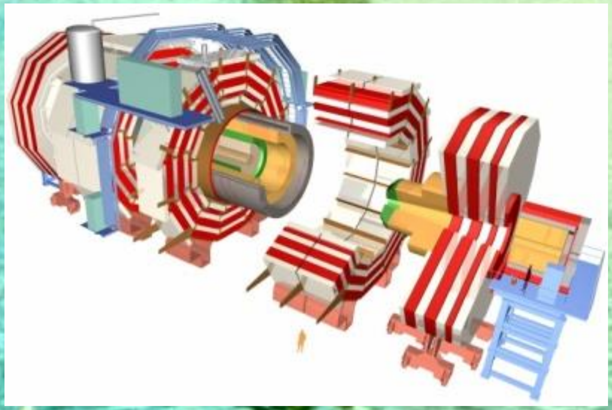
Measurement of jet production in association with a Z boson at the LHC & Jet energy correction & calibration at HLT in CMS

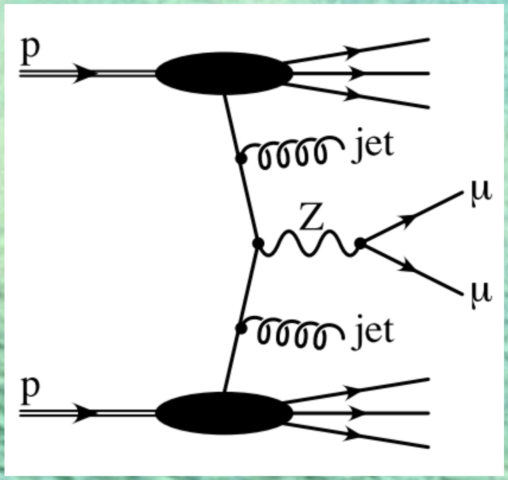
Fengwangdong Zhang

Peking University (PKU) & Université Libre de Bruxelles (ULB)



PhD Public Defense
@PKU

08/06/2017



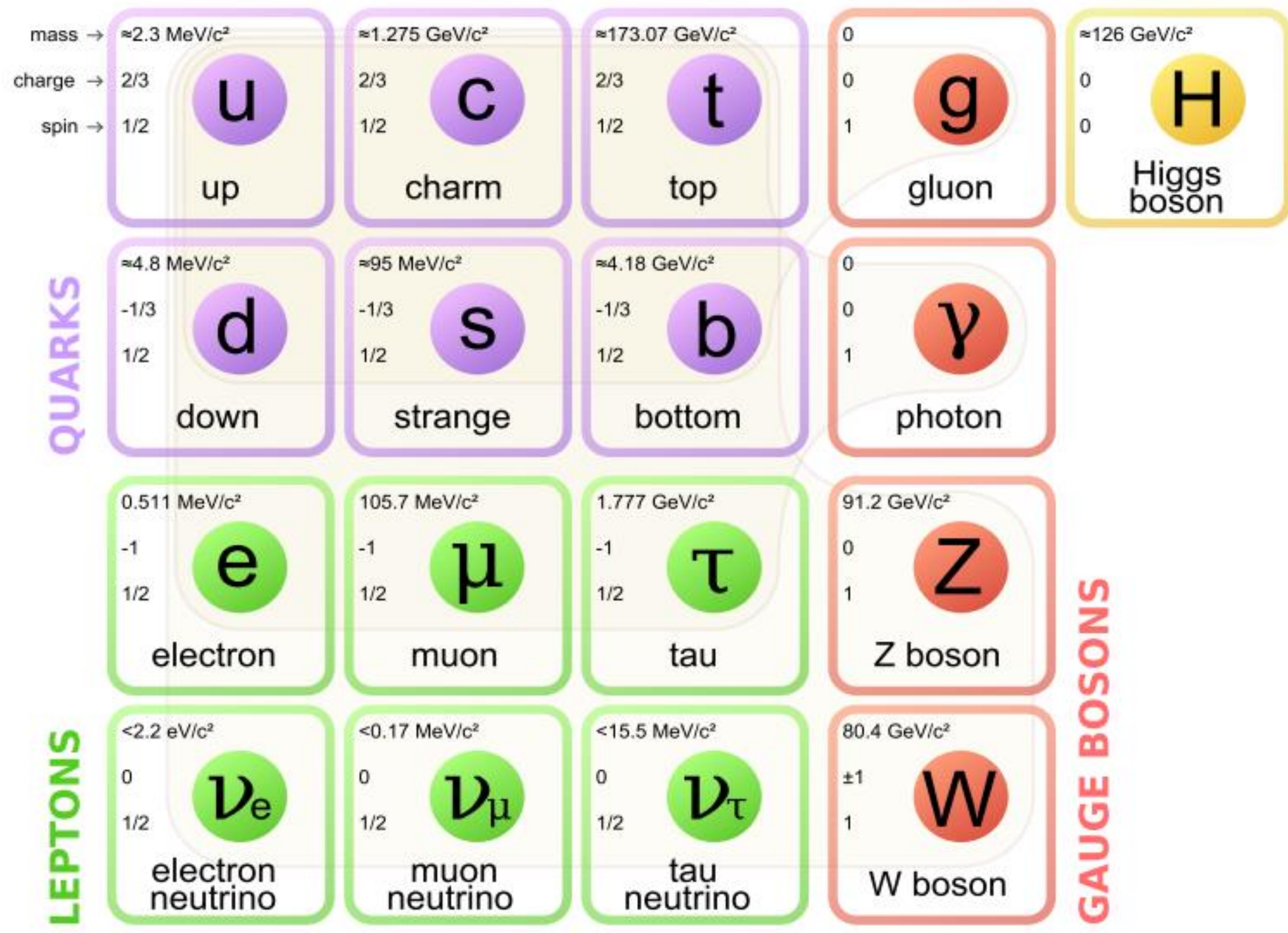
Outline

- Introduction
  - Standard Mode theory of particle physics
  - Quantum Chromodynamics coupling
  - Factorization & parton distribution function
  - Hard process & radiation correction
  - Parton shower
  - Hadronization & jet production
  - Multiple parton interaction
  - Importance of Z + jets measurement
  - Monte Carlo simulation
  - Large Hadron Collider & Compact Muon Solenoid

- Main data analysis (Z+jets measurement)
  - Event display
  - Data & Simulation preparation
  - Methodology for the cross section measurement
  - Uncertainties
  - Channel combination approach
  - Theoretical predictions
  - Results
  - Publication status (+ Jet energy correction & calibration at high level trigger)

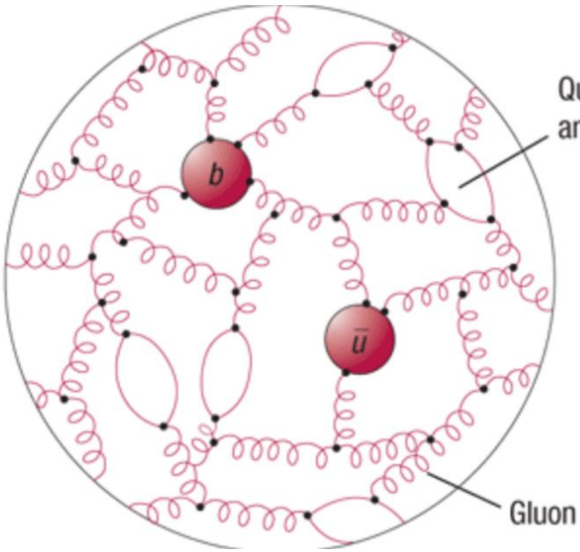
- Conclusion

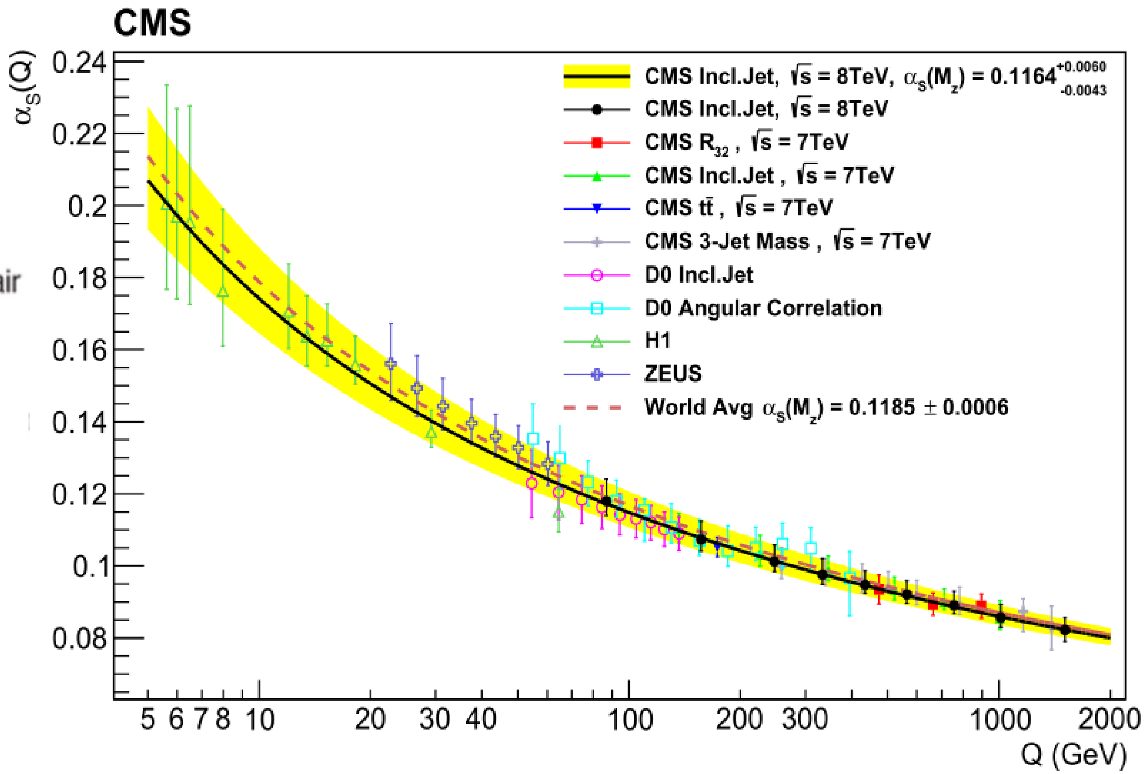
Standard Model of particle physics



- Comprehensive theory to encode strong, electromagnetic, weak interactions in the universe
- Successful prediction of the quantification of Higgs field in nature: Higgs boson

Quantum Chromodynamics Coupling





- ➢ QCD interactions between quarks and anti-quarks are realized by gluon exchange
- ➢ Self-interaction of gluons give rise to the running coupling strength
- ➢ Higher energy scale results in smaller coupling strength
- ➢ α_s becomes small enough - observables can be treated in perturbative way:

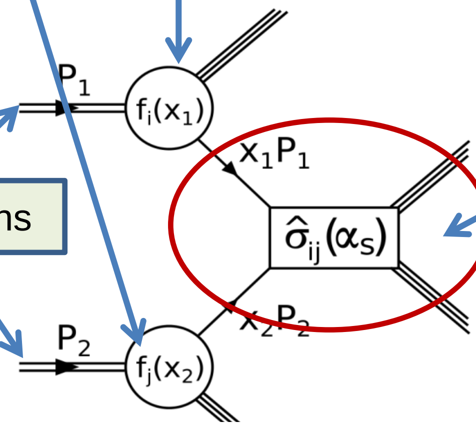
$$X = \sum_{n=0}^{N} a_n \alpha_s^n = a_0 + a_1\alpha_s + a_2\alpha_s^2 + \ldots$$

How can we describe proton-proton interaction

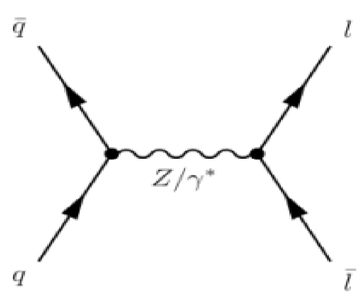
Probability of finding a parton carrying a momentum fraction (x1 / x2) in the proton (PDF)

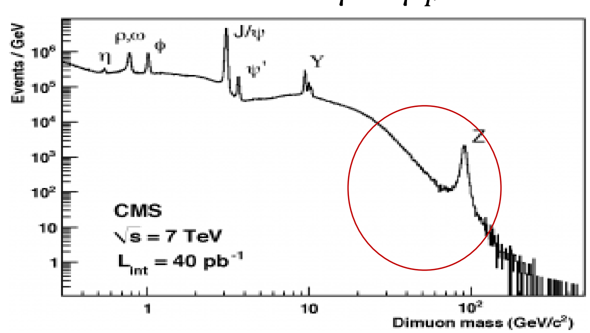
Cross section of parton interaction (matrix element computation)

momenta of the incident protons



- Factorization: $\sigma(P_1, P_2) = \sum_{ij} \int dx_1 dx_2 \left[f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \times \sigma_{ij} \left(p_1, p_2, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_R^2} \right) \right]$
  - The parton cross section σ_{ij} is computed in perturbative way
  - Factorization holds for large scale $\mu^2 = Q^2$
- PDF f(x, μ) evolves from a designed energy scale to factorization scale μ_F
  - Described by DGLAP equation (see next page)
  - PDF contains non-perturbative part of proton structure with $\mu < \mu_F$
- Drell-Yan process:

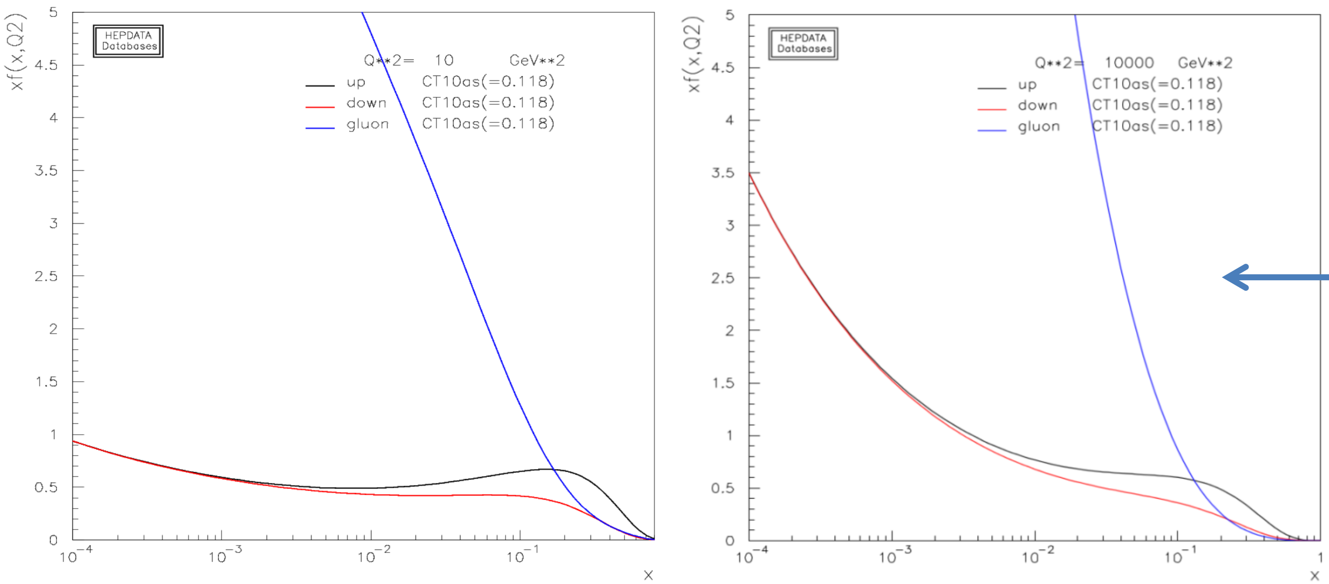




PDF evolution

- DGLAP equation:

$$Q^2 \frac{\partial}{\partial Q^2} \begin{pmatrix} q_i(x,Q^2) \\ g(x,Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \sum_{q_j,\bar{q}_j} \int_x^1 \frac{d\xi}{\xi} \times \begin{pmatrix} P_{q_i q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) & P_{q_i g}(\frac{x}{\xi}, \alpha_s(Q^2)) \\ P_{g q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) & P_{gg}(\frac{x}{\xi}, \alpha_s(Q^2)) \end{pmatrix} \begin{pmatrix} q_j(\xi,Q^2) \\ g(\xi,Q^2) \end{pmatrix}$$

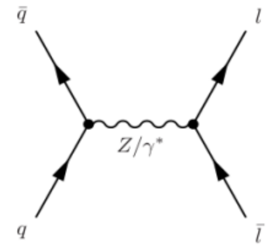


- PDF sets used:
  - CTEQ6L1
  - CT10
  - CT14
  - NNPDF3.0

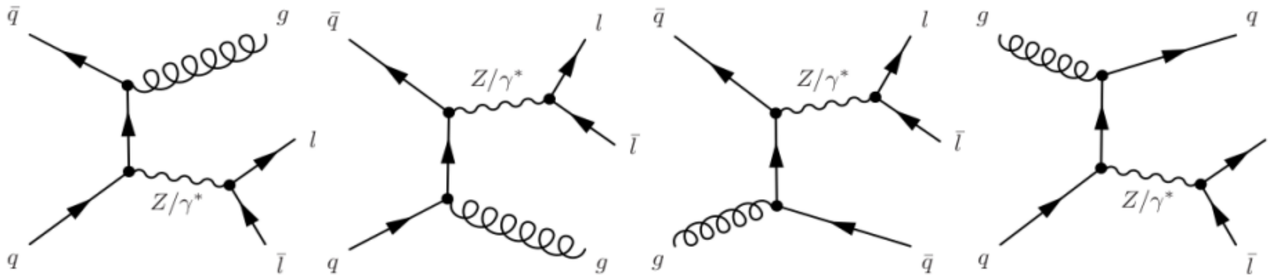
- PDF of proton can be extracted from the data of lepton-proton interactions
- Higher scale gives rise to larger gluon and sea quark distributions in small x region

Hard process matrix element of Drell-Yan process

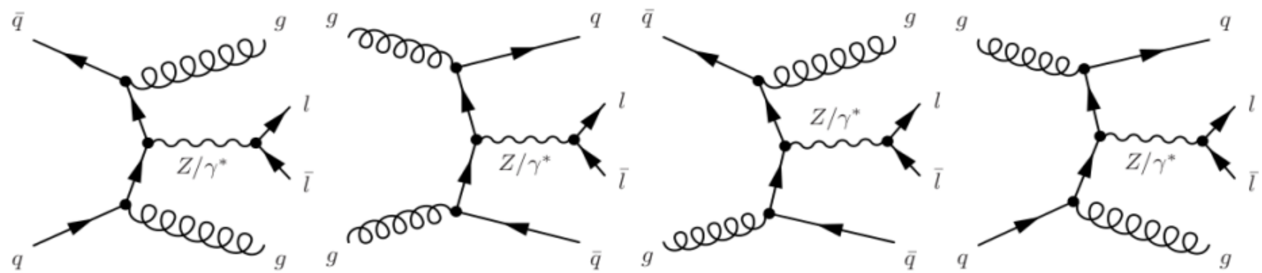
Leading order (pure electroweak interaction)



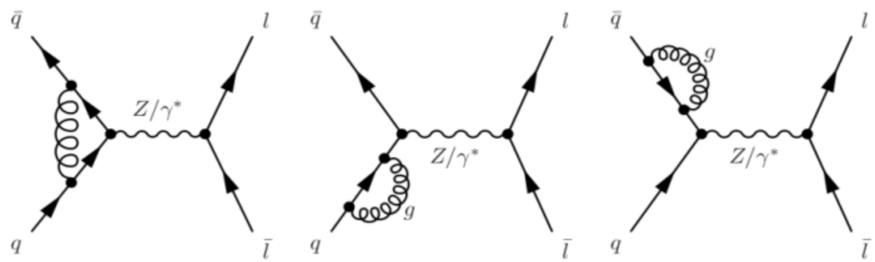
First order of QCD correction (one parton in the final state)



Second order of QCD correction (two partons in the final state)

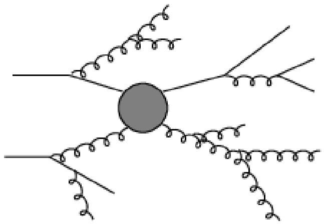


First order of QCD correction (no additional parton in the final state)



Parton shower

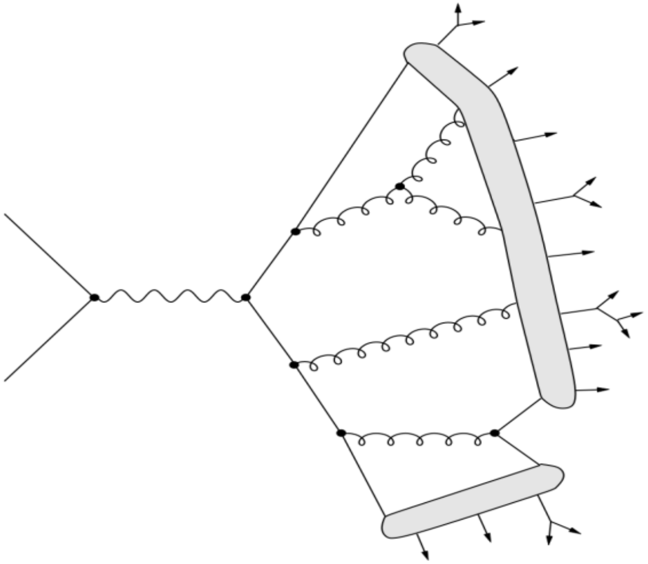
- Matrix elements are complemented by parton shower : modelized shower with quark/gluons radiations
- The physical phase space of parton emission is splitted into two parts in k_T:
  - ME contribution : $p_T > k_T$
  - Parton shower domain : $p_T < k_T$

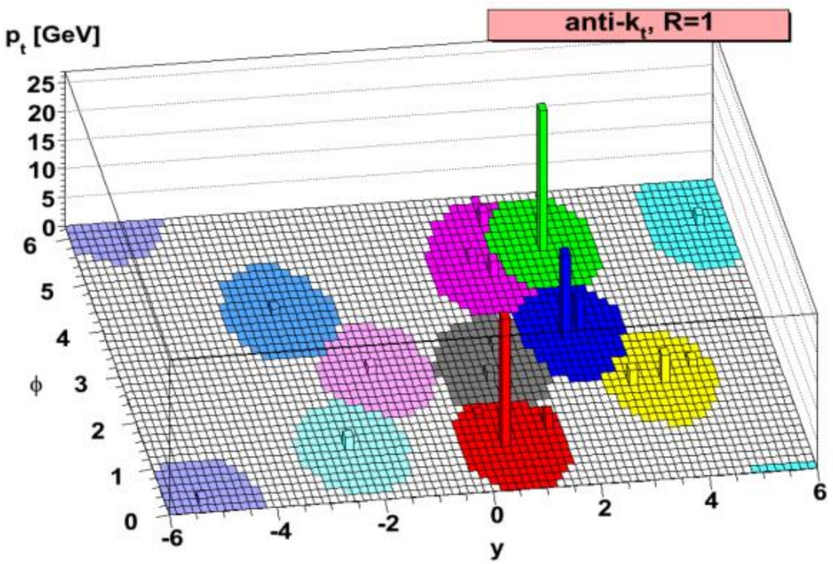


- Parton shower happens to all partons of initial and final states
  - Overcome the infrared divergence of real gluon emission corrections in the matrix element
  - Mimic missing higher order QCD radiation in the matrix element (always limited order calculation for matrix element)
- Parton shower starts from $Q^2 = {k_T}^2$ and stops at hadronization scale
- Matching scheme between matrix element and parton shower are needed to avoid double counting

- List of matching scheme of ME-PS in this analysis:
  - KT-MLM (LO)
  - FxFx (NLO)
  - CKKW (LO)
  - MEPS@NLO (NLO)

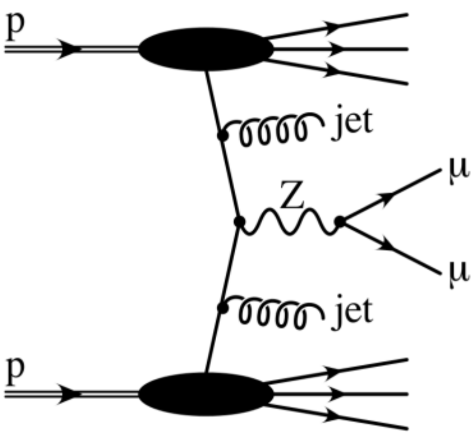
Hadronization & Jet Production

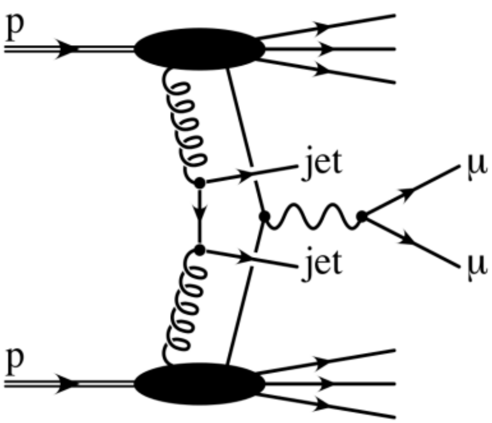




- Hadronization happens after parton shower due to quark-antiquark/gluon confinement
- Hadrons are clustered in jets through specific algorithm (e.g. anti-kt)

Multiple Parton Interaction (MPI)





- Secondary parton interaction dominated by low energy scale while becoming negligible (double parton scattering) at high energy scale

Importance of Z+jets process measurement

- Z + jets process prediction requires all of the above steps:
  - It allows precision study of each step
  - In this analysis we focus on the jet production
  - High cross section
  - Low background contamination

- Important for the modeling of the production mechanism involved in new physics searches (e.g. Supersymmetry)

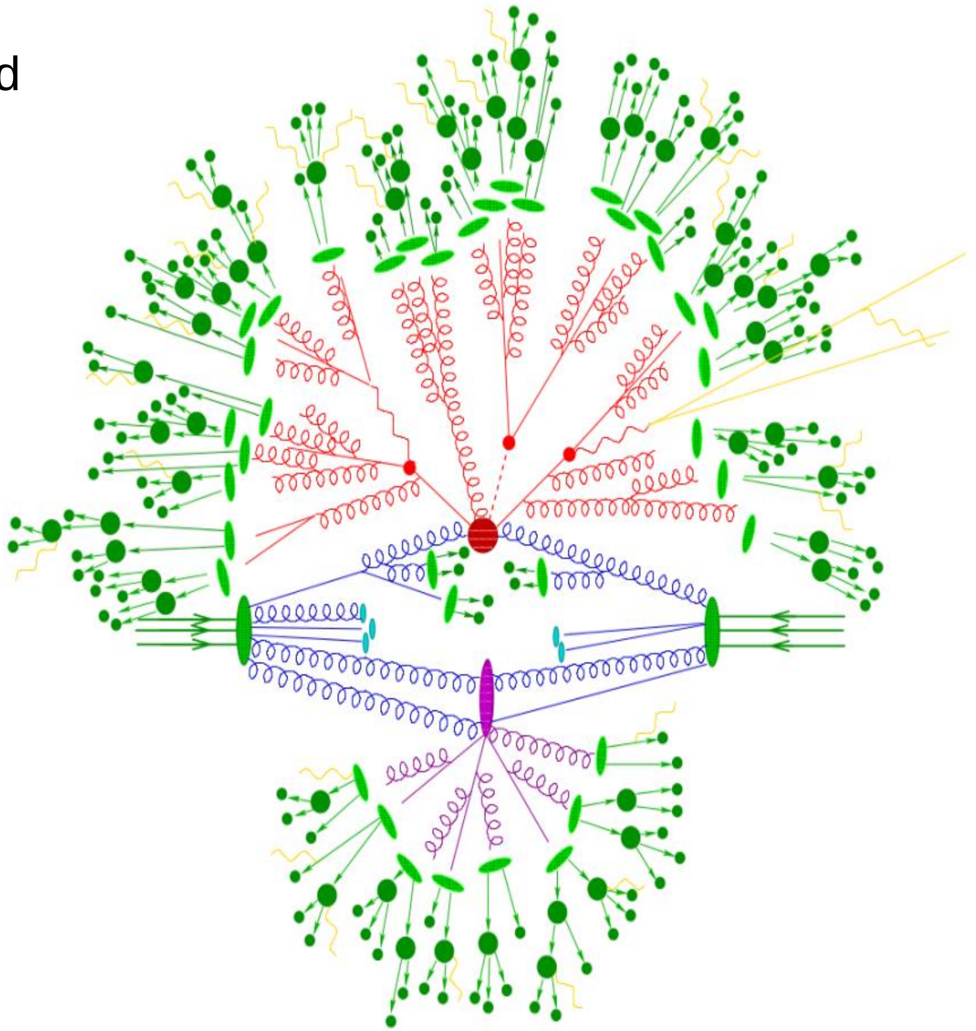
- Z+jets is a dominant background for:
  - Top-quark measurements
  - Vector boson fusion measurement
  - Precision measurement of Higgs physics in VH(->bb) channel

- Measurement on the cross section of Z+jets as a function of different kinematical observables is crucial with highest possible precision

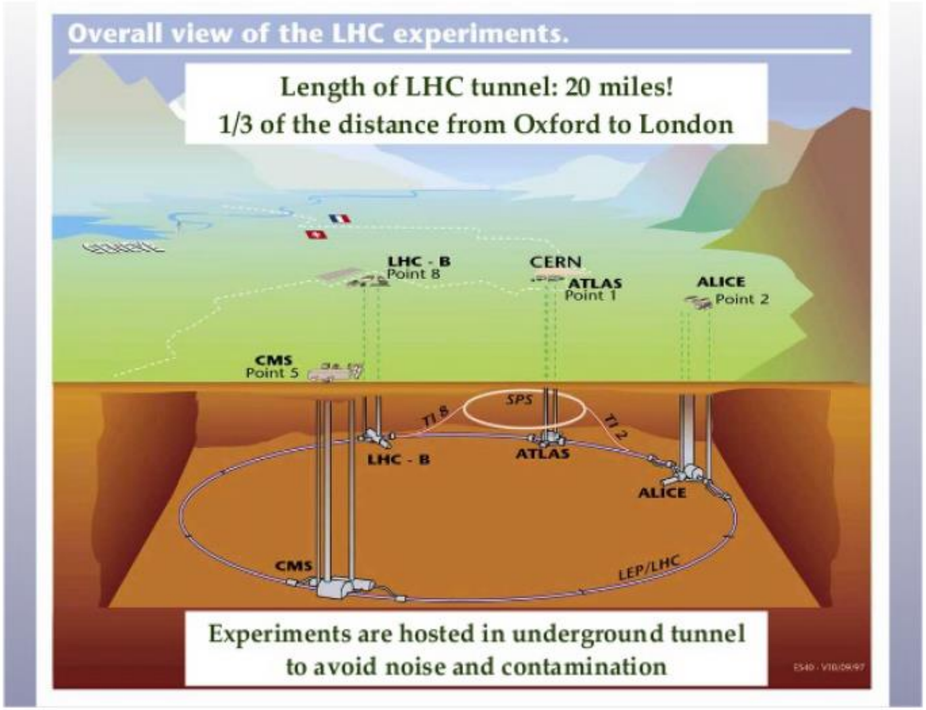
Monte Carlo simulation

- Special Monte Carlo tools are developed to generate pseudo-data, mimicking the parton interactions when hadrons are colliding

- Different models are implemented for:
    - PDF
    - Partonic cross section (matrix element)
    - Parton shower
    - Hadronization
    - Multiple parton interactions

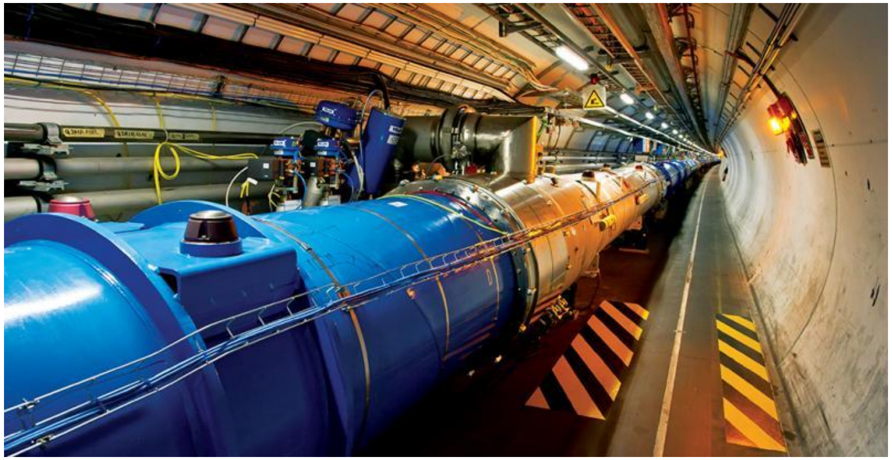


Large Hadron Collider

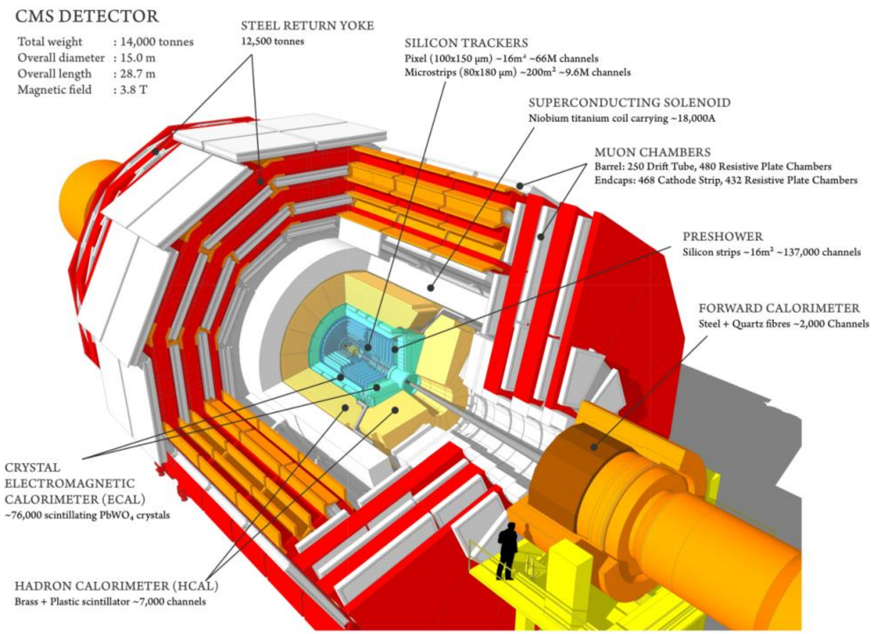


- Major hadron/ion accelerator that human being has made so far
- Four main experiments: ATLAS, CMS, LHCb, ALICE
- The designed collision energy of protons is 14 TeV

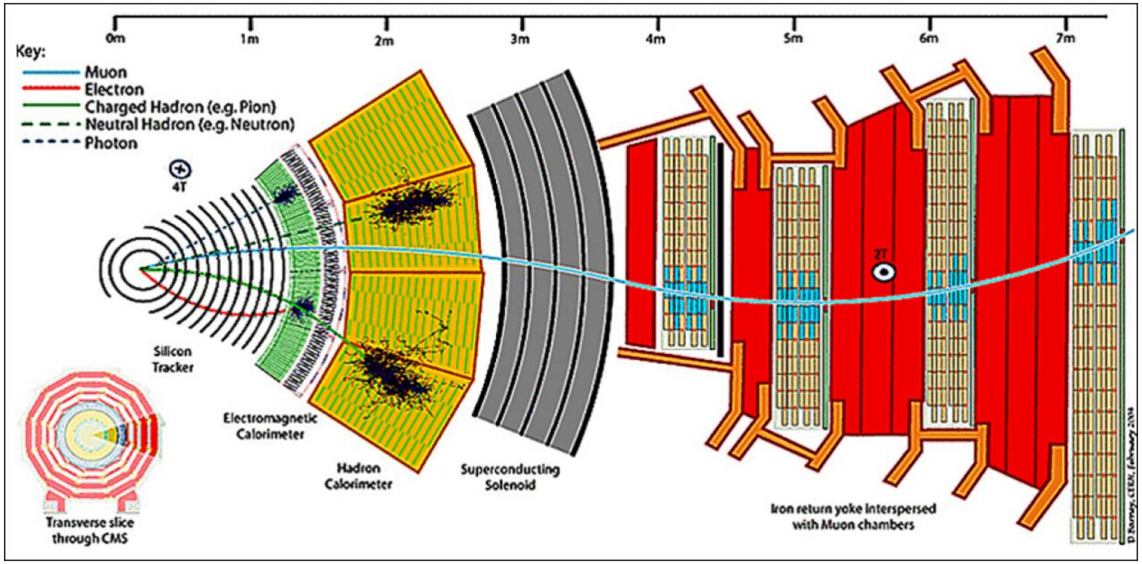
- Major achievement on physics: discovery of Higgs boson (main goal)
- Searches for new physics
- Precise measurement of Standard Model theory of particle physics

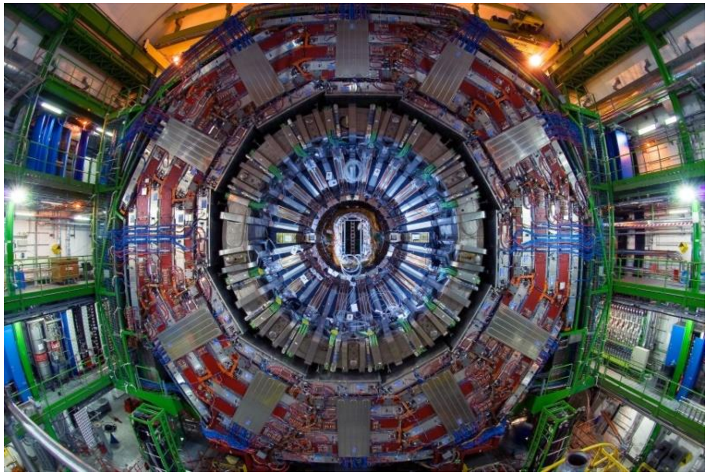


Compact Muon Solenoid

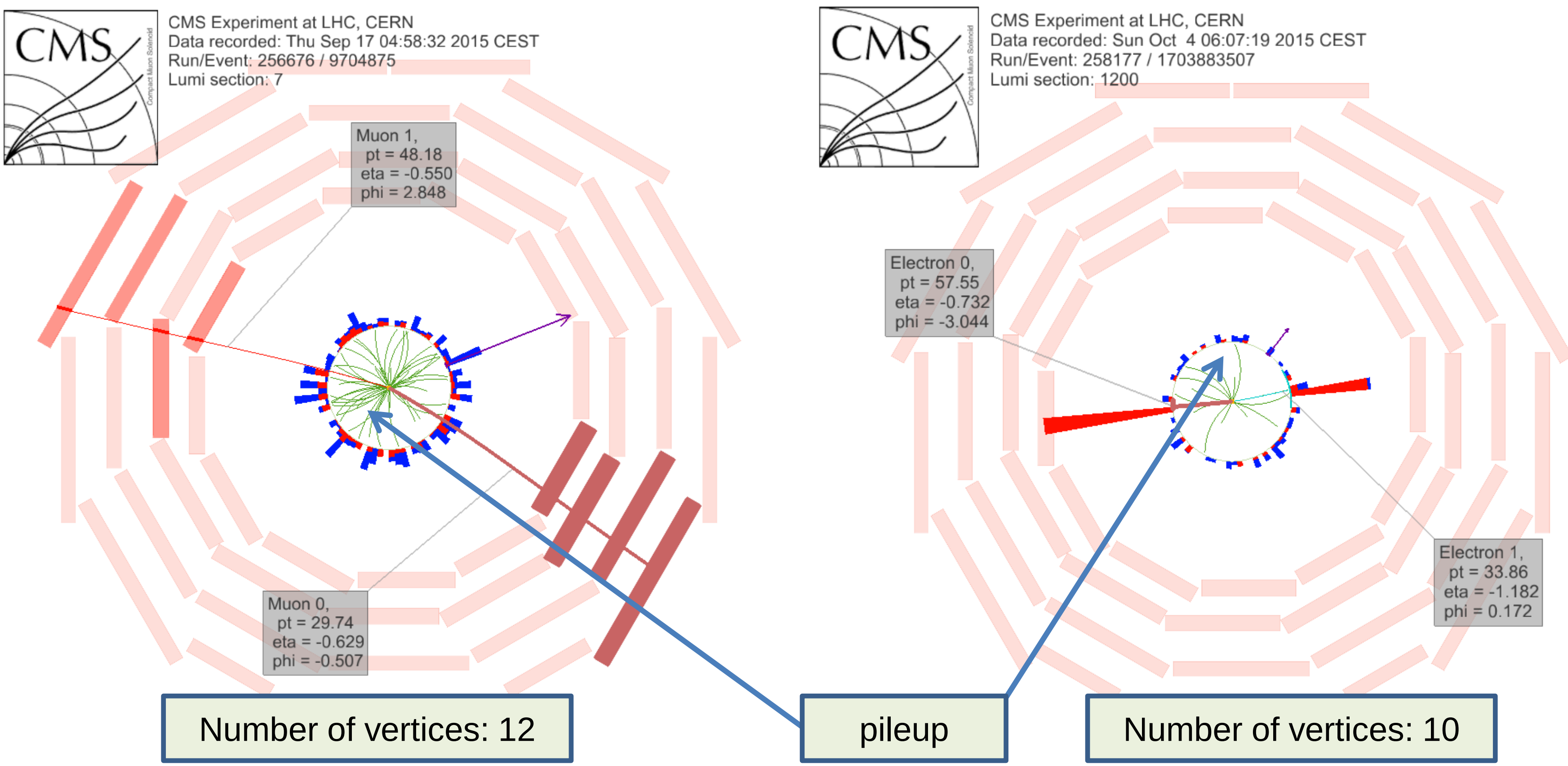


- Major detector at the LHC for general purpose of particle physics
- Implemented with the largest super conducting magnet in the world
- Contribute to the discovery of Higgs boson



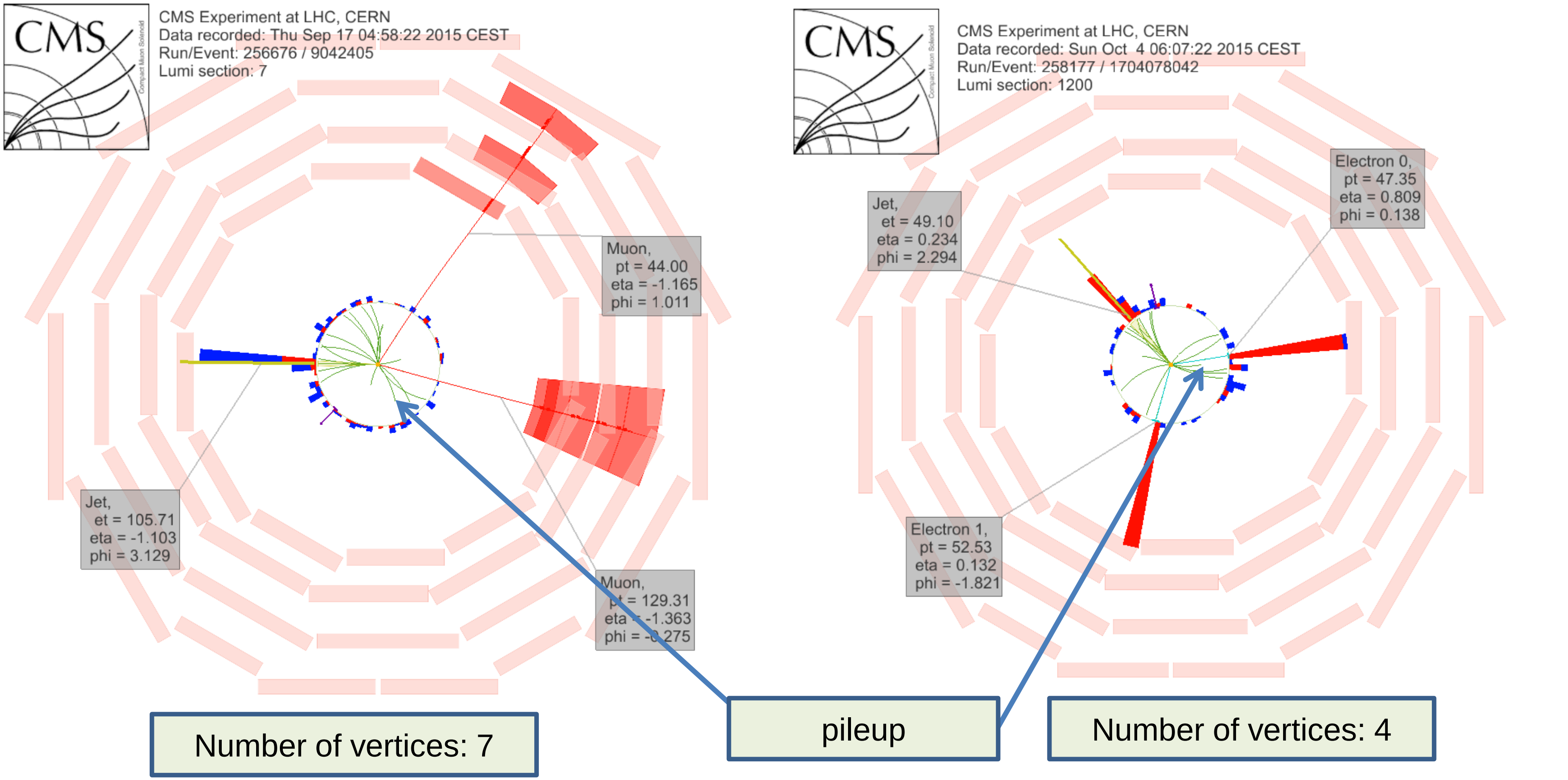


Event candidate of Drell-Yan process



- ➢ Clean signature due to Z boson decaying to two oppositely charged leptons
- ➢ High reconstruction efficiency in the CMS detector
- ➢ Z boson is reconstructed through its leptonic decays
  - ◆ Muons are reconstructed according to their tracks in the detector
  - ◆ Electrons are reconstructed according to their tracks in the tracker and energy deposition in the electromagnetic calorimeter

Event candidate of Z+jet process



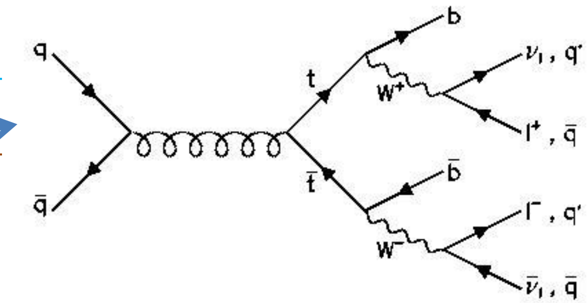
- An ideal laboratory for jet production study without jet trigger selection bias (trigger is implemented on muons/electrons)
- Jets are reconstructed according to the tracks and hadronic energy deposition in the calorimeters
  - Pileup subtraction is implemented for jets

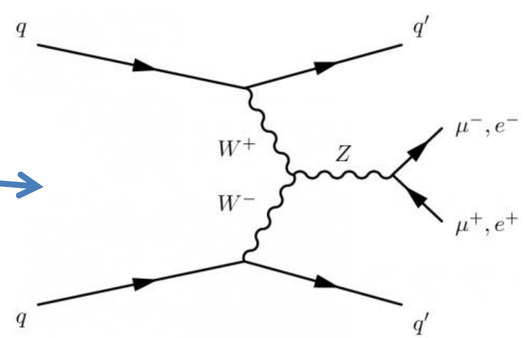
Data sample and background simulation

- 8TeV (center-of-mass energy of proton-proton collision):
  - Collected during 2012
  - Bunch spacing: 50 ns
  - Integrated luminosity : 19.6 /fb

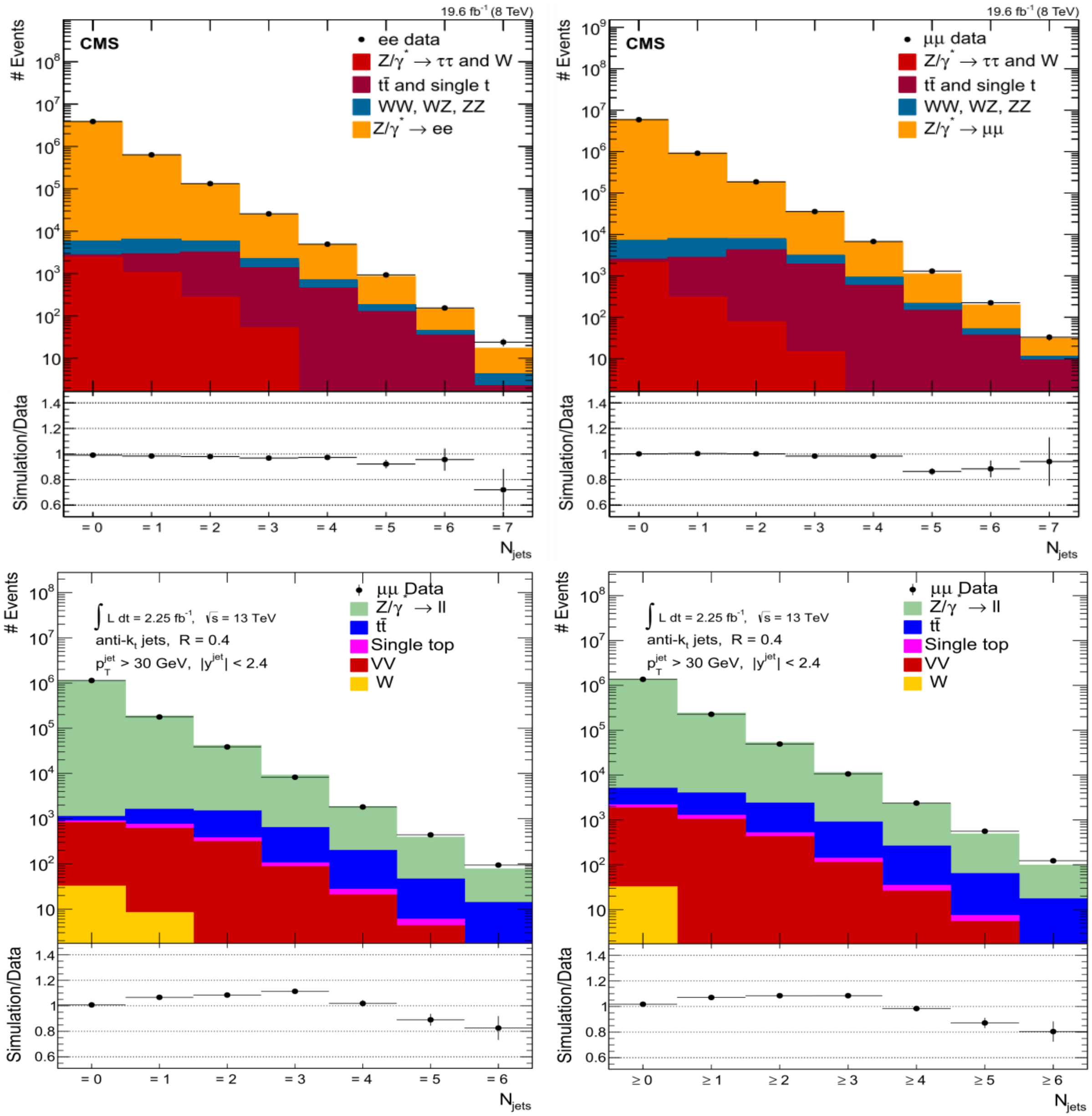
- 13TeV (center-of-mass energy of proton-proton collision):
  - Collected during 2015
  - Bunch spacing: 25 ns
  - Integrated luminosity : 2.25 /fb

- Main background:
  - TTbar
  - Double boson
    - WW
    - WZ
    - ZZ





Data & simulation at detector level



- **Phase space:**
  - **Lepton selection:**
    - p_T (l) > 20 GeV
    - |η (l)| < 2.4
    - 71 < M (ll) < 111 GeV
  - **Jet selection:**
    - p_T (j) > 30 GeV
    - |y (j)| < 2.4

◆**Significant statistics up to large jet multiplicity!!!**

◆**Low background contamination at low jet multiplicity (Njets < 3)**

Measured observables

- 8TeV:
  - Jet Multiplicity
  - Kinematics of the five leading jets:
    - Jet transverse momentum
    - Jet rapidity
  - scalar sum of the jets transverse momenta, (for Njets >= 1, 2, 3, 4, 5)
  - Rapidity correlation between Z boson and jet (for Njets >= 1, 2)
- 13 TeV:
  - Jet Multiplicity
  - Kinematics of the three leading jets:
    - Jet transverse momentum
    - Jet rapidity
  - scalar sum of the jets transverse momenta, (for Njets >= 1, 2, 3)

From detector reconstructed events to cross section

- Background subtraction:
  - Data signal = Data – background (from MC simulation)

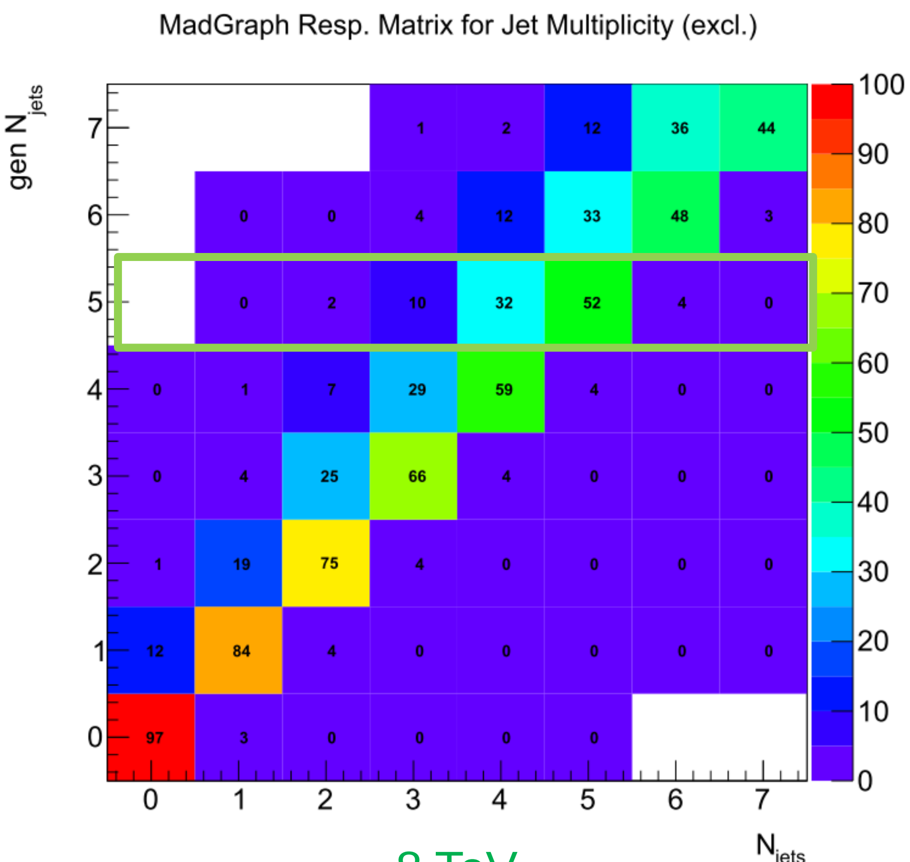
- Fake event subtraction:
  - the events passing the selection after reconstruction while failing the selection at produced level
  - Data signal – fakes (from Z+jets MC simulation)

- Detector effect corrections:
  - Data points differ from their true values due to detector effects:
    - Reco = R $\times$ Gen
  - The Gen (produced) level needs to be estimated from the reconstruction level (Reco) through unfolding program

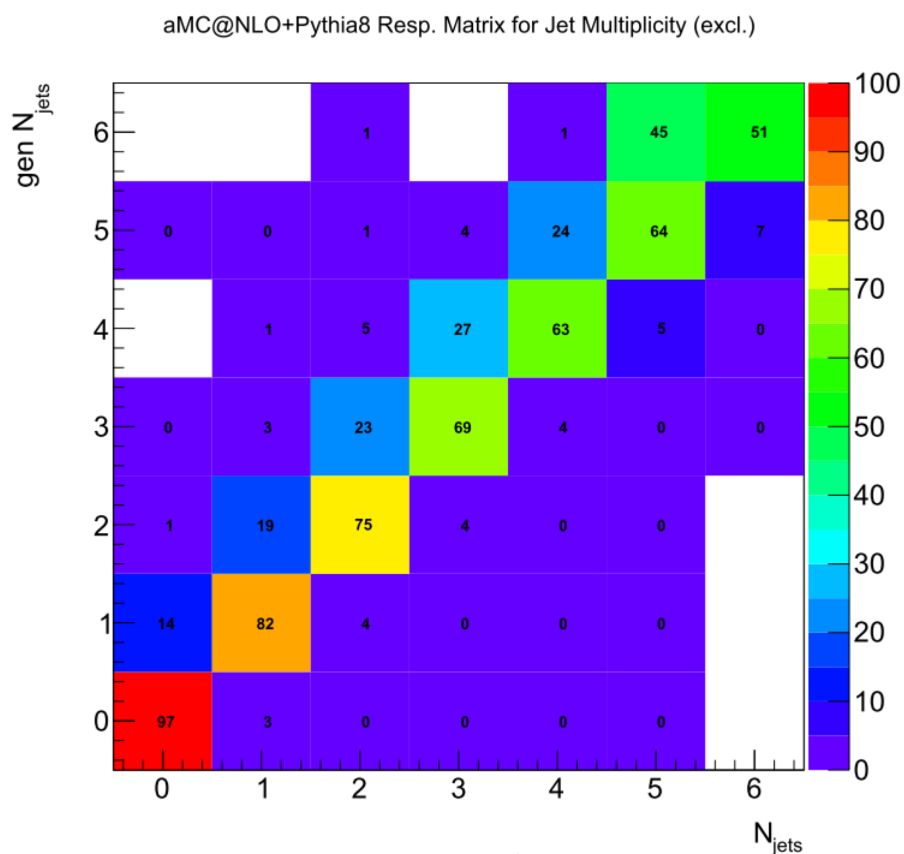
- Unfolding using D'Agostini approach
  - Acceptance correction included
  - Using Bayes statistical method
  - Iterative procedure: to be less model dependent (choice of number of iterations)
  - Statistical and systematic uncertainties are propagated through covariance matrices

Example of Response Matrix

- Percentage of Reco events in bin i from Gen events in bin j (i.e. row normalization)
- Eg. From all reconstructed events generated with 5 jets, 52% are indeed reconstructed with 5 jets, 32% with 4 jets, ...
- The response matrix is constructed using the selected events generated by MADGRAPH 5 (8 TeV) and AMC@NLO (13 TeV) MC generator

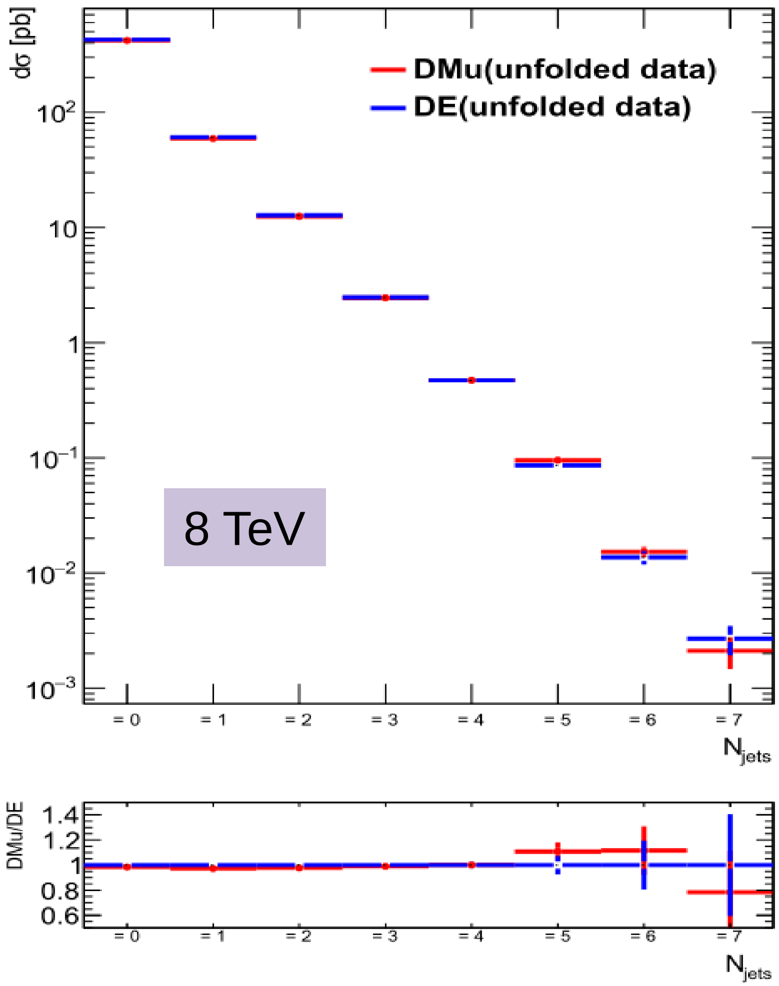


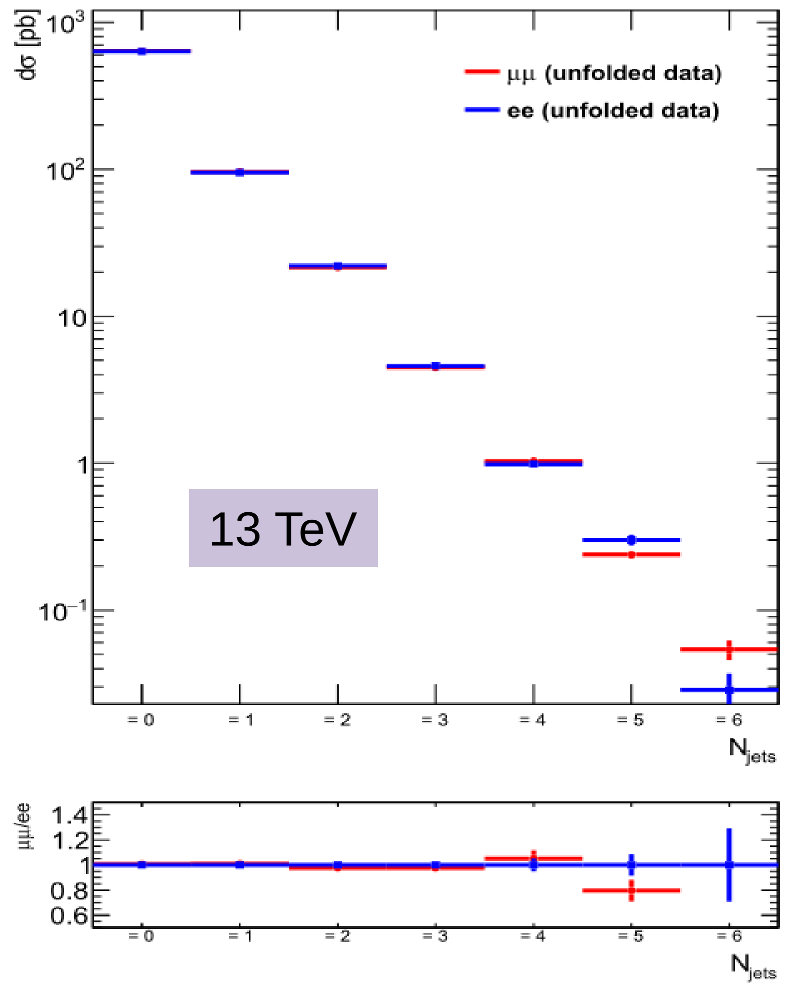
8 TeV



13 TeV

Channel Comparison





- The unfolded data of jet multiplicity with two leptonic decay channels are compatible
- The error bar in the plots stands for statistical uncertainty

Main Systematic Uncertainties

- Jet energy correction and resolution uncertainty
  - 5% ~ 28% (8 TeV)
  - 1% ~ 17% (13 TeV)

- Cross section uncertainty of backgrounds
  - < 5% (8 TeV & 13 TeV)

- Pileup uncertainty
  - 0.2% ~ 5.6% (8 TeV)
  - 0.2% ~ 2.2% (13 TeV)

- Unfolding uncertainty
  - 1% ~ 7% (8 TeV)
  - 0.1% ~ 20% (13 TeV)

- Integrated luminosity uncertainty
  - 2.6% (8 TeV)
  - 2.7% (13 TeV)

- Lepton trigger/identification/reconstruction uncertainty
  - 1.5% ~ 2.8%
  - 1.8% ~ 2.4%

Channel Combination

- Combine differential cross sections of two channels with weighted average approach
- The coefficient (weight) for each channel is computed according to the total uncertainty of each channel

$$\sigma_\alpha^{comb.} = \lambda_{\alpha\alpha}\sigma_\alpha^{ee} + \lambda_{\alpha\alpha+n}\sigma_\alpha^{\mu\mu} = \frac{(\Delta\sigma_\alpha^{ee})^{-2}}{(\Delta\sigma_\alpha^{ee})^{-2} + (\Delta\sigma_\alpha^{\mu\mu})^{-2}}\sigma_\alpha^{ee} + \frac{(\Delta\sigma_\alpha^{\mu\mu})^{-2}}{(\Delta\sigma_\alpha^{ee})^{-2} + (\Delta\sigma_\alpha^{\mu\mu})^{-2}}\sigma_\alpha^{\mu\mu}$$

- Combine the uncertainties of two channels with the full measurement covariance matrix

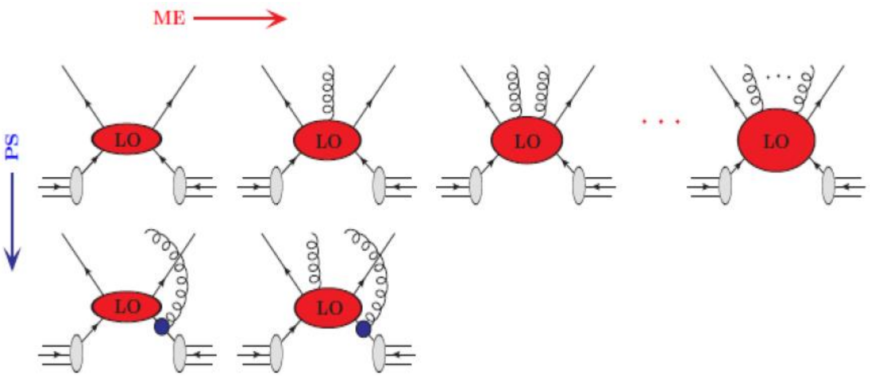
$$\Delta\sigma_{\alpha\beta}^{comb.} = \sum_{i=1}^{2n}\sum_{j=1}^{2n}\lambda_{\alpha i}M_{ij}\lambda_{\beta j}$$

$$\begin{aligned} M &= \sum_{s\in unc.sources} M^s \\ &= \sum_{s\in\{JES,JER,PU,Bgnd,Lumi\}} \left(\begin{array}{c|c} (\Delta\sigma^{ee})^s & \pm\sqrt{(\Delta\sigma_\alpha^{ee})^s \times (\Delta\sigma_\beta^{\mu\mu})^s} \\ \hline \pm\sqrt{(\Delta\sigma_\alpha^{ee})^s \times (\Delta\sigma_\beta^{\mu\mu})^s} & (\Delta\sigma^{\mu\mu})^s \end{array}\right) \quad \leftarrow \text{correlated unc.} \\ &+ \sum_{s\in\{Stat,Eff,LES,LER,Unf\}} \left(\begin{array}{c|c} (\Delta\sigma^{ee})^s & 0 \\ \hline 0 & (\Delta\sigma^{\mu\mu})^s \end{array}\right) \quad \leftarrow \text{uncorrelated unc.} \end{aligned}$$

Theoretical predictions of Z+jets cross section

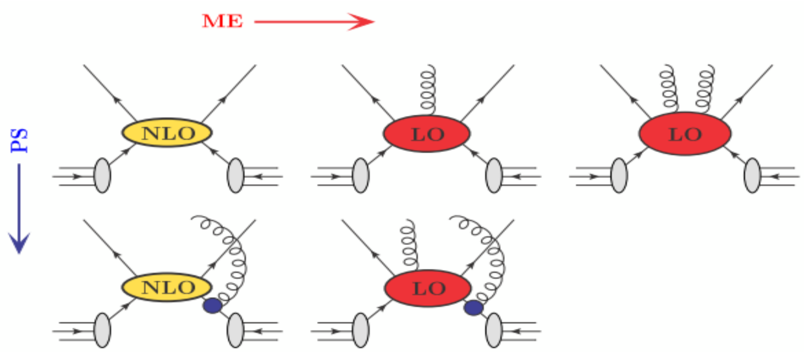
➢ **Multileg LO with 0~4 final partons in matrix element:**

- MG5 + PY6: KT-MLM PS matching, CTEQ6L1 PDF
- MG5_aMC + PY8: KT-MLM PS matching, CT10 PDF
- SHERPA 1.4: CKKW PS matching, CT10 PDF



➢ **Multileg with 0~2 partons at NLO, 3/4 partons at LO:**

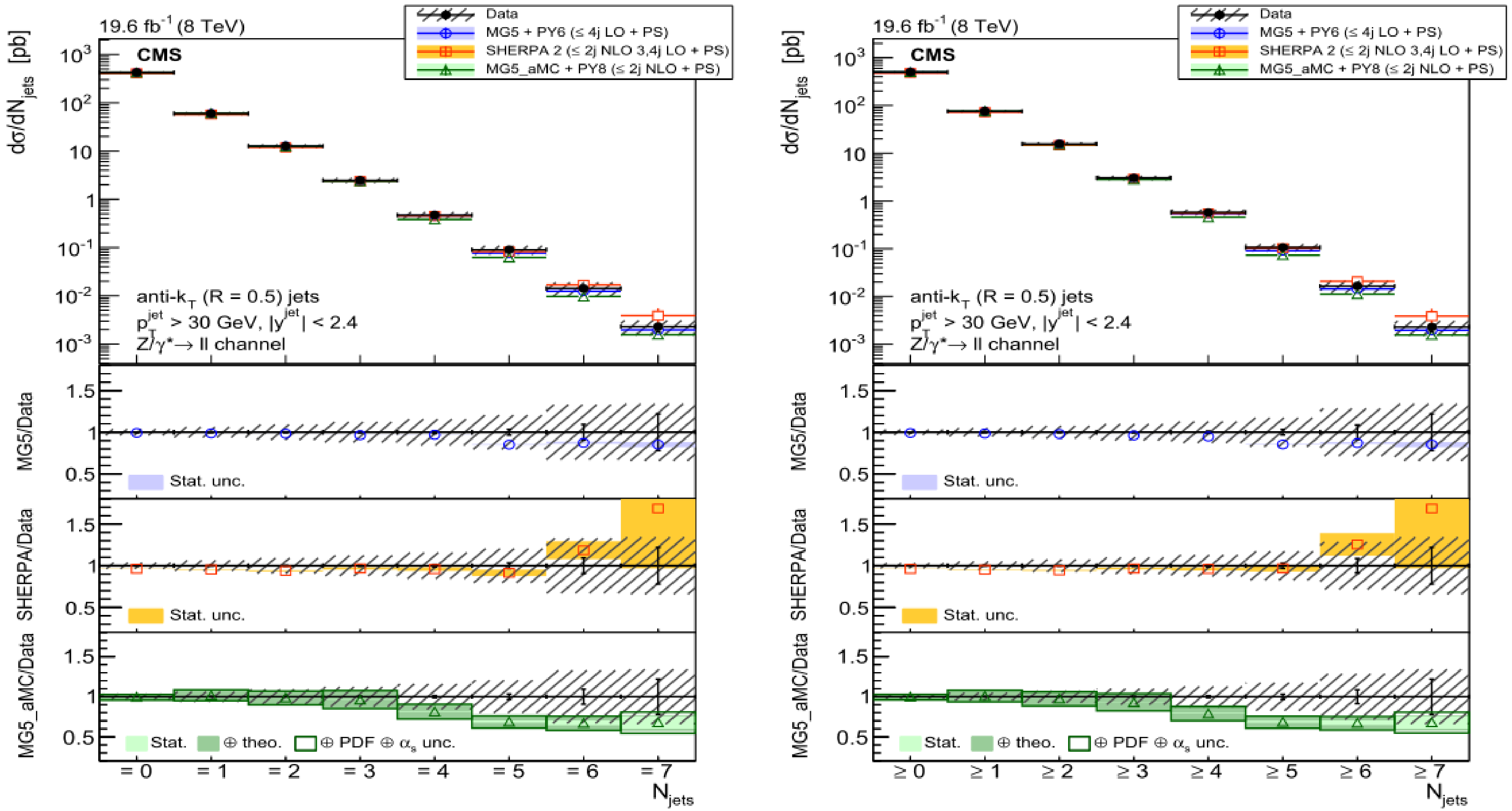
- SHERPA 2: MEPS@NLO merging, CT10 PDF
- MG5_aMC + PY8: FxFx jet merging scheme, NNPDF3.0, CUETP8M1 tune
  - Scale uncertainty : 3% ~ 15%
  - PDF uncertainty : 2% ~ 10%



➢ N_{jetti} NNLO (arXiv:1602.08140, arXiv:1512.01291)

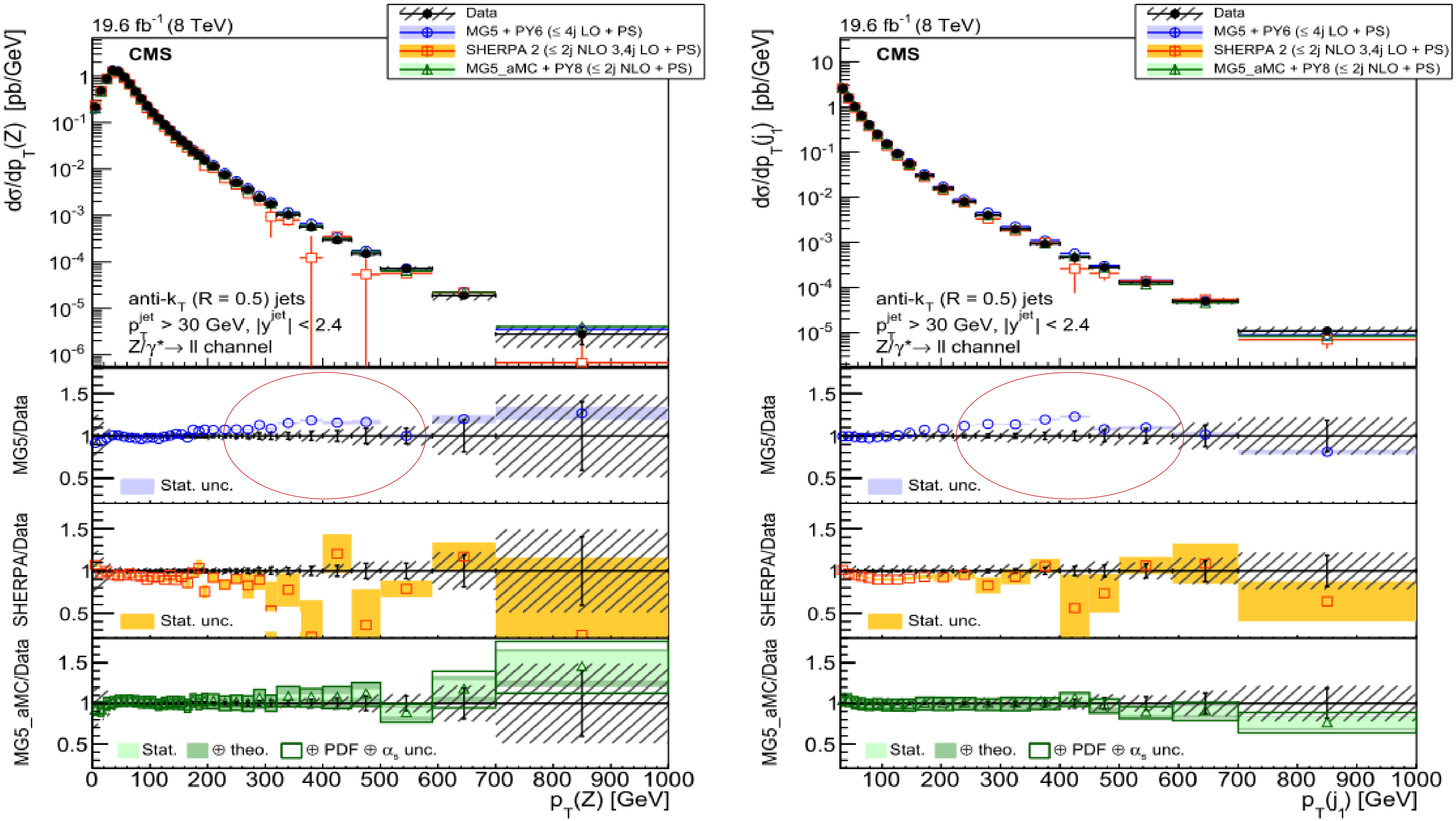
- **Z+1jet** with fixed order (Scale uncertainty < 10%)
- CT14 PDF set

Combined Results (8 TeV)



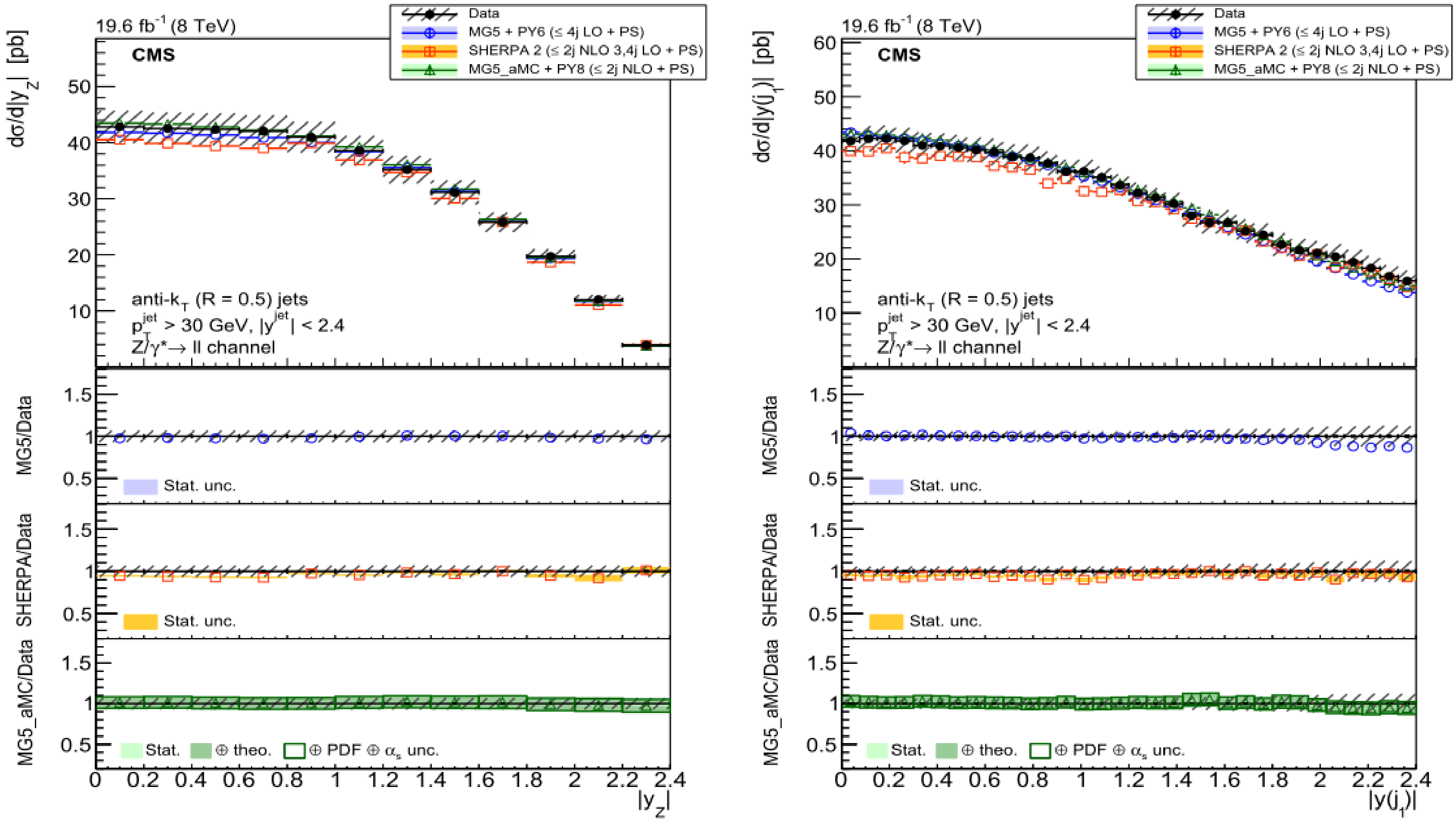
- Both MG5+PY6 (LO) and SHERPA2 (NLO) accuracy have good agreement with measurement
- MG5_aMC+PY8 differs at large jet multiplicities due to limited number of partons in matrix element calculation (only relying on parton shower)
- High precision measurement to high jet multiplicity benefits from large statistics of data

Combined Results (8 TeV)



- ➢ Discrepancy with LO computation has disappeared with NLO accuracy
- ➢ Sherpa2 NLO predictions have larger statistical fluctuations

Combined Results (8 TeV)



- Good agreements are seen for all the predictions

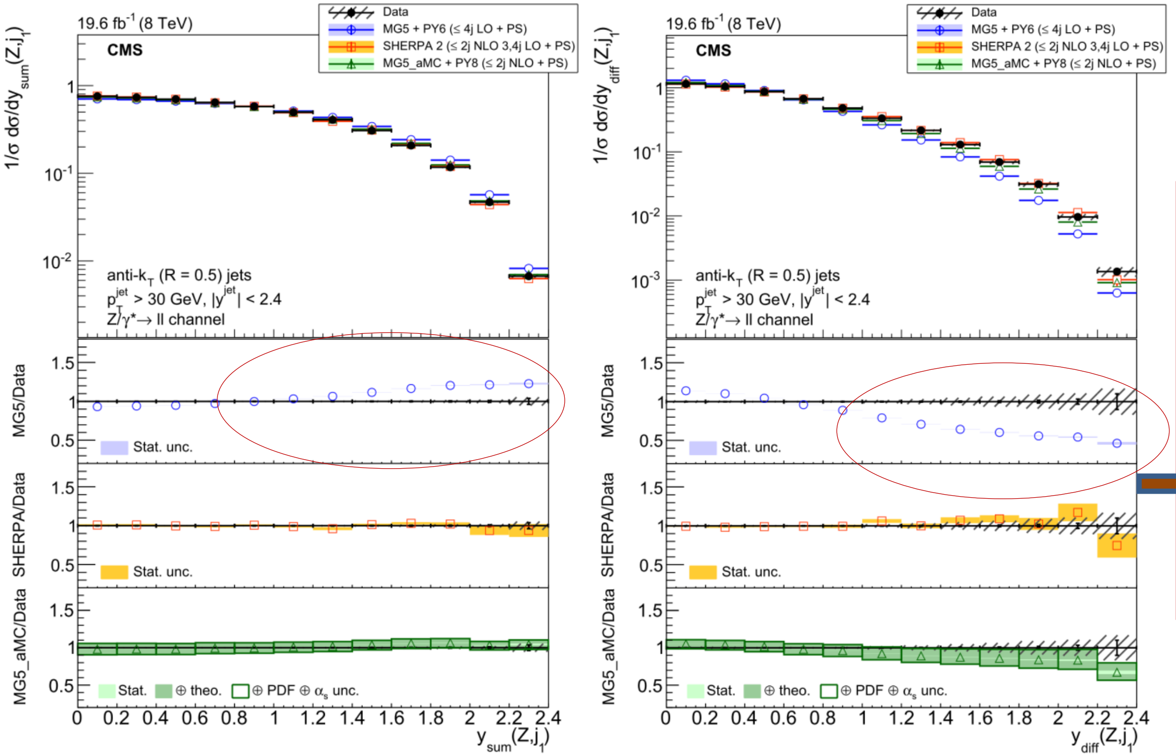
Rapidity correlation measurement of Z + jets (8 TeV)

- The measurement of jet angular correlation of Z+Jets can help understand QCD process much more accurately

$$y_{sum} = \frac{|y_Z + y_{jet}|}{2}$$

$$y_{diff} = \frac{|y_Z - y_{jet}|}{2}$$

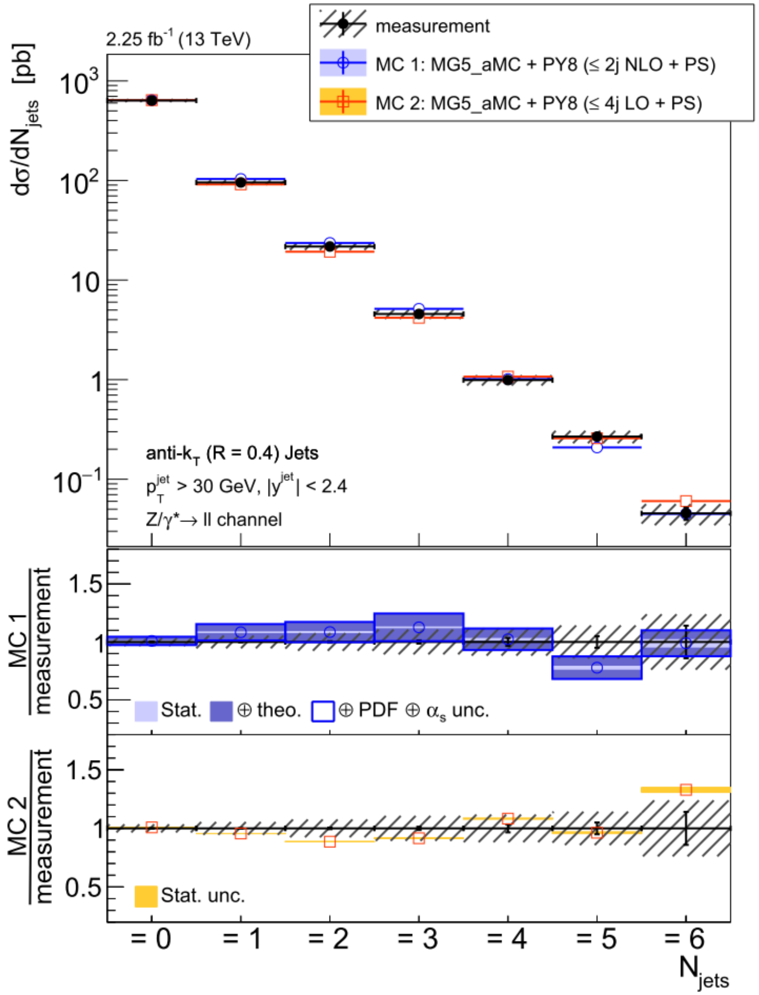
- Depends mainly on parton distribution functions
- Reflects the partonic differential cross section

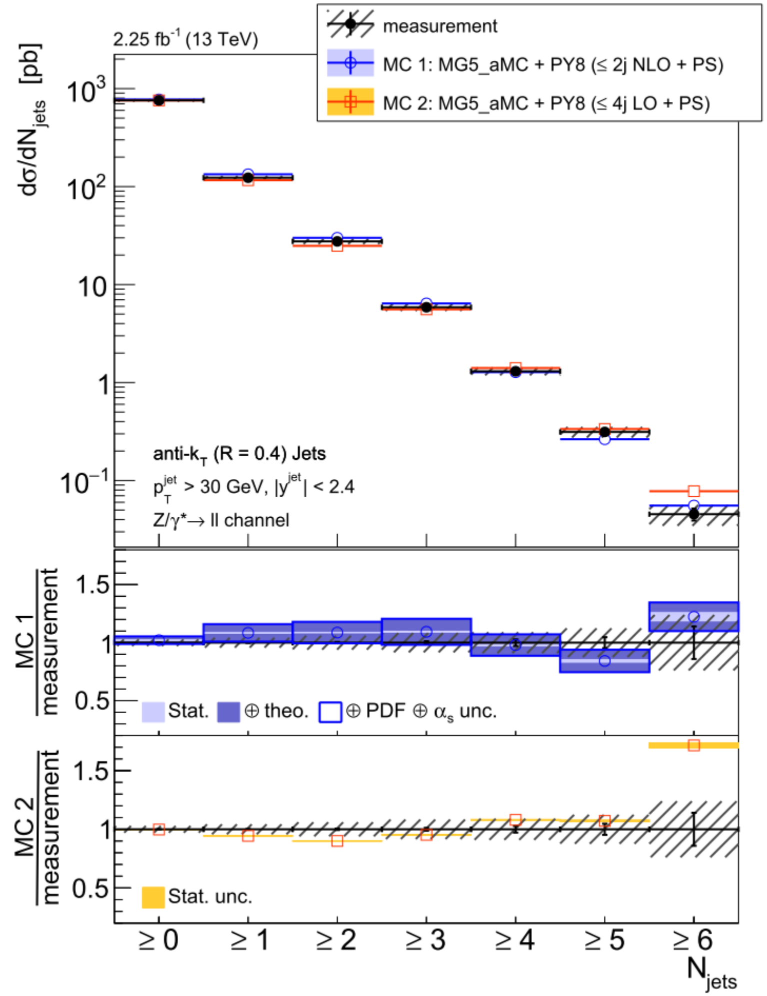


- LO calculation fails to describe the shape, confirms the observation at 7 TeV:
  - *10.1103/PhysRevD.88.112009*
- Discrepancy with LO computation has disappeared with NLO accuracy!

- The observed discrepancy of LO prediction results from the matching procedure between matrix element and parton shower

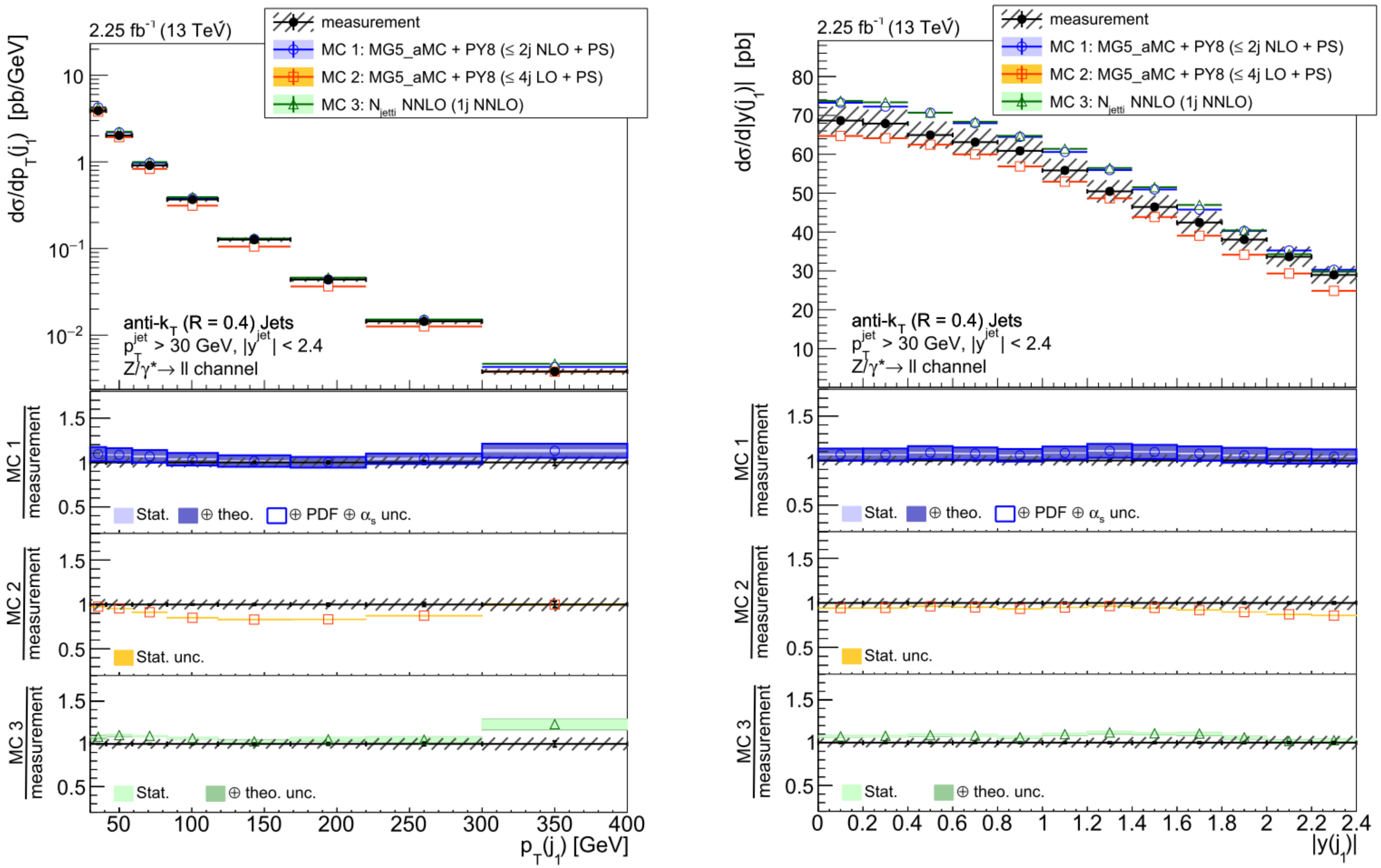
Combined Results (13 TeV)





- Good agreements between measurement and both predictions except the large jet multiplicity region (Njets >= 6)

Combined Results (13 TeV)



- Good agreements are seen with AMC@NLO and fixed NNLO predictions

Publication Status

- Z + jet (8 TeV):
  - The CMS collaboration, “Measurements of differential production cross sections for a Z boson in association with jets in pp collisions at $\sqrt{s}$ = 8 TeV”, J. High Energ. Phys. 2017(04):22

- Z + jet (13 TeV):
  - “Measurement of the differential cross section of Z boson production in association with jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV”, CMS-PAS-SMP-15-010, (CERN), 2015
  - Will be published in future

- Jet energy correction and calibration at high level trigger in CMS for RunII:
  - F. Zhang, “Performance of the CMS jets and missing transverse energy trigger for the upgraded LHC RunII”, Proceedings of Science, EPS-HEP2015/290 (2015)

- The technique of jet energy correction and calibration at high level trigger has been constructed and is still developing
- Optimal jet energy correction configurations have been used for data taking during 2015 and 2016!

Conclusion

- The measurement of Z boson plus jets process is quite important:
  - It deepens our understanding on QCD dynamics
  - It improves the modeling of background for scalar boson measurement and new physics searches

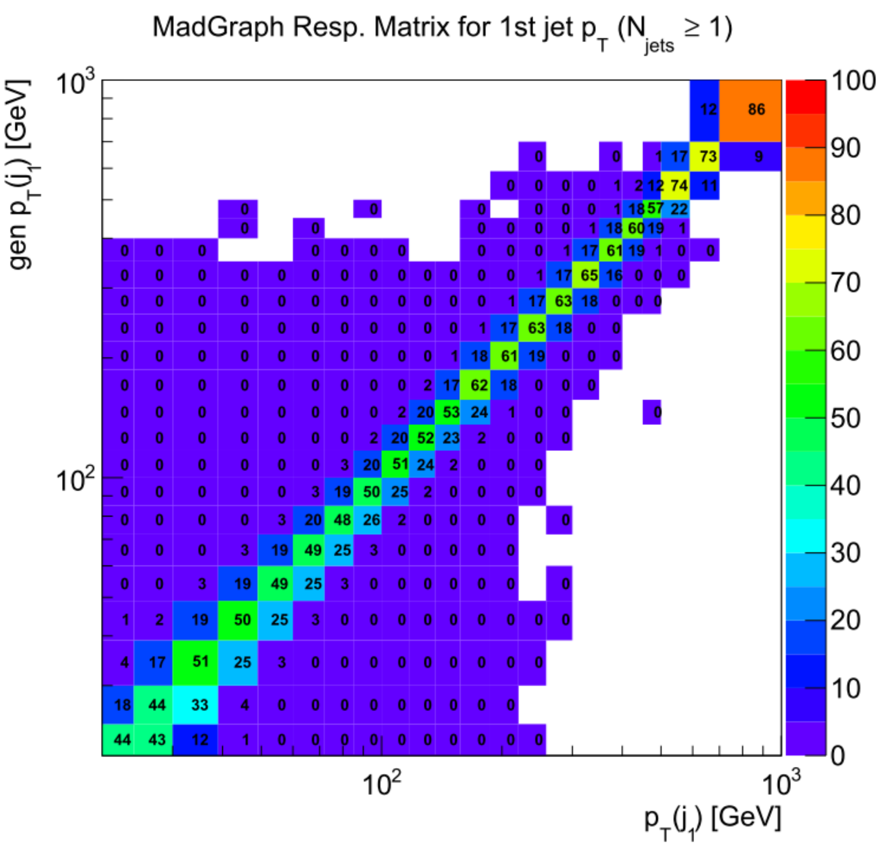
- The main contributions in this analysis:
  - Provide the systematics of jet production to give the precision for different theoretical models
  - Quantify the improvement with higher order of pQCD calculations on matrix elements
  - From the rapidity correlation study, validate the new matching scheme FxFx, MEPS@NLO for NLO matrix elements, showing improvements compared to the KT-MLM matching scheme for LO matrix element

- This analysis will benefit from more data statistics collected in 2016!

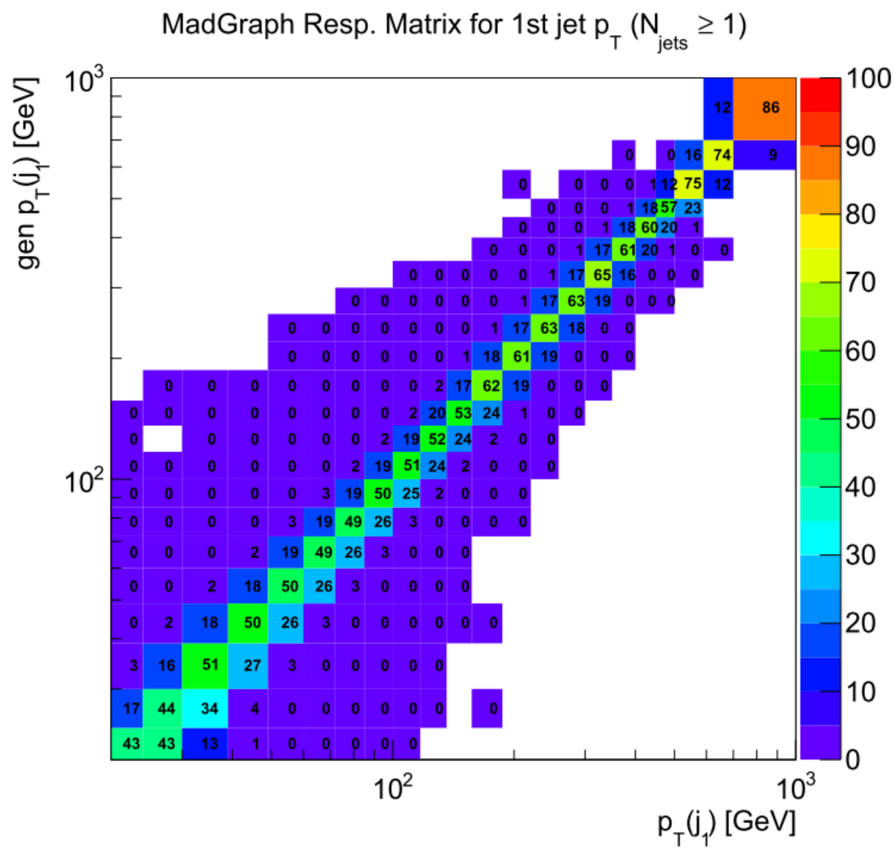
Backup

Response Matrix (8 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- There is no improvement for the diagonal purity with angular matching scheme for jet sorting



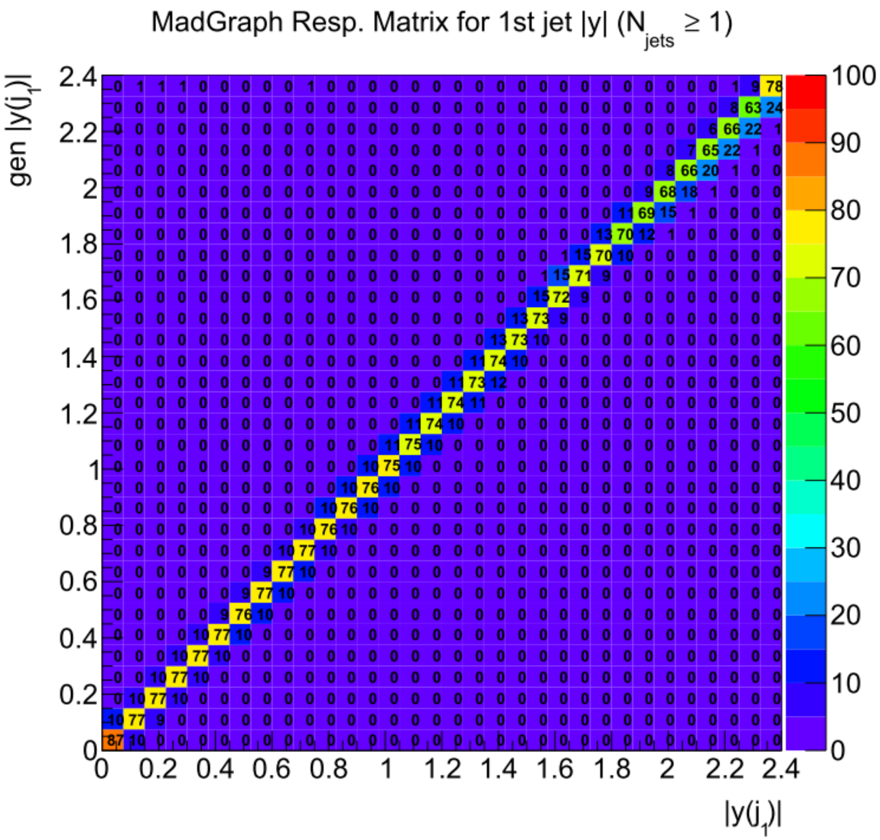
p_T Matching scheme (default)



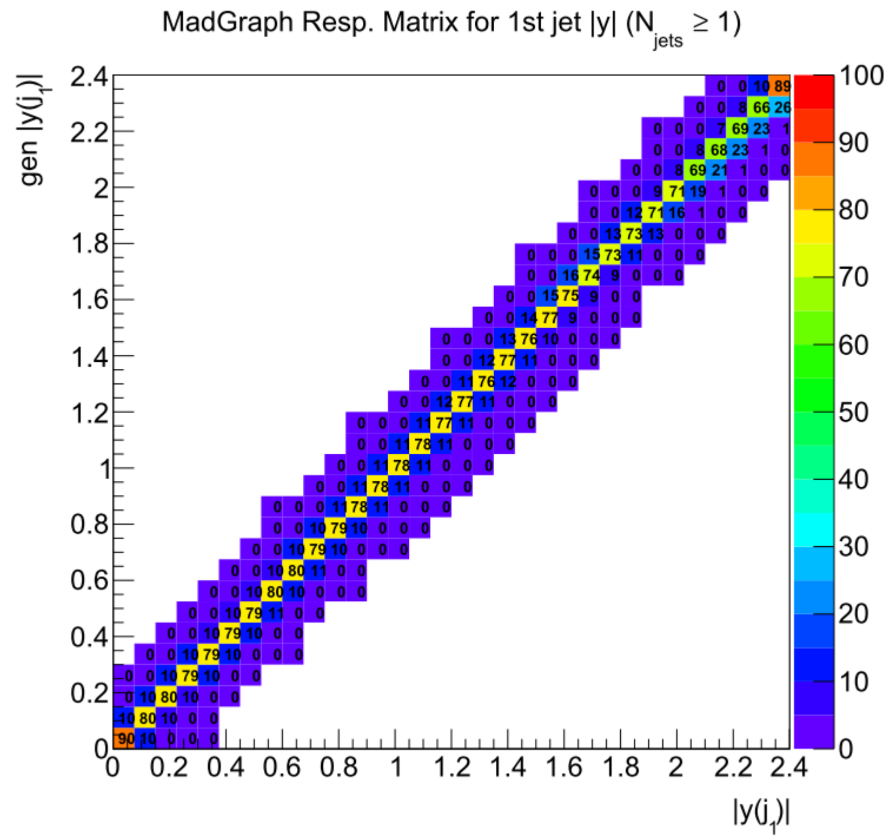
Angular Matching scheme

Response Matrix (8 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- Significant Improvement for the diagonal purity with angular matching scheme for jet sorting!



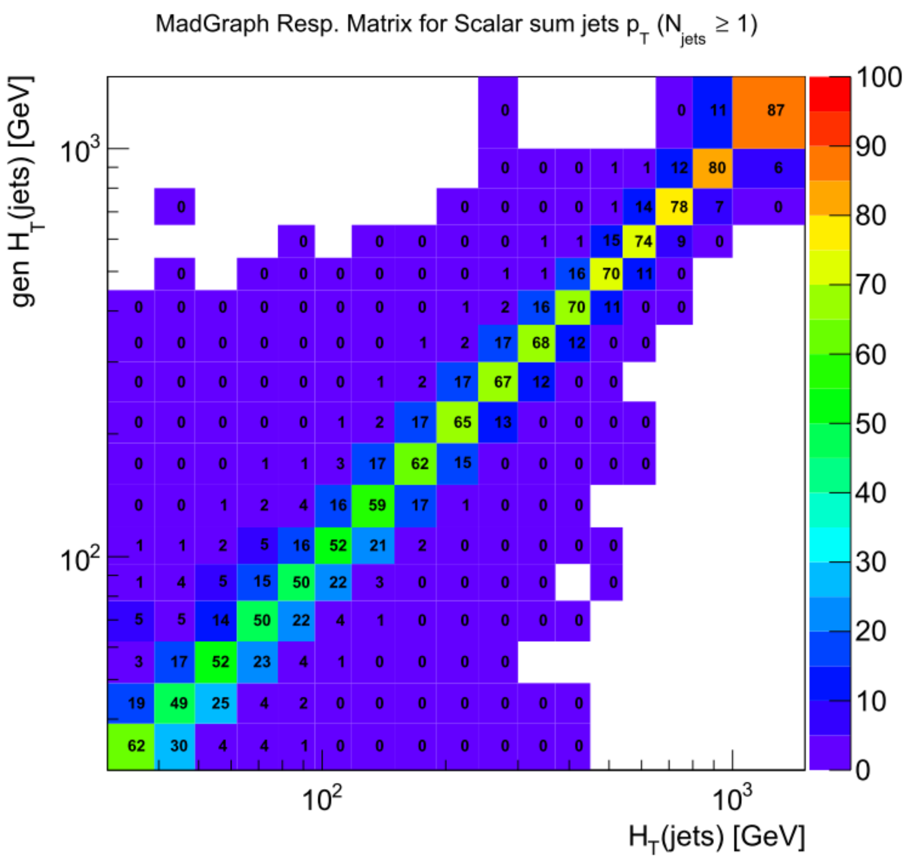
p_T Matching scheme (default)



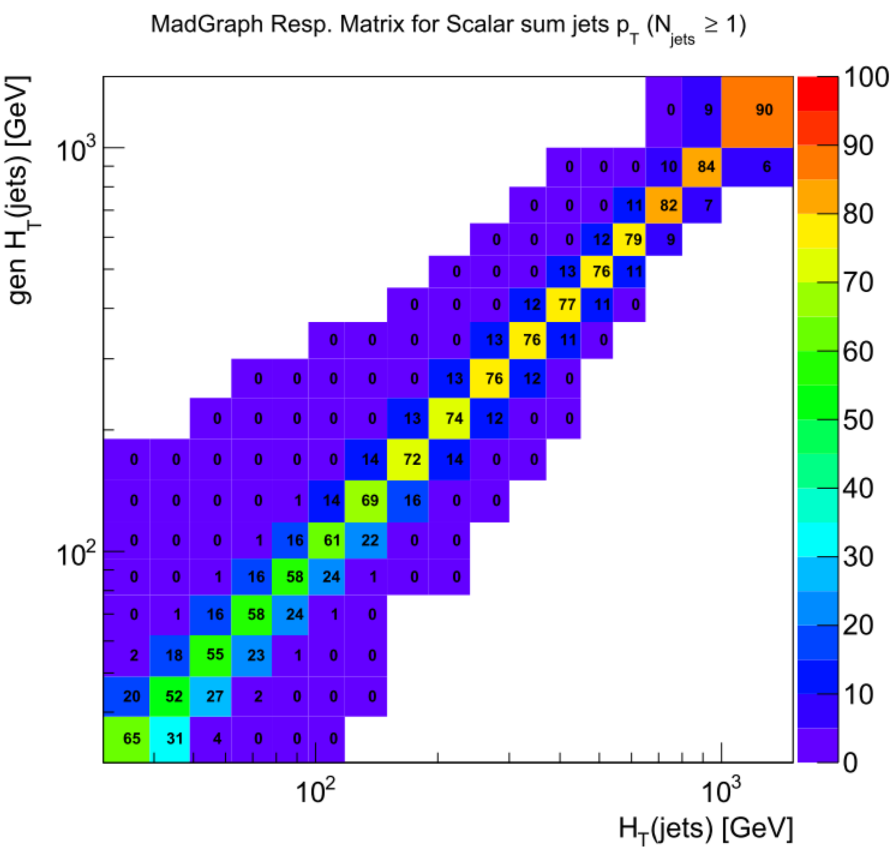
Angular Matching scheme

Response Matrix (8 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- A bit improvement for the diagonal purity with angular matching scheme for jet sorting



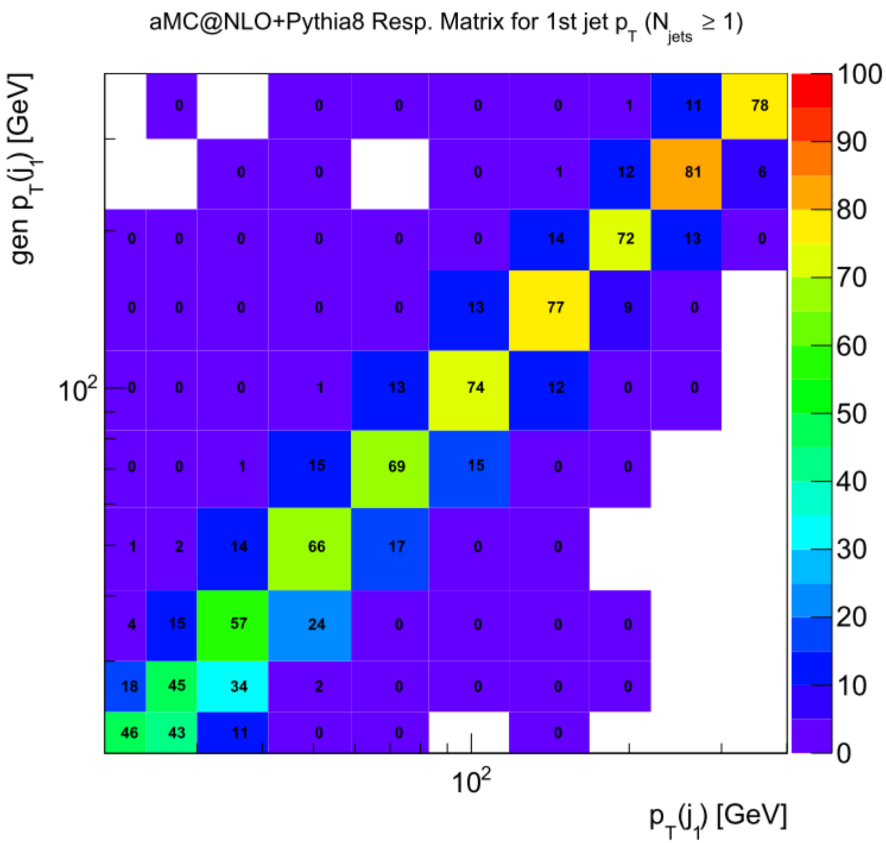
p_T Matching scheme (default)



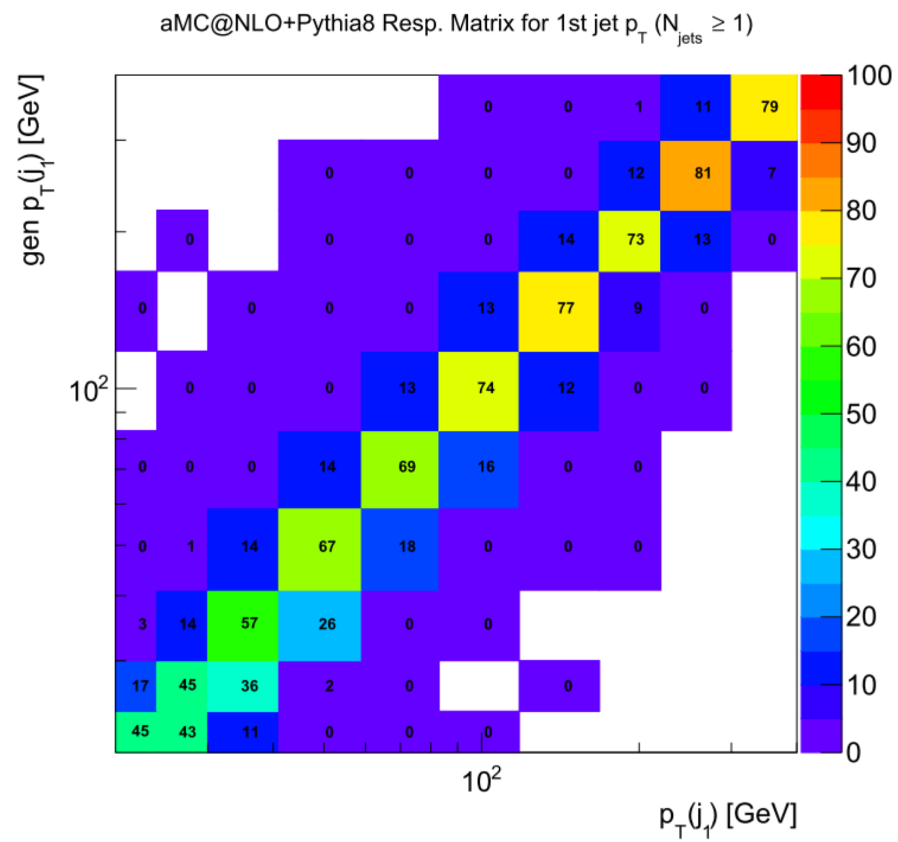
Angular Matching scheme

Response Matrix (13 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- There is no improvement for the diagonal purity with angular matching scheme for jet sorting



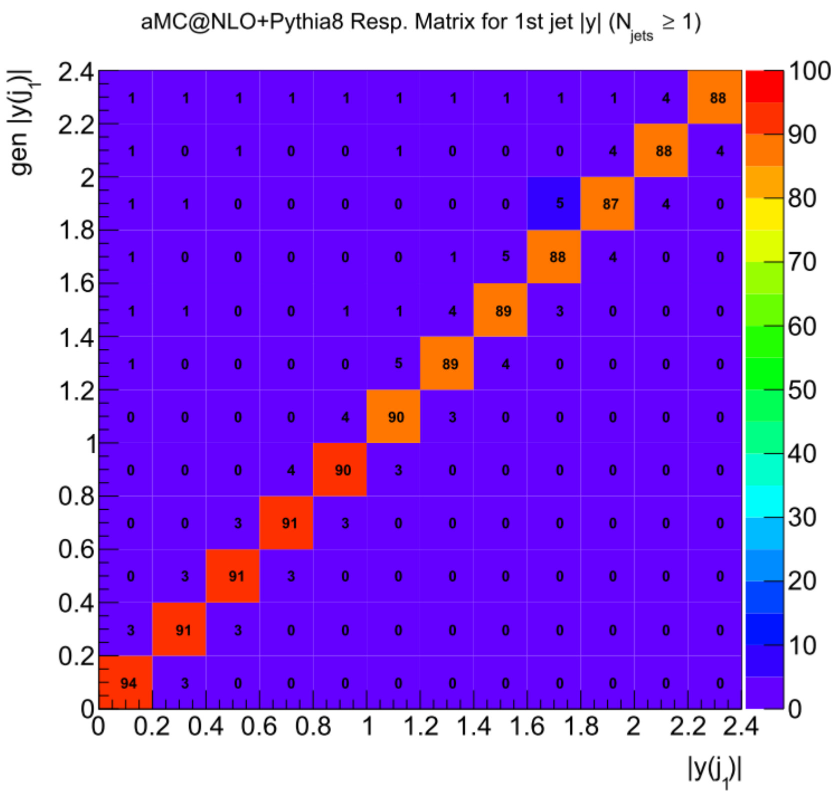
p_T Matching scheme (default)



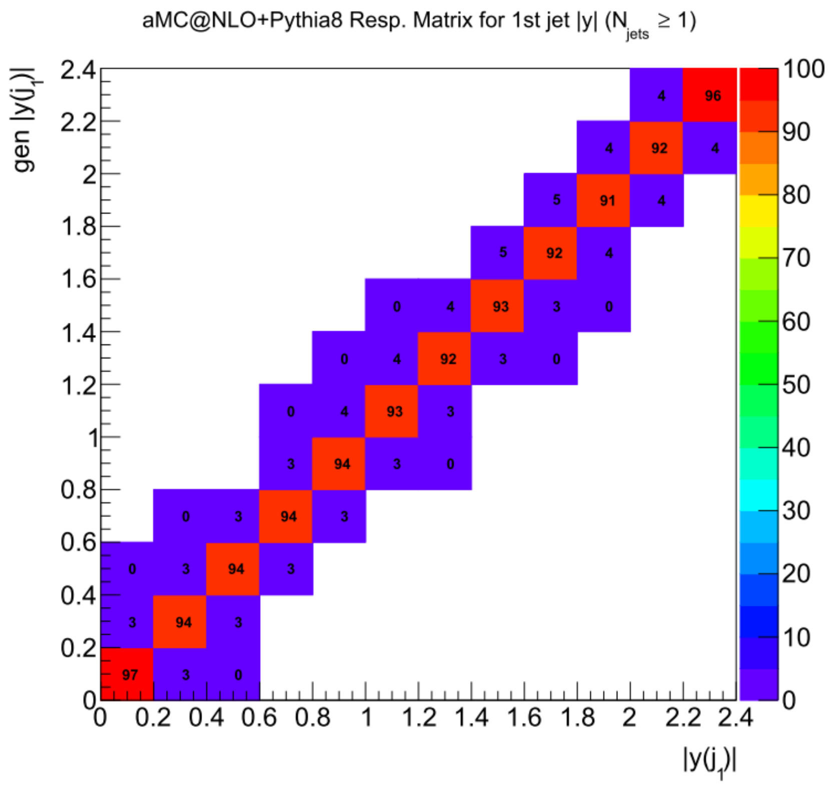
Angular Matching scheme

Response Matrix (13 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- Significant Improvement for the diagonal purity with angular matching scheme for jet sorting!



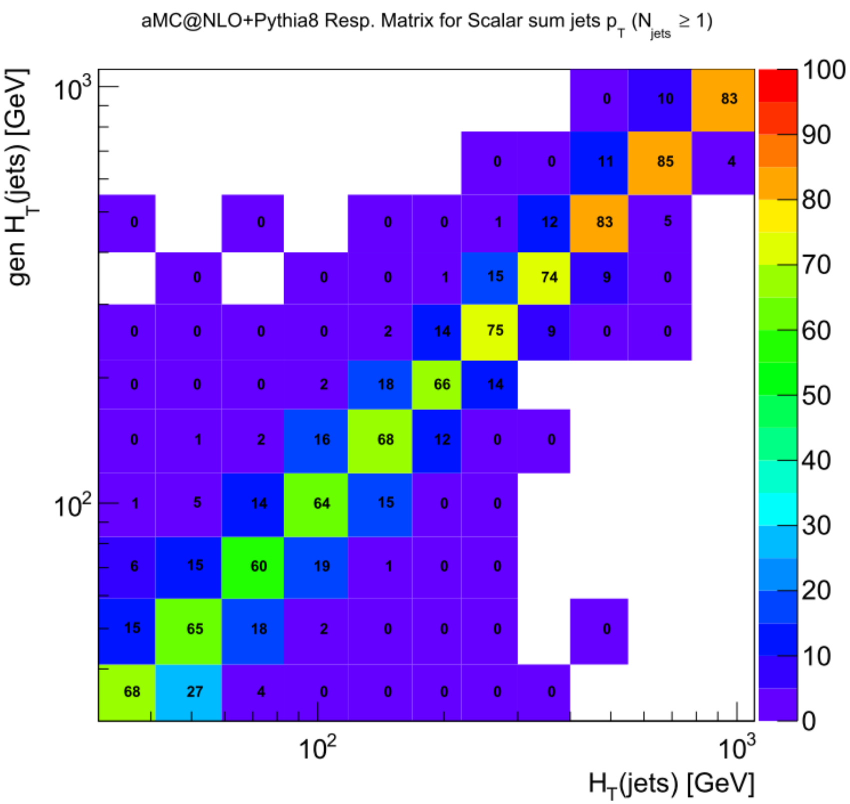
p_T Matching scheme (default)



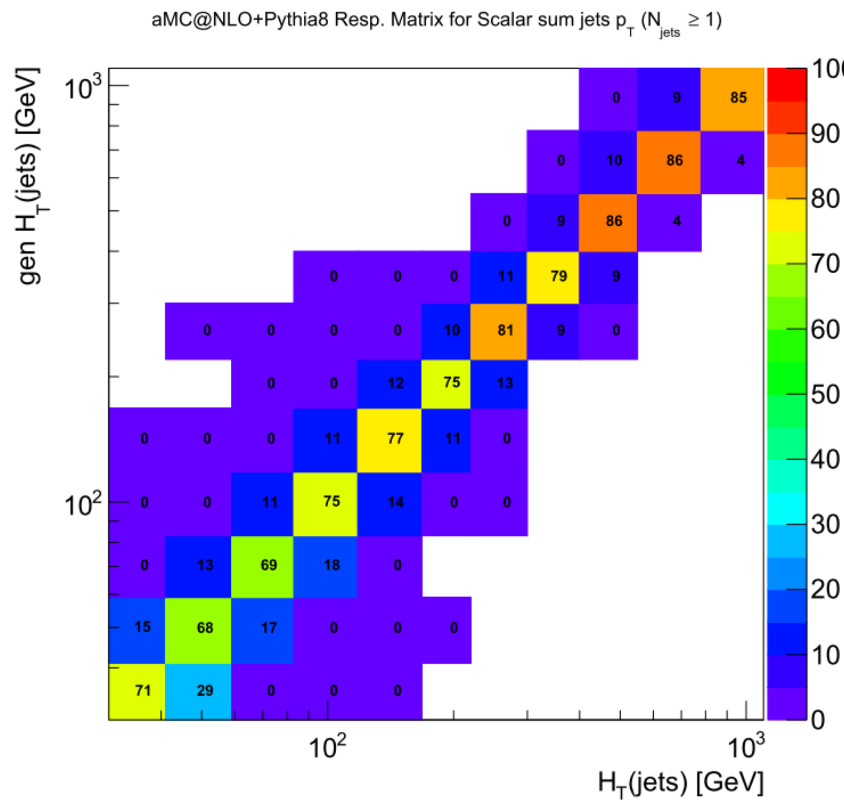
Angular Matching scheme

Response Matrix (13 TeV)

- Small fraction of events (close to zero) in the off-diagonal (far away from diagonal) elements
- A bit improvement for the diagonal purity with angular matching scheme for jet sorting

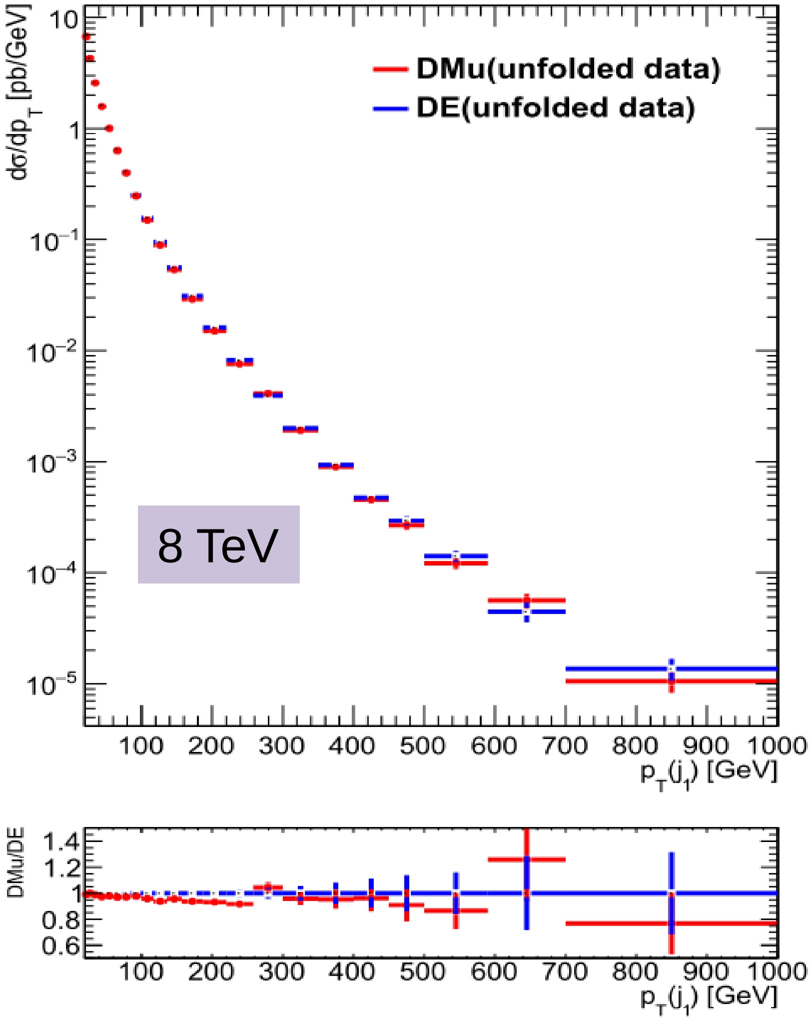


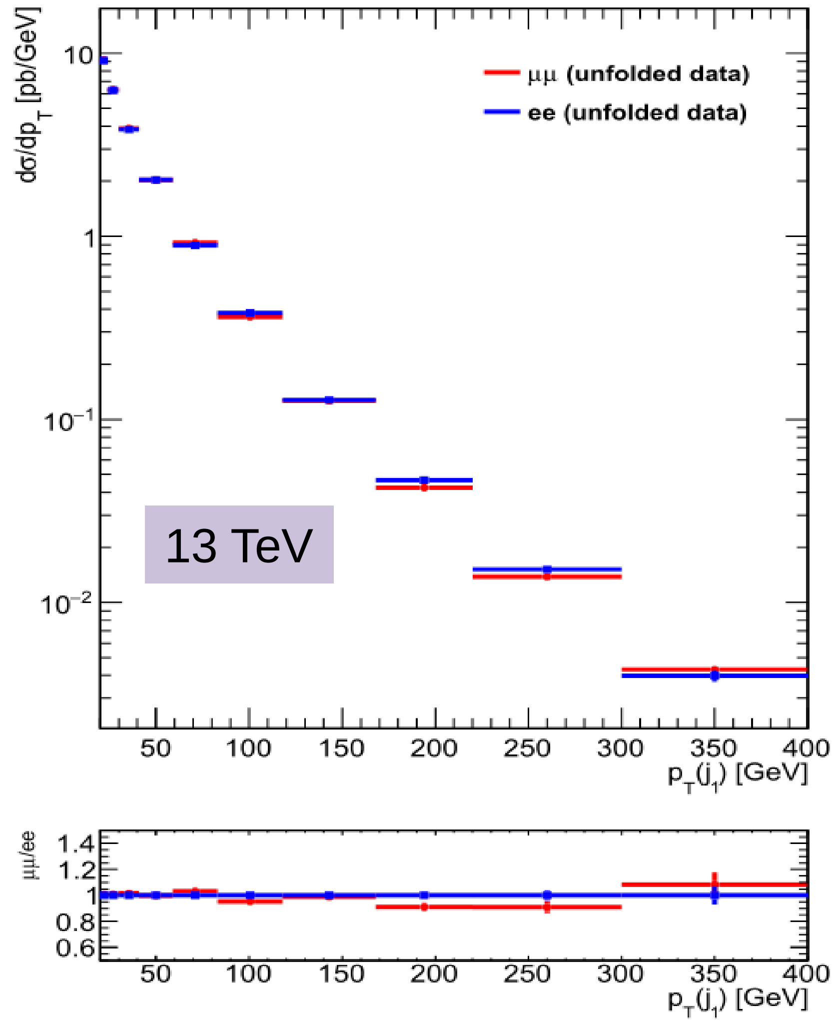
p_T Matching scheme (default)



Angular Matching scheme

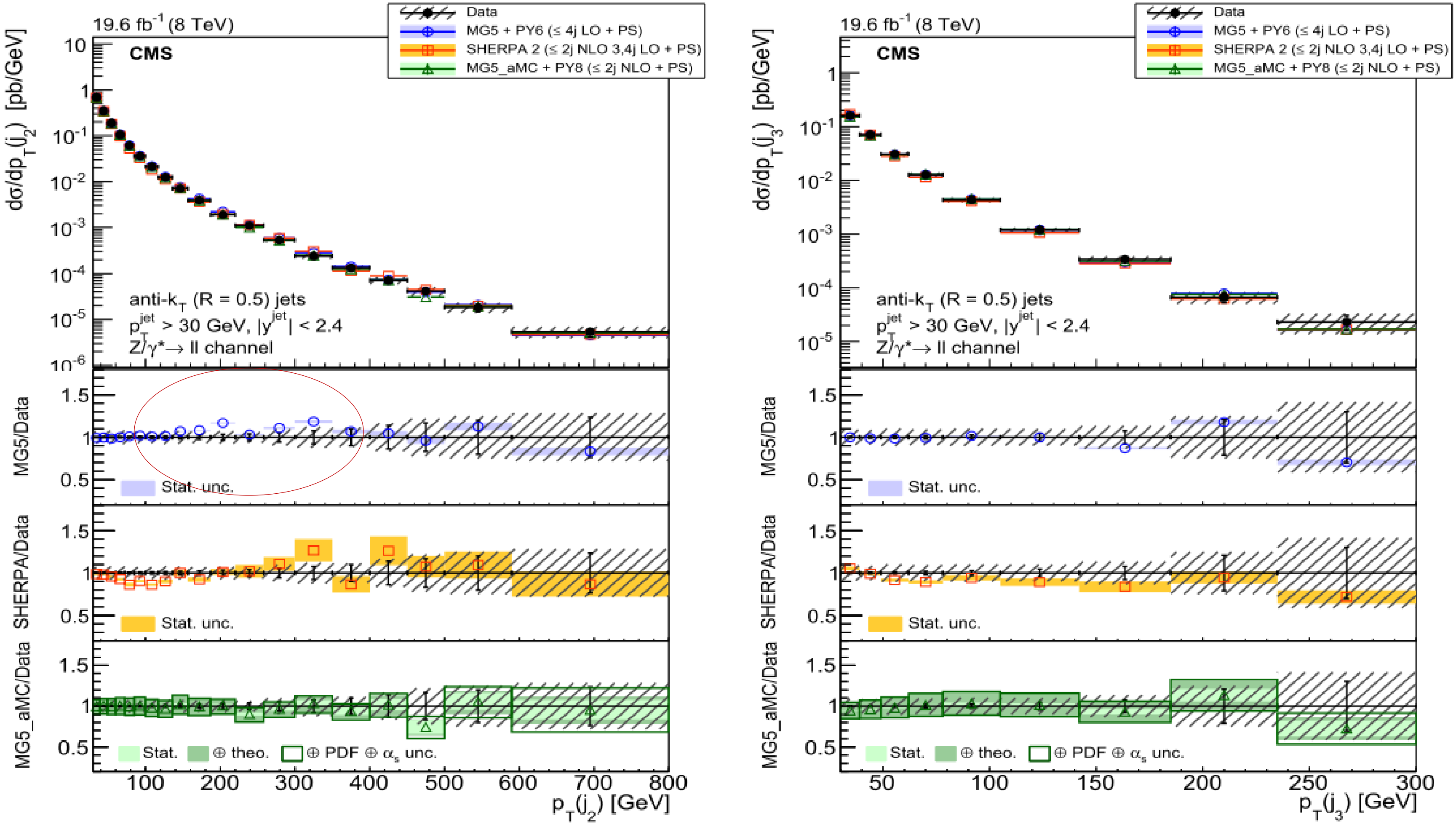
Channel Comparison





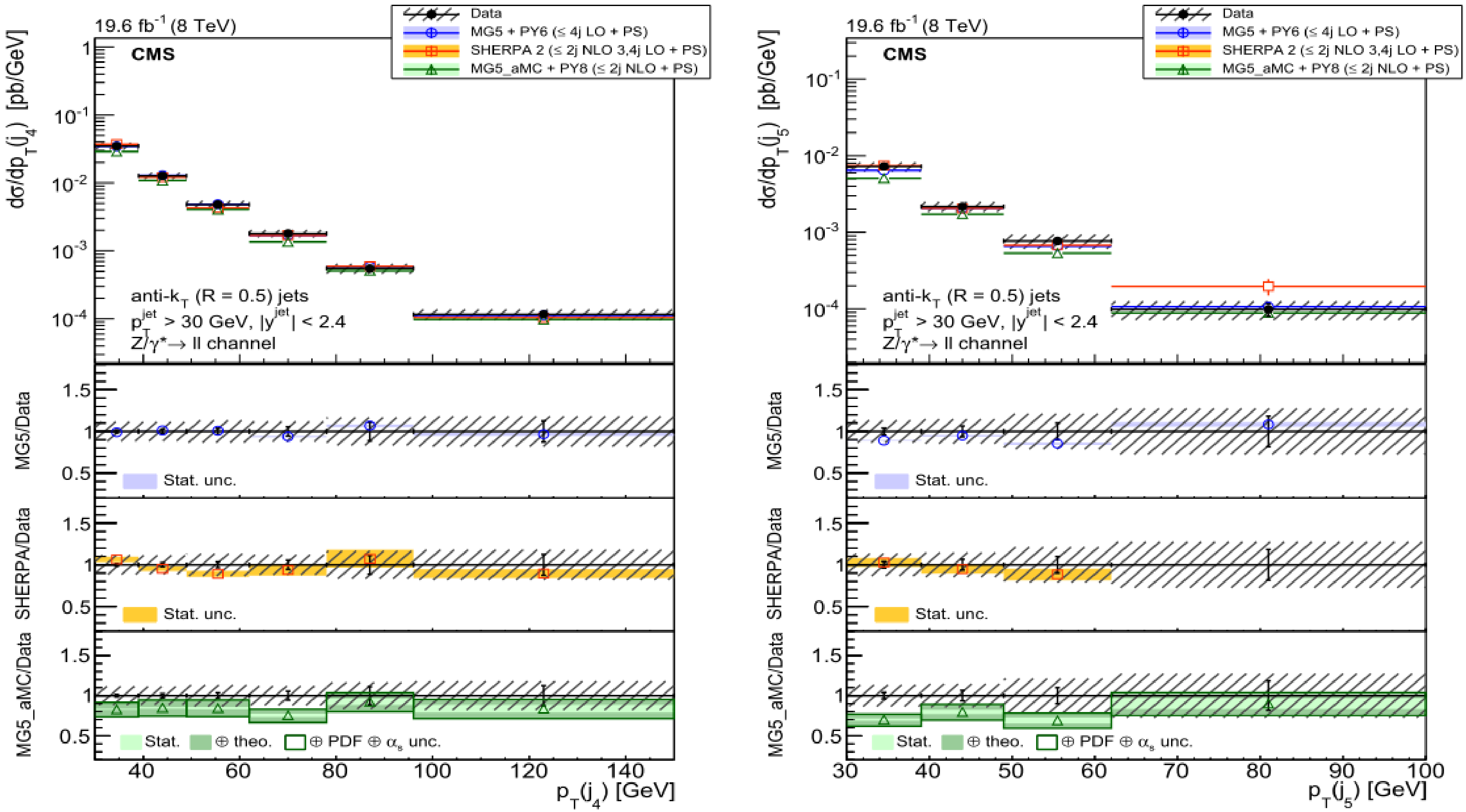
- The unfolded data of leading jet pt with two leptonic decay channels are compatible
- The error bar in the plots stands for statistical uncertainty

Combined Results (8 TeV)



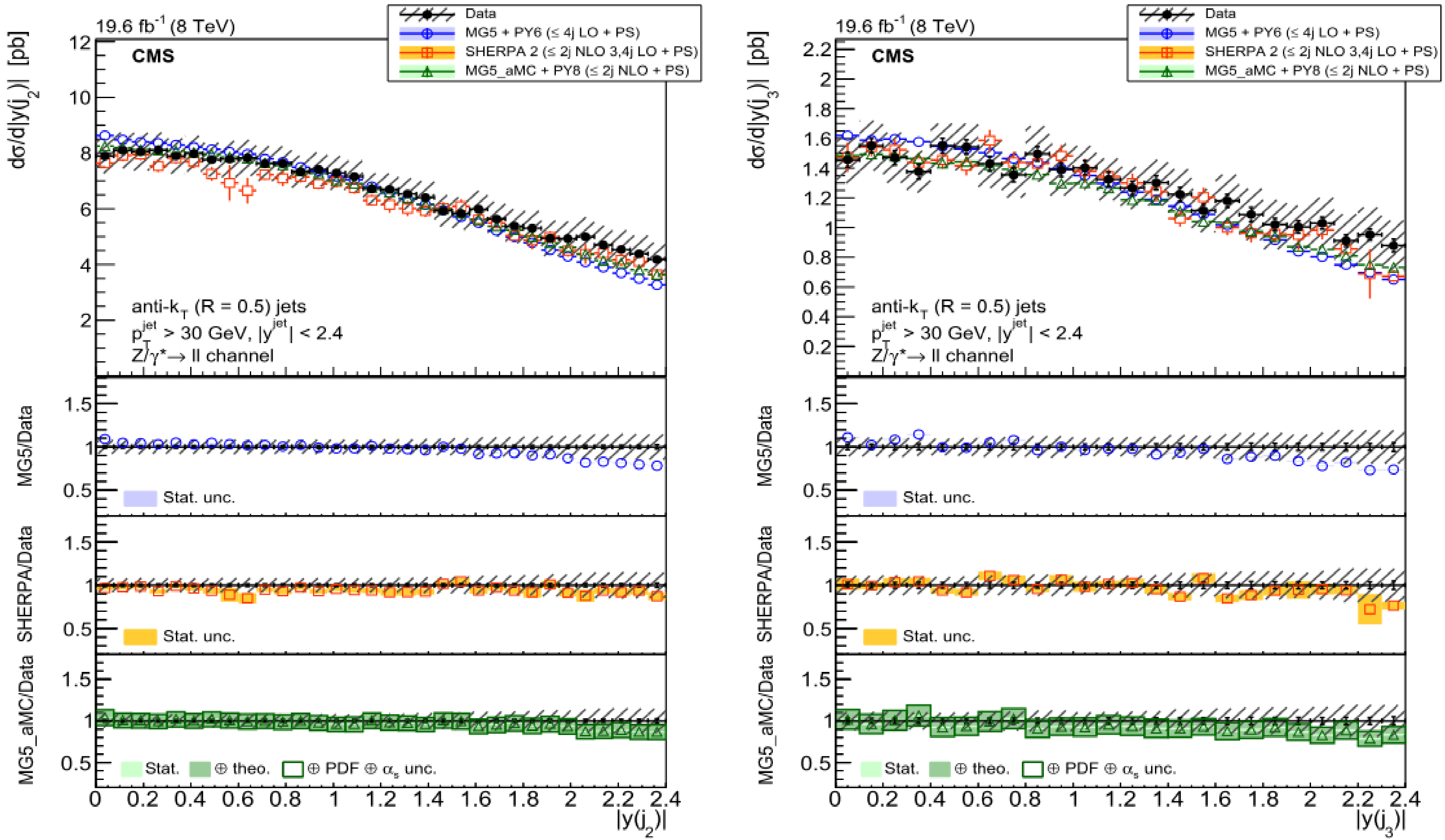
- Discrepancy with LO computation has disappeared with NLO accuracy
- Sherpa2 NLO predictions have a few statistical fluctuations

Combined Results (8 TeV)



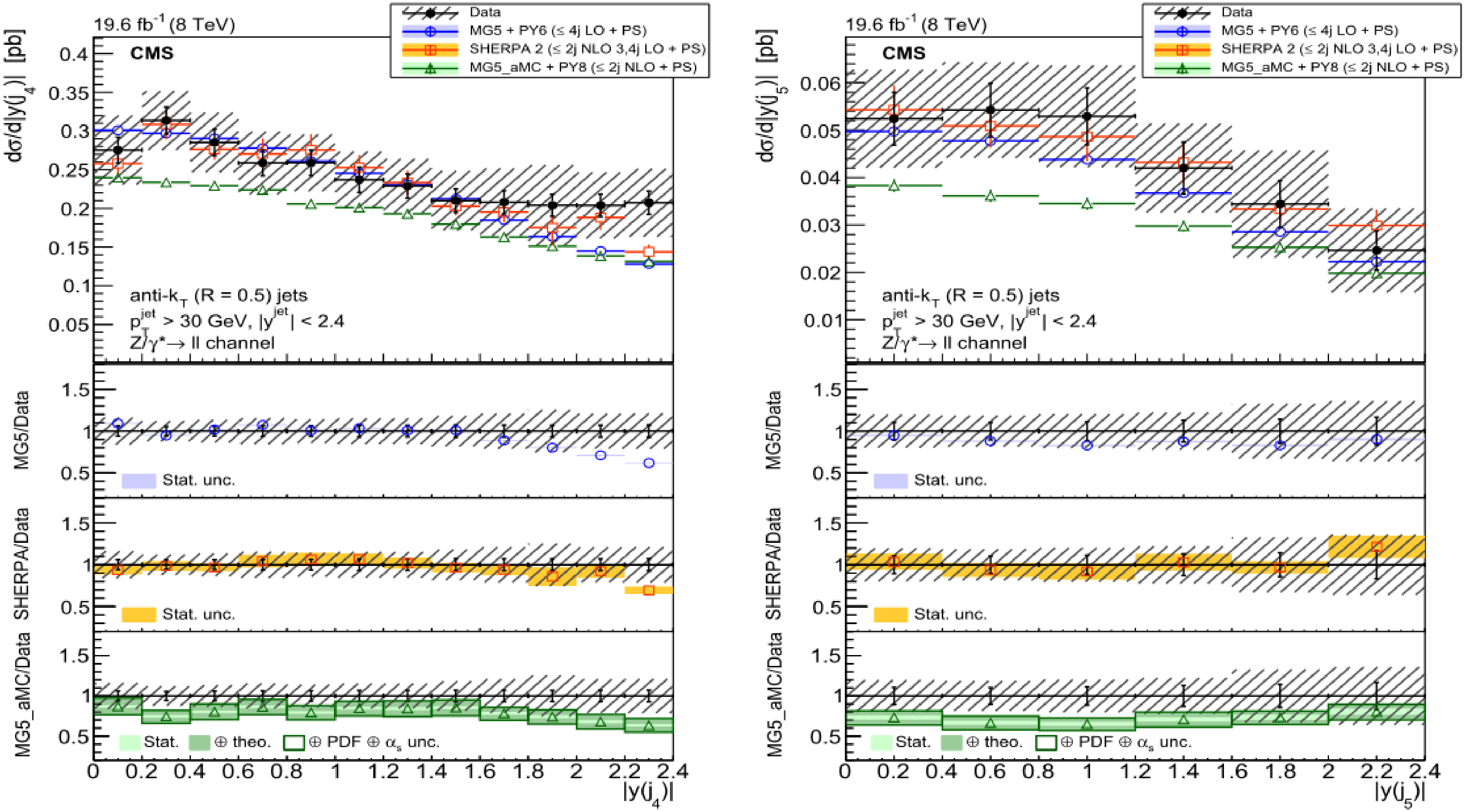
- Both MG5 (LO) and SHERPA2 (NLO) have good agreement with measurement
- MG5_aMC+PY8 differs due to limited number of partons in matrix element calculation (only relying on parton shower)

Combined Results (8 TeV)



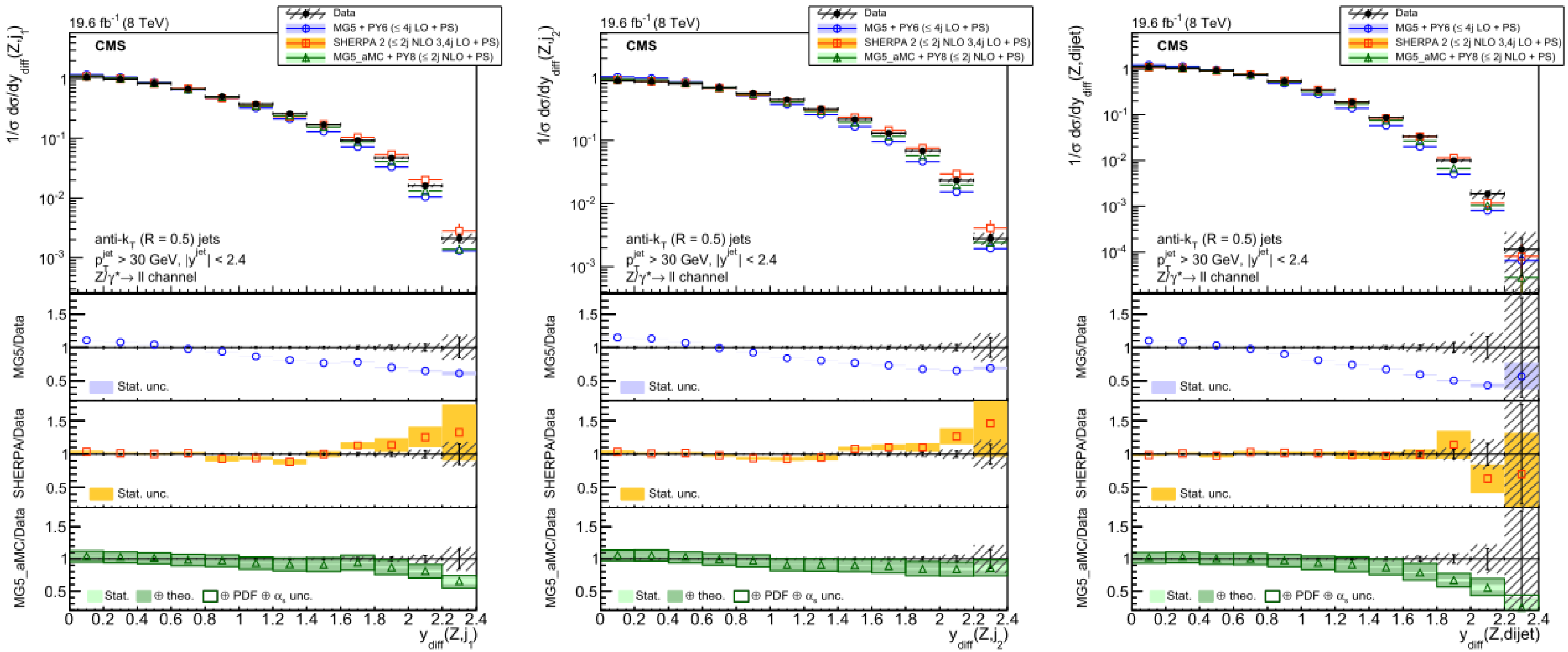
➤ Good agreements are seen for all the predictions

Combined Results (8 TeV)



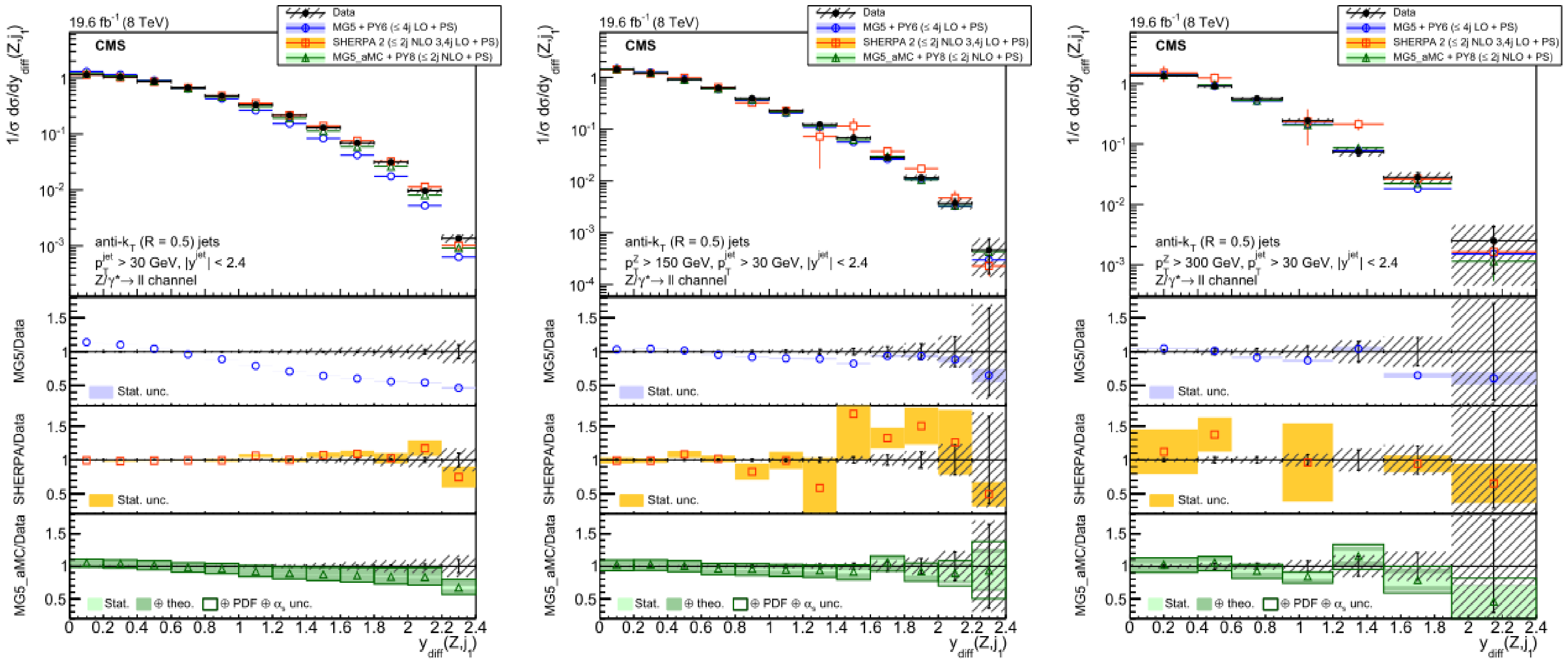
- Both MG5 (LO) and SHERPA2 (NLO) accuracy have good agreement with measurement
- MG5_aMC+PY8 differs due to limited number of partons in matrix element calculation (only relying on parton shower)

Rapidity correlation of Z + jets



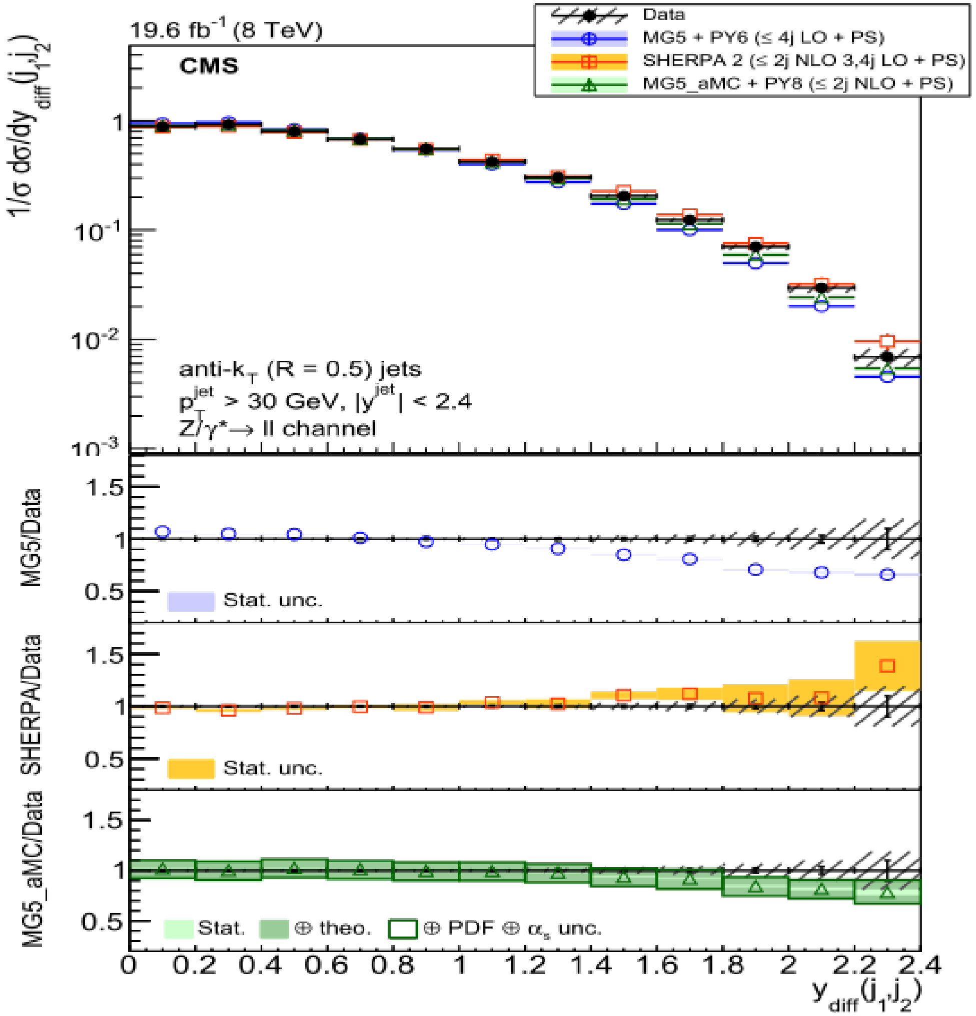
- The similar performance of rapidity difference at LO prediction when at least two jets are produced
- Discrepancy with LO computation has disappeared with NLO accuracy

Rapidity correlation of Z + jets



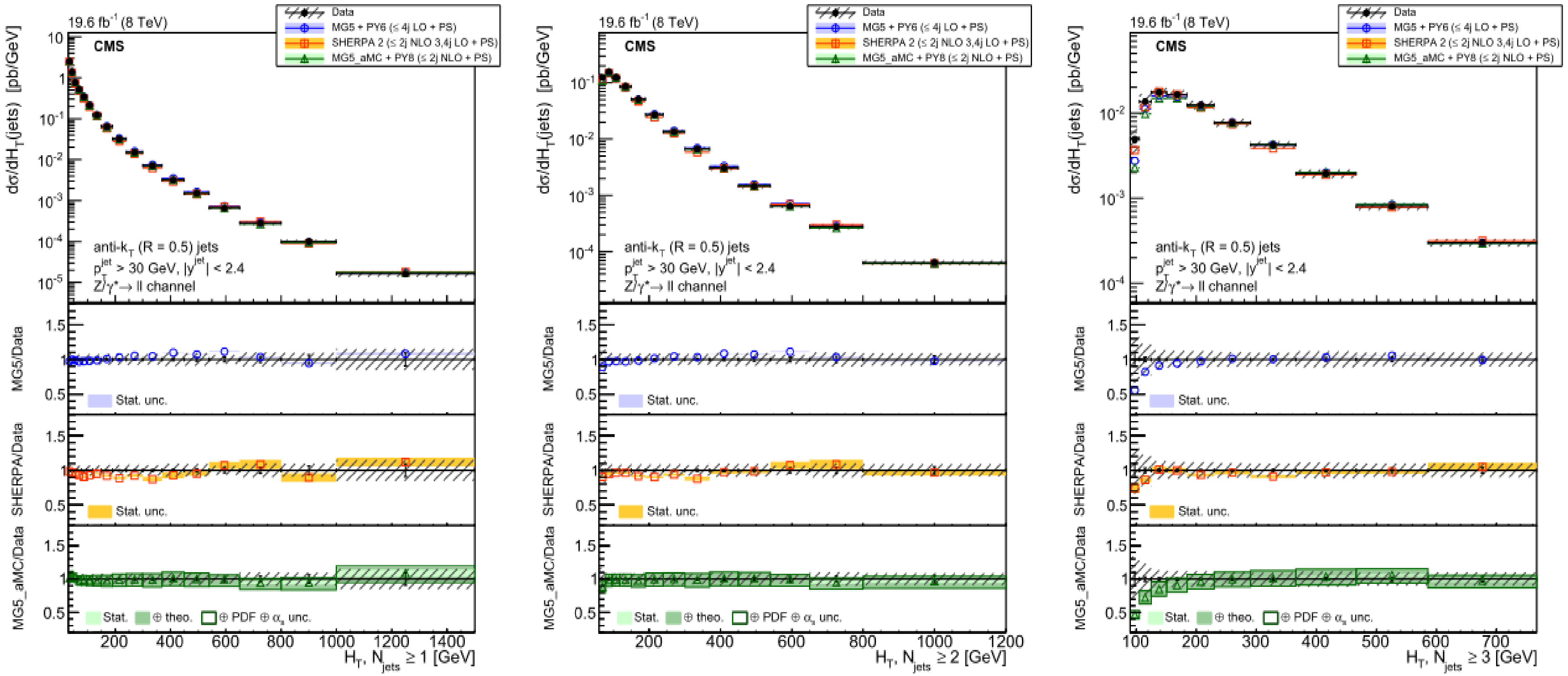
- Discrepancy with LO computation has disappeared with NLO accuracy
- The shape difference has decreased with large Z boson p_T cut

Rapidity correlation of Z + jets



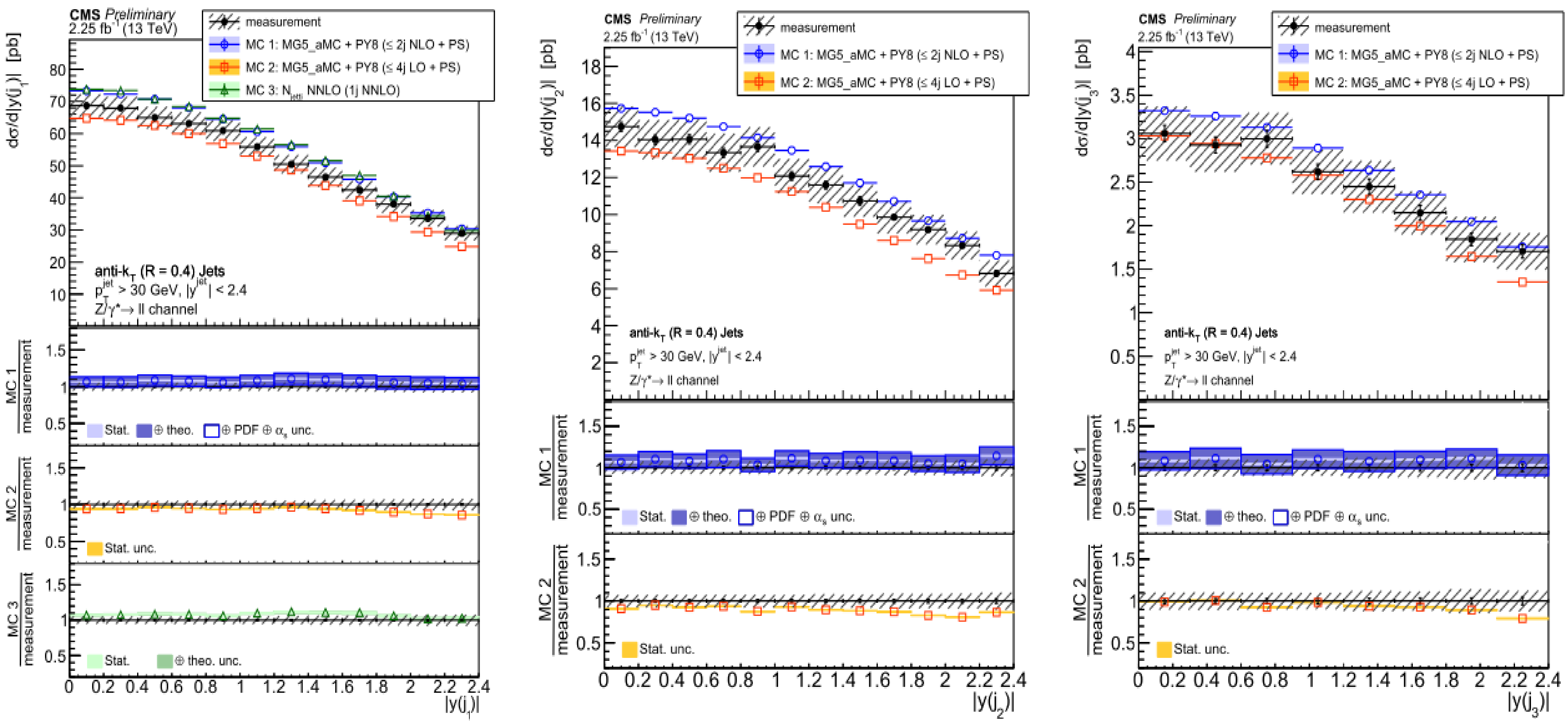
- Discrepancy with LO computation has disappeared with NLO accuracy

Combined Results (8 TeV)



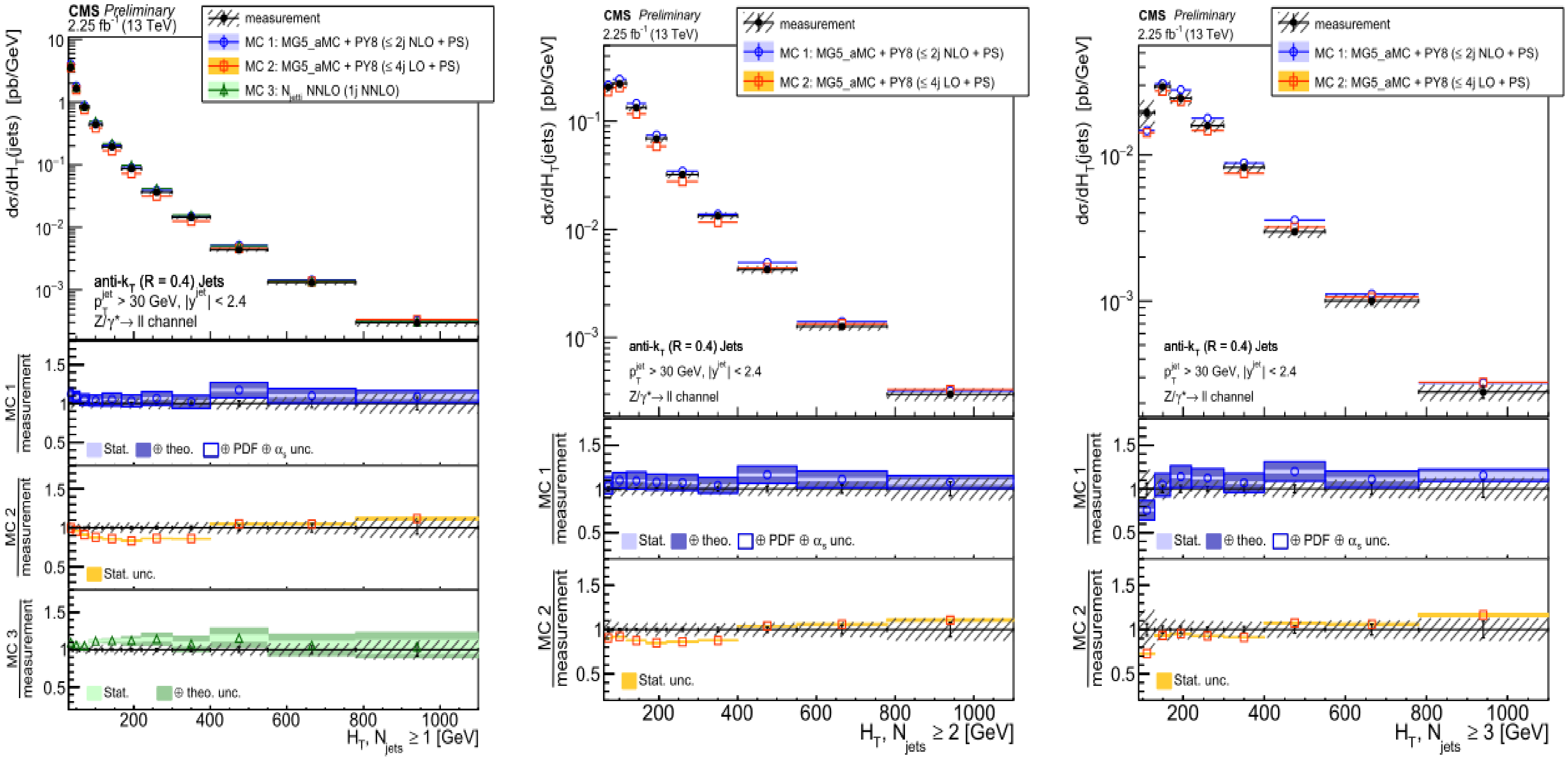
➢ Good agreements are seen for all the predictions

Combined Results (13 TeV)



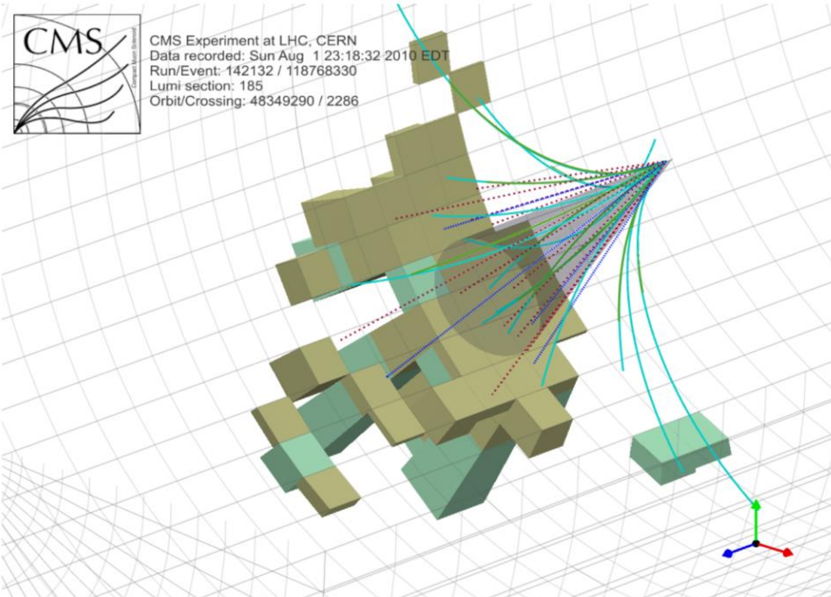
- ➢ Good agreements are seen with AMC@NLO and fixed NNLO predictions

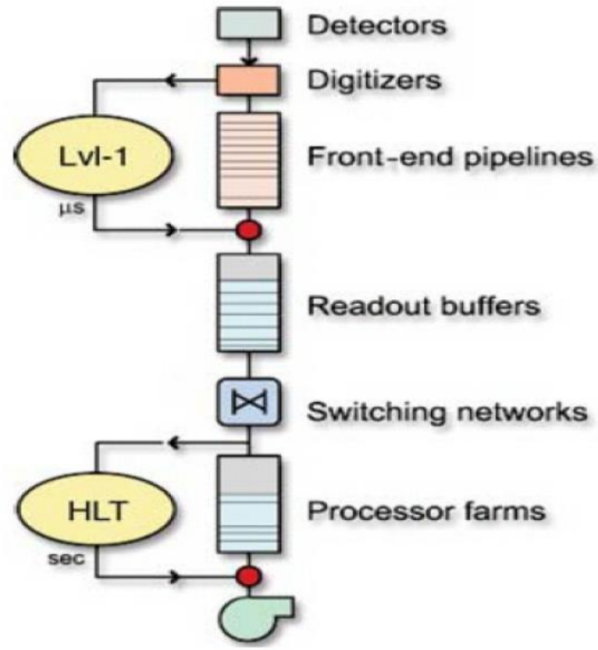
Combined Results (13 TeV)



- ➢ Good agreements are seen with AMC@NLO and fixed NNLO predictions

Jet energy correction & calibration at high level trigger in CMS





- RunII: much higher pileup contaminations on jets than RunI

- **jet p_T at reconstructed level ≠ jet p_T at produced level (the ratio is called jet response)**
  - Varying reconstruction efficiencies along the spatial coverage of the detector
  - Limited detector resolution

- Good level of jet response is crucial:
  - Efficient jet triggers
  - Proper jet trigger rate

Methodology of JEC

- L1 correction:
  - Subtract the pileup energy from jet clusters
- L2 corrections:
  - Correct the non-uniformity of reconstructed jet response along varying spatial angle
- L3 corrections:
  - Correct the jet response back to unity
- Calibration:
  - Jet response
  - Jet transverse energy resolution
  - Perform the residual correction for data if any difference between data and simulation (using the events of Z + jet production)

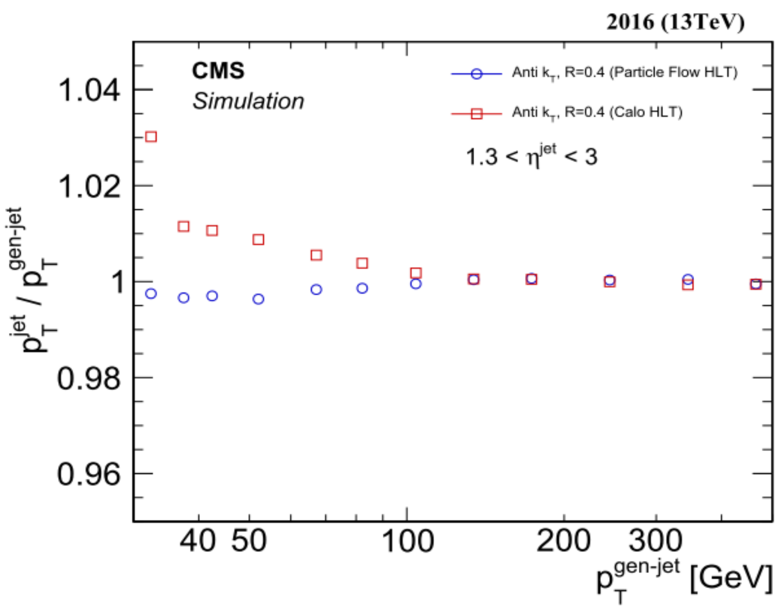
Methodology of JEC

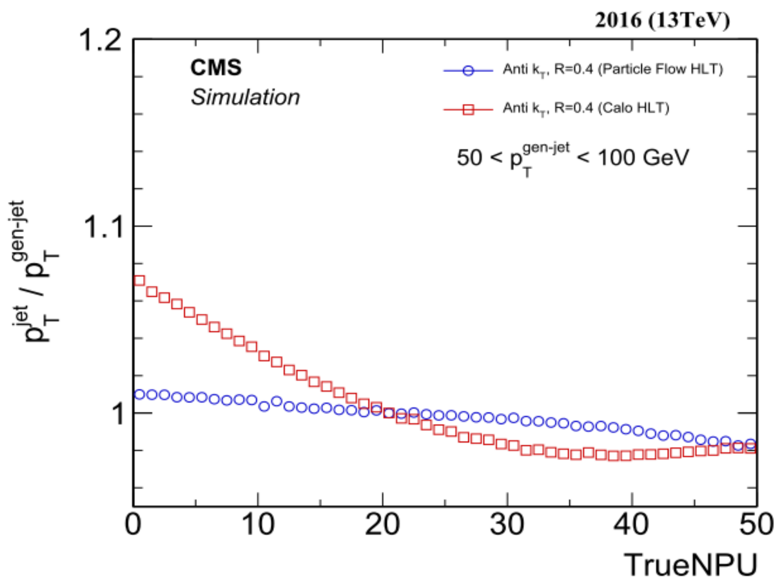
- Pileup (L1) correction:
  - Two jet enriched samples with the same pseudo-events in particle level while with/without pileup condition during detector simulation
  - Match the reconstructed jet event-by-event in between the two samples
  - Compute the transverse momentum ratio of each matched couple
  - Parametrize the ratios as correction factors

- Eta (L2) + Pt (L3) corrections:
  - The jet enriched sample with pileup condition with L1 correction applied to the reconstructed jet
  - Match the reconstructed jet with the jet in particle level according to their spatial distance
  - Derive the transverse momentum ratio of each matched couple
  - Parametrize the ratios as correction factors versus jet pt in fine eta ranges

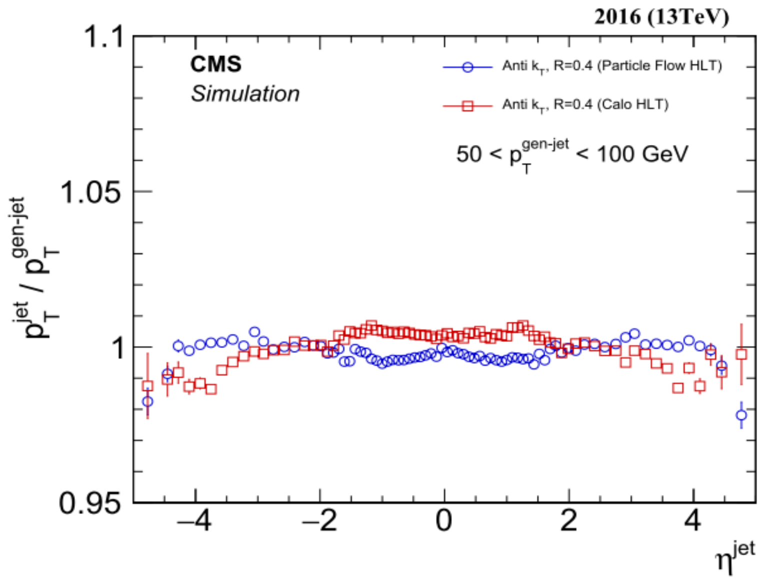
- Calibration:
  - Check the differential jet response and jet energy resolution versus number of pileup, eta, and pt after full corrections applied
  - Perform the residual correction for data if apparent difference between data and simulation (using the events of Z + jet production)

Calibration Results in simulation



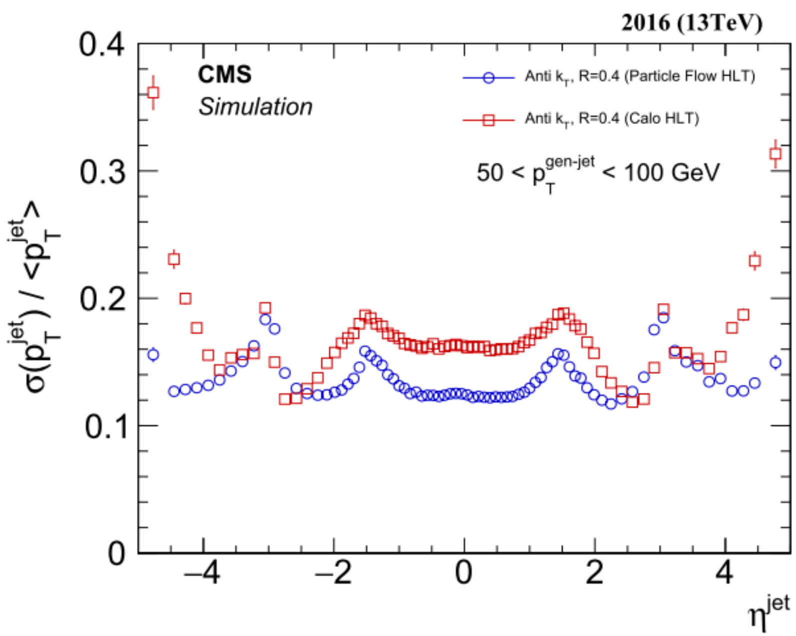


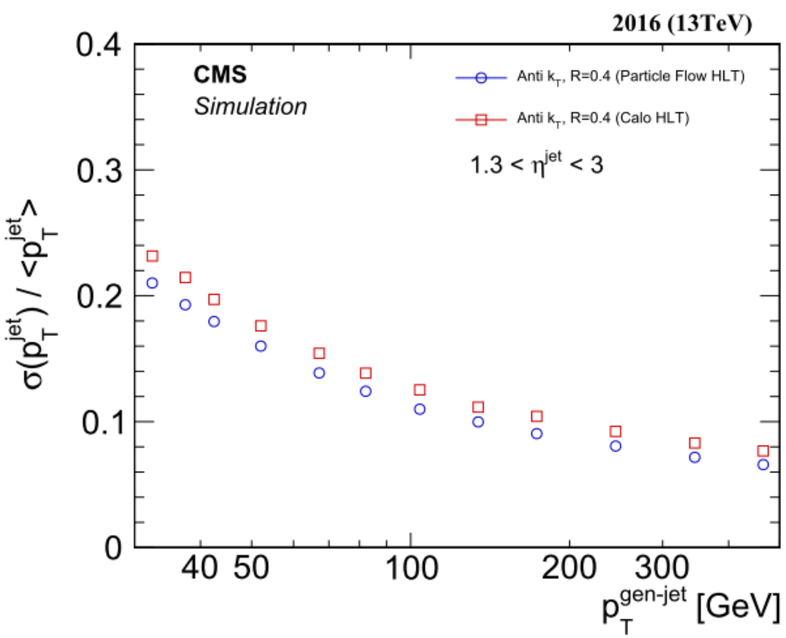
- ➢ Particle flow jet reconstruction using track + calorimetry
- ➢ Calorimeter jet reconstruction using calorimetry



- The remained discrepancy (away from unity) has little influence on the trigger efficiency and trigger rate

Calibration Results in simulation





- ➢ In general, particle flow jet at HLT has better transverse momentum resolution than calorimeter jet at HLT after full JEC at simulation level