Karlsruhe Institute of Technology

Institute of Experimental Particle Physics

3D Z+Jet Cross-Section Measurement at $\sqrt{s} = 13\,\text{TeV}$

ETP Meeting – Master Thesis Presentation

Cedric Verstege | 07. November 2022

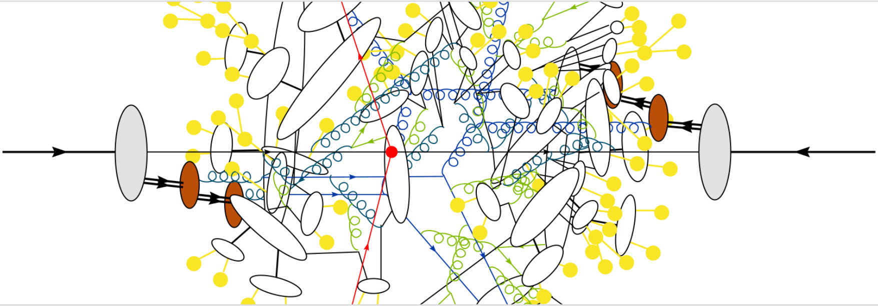


Table of Contents

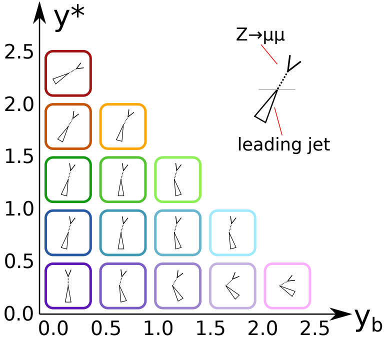
Why . . . ?

. . . Z $\rightarrow \mu^+\mu^-$

- Adequate number of signal events with low background
- Muons efficiently reconstructable and identifiable

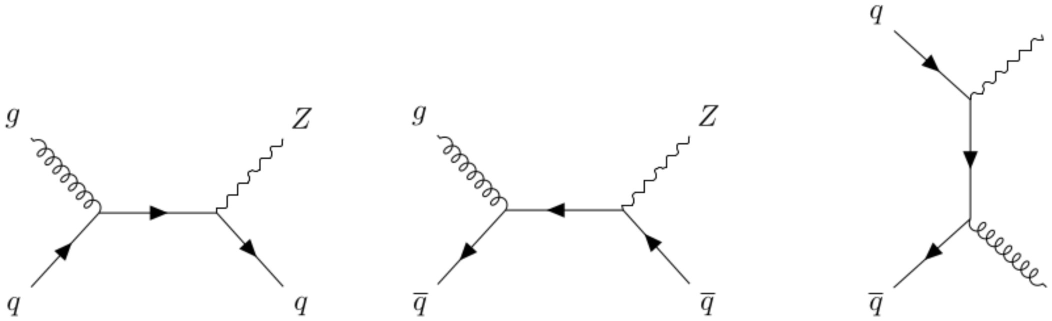
. . . **Triple Differential**

- Transverse momentum of dimuon-system $\mathbf{p}_\mathrm{T}^\mathrm{Z}$
  - information about momentum transfer of the hard interaction
- $\mathbf{y_b} = \frac{1}{2}|y^Z + y^\mathrm{Jet1}|$
  - boost of center-of-mass system
  - information about the initial state parton-momentum-fractions
- $\mathbf{y}^* = \frac{1}{2}|y^Z - y^\mathrm{Jet1}|$
  - Lorentz-invariant "scattering angle"
  - information about contributing parton luminosities



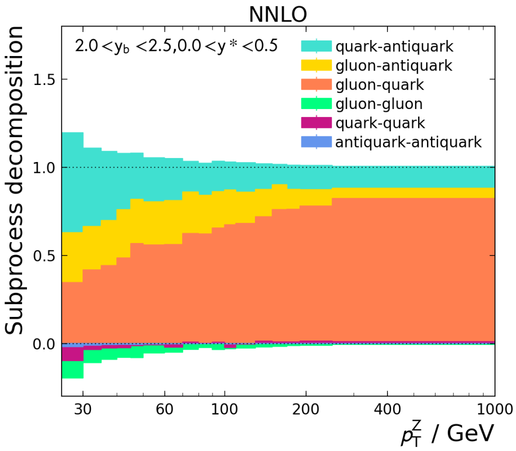
M. Schnepf 2022 [2]

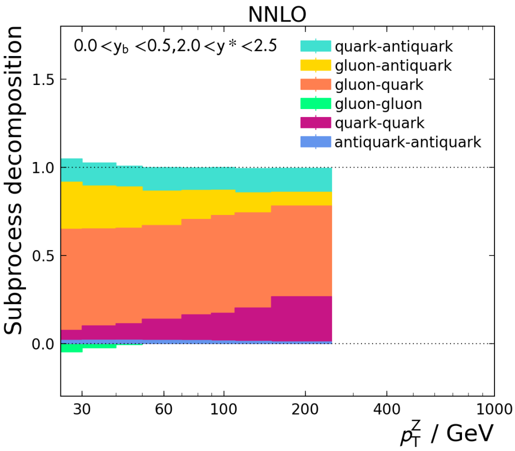
Z+Jet Production Channels at LO



(a) quark-gluon (b) antiquark-gluon (c) quark-antiquark

Variations of Parton Lumis in the Analysis Phase-Space



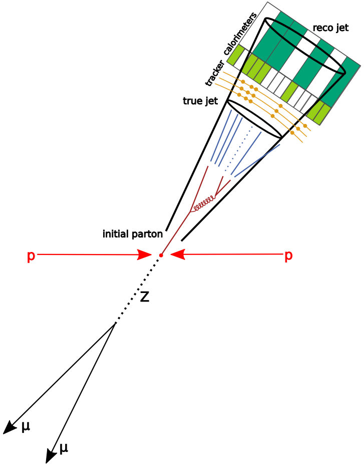


T. Berger 2019 [1]

Event Selections and Corrections

- Muon events selected with single muon trigger, corrected for L1 Prefiring
- Two muons passing tight ID and ISO above trigger threshold inside muon system coverage, dressed $\rightarrow$ Compatible with Z-boson
- At least one jet passing tight ID inside roughly same detector coverage
- Lepton veto in Jet ID and muon-jet-overlap removal

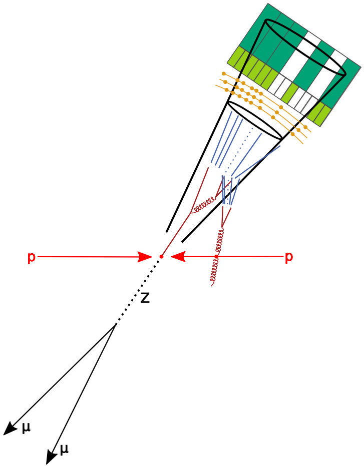
Detailed selections and corrections in Backup



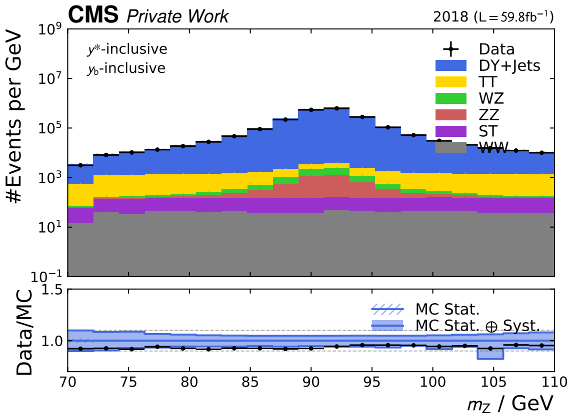
Event Selections and Corrections

- Muon events selected with single muon trigger, corrected for L1 Prefiring
- Two muons passing tight ID and ISO above trigger threshold inside muon system coverage, dressed
  $\rightarrow$ Compatible with Z-boson
- At least one jet passing tight ID inside roughly same detector coverage
- Lepton veto in Jet ID and muon-jet-overlap removal
- Reduced sensitivity to PU with PUJetID and Jet-p_T cut

Detailed selections and corrections in Backup



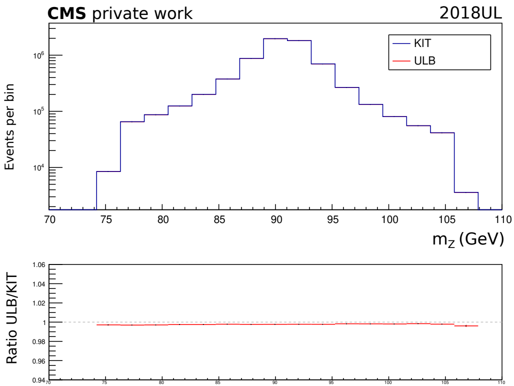
Detector Level Comparison of MC and Data



- Data and MC in good agreement within uncertainties
- Dominated by signal events
- MC overshoots by a small constant factor
- Inclusive NNLO cross-section
  FEWZ NNLO from 2019: 6077.22 pb
  FEWZ NNLO from 2017: 5818.37 pb
  - Theory cross section NNLO for inclusive Z production
    → NLO for Z+Jet
  - Dependent on y^*-y_b-bins
  → Results may help improve theory predictions

Framework Validation

- Updated code to UL
- Complete code review
  - → Found and fixed some bugs
- Framework cross check with Brussels for 2018 data
- Further updates on unfolding, uncertainty handling, . . .



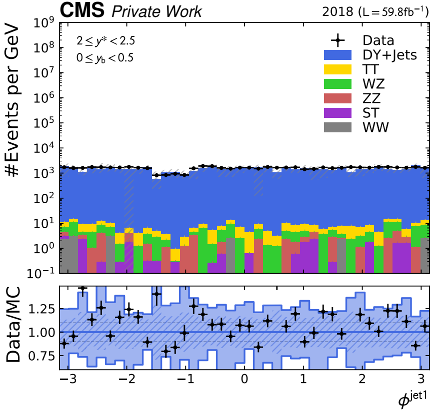
Combination of Inclusive and N_{jet}-Exclusive MC

- Systematic uncertainties through limited number of events in MC samples

⇒ Gather as much MC as available

- DYJetsToLL signal MC ∼3.9M events after selection

Inclusive DYJetsToLL



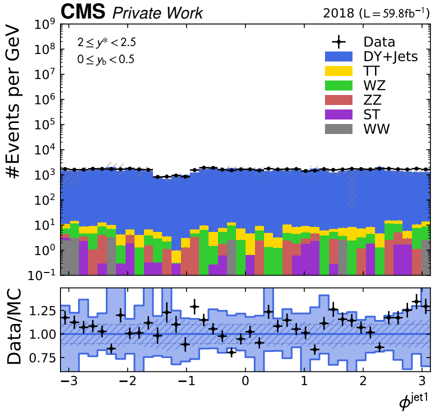
Combination of Inclusive and N_{jet}-Exclusive MC

- Systematic uncertainties through limited number of events in MC samples

$\Rightarrow$ Gather as much MC as available

- DYJetsToLL signal MC $\sim$3.9M events after selection
- Add jet-binned "high-stat." samples DYJetsToLL_0J, DYJetsToLL_1J, DYJetsToLL_2J (0.2M + 4.7M + 2.7M events after selection)
  - Reweight each exclusive sample to corresponding contribution in inclusive

Jet-binned DYJetsToLL



Combination of Inclusive and N_{jet}-Exclusive MC

Karlsruhe Institute of Technology

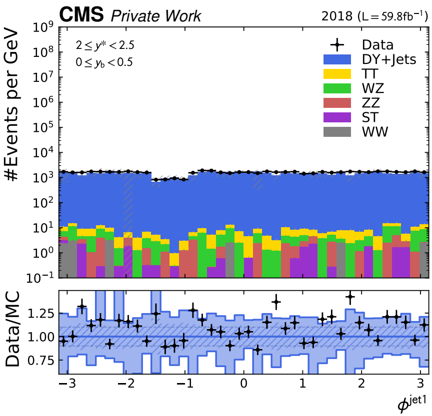
- Systematic uncertainties through limited number of events in MC samples

⇒ Gather as much MC as available

- DYJetsToLL signal MC ∼3.9M events after selection
- Add jet-binned "high-stat." samples DYJetsToLL_0J, DYJetsToLL_1J, DYJetsToLL_2J (0.2M + 4.7M + 2.7M events after selection)
  - Reweight each exclusive sample to corresponding contribution in inclusive
  - Reweight exclusive and inclusive samples according to effective number of events

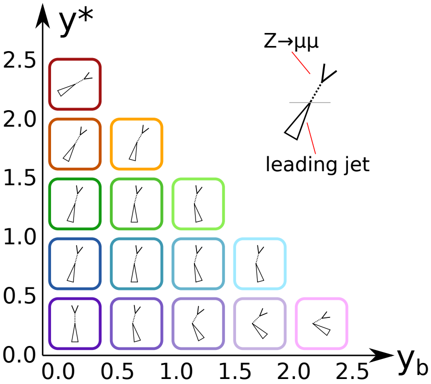
→ Reduced statistical uncertainty on MC

Combined signal samples

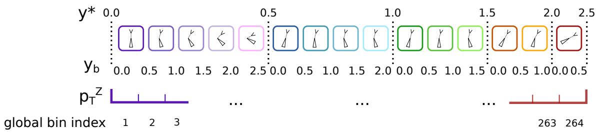


Parenthesis - Bin unraveling

3D phase-space → 1D visualization



→



From M. Schnepf 2022 [2]

Unfolding Basics

- Unfolding for detector effects of observation y to true spectrum x
  - Detector resolution $\longrightarrow$Migration between generator and reconstruction bins
  - Detector efficiency $\longrightarrow$Less events on reconstruction level than generator level

Statistical fluctuations $\tilde{y}$ and $\tilde{x}$ of true spectrum

Unfolding Basics

- Unfolding for detector effects of observation y to true spectrum x
  - Detector resolution $\rightarrow$Migration between generator and reconstruction bins
  - Detector efficiency $\rightarrow$Less events on reconstruction level than generator level
- Usually discretized observations and predictions in histograms $\rightarrow$ “invert” response matrix A (i.e. `TUnfold`)
  - Ill-conditioned matrix $\rightarrow$ Regularize “unphysical” high-frequency oscillations
  - Estimate response matrix from MC $\rightarrow$ Systematic and statistical uncertainties

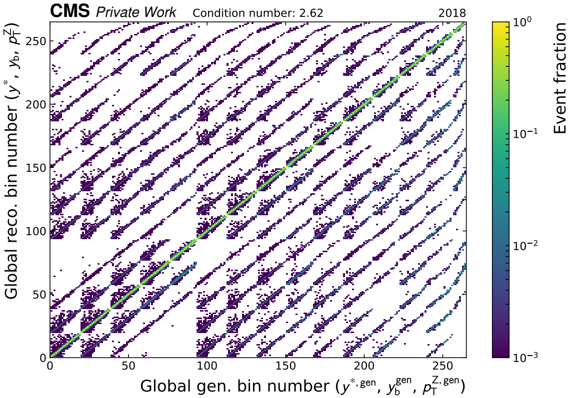
Statistical fluctuations $\tilde{y}$ and $\tilde{x}$ of true spectrum

We have: $\tilde{y}_i = \sum_i A_{ij}\tilde{x}_j + b_i$

We want: $x_j \rightarrow$ `TUnfold`

Response Matrix

2018

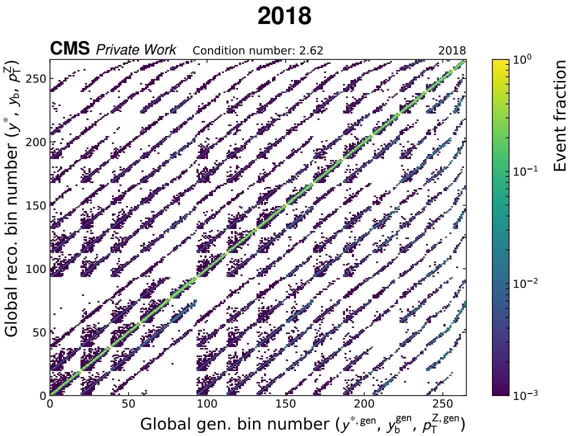


- Low condition number < 10

$\rightarrow$ Regularization not necessary

- Stat. uncertainty on data propagated through unfolding
- Unfolding uncertainties from limited MC precision propagated internally by `TUnfold`
- Systematic uncertainty propagated separately by new unfolding for each variation

Response Matrix

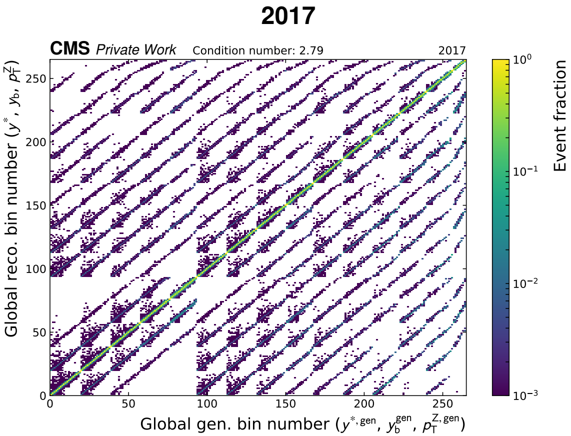


- Low condition number < 10

$\rightarrow$ Regularization not necessary

- Stat. uncertainty on data propagated through unfolding
- Unfolding uncertainties from limited MC precision propagated internally by `TUnfold`
- Systematic uncertainty propagated separately by new unfolding for each variation
- Similar for all data periods (2018)

Response Matrix



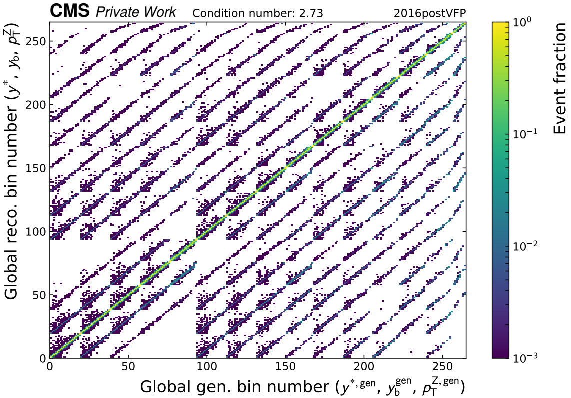
- Low condition number < 10

$\rightarrow$ Regularization not necessary

- Stat. uncertainty on data propagated through unfolding
- Unfolding uncertainties from limited MC precision propagated internally by `TUnfold`
- Systematic uncertainty propagated separately by new unfolding for each variation
- Similar for all data periods (2017)

Response Matrix

2016postVFP



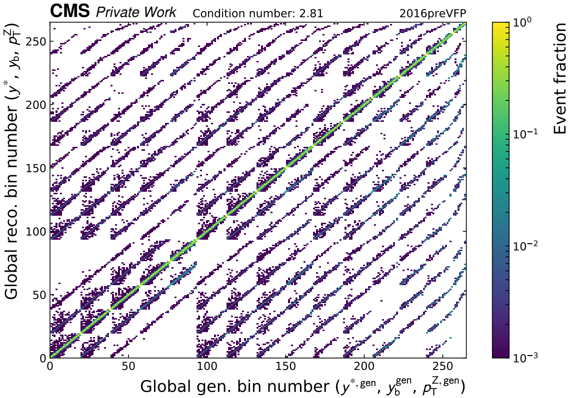
- Low condition number < 10

$\rightarrow$ Regularization not necessary

- Stat. uncertainty on data propagated through unfolding
- Unfolding uncertainties from limited MC precision propagated internally by `TUnfold`
- Systematic uncertainty propagated separately by new unfolding for each variation
- Similar for all data periods (2016postVFP)

Response Matrix

2016preVFP



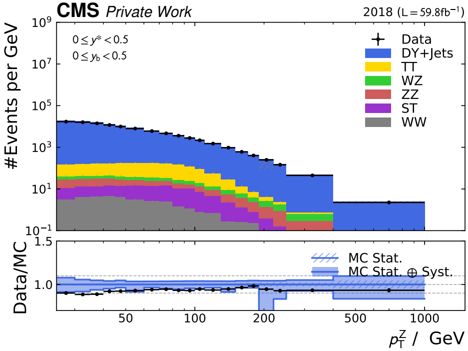
- Low condition number < 10

$\rightarrow$ Regularization not necessary

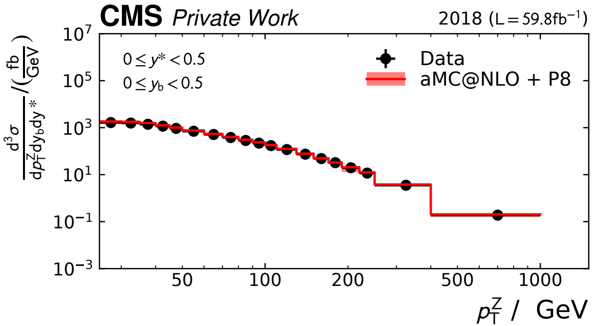
- Stat. uncertainty on data propagated through unfolding
- Unfolding uncertainties from limited MC precision propagated internally by `TUnfold`
- Systematic uncertainty propagated separately by new unfolding for each variation
- Similar for all data periods (2016preVFP)

Unfolded Result

Detector Level



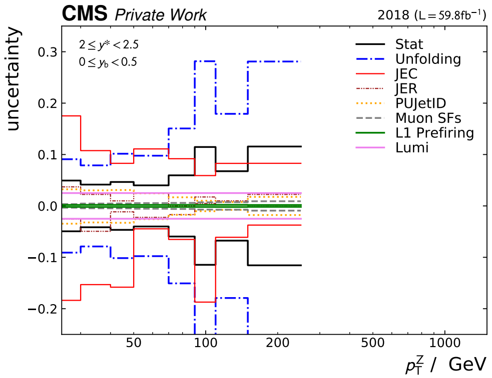
Unfolded



Unfolding - Statistical Uncertainty

- Derived through uncertainty propagation in the `TUnfold` method
- Systematic uncertainty through limited number of events in MC sample
- DYJetsToLL signal MC $\sim$3.9M events after selection

→ High statistical unfolding uncertainty (pseudo MC generation?)



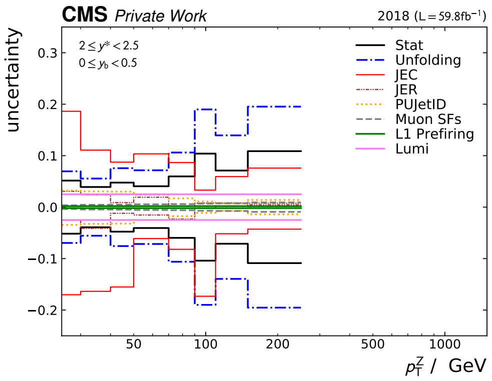
Unfolding - Statistical Uncertainty

- Derived through uncertainty propagation in the `TUnfold` method
- Systematic uncertainty through limited number of events in MC sample
- DYJetsToLL signal MC $\sim$3.9M events after selection

→ High statistical unfolding uncertainty (pseudo MC generation?)

- Add jet-binned "high-stat." samples DYJetsToLL_0J, DYJetsToLL_1J, DYJetsToLL_2J (0.2M + 4.7M + 2.7M events after selection)

→ Significant improvement in statistical unfolding uncertainty (no pseudo generation needed)

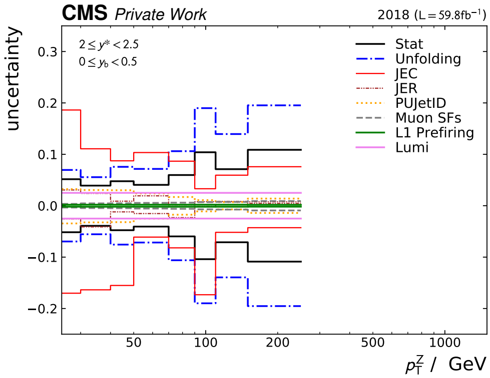


Systematic Uncertainties

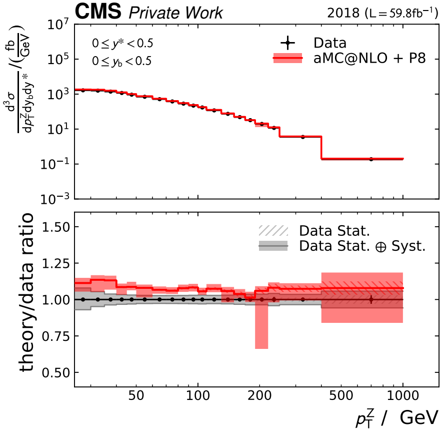
- Various systematic uncertainty sources, e.g. luminosity, JEC, trigger & muon scale factors, . . .
- Uncertainty propagation by creating new response matrices for each uncertainty variation and repeat unfolding

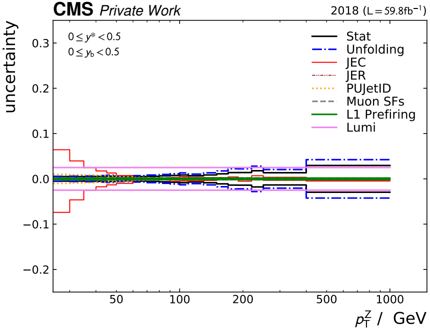
→ JEC dominant in low, statistical uncertainty in high p_{T}^{Z}-region

- **Crucial:** Waiting for SHERPA DY-MC for estimation of modelling bias!



Unfolded Cross-Sections 2018: Central Region

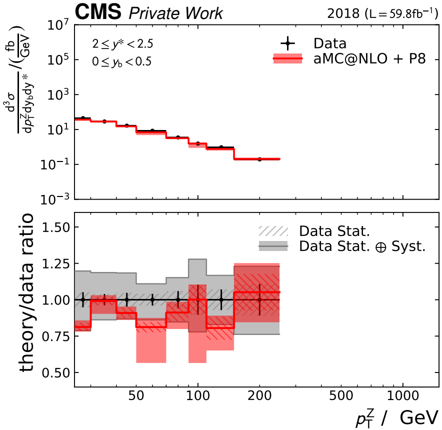




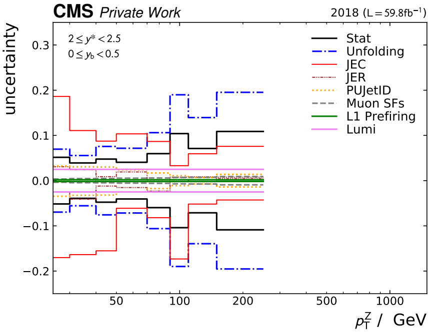
- Gen level at $k_{\mathrm{NNLO}} \times$NLO (0,1,2 jets)
- Inclusive XS too high, shape okay

- JEC uncertainty dominant at low p_T^Z
- Stat. + unfolding uncertainty dominant at high p_T^Z

Unfolded Cross-Sections 2018: Forward-Backward

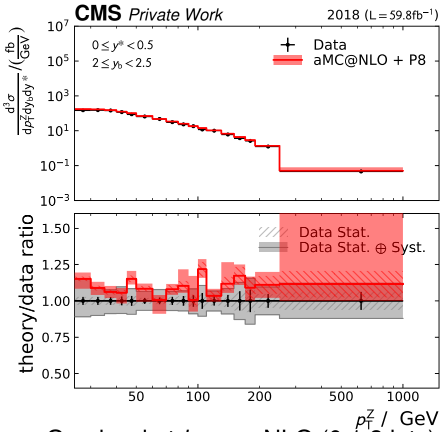


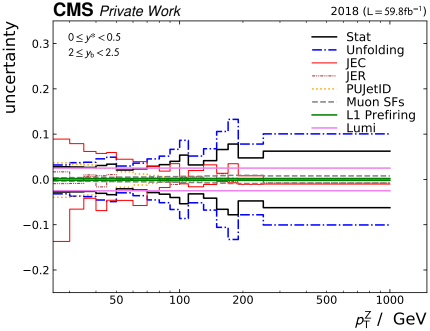
- Gen level at $k_{NNLO} \times$NLO (0,1,2 jets)
- No significant deviation, XS slightly too low



- JEC uncertainty dominant at low p_T^Z, stat. + unfolding at high p_T^Z
- High uncertainties: high η, low stats

Unfolded Cross-Sections 2018: High Boost



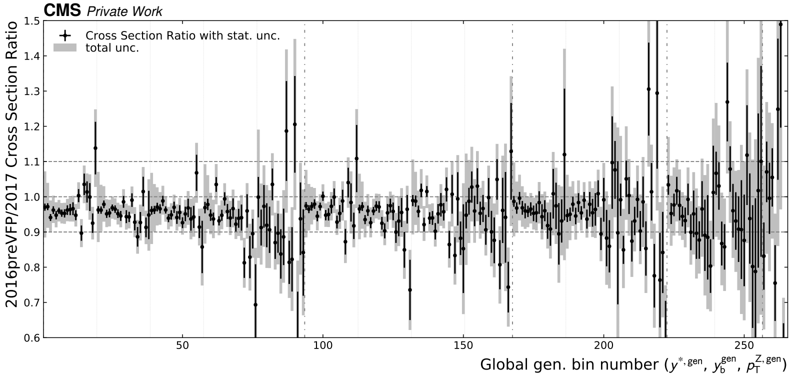


- Gen level at $k_{\mathrm{NNLO}} \times$NLO (0,1,2 jets)
- No significant deviation, XS slightly too high
- JEC uncertainty dominant at low p_{T}^{Z}, stat. + unfolding at high p_{T}^{Z}
- Higher uncertainties: high η

Compatibility between Years

- Overall cross-section for 2017 data $\sim 5.0 \pm 2.4\,(\text{Lumi.})\,\%$ significantly higher than for 2016preVFP data

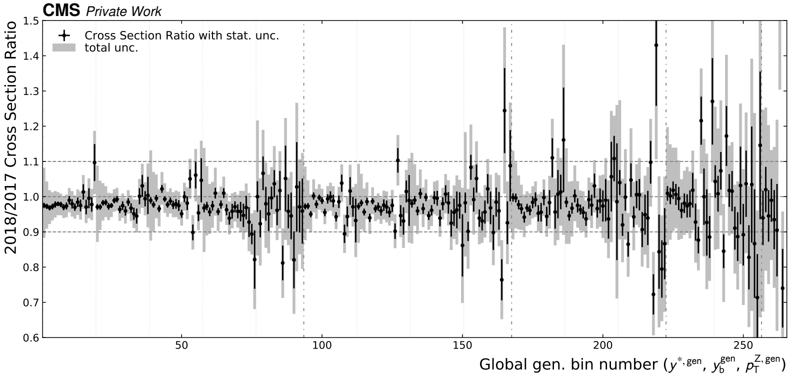
2016preVFP/2017 UL



Compatibility between Years

- Overall cross-section for 2017 data $\sim 5.0 \pm 2.4\,(\text{Lumi.})\,\%$ significantly higher than for 2016preVFP data
- Discrepancy between 2017 and 2018 data $(\sim 2.0 \pm 2.8\,(\text{Lumi.})\,\%)$ insignificant
- **Similar observations in independent analyses** (Z-Counting, Brussels)

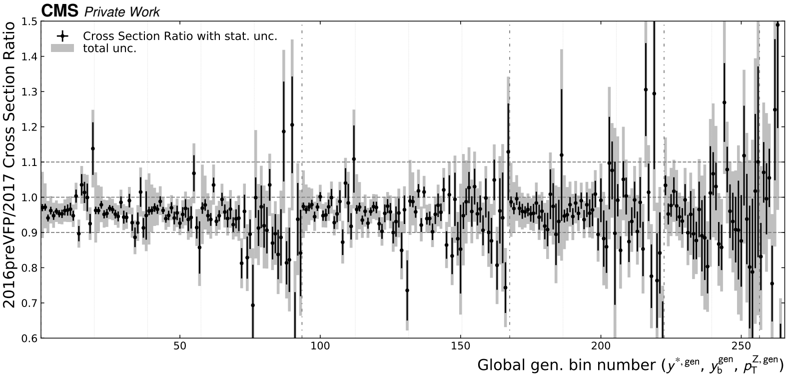
2018/2017 UL



Compatibility of 2016 Data

- Overall cross-section for 2017 data $\sim 5.0 \pm 2.3\,(\text{Lumi.})\,\%$ significantly higher than for 2016preVFP data

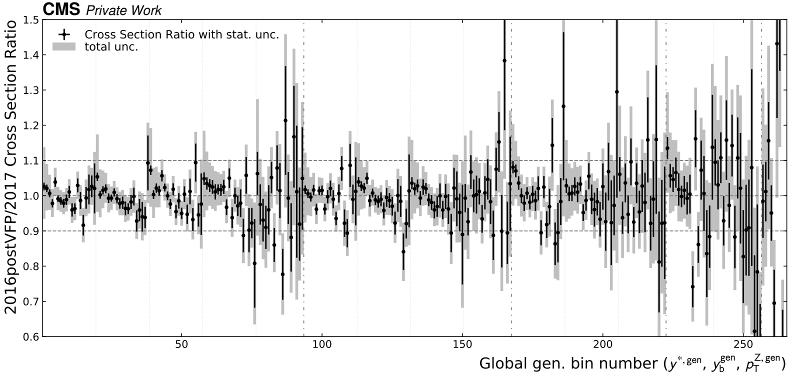
2016preVFP/2017 UL



Compatibility of 2016 Data

- Overall cross-section for 2017 data $\sim 5.0 \pm 2.3\,(\text{Lumi.})\,\%$ significantly higher than for 2016preVFP data
- Less discrepancy in normalization of 2016postVFP compared to 2017 data
- Weird p_T^Z-dependent trend in 2016postVFP data (compared to 2017)

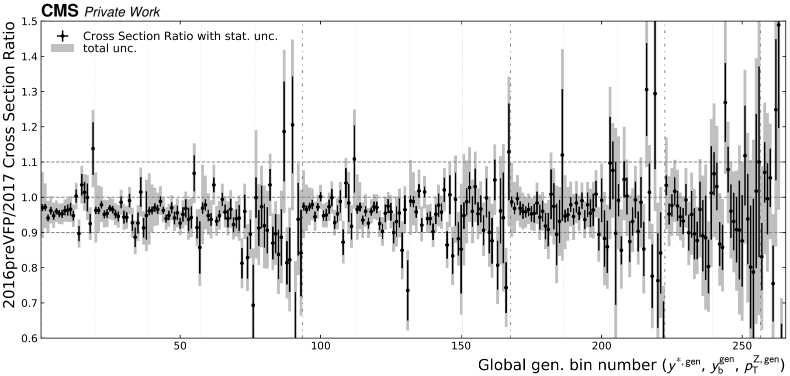
2016postVFP/2017 UL



Compatibility of 2016 Data

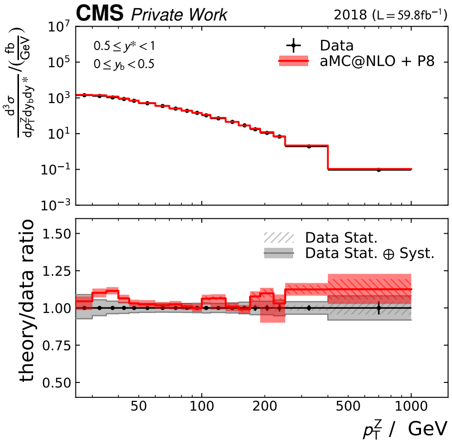
- Overall cross-section for 2017 data $\sim 5.0 \pm 2.3\,(\text{Lumi.})\,\%$ significantly higher than for 2016preVFP data
- Less discrepancy in normalization of 2016postVFP compared to 2017 data
- Weird p_T^Z-dependent trend in 2016postVFP data (compared to 2017)
  - Saw-tooth pattern less pronounced in 2016preVFP data

2016preVFP/2017 UL



Conclusions and Outlook

- First 3D Z+Jet cross section measurement of full Run II data presented
- Software framework updated to UL and validated with ULB
- Discrepancies of 2016 data confirmed and made CMS Collaboration aware of it
- By now also seen in independent analyses:
  - DY in a wide mass range by Brussels (ULB)
  - Z counting by LumiPOG
- Aiming for **publication by the end of 2023**
- I'm excited to stay at ETP for my PhD starting in January!



Backup

Datasets

- Data 2016preVFP: `/SingleMuon/Run2016[B-ver1,B-ver2,C-F]_HIPM_UL2016_MiniAODv2-v2/MINIAOD`
- Data 2016postVFP: `/SingleMuon/Run2016[F-H]_UL2016_MiniAODv2-v2/MINIAOD`
- Data 2017: `/SingleMuon/Run2017[B-F]-UL2017_MiniAODv2-v1/MINIAOD`
- Data 2018: `/SingleMuon/Run2018[A-D]-UL2018_MiniAODv2-v[2,3]/MINIAOD`
- MC
  - `DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8`
  - `TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8`
  - `ST_t-channel_(anti)?top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8`
  - `ST_tW_(anti)?top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8`
  - `[WW,WZ,ZZ]_TuneCP5_13TeV-pythia8`
- Global Tags
  - Data: `106X_dataRun2_v35`
  - MC 2016preVFP: `106X_mcRun2_asymptotic_preVFP_v11`
  - MC 2016postVFP: `106X_mcRun2_asymptotic_v17`
  - MC 2017: `106X_mc2017_realistic_v9`
  - MC 2018: `106X_upgrade2018_realistic_v16_L1v1`

Detailed Event Selection

One $Z \rightarrow \mu\mu$ 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

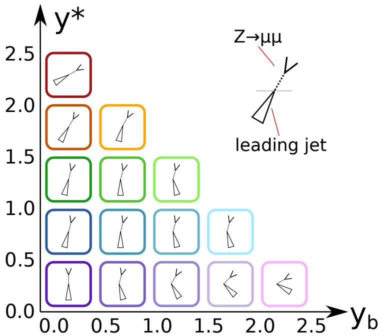
Selection	Value
Jet ID	Tight + Lepveto
PUJetID	Tight
$\Delta R(\mu_Z, \mathrm{Jet})$	> 0.4
Jet p_T	> 20 GeV
Jet $\lvert y\rvert$	< 2.4
Jet Veto Maps	✓

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

Analysis Overview

- Measurement of the jet associated Z-boson production cross-section differentially in **three obervables** p_T^Z, y_b, y^* for full Run 2 data
- $Z(\to \mu\mu) + \mathrm{Jet}$ analysis for 2016 and 2017 with preliminary data (T. Berger 2019 [1] and M. Schnepf 2022 [2])

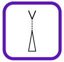


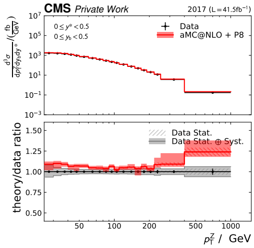
M. Schnepf 2022 [2]

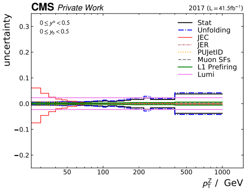
Updates Compared to pre-UL Analysis

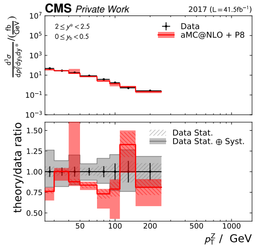
- Updated code/inputs to UL (IDs, SFs, corrections, ...)
- Extensive code review of all modules
  - $\rightarrow$ Found and fixed some minor bugs, no significant effects
- Framework cross check with Brussels with 2018 data $\rightarrow$ Conclude agreement withn numerical uncertainties
- Updated home-brewed unfolding and uncertainty handling to CMS (UL-)recommendations
  - $\rightarrow$ Preparation for the paper
- Verified higher event count per lumi in 2017 for UL data, as previously observed
  - Talk at the SMP Meeting during CMS week (28.06.2022) `https://indico.cern.ch/event/1171502/`
  - Similar results by Lumi POG
  - Confirmed by Brussels group ($Z \rightarrow ee$ results still pending)
  - However, effect almost within uncertainties $\rightarrow$ No show stopper for the publication

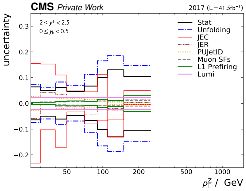
Unfolded Cross-Sections 2017

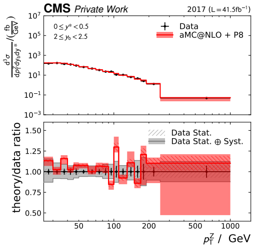


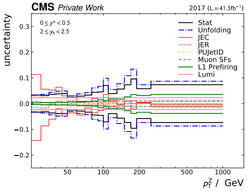




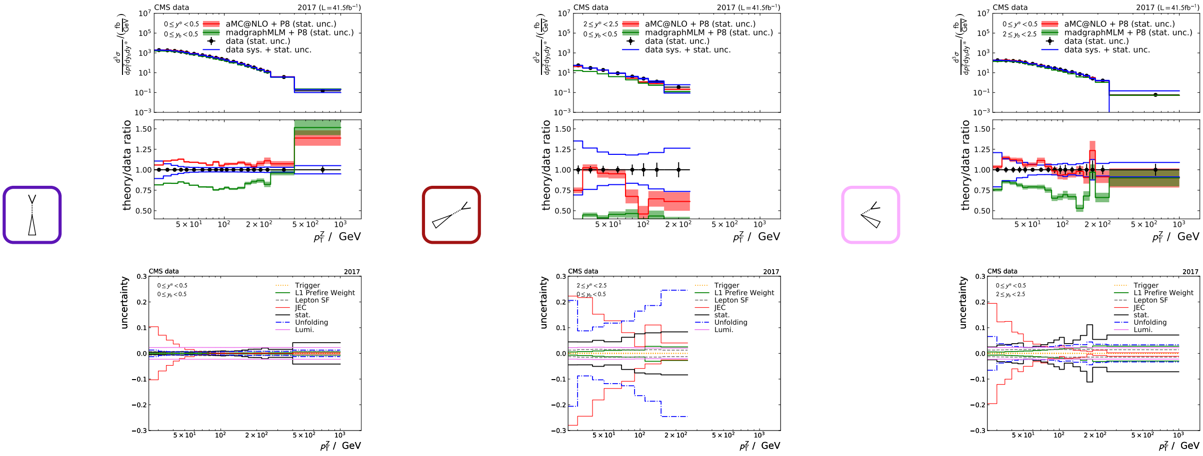






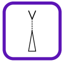


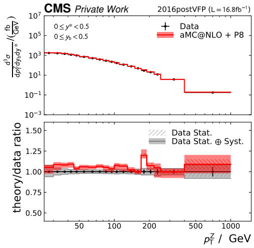
Unfolded Cross-Sections Legacy EOY 2017

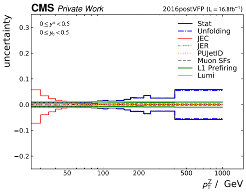


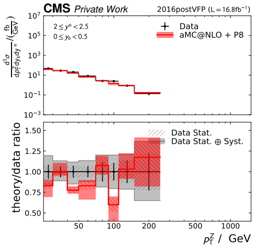
Karlsruhe Institute of Technology

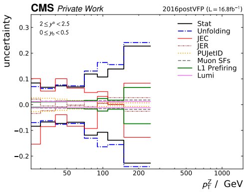
Unfolded Cross-Sections 2016postVFP

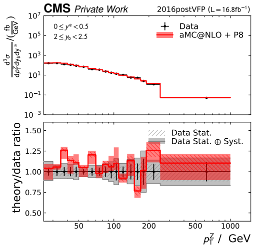


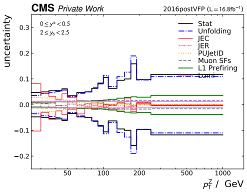




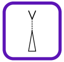


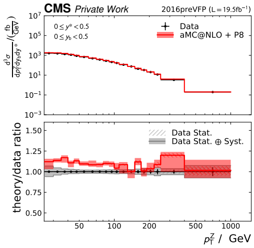


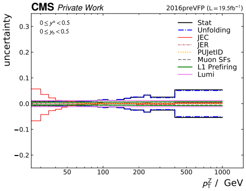


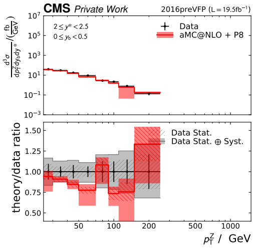


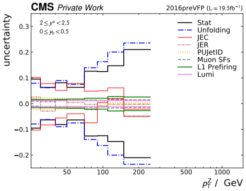
Unfolded Cross-Sections 2016preVFP

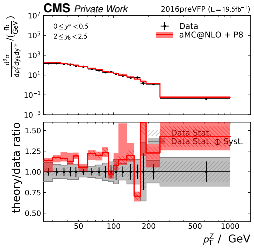


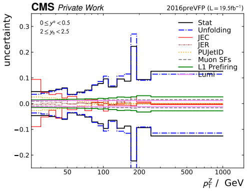




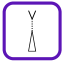


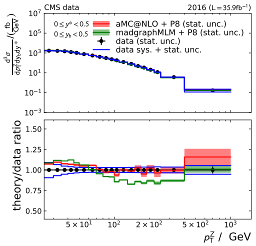


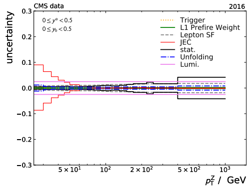


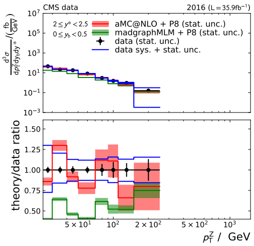


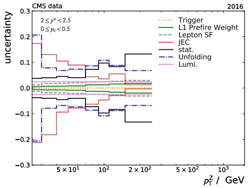
Unfolded Cross-Sections Legacy EOY 2016

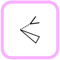


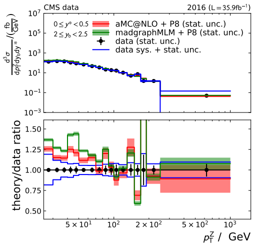


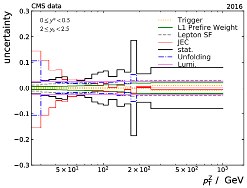








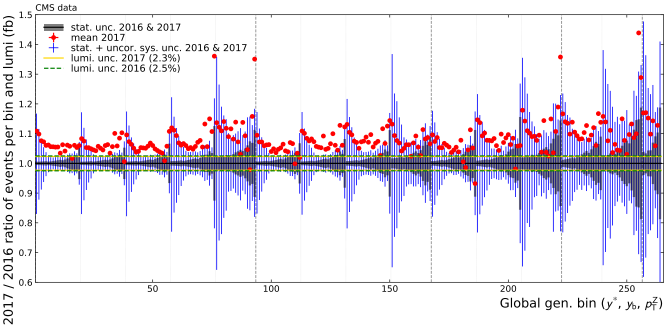




Previous Results - Reco level

- Comparison of Run II 2016 and 2017 end-of-year data at reconstruction level

- Expectation:
  2016 and 2017 data yield same cross-sections within uncertainties
- Observation:
  Systematic shift in 2017 data towards higher cross-sections

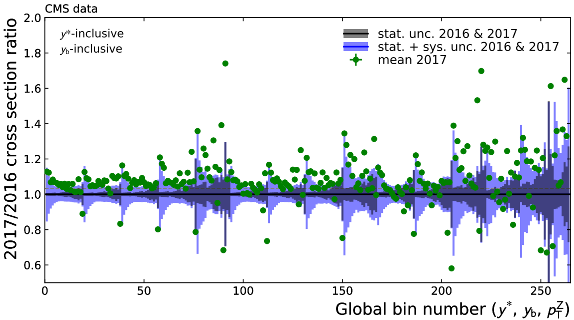


Taken from Matthias Schnepf [2]
first presented in SMP V+Jet Meeting 23.07.2021

→ If effect is understood in MC, unfolded cross-sections expected to be clean

Previous Results - 2017/2016 Unfolded

- Comparison of Run II 2016 and 2017 end-of-year unfolded data



From Matthias Schnepf
first presented in SMP V+Jet Meeting 23.07.2021

- Expectation:
  Same cross-section for 2016 and 2017 within uncertainties
- Observation:
  Systematic shift in 2017 data towards higher cross-section

→ Detailed Ultra-Legacy reevaluation of full Run 2 data (→ this Thesis)

References

[1] Thomas Berger. “Jet energy calibration and triple differential inclusive cross section measurements with Z (→μμ) + 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.

[2] 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.