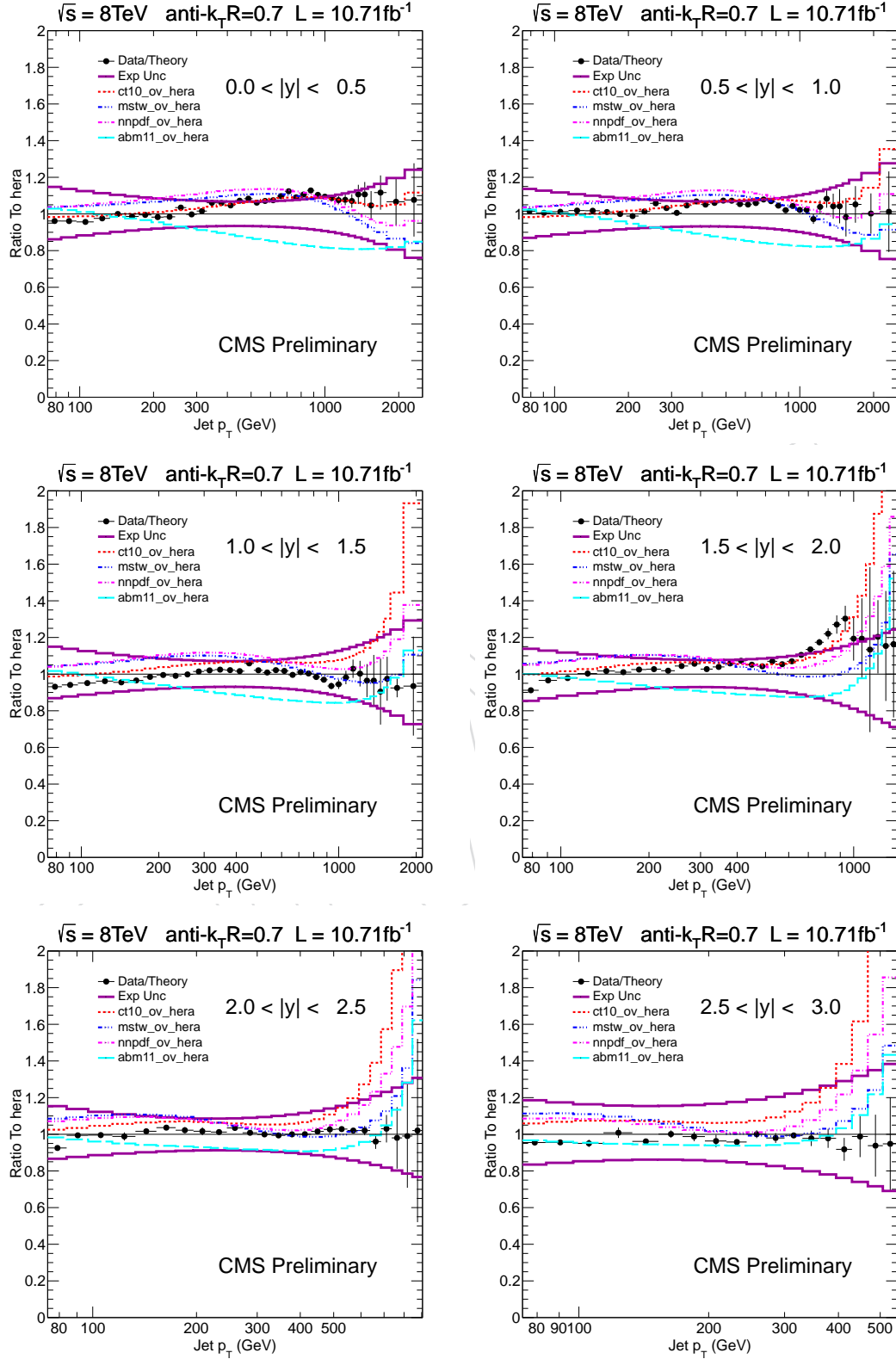


Figure 38: Ratio to HERA

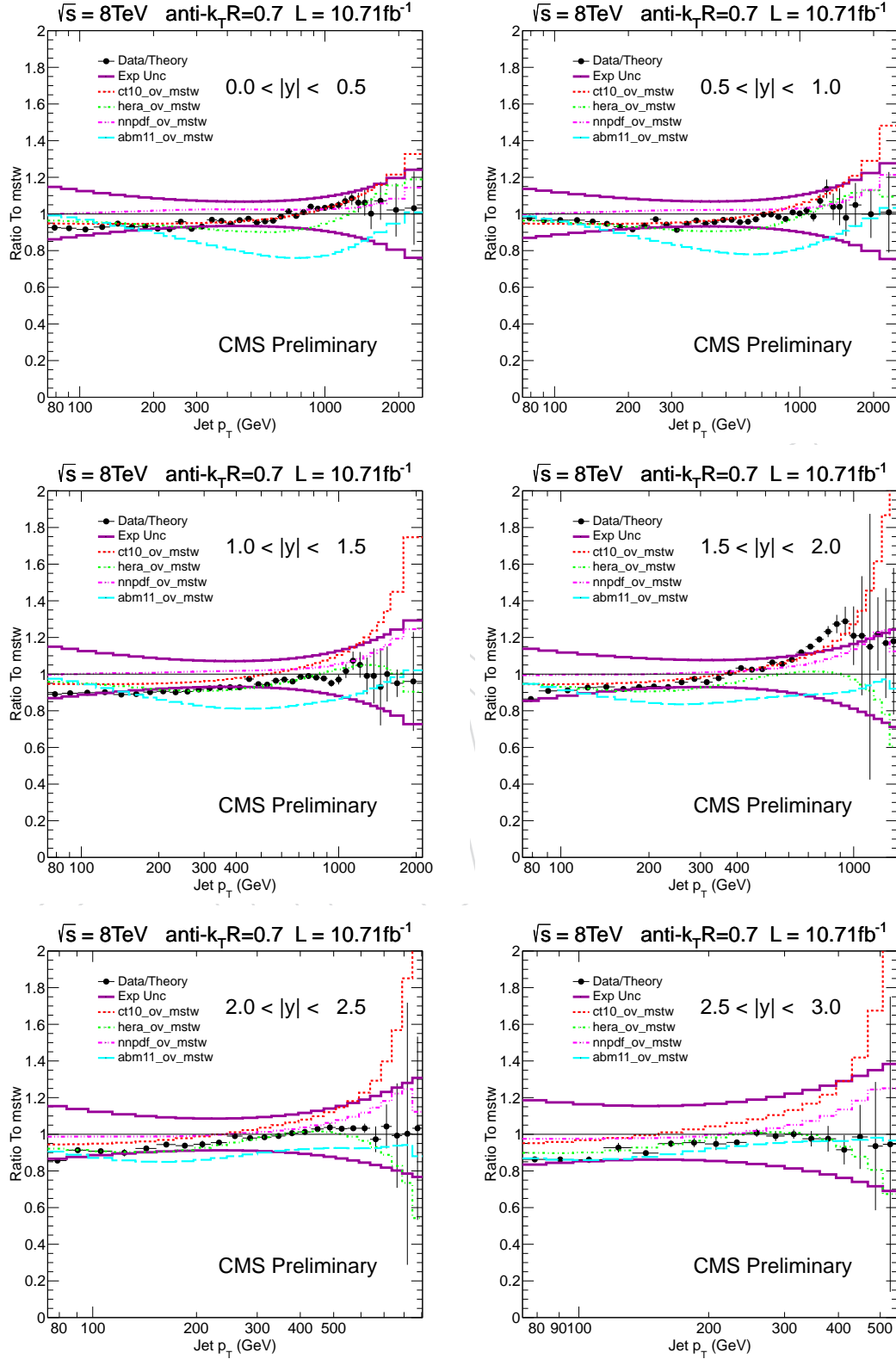


Figure 39: Ratio to MSTW

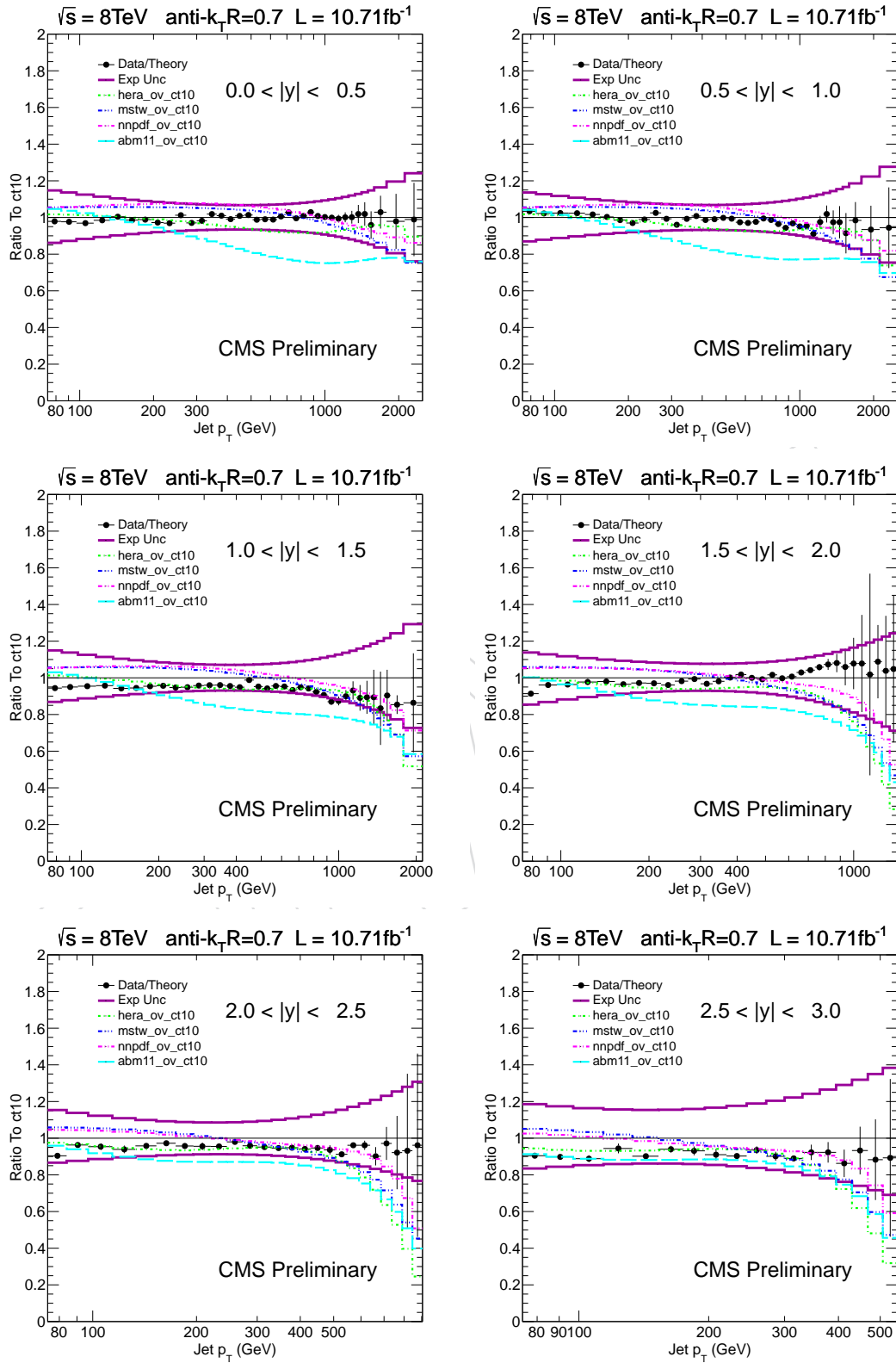


Figure 40: Ratio to CT10

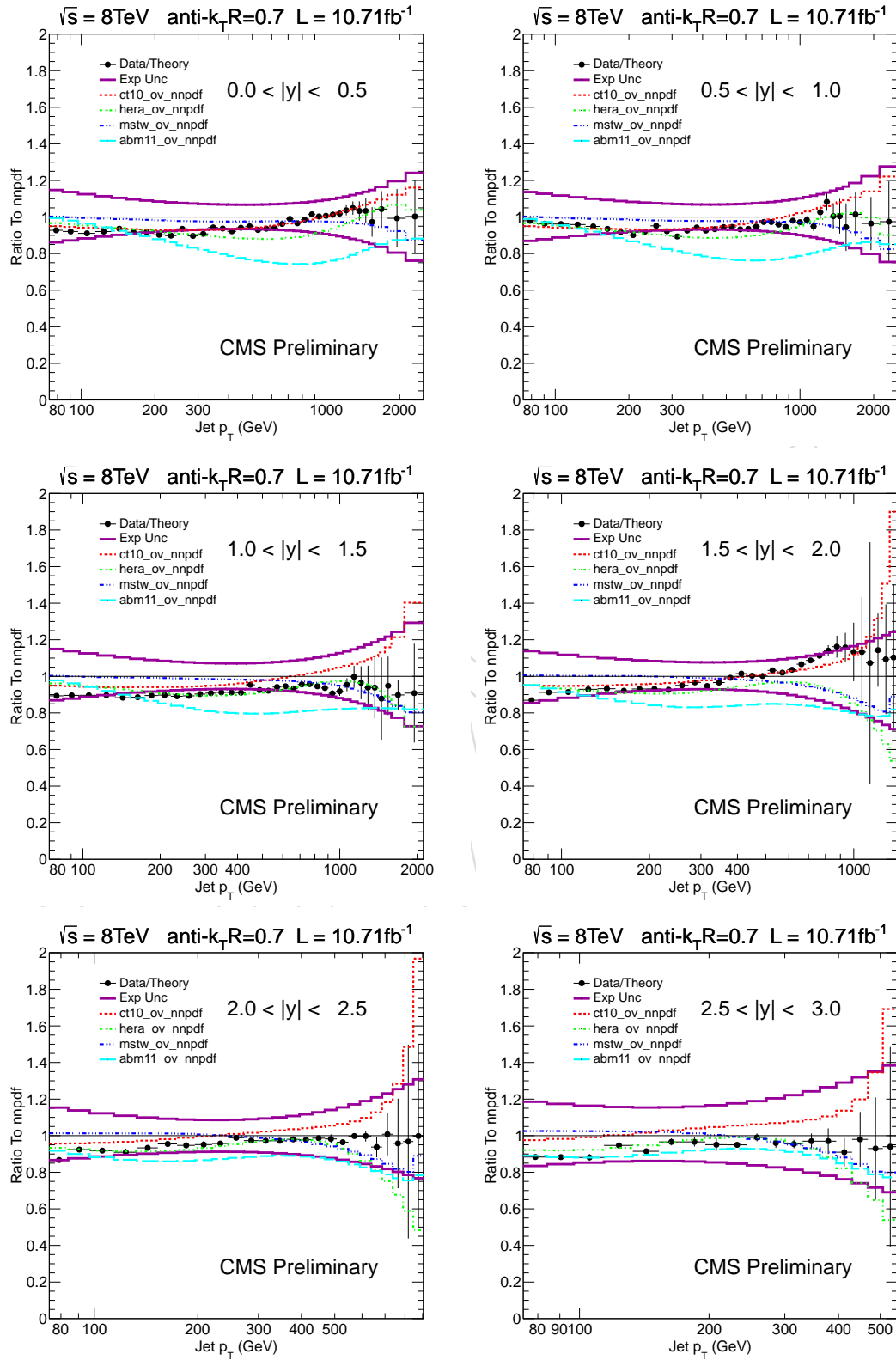


Figure 41: Ratio to NNPDF

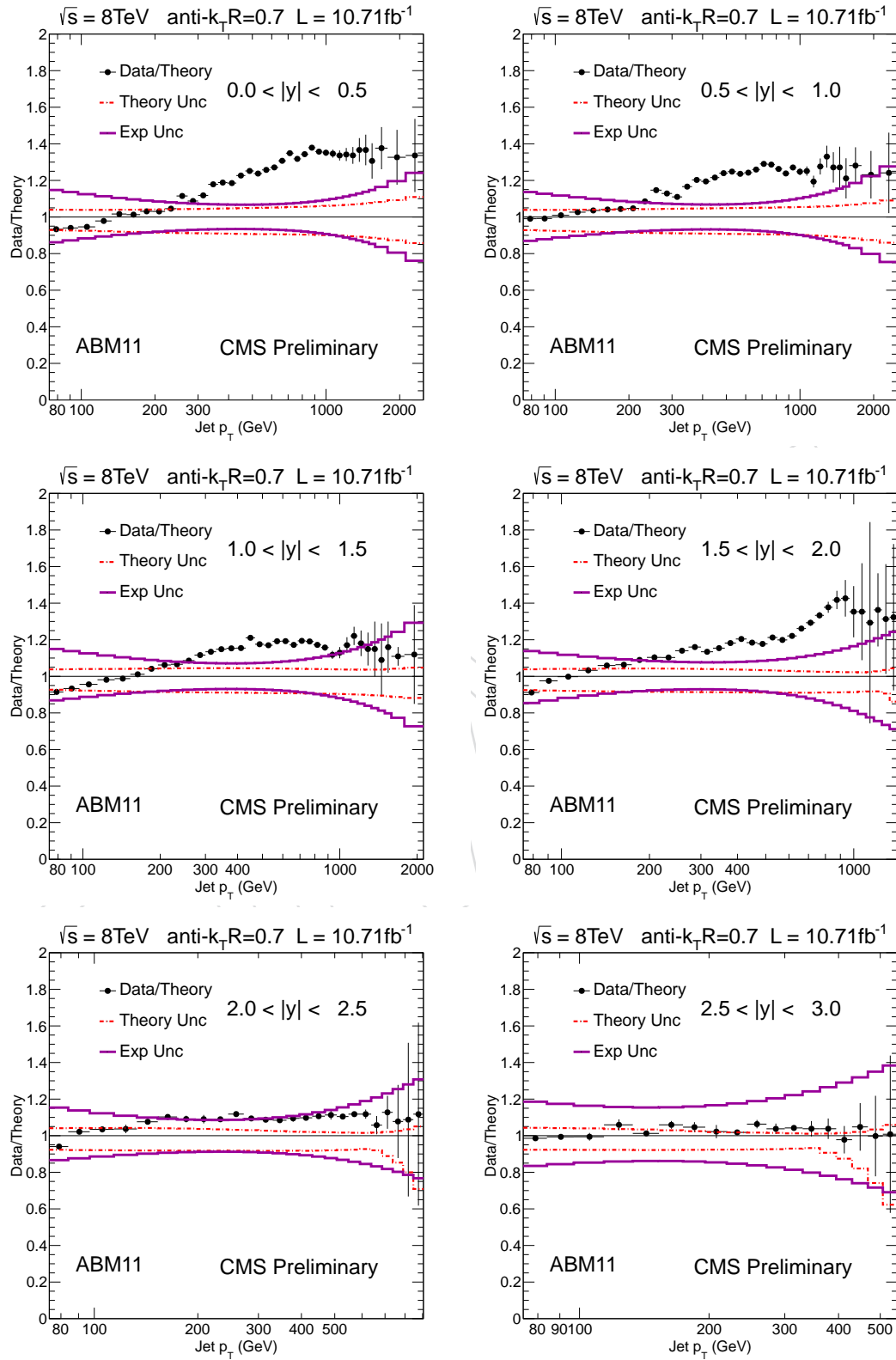


Figure 42: Data Over Theory Plot for ABM11

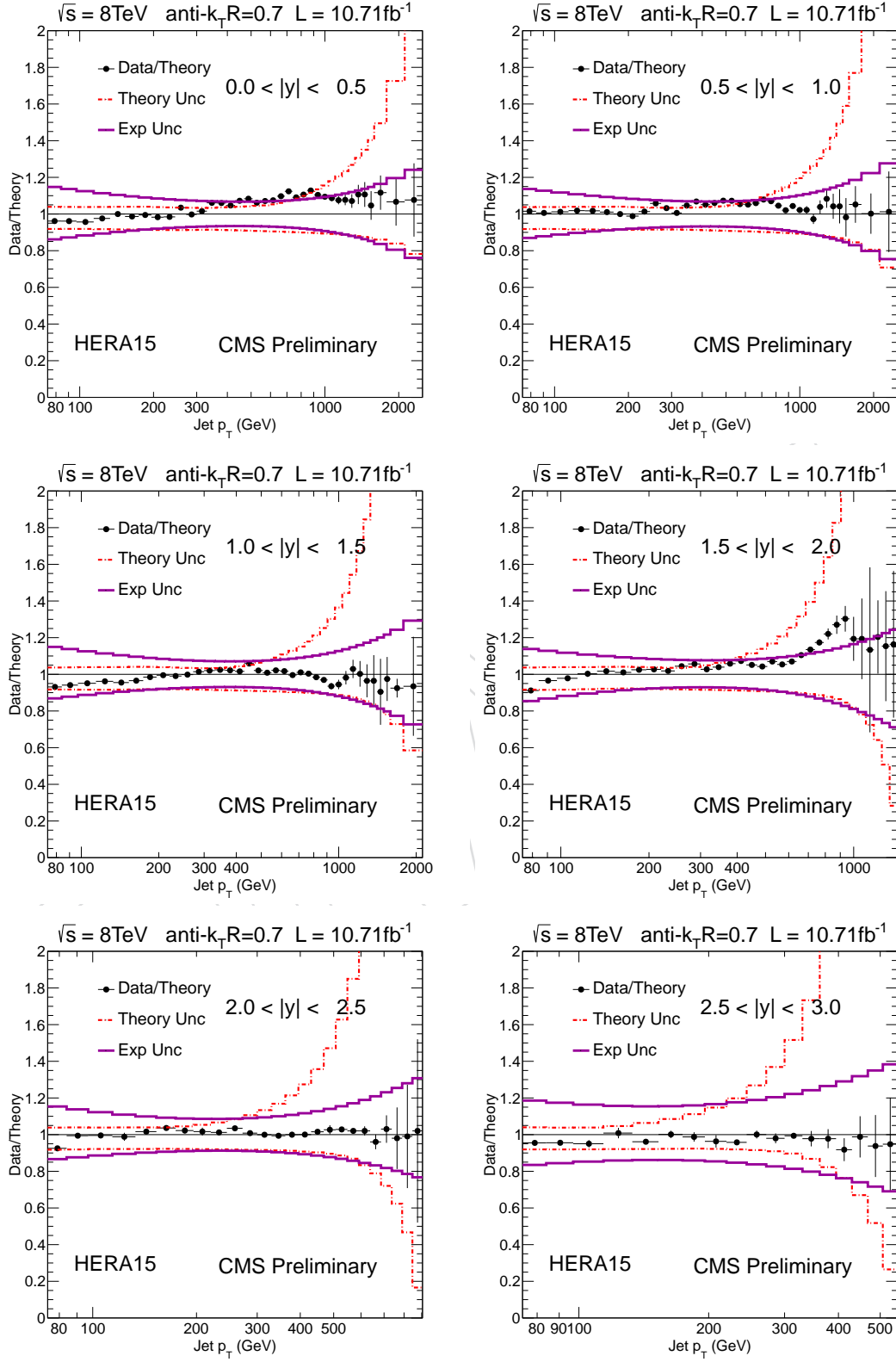


Figure 43: Data Over Theory Plot for HERA

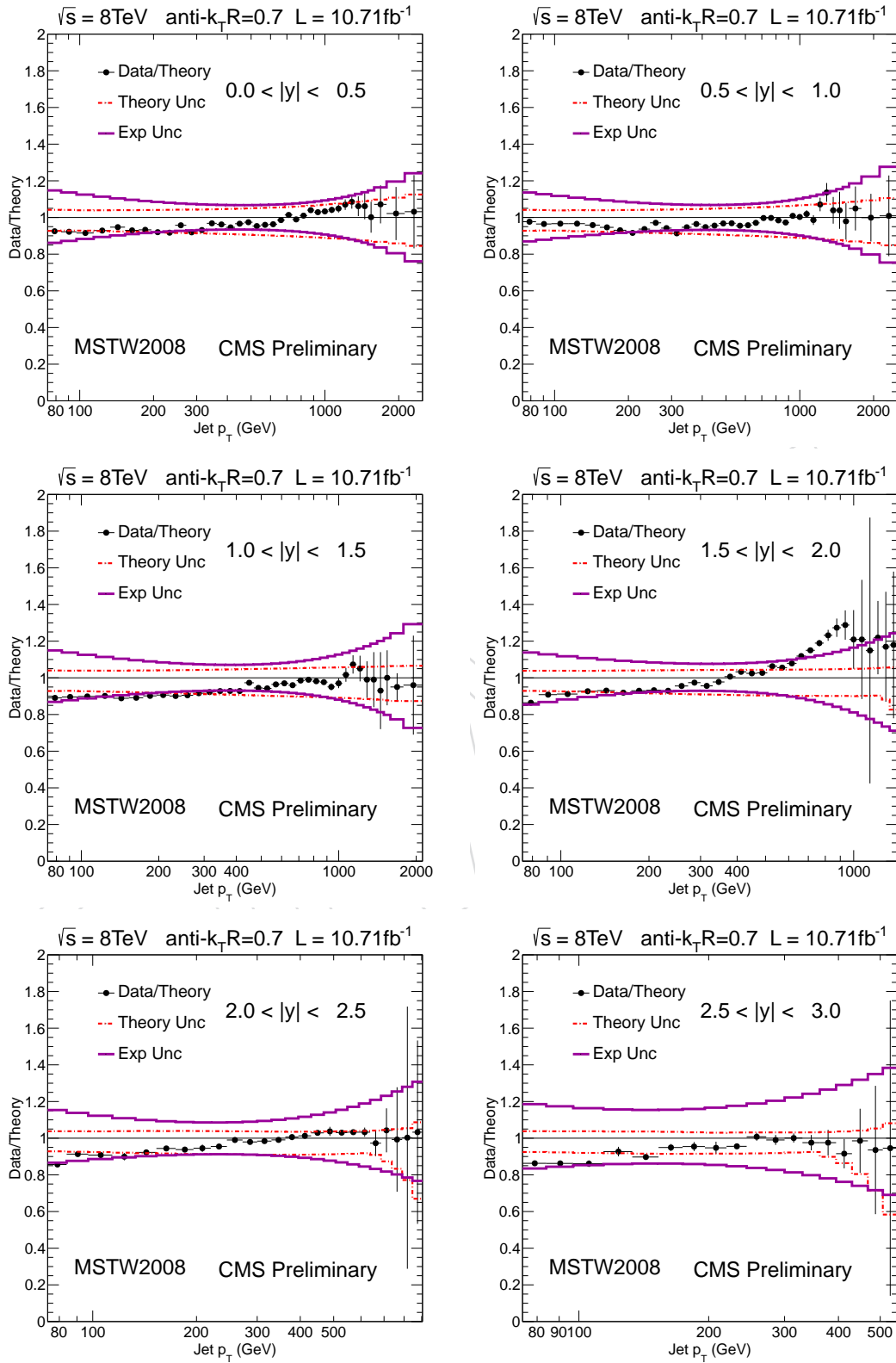


Figure 44: Data Over Theory Plot for MSTW

6 Spectrum Fitting

The raw and the unfolded spectrum are fitted to study how much bin to bin fluctuation is present in the measured spectrum. For this we choose the following eight parameter function to fit both the raw and the unfolded spectrum:

$$f(x) = A_0 \left(1 - \frac{2}{\sqrt{s}} \frac{x}{A_6}\right)^{A_7} \times 10^{F(x)},$$

where

$$F(x) = \sum_{i=1}^5 A_i \left(\log\left(\frac{x}{A_6}\right)\right)^i$$

The same equation was used by CDF collaboration, to fit the inclusive jet cross section measurement at $\sqrt{s} = 1.8 \text{ TeV}$ [17]. Here the parameter A_6 has been fixed to $\frac{\sqrt{s}}{2 \cosh(y_{min})}$, where $\sqrt{s} = 8000 \text{ GeV}$ and y_{min} is the minimum rapidity of the bin. The other parameters are derived from the fitting.

Here we mainly study the variable $(\text{Fit} - \text{Data})/(\text{Error})$ which is commonly known as pull variable, to see how much the entries are fluctuating in each bin from fitted spectrum.

The χ^2 for the unfolded distribution is derived for each rapidity bin using the formula:

$$\chi^2 = \sum_{i,j=1}^{NBINS} (\text{Data} - \text{Fit})/(\text{Error})_i C_{ij}^{-1} (\text{Data} - \text{Fit})/(\text{Error})_j;$$

Where the covariance matrix C is derived from unfolding. Figure 20 shows the correlation matrix for each rapidity bin for varying p_T bin. Here $NBINS$ denote the number of p_T bin for each rapidity bins. The trigger paths fire at p_T threshold values 74, 133, 220, 300, 395, 507 GeV respectively.

In figures 47 and 48, in the pull distribution we observe atmost 4σ fluctuation in the reco spectrum across varying rapidity bins, whereas the same fluctuation is atmost upto 2.5σ for the unfolded distribution.

7 Comparison With 7TeV

To check the consistency with earlier 7TeV CMS measurement, we study the quantity double ratio which is defined as

$$\text{DoubleRatio} = \frac{(\text{Data/Theory})_{7\text{TeV}}}{(\text{Data/Theory})_{8\text{TeV}}}$$

In this study the correlated systematic experimental uncertainties get cancelled from the numerator and denominator. The double ratio is obtained as a function of jet p_T . The amount of deviation from unity for the double ratio is studied to study the consistency with 7TeV measurement.

8 PileUp Stability

The pileup dependency on the measured spectrum is studied. For this purpose we choose two different fill, viz. fill 2513 and fill 2718. which have distinct pileup scenario as shown in figure 54.

Fill 2513 has only one run with run number 191226 and luminosity 93.55/pb. Fill 2718 has three runs with run numbers 195947, 195948 and 195950, where the corresponding luminosities are 6.51/pb, 62.447/pb and 111.138/pb respectively.

The ratio of unfolded spectrum from two different fills are studied with the total pileup uncertainty band as shown in figure 55. The figure shows that for all rapidity bins, at lower p_T , the statistical fluctuations are covered within the PU uncertainty. Hence there is no major effect of PU on the measured spectrum.

9 Summary

A measurement of the double-differential inclusive jet cross section has been presented. Using 10.71 fb^{-1} of data from proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$ collected with the CMS detector. The measurement covers the jet p_T range from 74 GeV to 2.5 TeV, in six rapidity bins, up to $|y| = 3.0$.

A detailed studies of various experimental and theoretical uncertainties, and their relative comparison are carried out. We find that the dominant source of systematic uncertainty arises due to the JES which results about 10-40% relative uncertainty in the cross section measurement across varying rapidity bins. On the other hand in theoretical calculations, PDF variation causes 10% to 50% uncertainty depending on the rapidity bins, where as choices of renormalization and factorization scales contributes 5% to 10% for central rapidity bin, 40% for the outer $|y|$ bins. Additional uncertainties due to the MPI and hadronization effects are also considered. Eventually, the measured double differential cross section is compared with the $NLO \times NP$ theory predictions based on perturbative QCD predictions corresponding to five different PDF sets. Detailed comparisons show a very good agreement between the two predictions within 5-10% relative fluctuation, which is within total theoretical and experimental uncertainty limits.

For the sake of comparison of measurements with the theory predictions, a ratio of double differential cross sections for both data and theory are computed for all accessible p_T range, for each rapidity bins (Fig. 42 - 46). Moreover, a comparison of theory predictions based on different PDF sets are also presented by comparing the ratios of double differential cross section between a given PDF set and with the other all PDF sets (Figs. 37 - 41) In all cases the comparison with data over theory show a very nice agreement 5-10% and within the theoretical and experimental uncertainty limits and also consistent with the results for the ratio between different PDF sets as well. It is to be noted that the parton momentum fractions x probed in this measurement cover the range $0.019 < x < 0.625$.

A cross check with earlier 7TeV measurement was done. Within the uncertainty limit, this measurement is in a very good agreement with the previous result.

This measurement confirms once more the fundamental principles of QCD used to describe the strong interactions between the hadrons. This new inclusive jet cross sections measurement probes wide ranges in x and hence can be used widely to cross check many aspects of QCD which directly involves jet observables. For instance, using this result, one can constrain PDFs

in different x regions and as well as can extract the strong coupling constant(α_s).

References

- [1] CMS Collaboration, “Jet cross sections and PDF constraints”,.
- [2] CMS Collaboration, “Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss} ”, *CMS Physics Analysis Summary CMS-PAS-PFT-09-001* (2009).
- [3] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063.
- [4] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026.
- [5] S. Agostinelli et al., “Geant 4 – A Simulation Toolkit”, *Nucl. Inst. Meth.* **A506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [6] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002.
- [7] T. Adye, “Unfolding algorithms and tests using RooUnfold”, arXiv:physics.data-an/1105.1160.
- [8] T. Adye, “Unfolding algorithms and tests using RooUnfold”, *ArXiv e-prints* (May, 2011) arXiv:1105.1160.
- [9] G. D’Agostini, “Improved iterative Bayesian unfolding”, *ArXiv e-prints* (October, 2010) arXiv:1010.0632.
- [10] Z. Nagy, “Next-to-leading order calculation of three jet observables in hadron-hadron collision”, *Phys. Rev.* **D68** (2003) 094002, doi:10.1103/PhysRevD.68.094002.
- [11] T. Kluge, K. Rabbertz, and M. Wobisch, “fastNLO: Fast pQCD calculations for PDF fits”, arXiv:hep-ph/0609285v2.
- [12] H.-L. Lai et al., “New parton distributions for collider physics”, *Phys. Rev.* **D82** (2010) 074024, doi:10.1103/PhysRevD.82.074024.
- [13] A. D. Martin et al., “Parton distributions for the LHC”, *Eur. Phys. J.* **C63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5.
- [14] R. D. Ball et al., “A first unbiased global NLO determination of parton distributions and their uncertainties”, *Nucl. Phys.* **B838** (2010) 136, doi:10.1016/j.nuclphysb.2010.05.008.
- [15] H1 and ZEUS Collaboration, “Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA”, *JHEP* **01** (2010) 109, doi:10.1007/JHEP01(2010)109.
- [16] S. Alekhin, J. Blümlein, and S. Kleinetal, “The3-,4-,and 5-flavor NNLO Parton from Deep-Inelastic-Scattering Data and at Hadron Colliders”, *Phys. Rev.* **D81** (2010) 014032, doi:10.1103/PhysRevD.81.014032.

- 448 [17] CDF Collaboration Collaboration, “Measurement of the inclusive jet cross section in $\bar{p}p$
449 collisions at $\sqrt{s} = 1.8$ TeV”, *Phys.Rev.* **D64** (2001) 032001,
450 doi:10.1103/PhysRevD.65.039903, 10.1103/PhysRevD.64.032001,
451 arXiv:hep-ph/0102074.

DRAFT

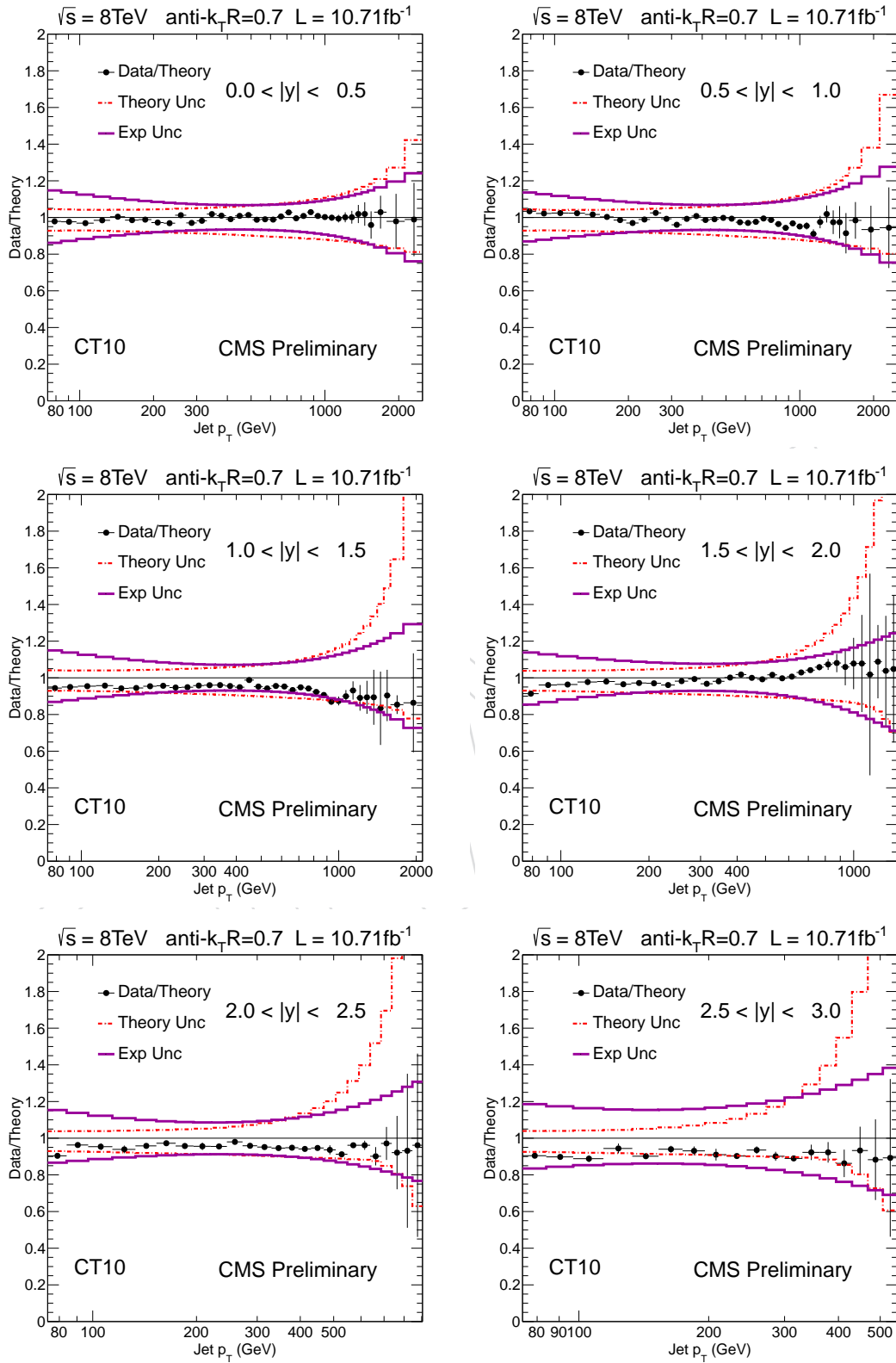


Figure 45: Data Over Theory Plot for CT10

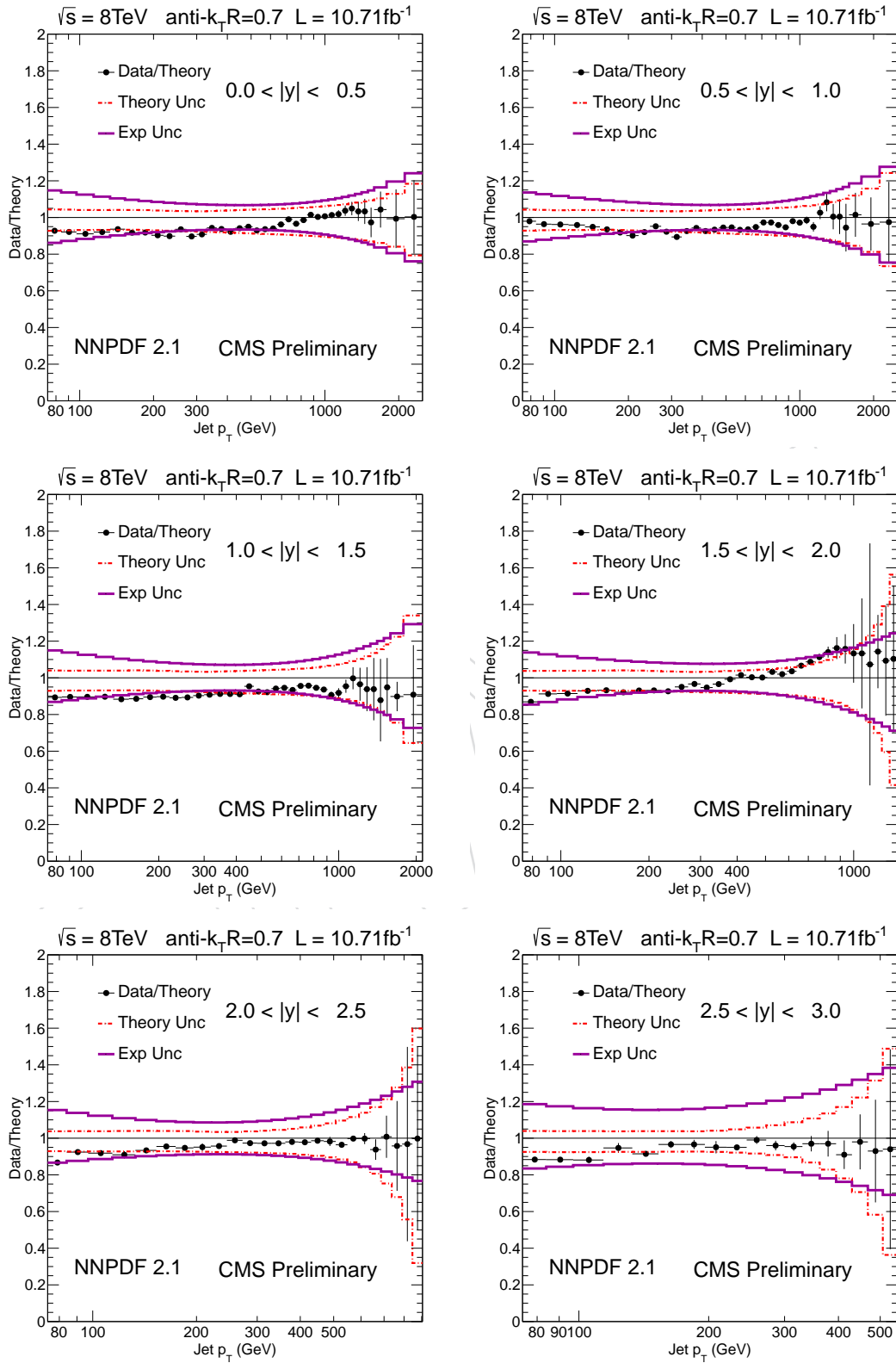


Figure 46: Data Over Theory Plot for NNPDF

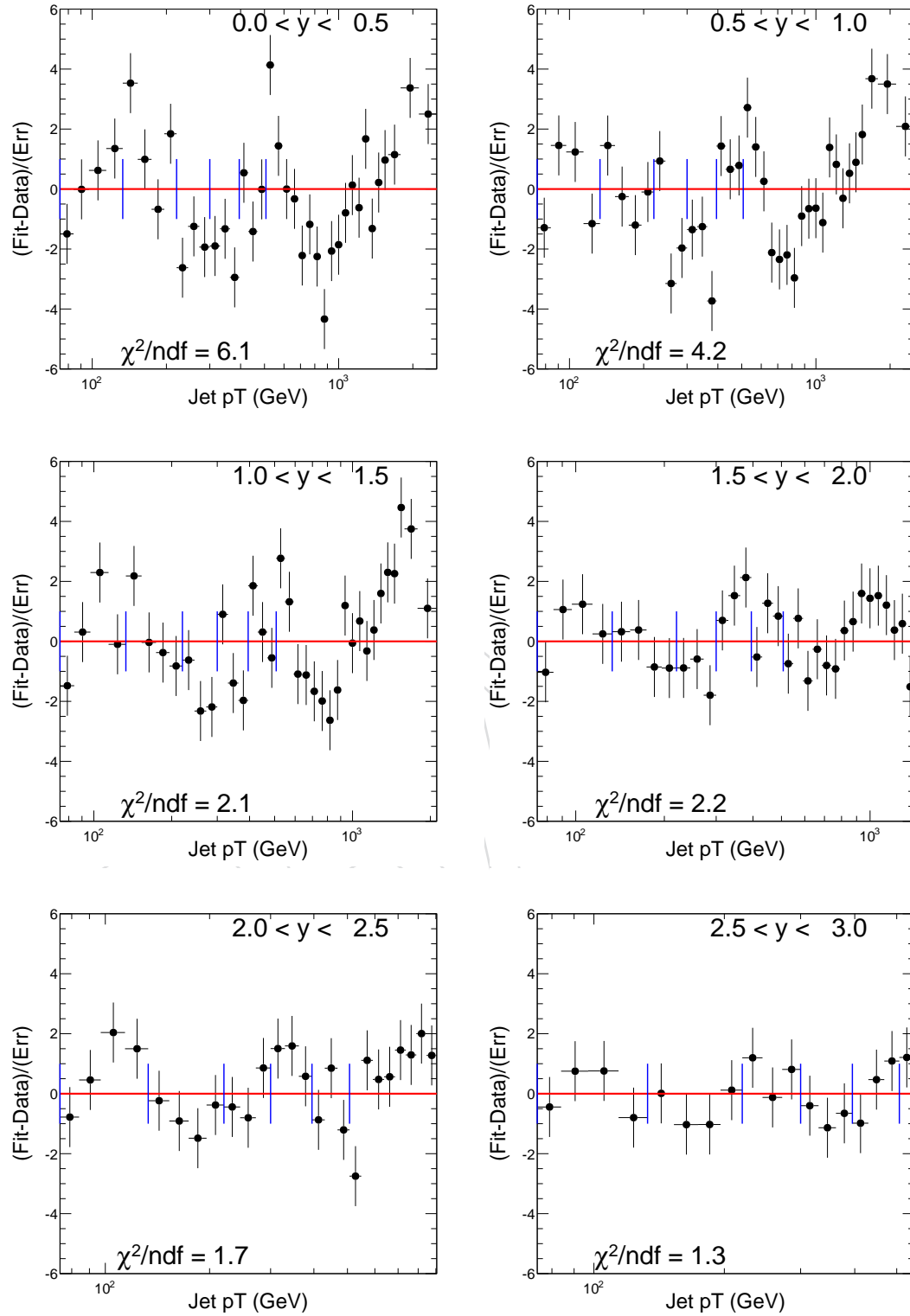


Figure 47: The pull variable distribution for the measured spectrum. The error bars shown here denotes 1 standard deviation fluctuation. Across different rapidity bins, we see atmost 4σ fluctuation. The blue line shows the thresholds where each individual trigger path takes on.

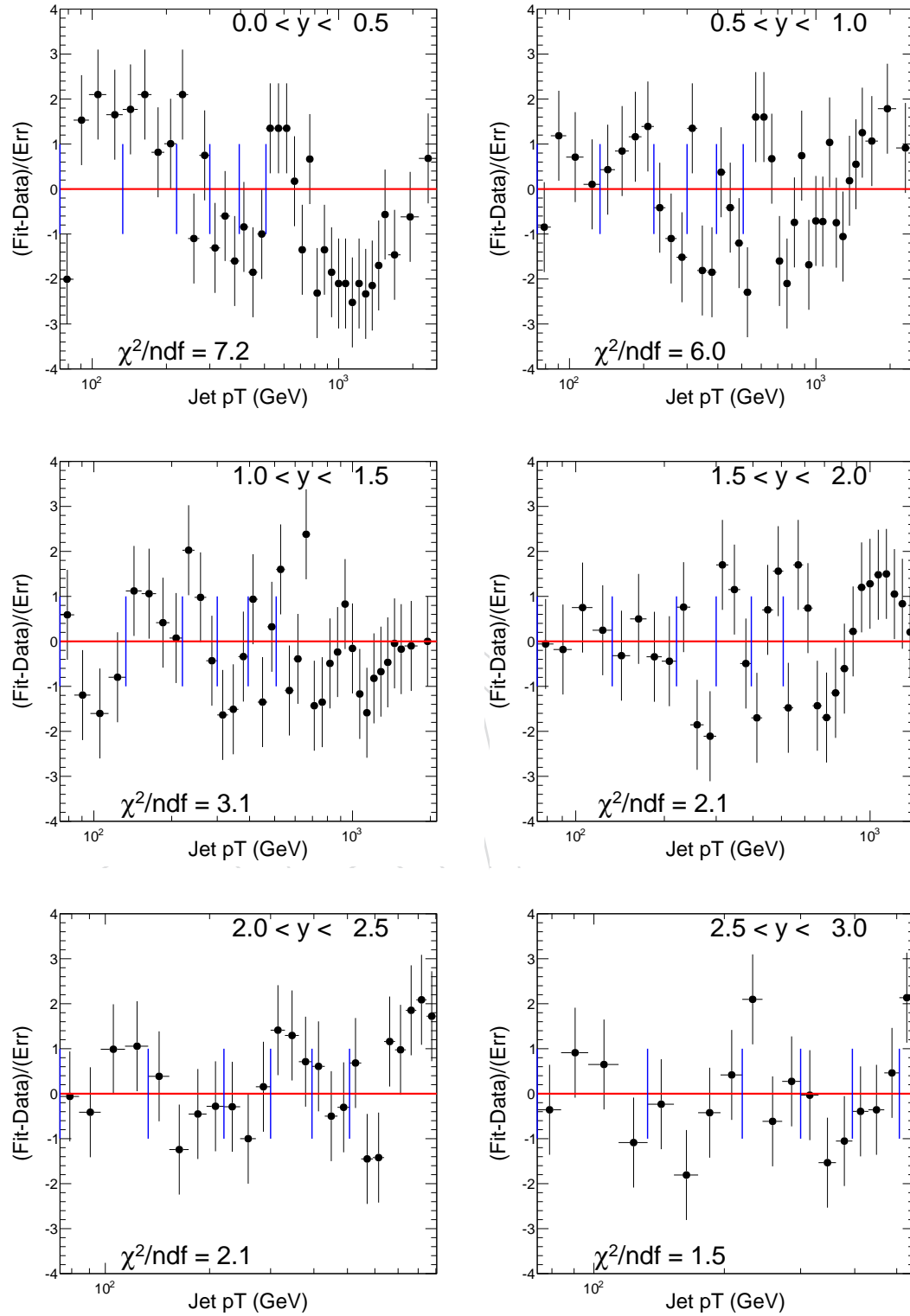


Figure 48: The pull variable distribution for the unfolded spectrum. The error bars shown here denotes 1 standard deviation fluctuation. Across different rapidity bins, we see at most 2.5σ fluctuation. The blue line shows the thresholds where each individual trigger path takes on

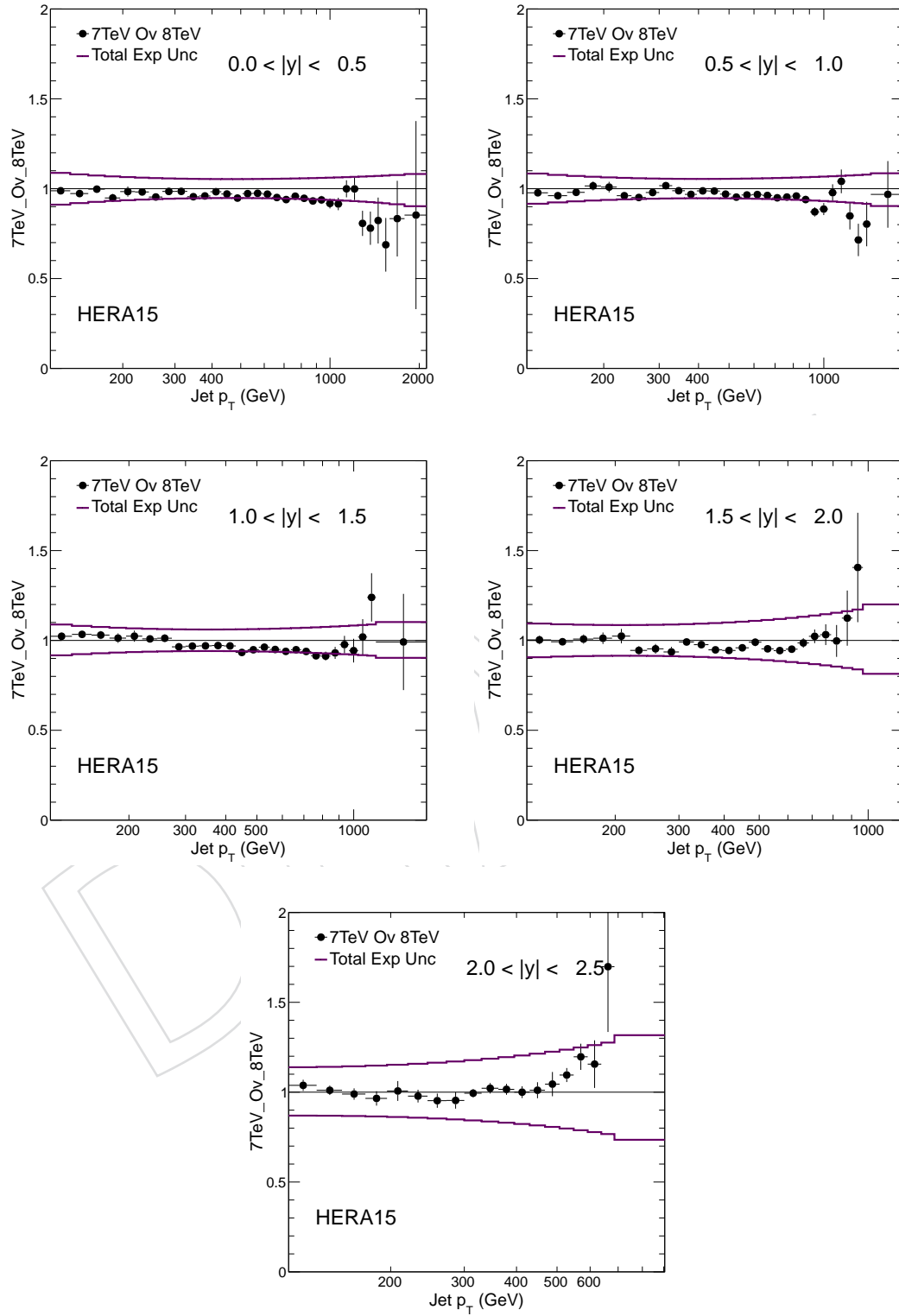


Figure 49: Double Ratio Plot for HERA PDF. The uncertainty band is computed taking the correlations into account.

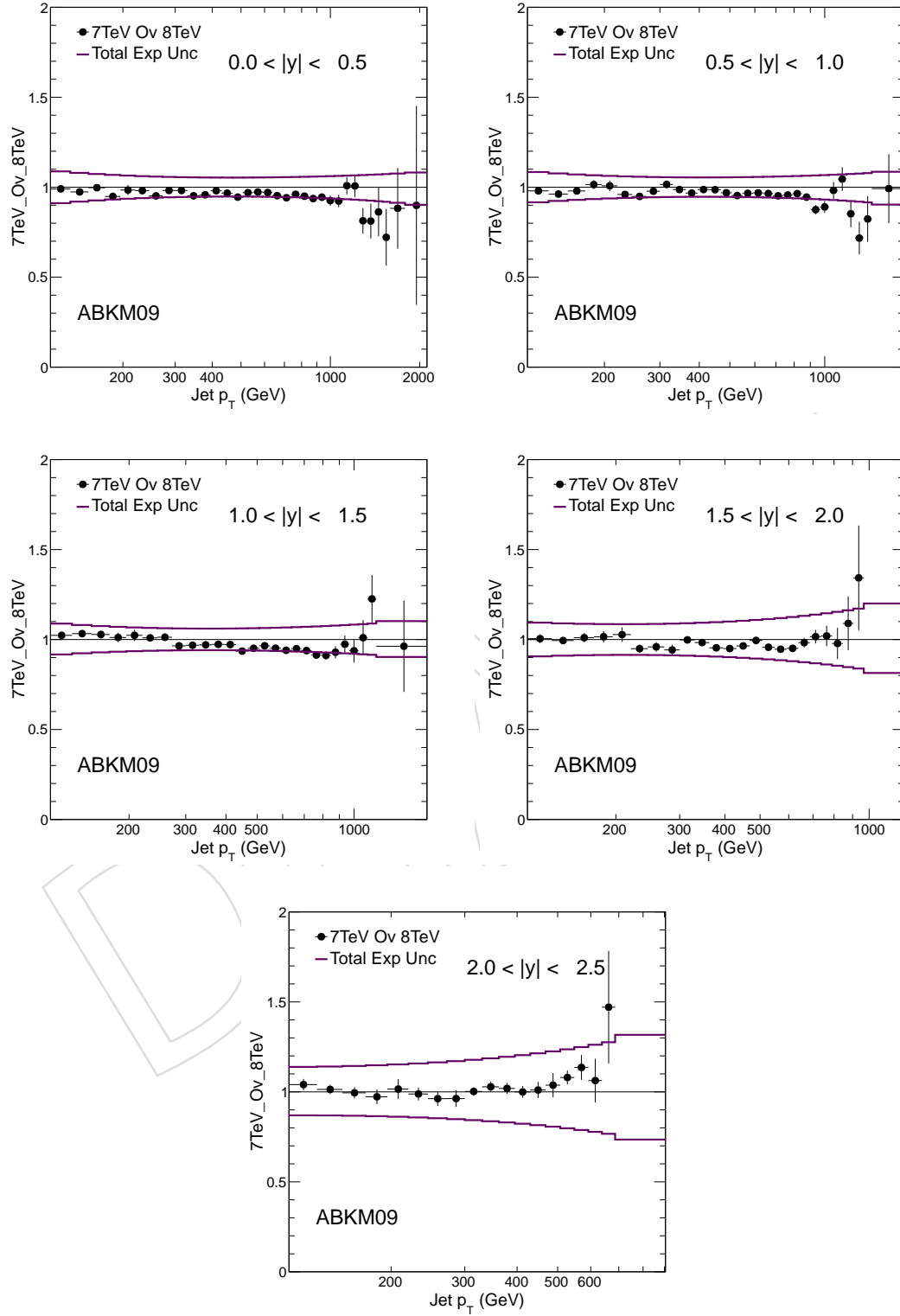


Figure 50: Double Ratio Plot for AKM09 PDF. The uncertainty band is computed taking the correlations into account.

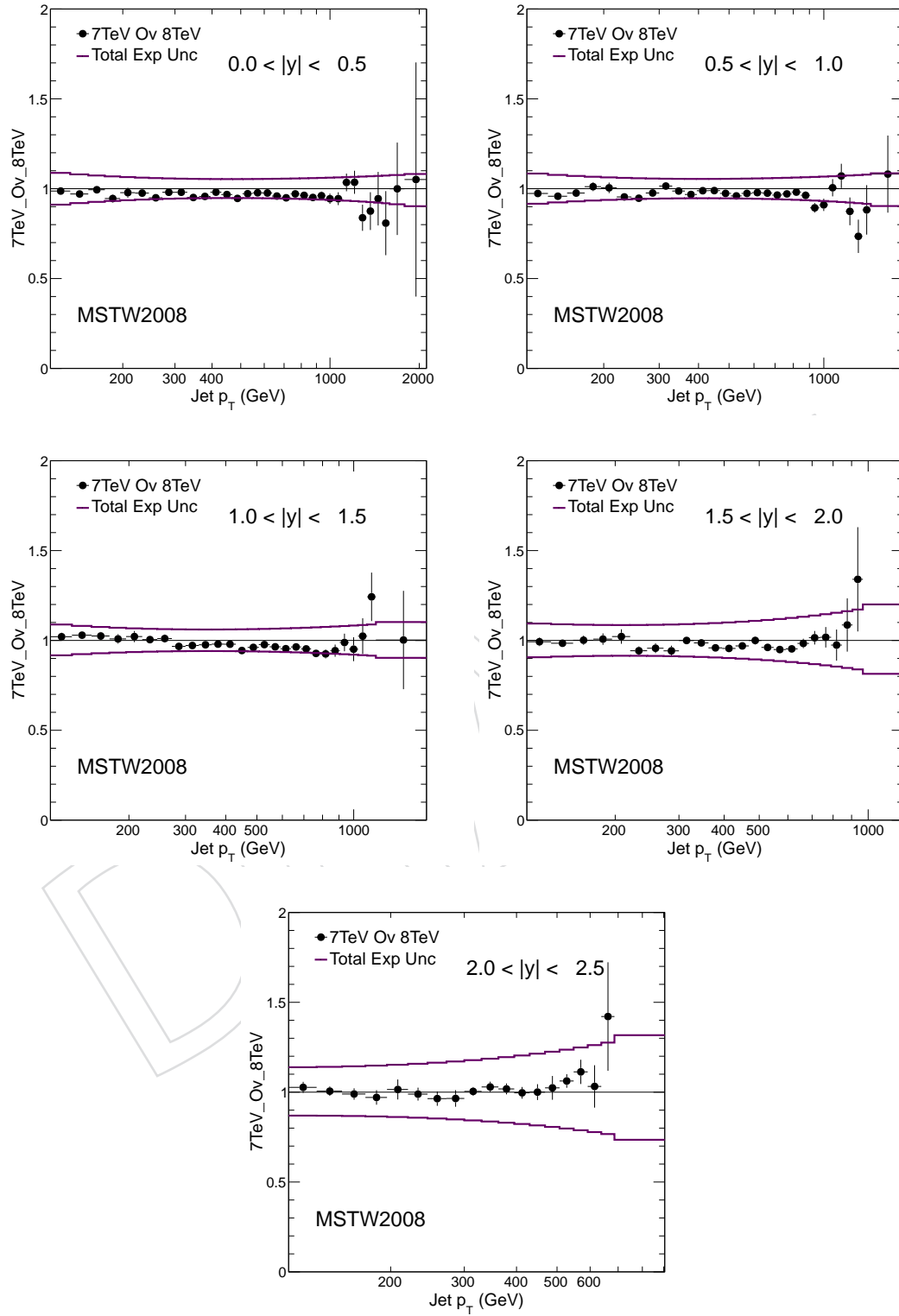


Figure 51: Double Ratio Plot for MSTW PDF. The uncertainty band is computed taking the correlations into account.

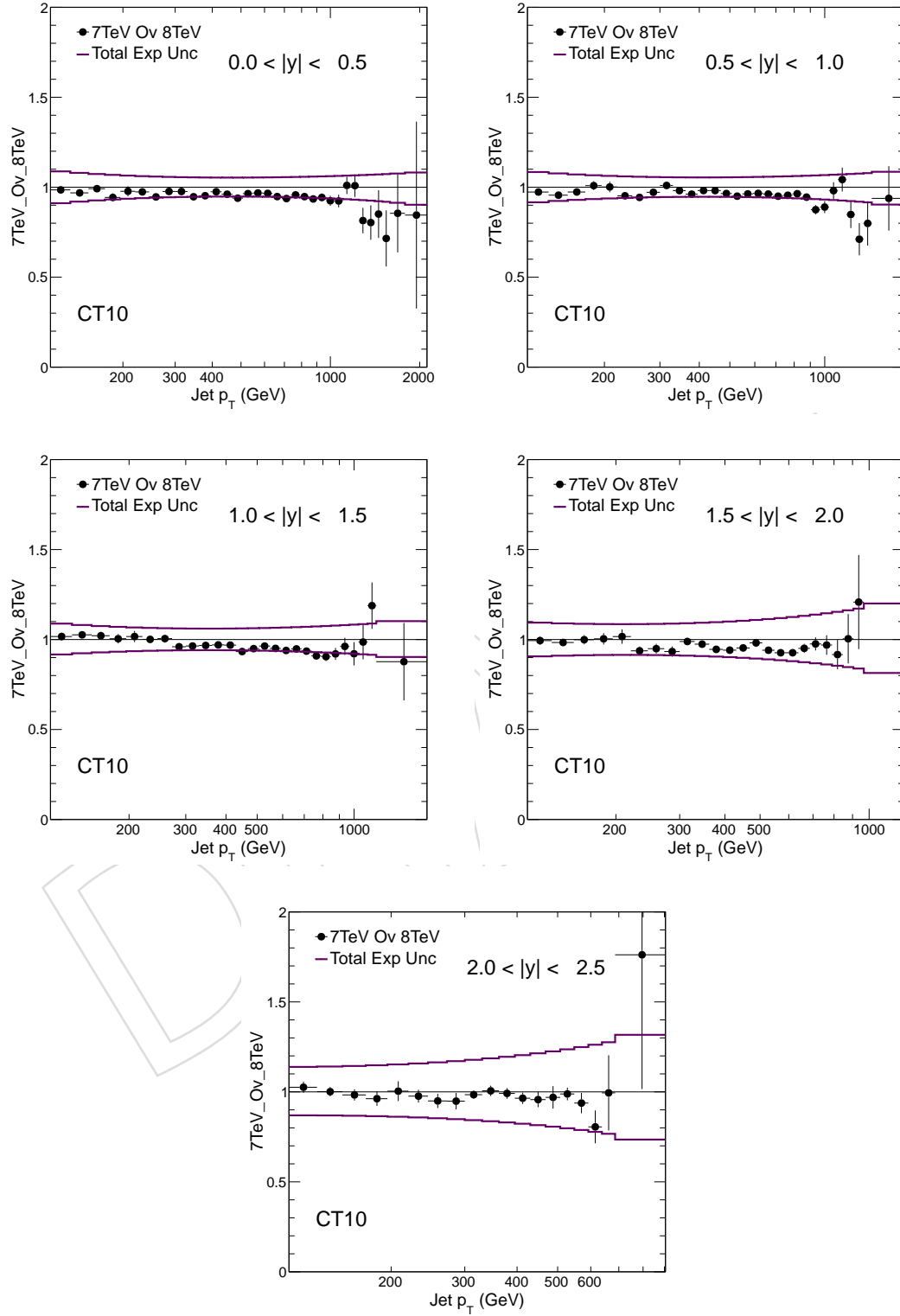


Figure 52: Double Ratio Plot for CT10 PDF. The uncertainty band is computed taking the correlations into account.

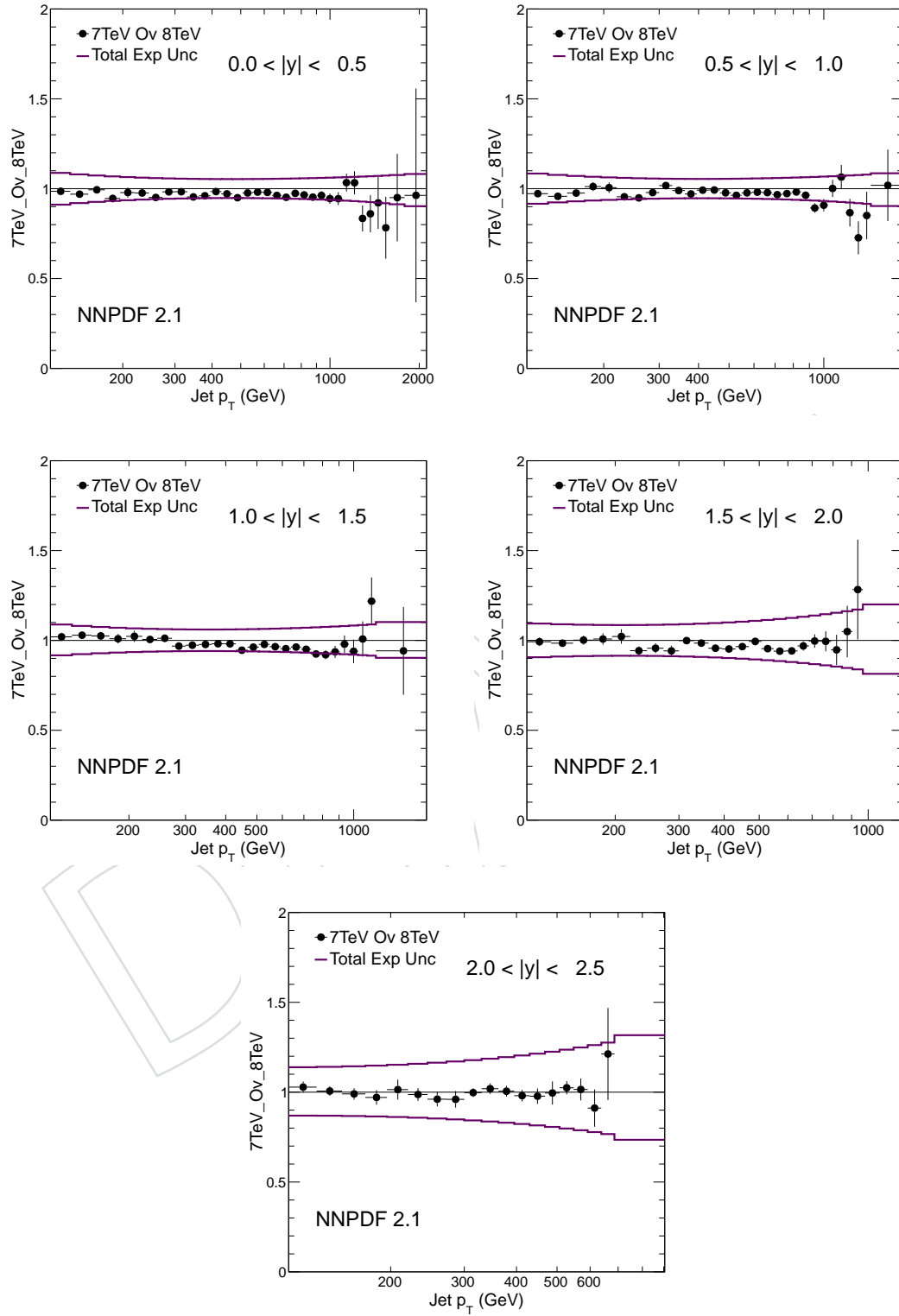


Figure 53: Double Ratio Plot for NNPDF PDF. The uncertainty band is computed taking the correlations into account.

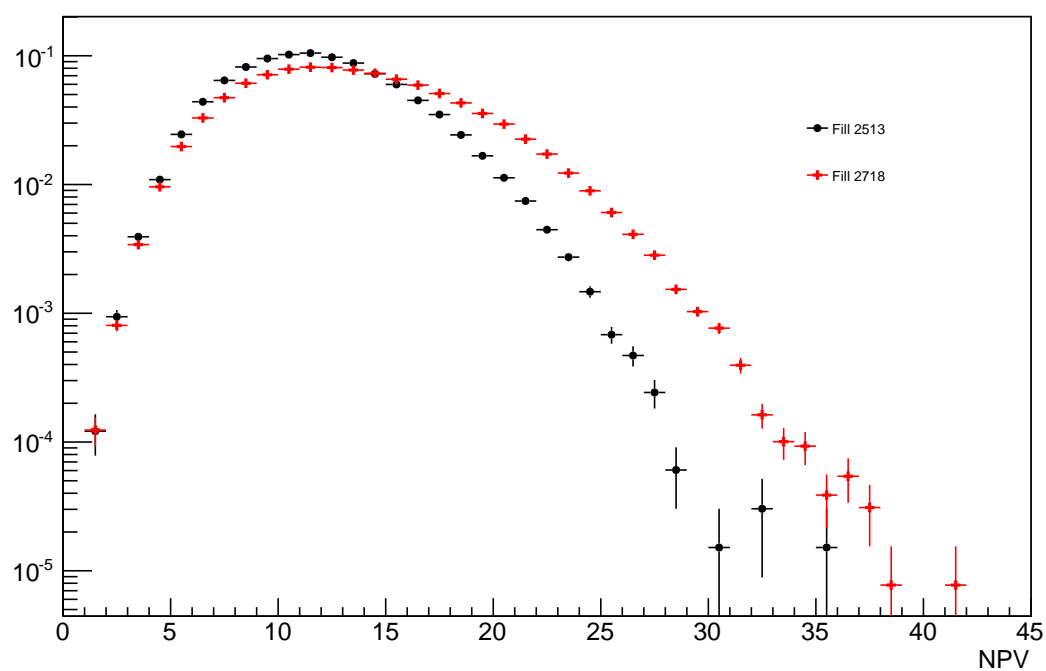


Figure 54: Number of primary vertex distribution for fill 2513 and 2718, showing the two different pileup scenario.

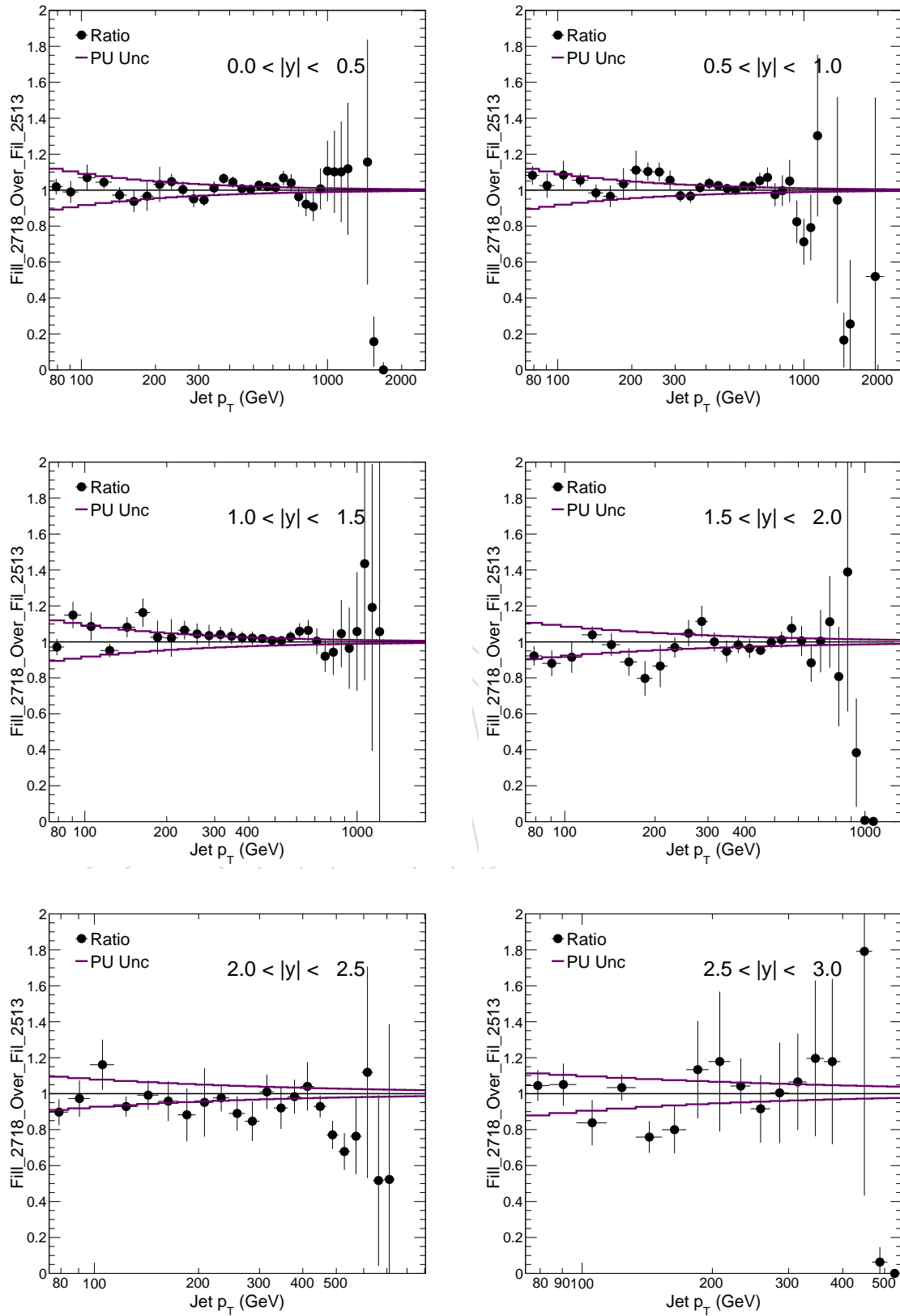


Figure 55: The ratio of unfolded spectrum for two different fills. The violet bands are showing the relative uncertainty on the measured cross section arising due to pileup. The total statistical fluctuations are covered by the PU uncertainty band in lower p_T .