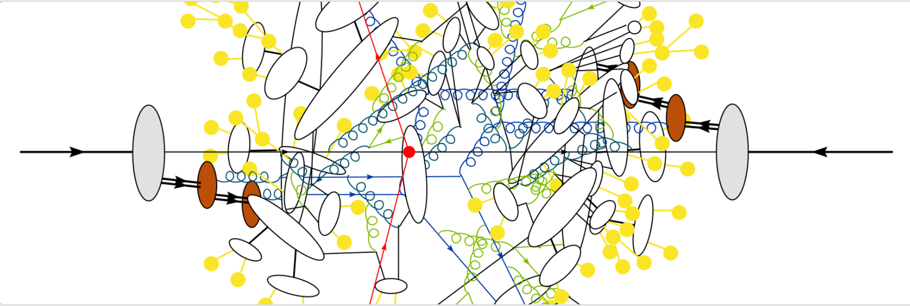
Triple-Differential Z+Jet Production at 13 TeV

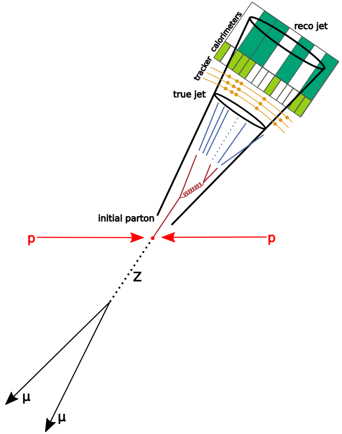
10th KSETA Plenary Workshop 2023

Robin Hofsaess, **Maximilian Horzela**, Günter Quast, Klaus Rabbertz, Cedric Verstege | 29. March 2023



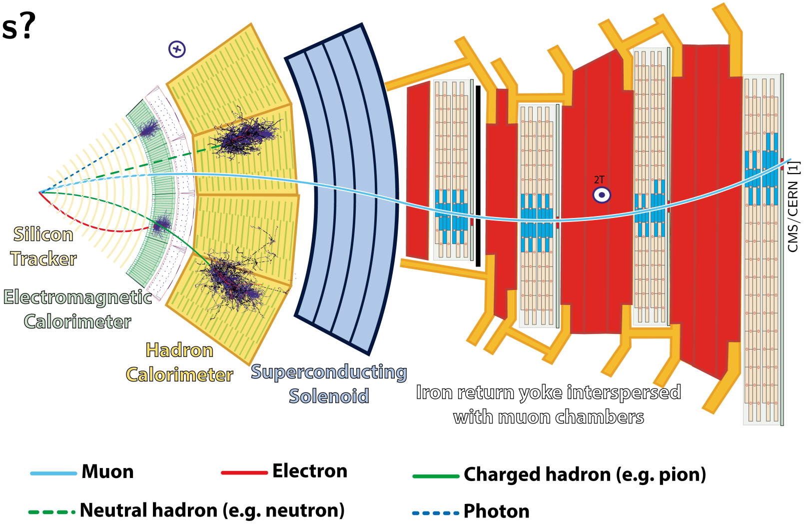
S. Gieseke

Why Z+Jets?

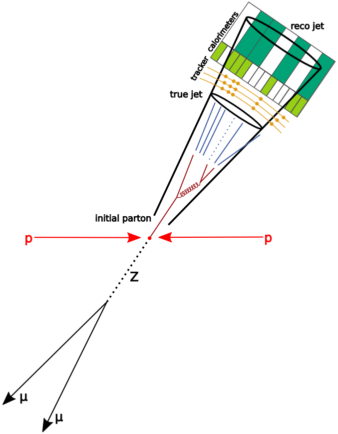


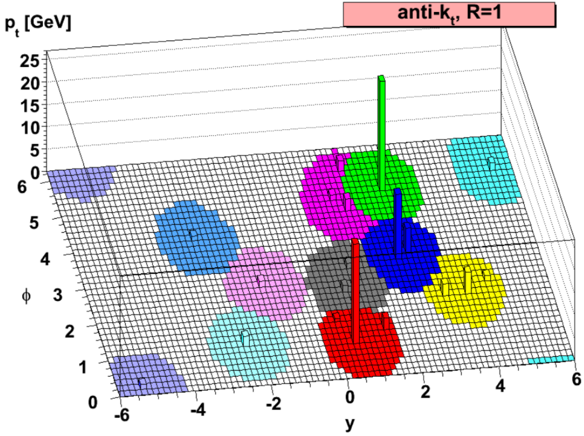
- Goal
  - Constraints on gluon (and other) parton distribution functions (PDFs)
  - Input for α_s fits
- $Z \to \mu^+\mu^-$
  - Good number of signal events
  - Low number of background events
  - Precise muon reconstruction and identification with CMS

Why Muons?


2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with muon chambers
CMS/CERN [1]
Muon
Electron
Charged hadron (e.g. pion)
Neutral hadron (e.g. neutron)
Photon

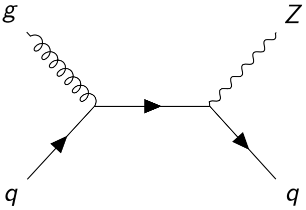
What is a Jet?



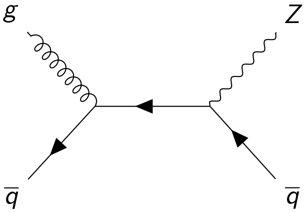


JHEP04(2008)063 [3]

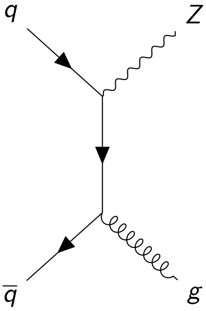
Why plus Jet?



(a) quark-gluon

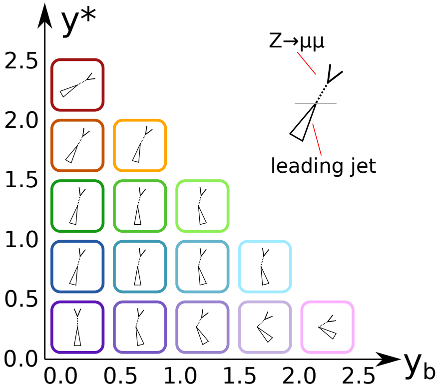


(b) antiquark-gluon



(c) quark-antiquark

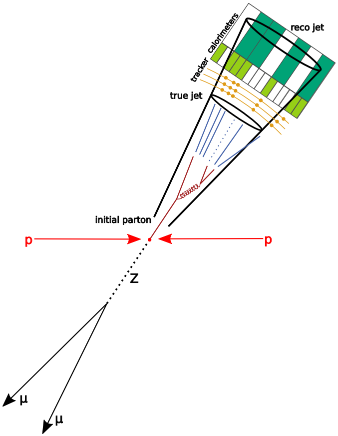
Why Triple-Differential?



M. Schnepf 2022 [5]

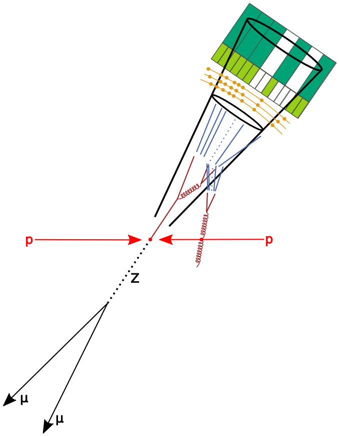
- Transverse momentum of di-muon system
  - p_{T}^{Z}
  - Scale of the hard interaction
- Boost of center-of-mass system
  - $y_b = \frac{1}{2}|y^Z + y^{\mathrm{Jet1}}|$
  - Parton momentum fractions of the protons
- Rapidity separation
  - $y^* = \frac{1}{2}|y^Z - y^{\mathrm{Jet1}}|$
  - Scattering angle in center-of-mass system

Event Selections and Corrections (Muons)



- Events passing single muon trigger
  - p_T^μ above trigger threshold
  - Corrected for trigger efficiency
- Muon selection
  - Within muon system coverage $|\eta| < 2.4$
  - Tight identification and isolation criteria
  - Cluster final state radiation (dressed muons)
- Z-boson reconstruction
  - $\mu^+\mu^-$ pair compatible with Z-boson mass

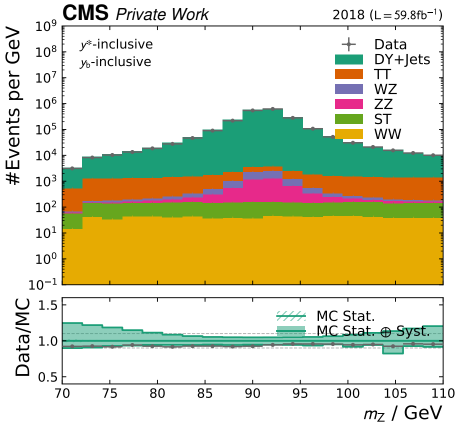
Event Selections and Corrections (Jets)



- Jet selection
  - Within similar detector coverage $|y| < 2.4$
  - Tight identification incl. lepton veto
  - Muon-jet overlap removal
  - Tight pile-up jet identification
  - p_T cut to reduce pile-up contributions

Datasets and MC

- Proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ recorded from 2016 to 2018
- Total integrated luminosity $138\,\text{fb}^{-1}$
- Signal MC:
  - $Z(\rightarrow \ell^+\ell^-) + 0, 1, 2$ jets MadGraph+Pythia8 aMC@NLO FXFX
- Background MC:
  - $t\bar{t} \rightarrow 2b2\ell2\nu$ Powheg+Pythia8 NLO
  - Single top quark t-channel and tW Powheg+MadSpin+Pythia8 NLO
  - Di-Boson WW, WZ, ZZ Pythia8 LO



Unfolding Basics

- Unfolding for detector effects of observation s to true spectrum t
  - Detector Resolution $\rightarrow$Migration between generator and reconstruction bins
  - Detector Efficiency $\rightarrow$Less events on reconstruction level than generator level

We have: $s(\vec{y}) = \int_X D(\vec{y}, \vec{x}) t(\vec{x}) \mathrm{d}\vec{x}$

We want: $t(\vec{x}) = \int_Y D'^{-1}(\vec{x}, \vec{y}) s(\vec{y}) \mathrm{d}\vec{y}$

Unfolding Basics

Karlsruhe Institute of Technology

- Unfolding for detector effects of observation s to true spectrum t
  - Detector Resolution $\rightarrow$Migration between generator and reconstruction bins
  - Detector Efficiency $\rightarrow$Less events on reconstruction level than generator level
- Variations below finite resolution $\leftrightarrow$ Ill-posed problem
- Usually s and t discretized in histograms $\rightarrow$ "invert" Response Matrix $\mathbf{R}$ (i.e. `TUnfold` [4])
  - Estimate Response Matrix from MC $\rightarrow$ Systematic and statistical uncertainties
  - If matrix ill-conditioned $\rightarrow$ Regularize "unphysical" oscillations

We have: $s(\vec{y}) = \int_X D(\vec{y}, \vec{x}) t(\vec{x}) \mathrm{d}\vec{x}$

We want: $t(\vec{x}) = \int_Y D'^{-1}(\vec{x}, \vec{y}) s(\vec{y}) \mathrm{d}\vec{y}$

$$s^i = \mathbf{R}^i_j t_j \rightarrow t_j = \mathbf{R^{-1}}^i_j s^i$$

$$\text{with} \quad \mathbf{R}^i_j = \frac{\int_{Y_i} \int_{X^j} D(\vec{y}, \vec{x}) t(\vec{x}) \mathrm{d}\vec{x} \mathrm{d}\vec{y}}{\int_{X^j} t(\vec{x}) \mathrm{d}\vec{x}}$$

TUnfold

- Algorithm for estimating truth $\mathbf{t}$ from measured observables $\mathbf{s}$
- Assumes Gaussian distribution of $\mathbf{s}$ with average $\tilde{\mathbf{s}} = \mathbf{R}\tilde{\mathbf{t}} \rightarrow$ least-square method
- Maximize likelihood

$$\mathcal{L} = (\mathbf{s} - \mathbf{Rt})^T \mathbf{V_{ss}} (\mathbf{s} - \mathbf{Rt}) + \mathcal{L}_{\text{reg}} + \mathcal{L}_{\text{norm}}$$

  with covariance matrix $\mathbf{V_{ss}}$

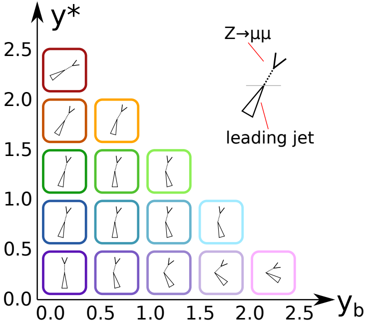
$\rightarrow$ General analytical solution $\mathbf{t_0}(\mathbf{s}, \mathbf{R}, \mathbf{V_{ss}})$ and covariance $\mathbf{V_{tt}}(\frac{\partial \mathbf{t_0}}{\partial \mathbf{s}}, \mathbf{R}, \mathbf{V_{ss}}) \rightarrow$ plug in and do the linear algebra

$\rightarrow$ Similar for contributions to $\mathbf{V_{tt}}$ due to statistical uncertainties on $\mathbf{R}$

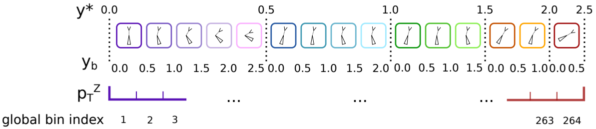
- Avoid regularization $\mathcal{L}_{\text{reg}}$, when $\mathbf{R}$ well-conditioned ($\delta\mathbf{s} \approx$ resolution)
- Avoid normalization $\mathcal{L}_{\text{norm}}$, when Gaussian approximation holds $\rightarrow$ true in this analysis

Bin Unraveling

3D phase-space → 1D visualization

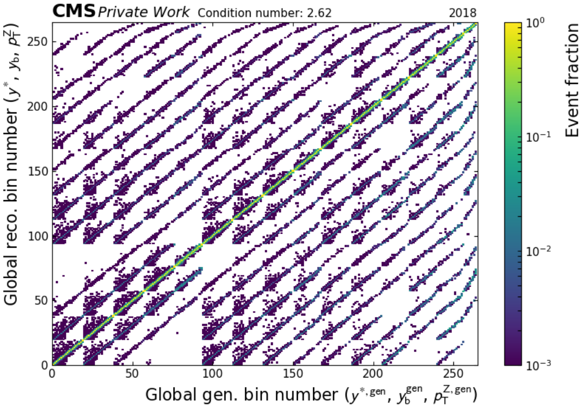


→



M. Schnepf 2022 [5]

Response Matrix



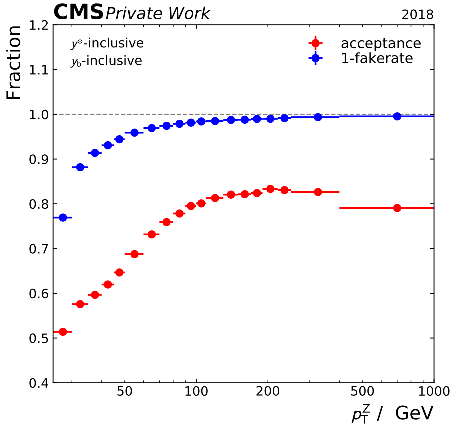
- Migrations from truth level to observation level due to detector effects

$$\mathcal{P}(\text{event in reco bin } i | \text{event in gen bin } j)$$

→ Fill MC events passing selections on reco- and gen-level for each analysis bin and normalize

- Small condition number, well-conditioned → Regularization not necessary
- Found to be similar for all data taking years

Acceptance & Fake rate

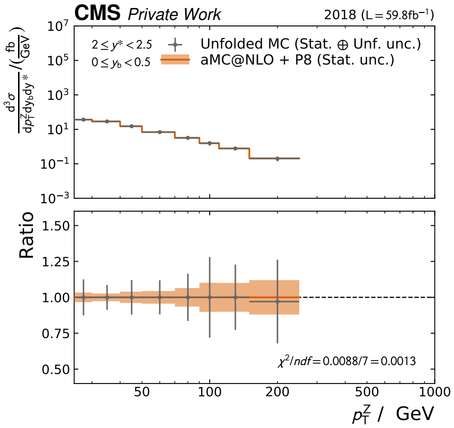


- Some events reconstructed in underflow or overflow of Response Matrix
- Acceptance
  - Events passing cuts on generator level but not on reconstruction level
  - Including detector & reconstruction inefficiencies

  → Treat as inefficiencies
- Fake rate
  - Events passing cuts on reconstruction level but not on generator level

  → Subtract as background
- Accounted for during unfolding

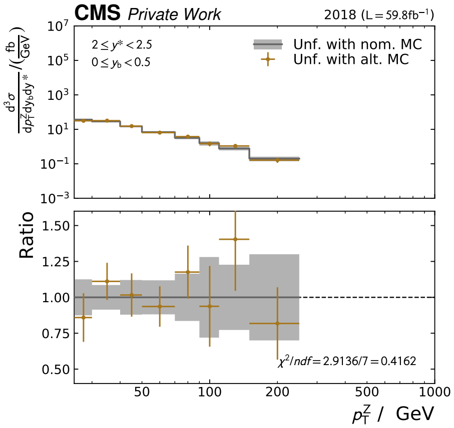
Unfolding Closure



- Consistency check
- Unfold simulated distribution with response using the same MC events
- Perfect agreement between unfolded MC and generator level

→ Unfolding works as expected

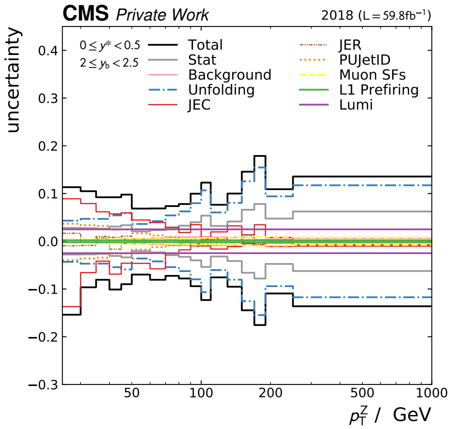
Model Dependence of Unfolding



- Choice of MC to fill Response Matrix might bias the results
- Estimate effect of alternative MC simulation
- Compare unfolded distribution using nominal & alternative response matrices
- Agreement within statistical uncertainties

→ No significant model dependence

Systematic Uncertainties

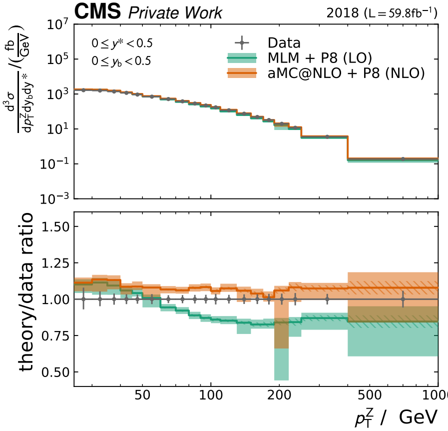


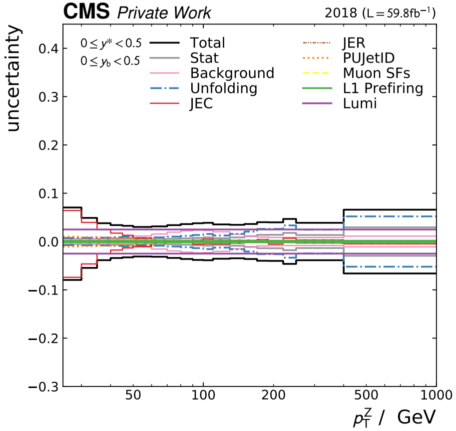
- Various systematic effects impacting unfolded event yields, e.g.
  - Jet energy calibration (JEC) and resolution correction on MC (JER)
  - Measured Luminosity
  - Limited number of events for creation of Response Matrix for Unfolding
  - Estimated Background contributions
  - ...
- Subject to systematical uncertainties

→ Adapt response matrix, acceptance, and fake rate for each systematical variation

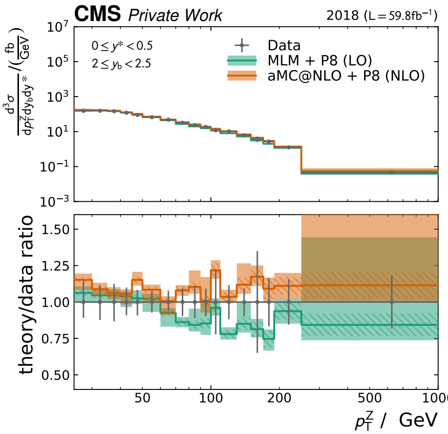
→ New unfolding for each uncertainty

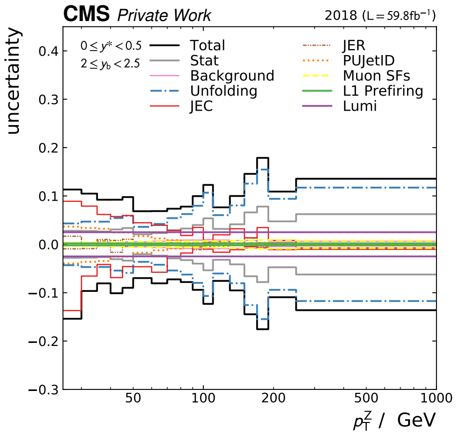
Unfolded Cross-Sections 2018: Central Region





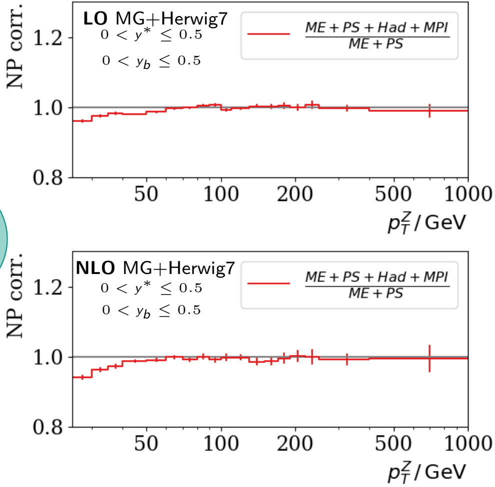
Unfolded Cross-sections 2018: High Boost





Comparison to NNLO Predictions and QCD Analysis

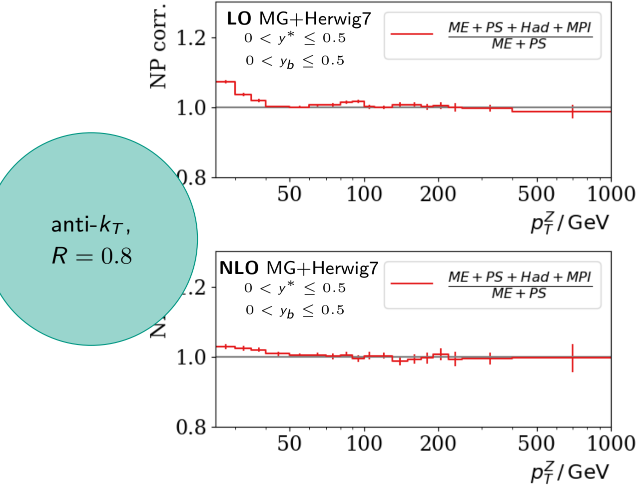
Karlsruhe Institute of Technology



anti-k_T, $R = 0.4$

- TBD: Compare and fit measured cross-sections to state-of-the-art theory predictions for Z+Jet $\rightarrow$ NNLO QCD $\otimes$ NLO-EWK $\otimes$ **non-perturbative (NP) corrections**
- NP corrections $\frac{ME+PS+Had+MPI}{ME+PS}$
  - dimish towards higher p_T^Z
  - change to slightly lower values from LO to NLO perturbative QCD

Comparison to NNLO Predictions and QCD Analysis

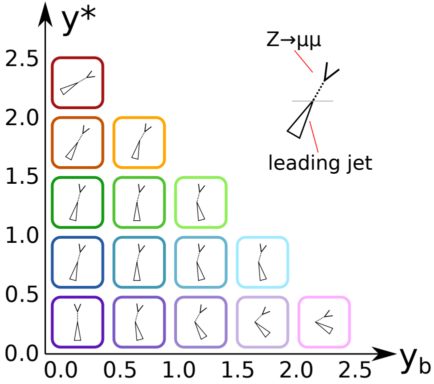


- TBD: Compare and fit measured cross-sections to state-of-the-art theory predictions for Z+Jet $\rightarrow$ NNLO QCD $\otimes$ NLO-EWK $\otimes$ **non-perturbative (NP) corrections**
- NP corrections $\frac{ME+PS+Had+MPI}{ME+PS}$
  - dimish towards higher p_T^Z
  - change to slightly lower values from LO to NLO perturbative QCD
  - depend on the jet type

Outlook

- First Z+Jet triple-differential cross-section measurement at $13\,\mathrm{TeV}$ with full Run II CMS data in progress
  - Combination of all data-taking periods from 2016 to 2018 into single measurement
  - Comparison to NNLO fixed-order predictions including electroweak and non-perturbative corrections
  - PDF and α_S fits

→ Aiming for publication early 2024



Backup

Detailed Event Selection

One $Z \to \mu\mu$ candidate with the following criteria

Selection	Value
Trigger	2016: HLT_IsoMu24 or HLT_TkMu24 2017: HLT_IsoMu27 2018: HLT_IsoMu24
Muon ID	Tight
Muon PF ISO	Tight
Muon p_T	> 29 GeV
Muon $\lvert\eta\rvert$	< 2.4
Z mass	$m_Z \pm 20$ GeV
Z p_T	> 25 GeV

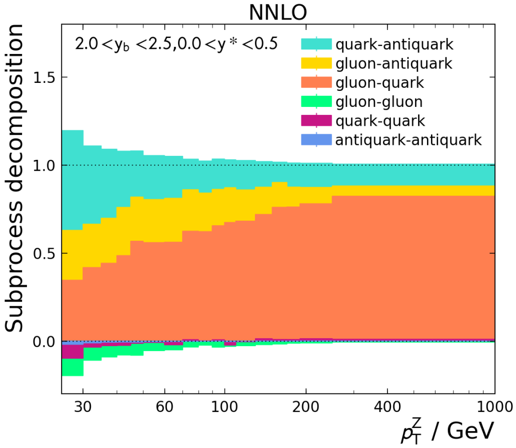
At least one Jet with the following criteria

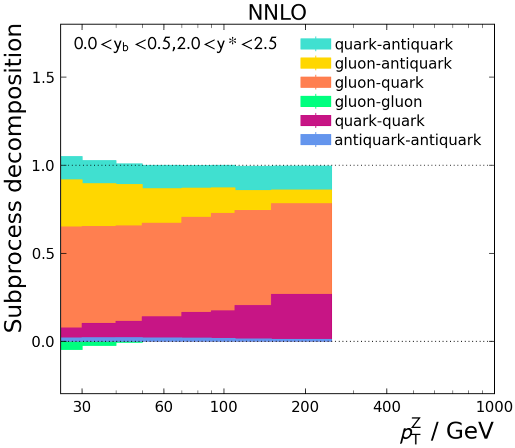
Selection	Value
Jet ID	Tight + Lepveto
PUJetID	Tight
$\Delta R(\mu_Z, \mathrm{Jet})$	> 0.4
Jet p_T	> 20 GeV
Jet $\lvert y\rvert$	< 2.4
Jet Veto Maps	✓

Corrections

Correction/SF	2016preVFP	2016postVFP	2017	2018
Muon RECO SFs	✓	✓	✓	✓
Muon ISO SFs	✓	✓	✓	✓
Muon ID SFs	✓	✓	✓	✓
Muon Trigger SFs	✓	✓	✓	✓
Muon Rochester	Data (kScaleDT) + MC (kSpreadMC)			
Muon Dressing	Data + MC with $\Delta R(\mu, \gamma) < 0.1$			
Muon L1Prefiring	✓	✓	✓	✓
ECAL L1Prefiring	✓	✓	✓	not needed
METFilters	Data + MC (All recommended for each year)			
PuJetID SFs	✓	✓	✓	✓
JEC	V7	V7	V5	V5
JER (hybrid)	V3	V3	V2	V2

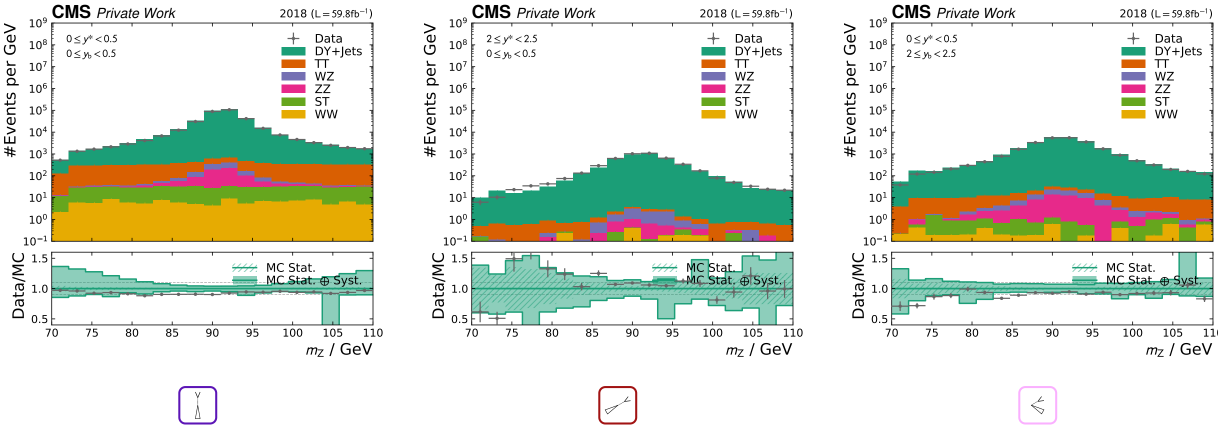
Dependence of Parton Luminosities on Phase-Space





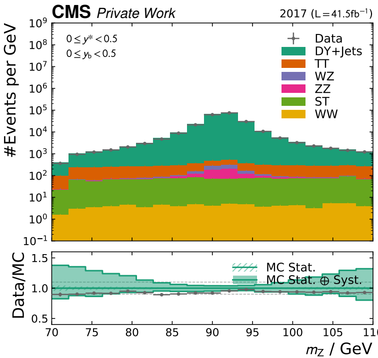
T. Berger 2019 [2]

Data/MC 2018

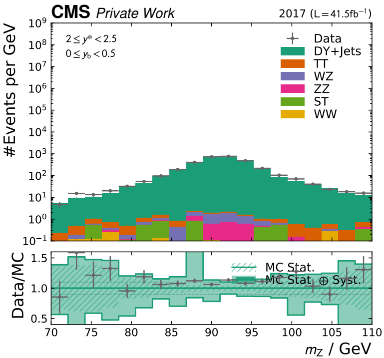


CMS Private Work
2018 (L = 59.8fb⁻¹)
0 ≤ y* < 0.5
0 ≤ y_b < 0.5
2 ≤ y* < 2.5
0 ≤ y_b < 0.5
0 ≤ y* < 0.5
2 ≤ y_b < 2.5
Data
DY+Jets
TT
WZ
ZZ
ST
WW
#Events per GeV
Data/MC
MC Stat.
MC Stat. ⊕ Syst.
m_Z / GeV

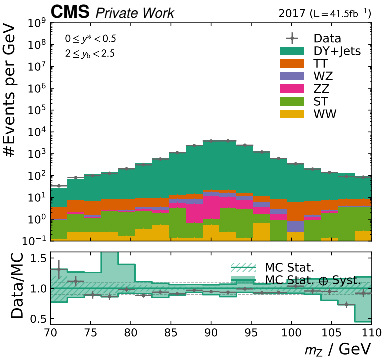
Data/MC 2017



CMS Private Work
2017 (L = 41.5fb⁻¹)
0 ≤ y* < 0.5
0 ≤ y_b < 0.5
Data
DY+Jets
TT
WZ
ZZ
ST
WW
#Events per GeV
Data/MC
MC Stat.
MC Stat. ⊕ Syst.
m_Z / GeV

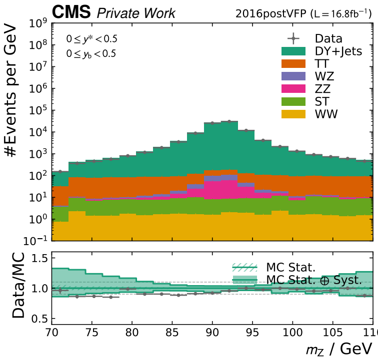


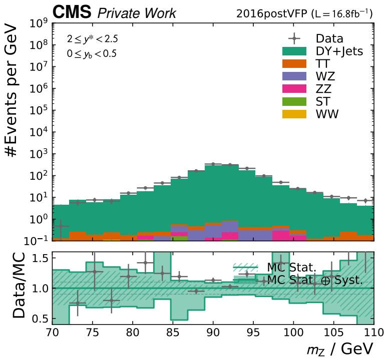
CMS Private Work
2017 (L = 41.5fb⁻¹)
2 ≤ y* < 2.5
0 ≤ y_b < 0.5
Data
DY+Jets
TT
WZ
ZZ
ST
WW
#Events per GeV
Data/MC
MC Stat.
MC Stat. ⊕ Syst.
m_Z / GeV

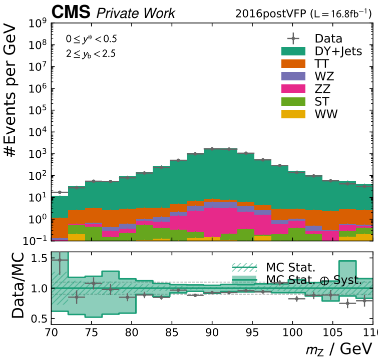


CMS Private Work
2017 (L = 41.5fb⁻¹)
0 ≤ y* < 0.5
2 ≤ y_b < 2.5
Data
DY+Jets
TT
WZ
ZZ
ST
WW
#Events per GeV
Data/MC
MC Stat.
MC Stat. ⊕ Syst.
m_Z / GeV

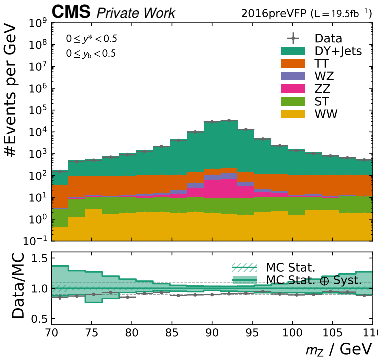
Data/MC 2016postVFP

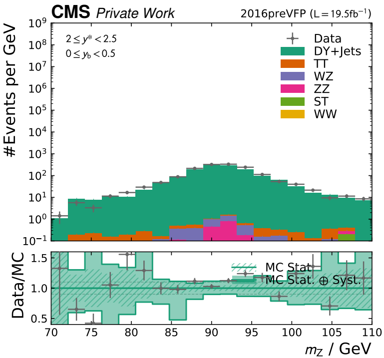


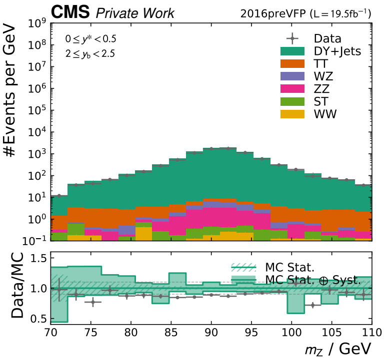




Data/MC 2016preVFP

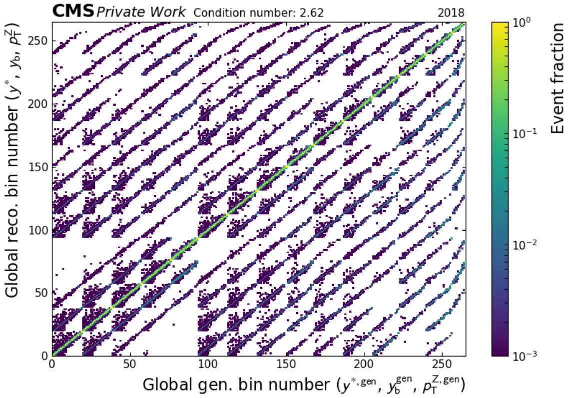




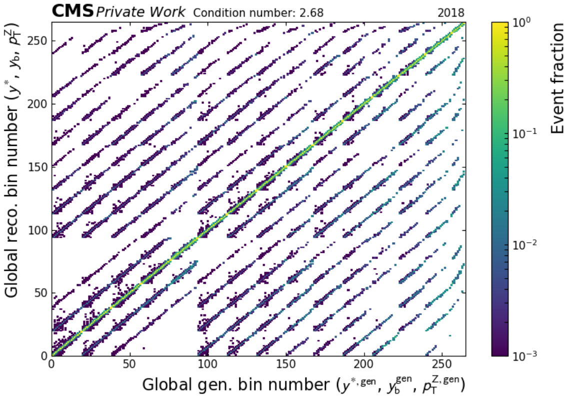


Response Matrices 2018

NLO

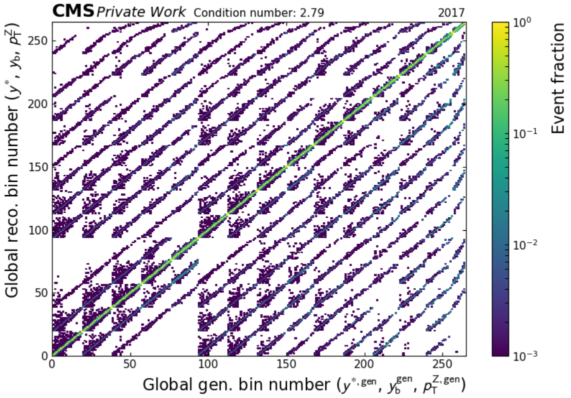


LO

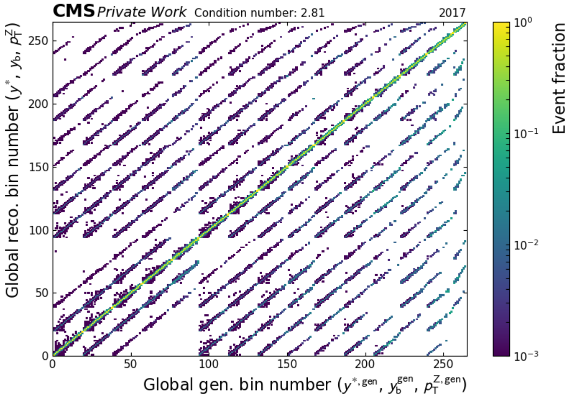


Response Matrices 2017

NLO

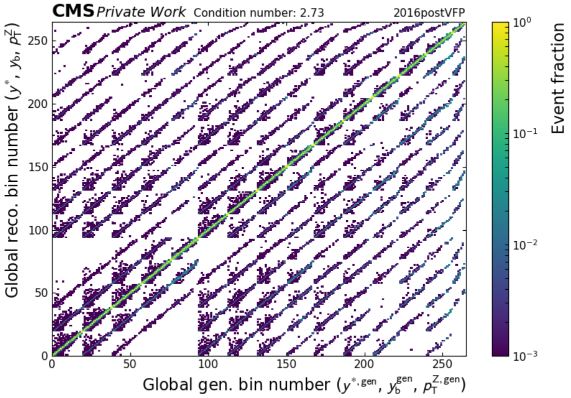


LO

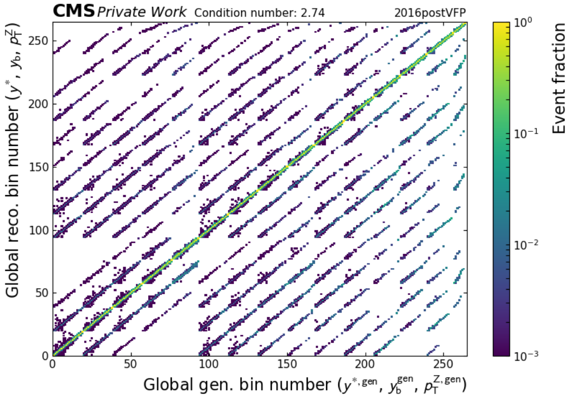


Response Matrices 2016postVFP

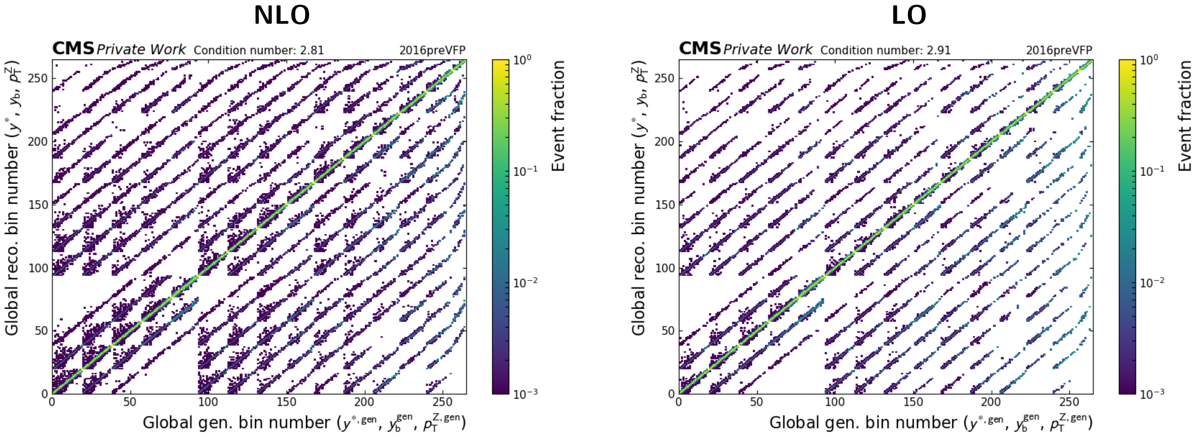
NLO



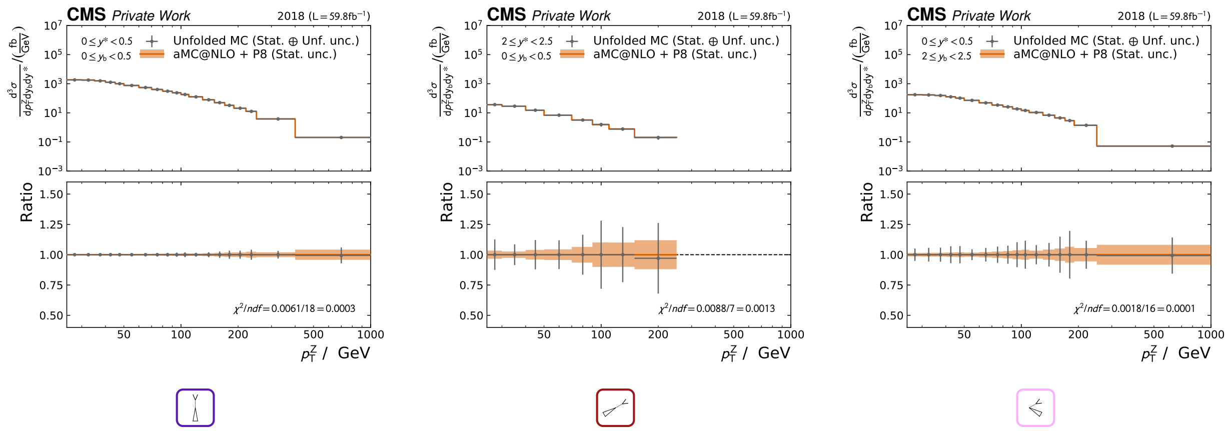
LO



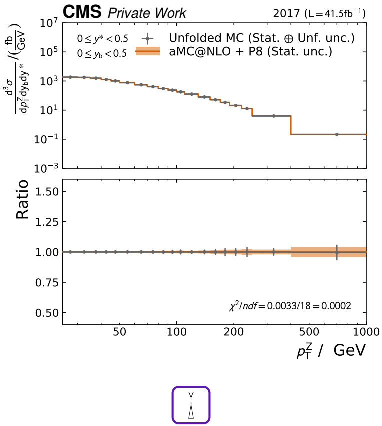
Response Matrices 2016preVFP

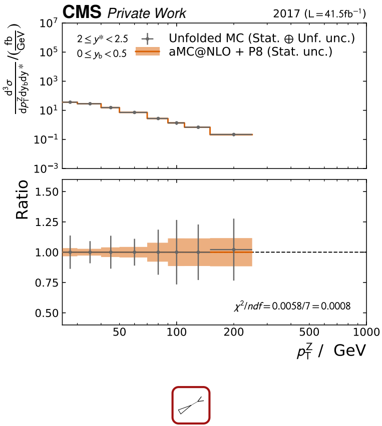


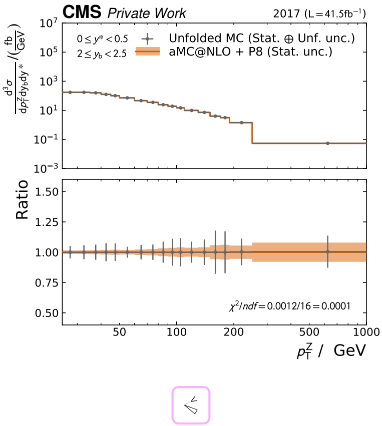
Unfolding Closure 2018



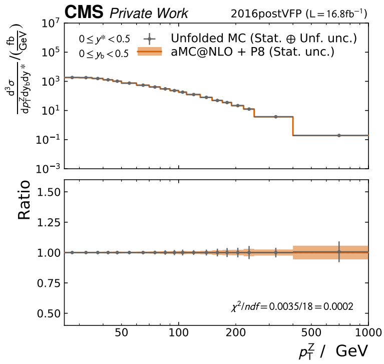
Unfolding Closure 2017

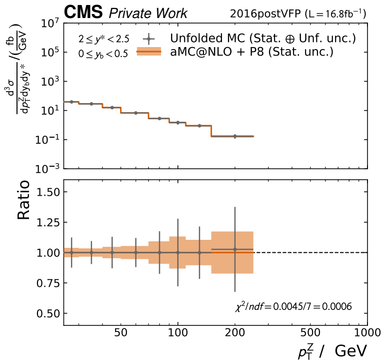


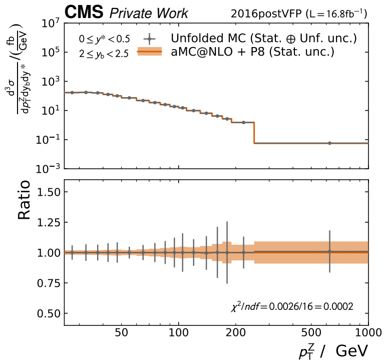




Unfolding Closure 2016postVFP

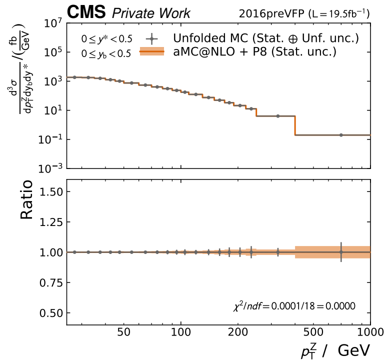


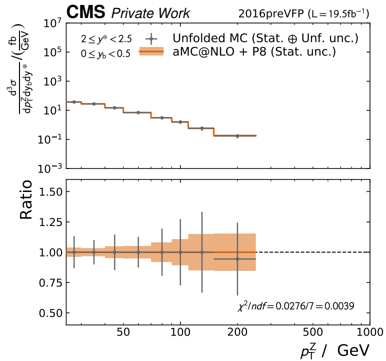


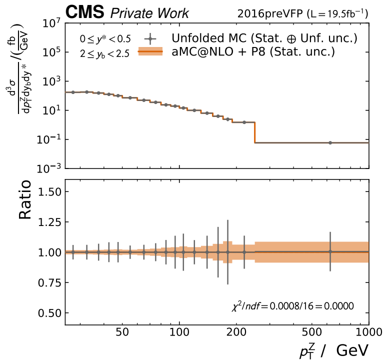




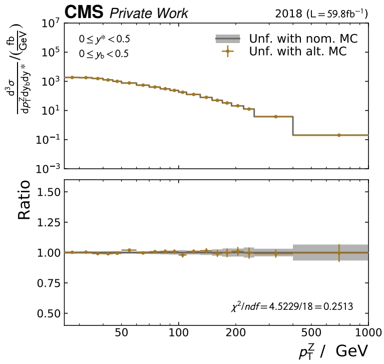
Unfolding Closure 2016preVFP

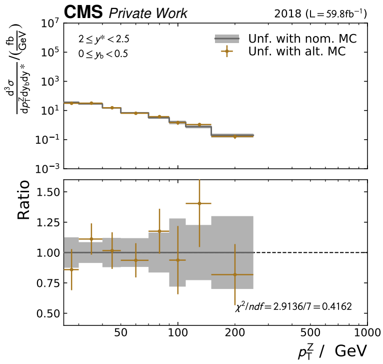


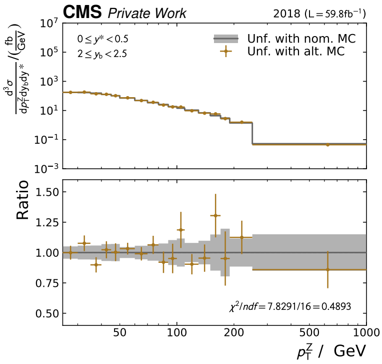




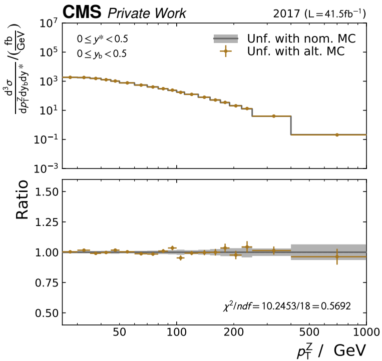
Unfolding Model Dependence 2018

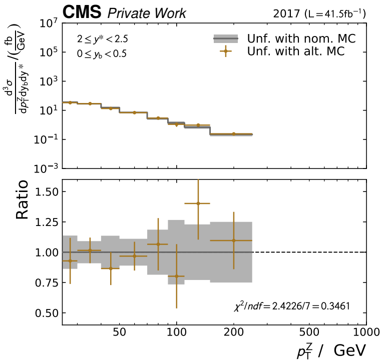


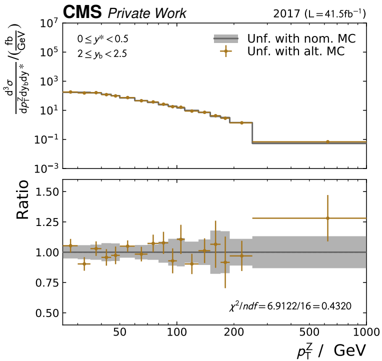




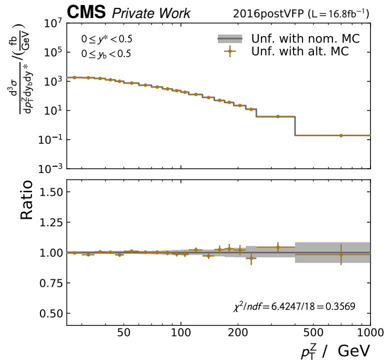
Unfolding Model Dependence 2017

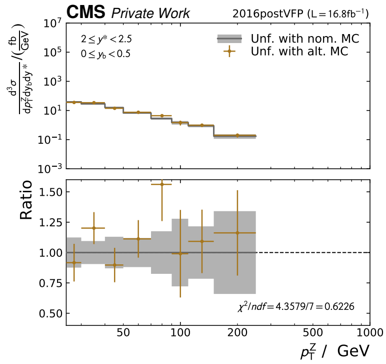


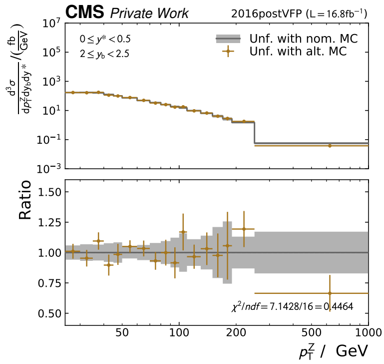




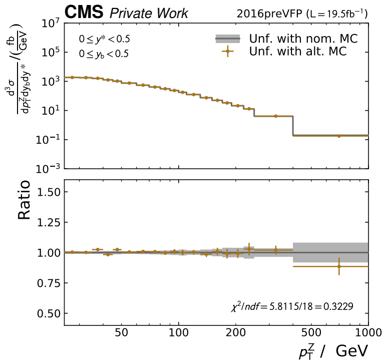
Unfolding Model Dependence 2016postVFP

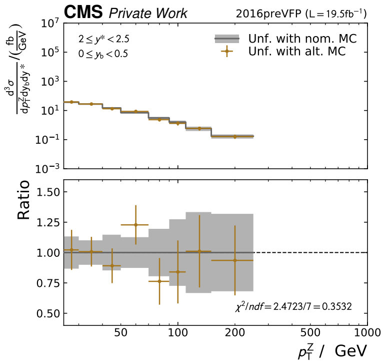


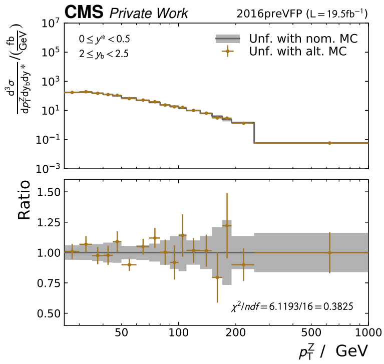




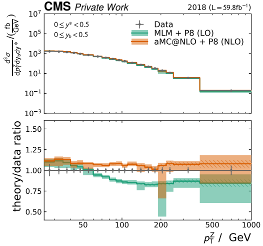
Unfolding Model Dependence 2016preVFP

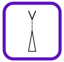


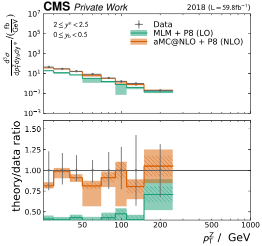


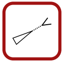


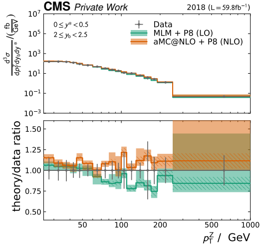
Unfolded Cross-Sections 2018

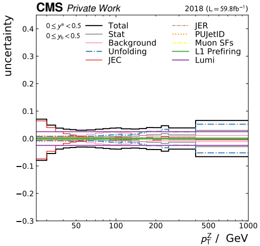


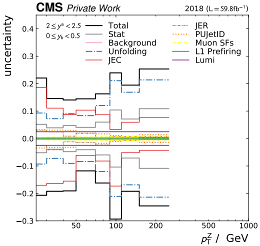


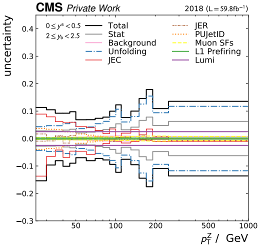




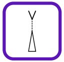


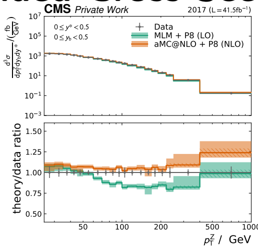


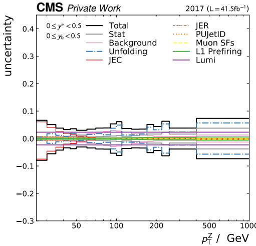


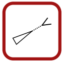


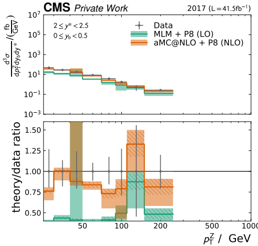
Unfolded Cross-Sections 2017

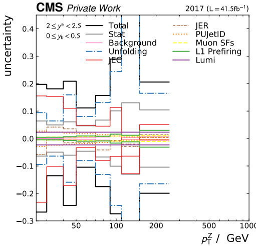


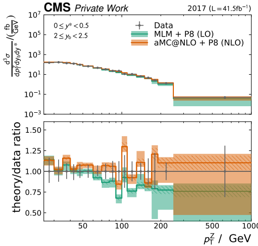


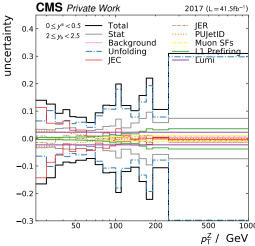




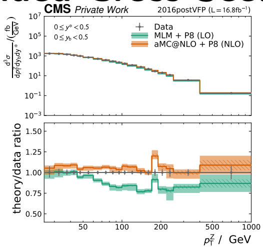


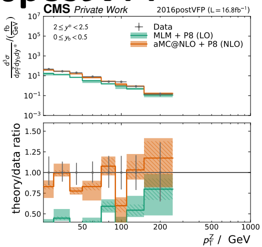


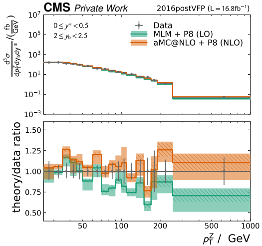


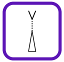


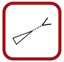
Unfolded Cross-Sections 2016postVFP

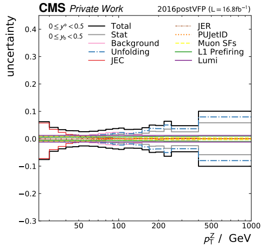


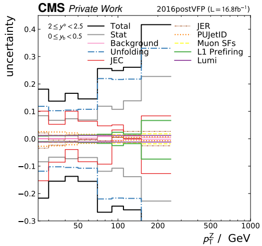


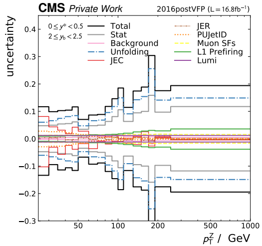




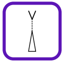


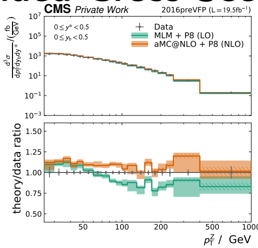


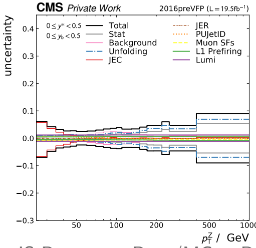


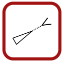


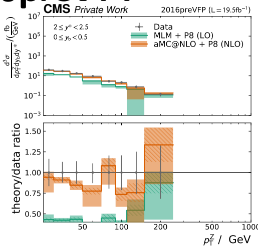
Unfolded Cross-Sections 2016preVFP

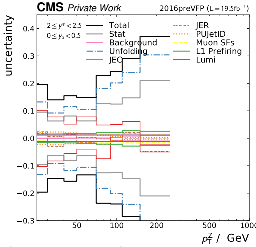


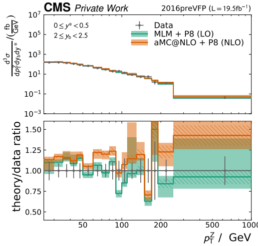


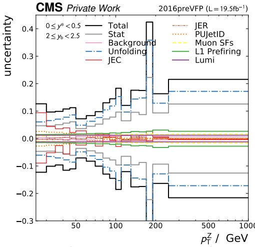




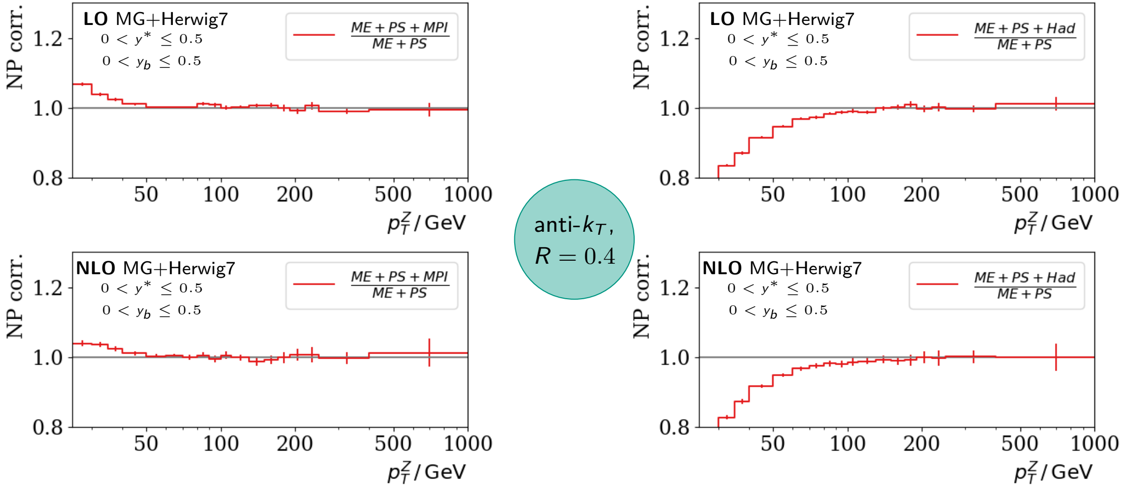




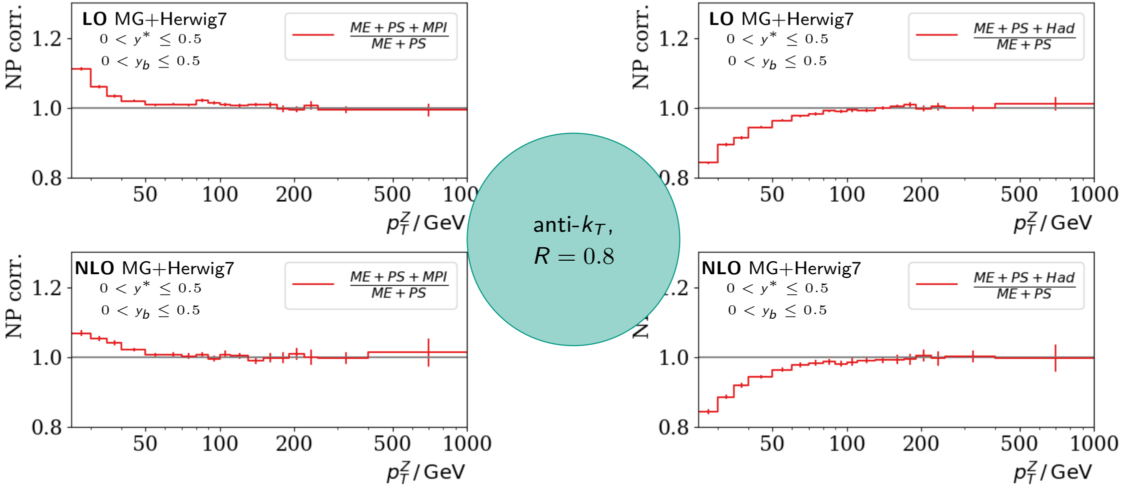




MPI and Hadronization effects



MPI and Hadronization effects



References I

[1] David Barney. "CMS Detector Slice". CMS Collection. 2016. URL: https://cds.cern.ch/record/2120661.

[2] Thomas Berger. "Jet energy calibration and triple differential inclusive cross section measurements with Z ($\rightarrow\mu\mu$) + jet events at 13 TeV recorded by the CMS detector". PhD thesis. Karlsruher Institut für Technologie (KIT), 2019. 139 pp. DOI: 10.5445/IR/1000104286.

[3] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-kt jet clustering algorithm". *Journal of High Energy Physics* 2008.04 (Apr. 2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. URL: https://dx.doi.org/10.1088/1126-6708/2008/04/063.

References II

[4] Stefan Schmitt. "TUnfold: an algorithm for correcting migration effects in high energy physics". *JINST* 7 (2012), T10003.
DOI: 10.1088/1748-0221/7/10/T10003. arXiv: 1205.6201 [physics.data-an].

[5] Matthias Schnepf. "Dynamic Provision of Heterogeneous Computing Resources for Computation- and Data-intensive Particle Physics Analyses". PhD thesis. Karlsruher Institut für Technologie (KIT), 2022. 129 pp.
DOI: 10.5445/IR/1000143165.