Optimization of the reconstruction of missing transversal momentum in the new software of the ATLAS experiment

Johannes Wünsche

Supervisors: Michael Kobel, Abhishek Nag, Joany Manjarres, Carsten Bittrich

26 June 2019

1. Introduction

1. Introduction

- consideration of vector boson scattering processes
- vector boson scattering at LHC:
  - quarks emit $W^{\pm}/Z$ bosons
  - interaction via triple or quartic boson vertices
  - outgoing vector bosons decay leptonically or hadronically

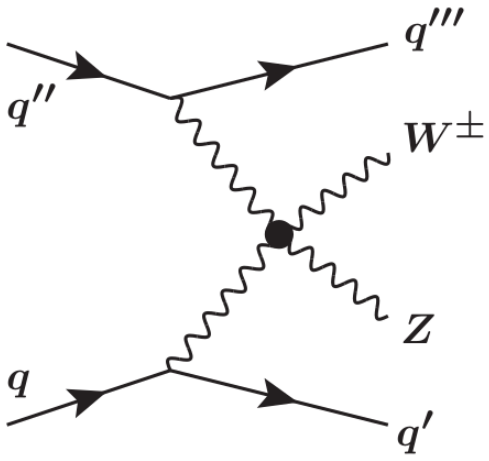
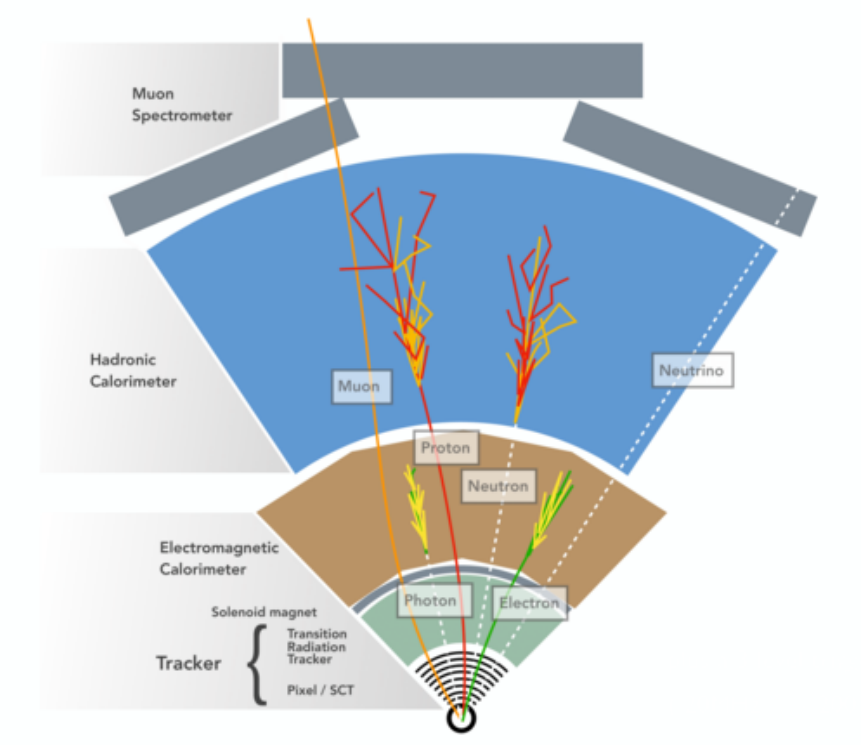


Figure 1: $WZ \to WZ$ scattering process as example for vector boson scattering (extracted from [1])

1. Introduction

- ATLAS experiment used to investigate such processes



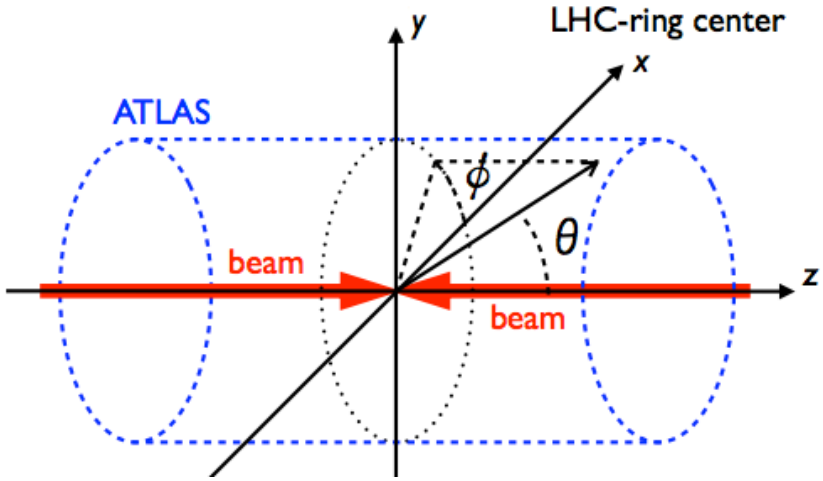


Figure 2: Visualization of inner structure of ATLAS detector (left side, taken from [2]) and coordinate system to describe positions and directions (right side, taken from [3]), pseudorapidity: $\eta = -\ln\tan(\theta/2)$

1. Introduction

- consideration of leptonical decay of $W^{\pm}$: $W^{\pm} \rightarrow l^{\pm} + \nu$
- neutrino carries information about process
- undetectable $\rightarrow$ no conclusions by direct observation
- momentum conservation $\rightarrow$ total momentum in transverse direction is zero
- neutrino missing in balance $\rightarrow$ missing transversal momentum

1. Introduction

- Missing transversal momentum (MET) = negative sum of transversal momentum vectors belonging to selected signals in detector (energy deposits and tracks)

- components: $$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = \vec{E}_{\mathrm{T,\,hard}}^{\mathrm{miss}} + \vec{E}_{\mathrm{T,\,soft}}^{\mathrm{miss}}$$

  - hard term: MET from reconstructed particles (electrons, muons, jets)
  - soft term: contributions from signals not associated with reconstructed particles (including tau leptons and photons)

2. MET Reconstruction

2.1. Reconstuction algorithm

- used MET reconstruction part of ELCore
- ELCore based on ATLAS software Athena
  - consideration of given detector signals
  - reconstruction and calibration of objects
  - calculation of quantities
- additional usable variations of MET reconstruction in Release 21 (latest version)

2.2. Possible Variations

- choice between EMTopo and EMPFlow jet reconstruction
  - EMTopo: usage of information about topology of energy deposits in calorimeter
  - EMPFlow: additional track information from inner detector
- 4 different jet selection working points: Loose, Tight, Tighter, Tenacious

2.2. Possible Variations

- jet selection working points for EMTopo:

Working point	p_T (all jets)	p_T (fw. jets)	JVT ($\|\eta\| < 2.4$)
Loose	> 20GeV	> 20GeV	> 0.59 (for $p_T < 60$GeV)
Tight	> 20GeV	> 30GeV	> 0.59 (for $p_T < 60$GeV)
Tighter	> 20GeV	> 35GeV	> 0.59 (for $p_T < 60$GeV)
Tenacious	> 20GeV	> 35GeV	> 0.91 (for $p_T < 40$GeV), > 0.59 (for $p_T < 60$GeV), > 0.11 (for $p_T < 120$GeV)

- for EMPFlow: only difference is JVT > 0.2 for working points Loose, Tight, Tighter
- application of forward jet vertex tagging (fJVT) algorithm possible: used to remove pile-up jets in forward-direction
  - two working points: Medium (cut: fJVT $= 0.4$), Tight (cut: fJVT $= 0.5$)

3. Practical Part

3.1. Samples

- for practical part: MET reconstruction applied to Monte Carlo Simulation of process $WZjj \rightarrow l\nu lljj$
- sample details:
  - DSID: 364284
  - Generator: Sherpa 2.2.2
  - PDF: NNLO NNPDF30
- considered subcampaigns: MC15 (2015+2016 data taking, Rel. 20.7), MC16a (2015+2016 data taking, Rel. 21), MC16d (2017 data taking, Rel. 21)

3.1. Samples

- differences:
  - different number of events
  - pile-up occurence ($N_{\text{Pile-up}}$ = number of pile-up interactions per bunch crossing, averaged over lumi block)

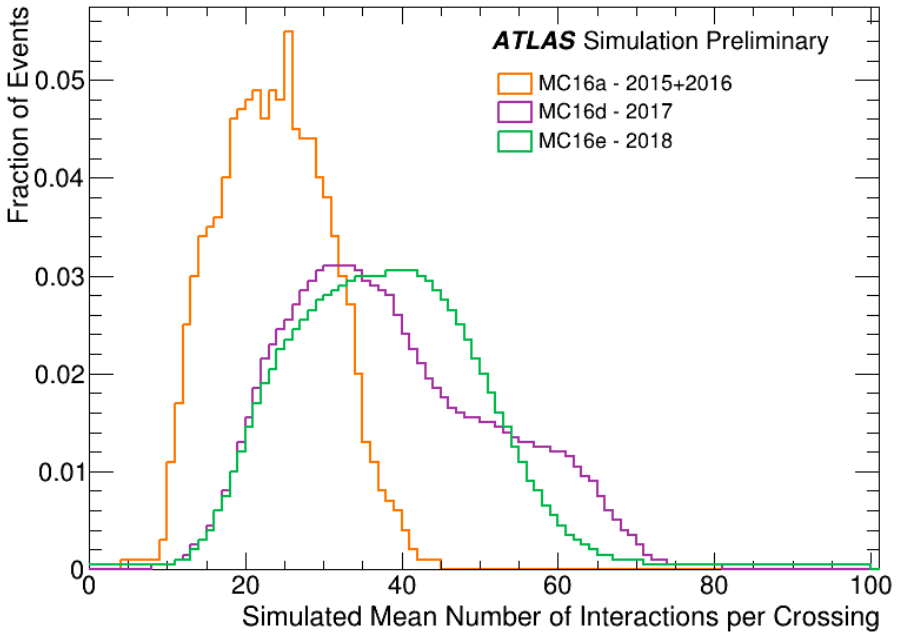


Figure 3: Occurrence of different average numbers of pileup vertices per bunch crossing for the MC16a and MC16d data (taken from [4])

3.2. Methodology

- MET studies based on consideration of Monte Carlo simulations
- properties of simulated particles are known (called truth)
- from this information truth MET can be calculated
- application of MET algorithm yields reconstructed value
- evaluation by comparison of truth and reconstructed value $\rightarrow$ analysis of distribution of difference values

3.3. MC16d Sample

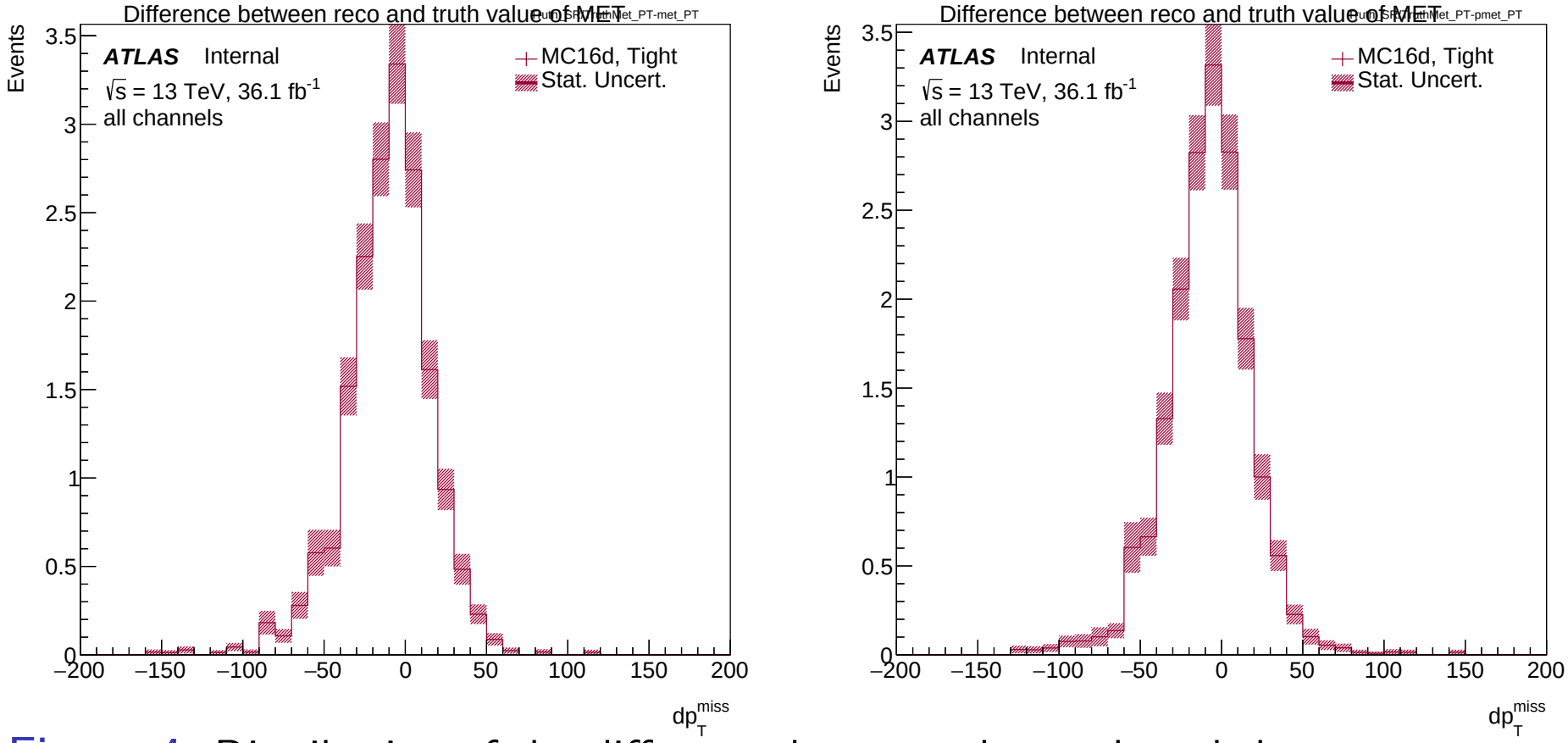


Figure 4: Distribution of the difference between the truth and the reconstructed value of the MET for the jet selction working point Tight without fJVT, used with EMTopo (left side) and EMPFlow (right side) reconstruction (MC16d)

3.3. MC16d Sample

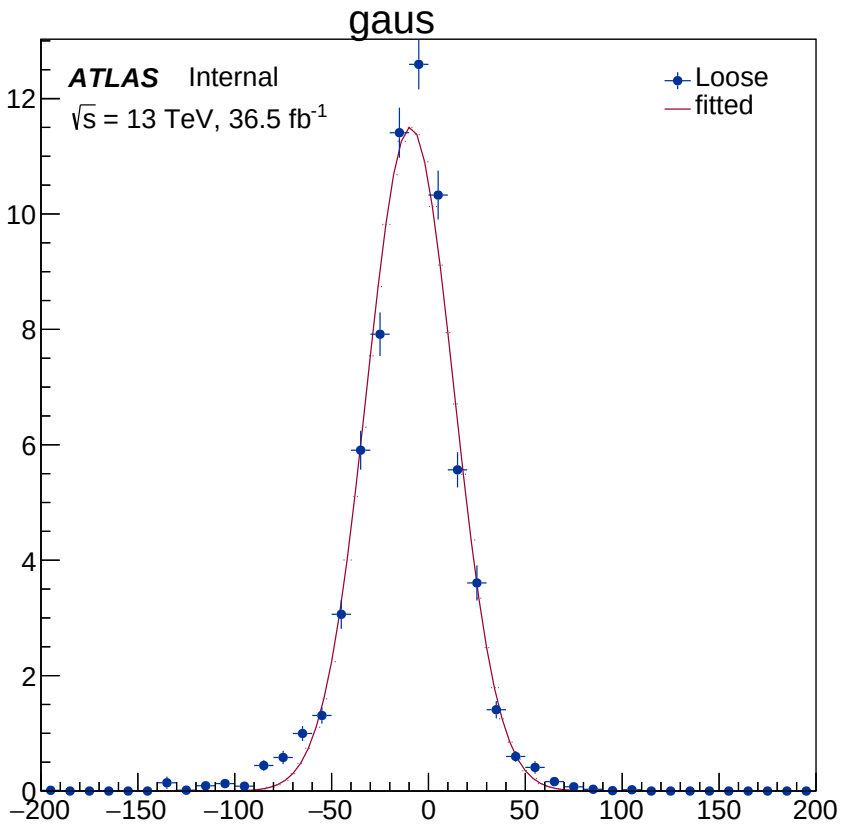


Figure 5: Gaussian fit applied to distribution of difference between truth and reconstructed value of MET for jet selection working point Loose without fJVT (sample MC16d)

3.3. MC16d Sample

- Fitting results (Gaussian fitting) for different working points:

Working points		EMTopo		EMPFlow	
Jet selection	fJVT	Res. / GeV	Shift / GeV	Res. / GeV	Shift / GeV
Loose	-	22.5 ± 0.3	-9.3 ± 0.4	22.8 ± 0.4	-7.9 ± 0.4
Tight	-	22.2 ± 0.4	-8.6 ± 0.4	21.6 ± 0.4	-7.0 ± 0.4
Tighter	-	22.5 ± 0.4	-9.1 ± 0.4	21.6 ± 0.4	-7.1 ± 0.4
Tenacious	-	22.5 ± 0.4	-9.1 ± 0.4	21.5 ± 0.4	-7.1 ± 0.4
Tight	Tight	22.1 ± 0.4	-8.4 ± 0.4	21.0 ± 0.4	-6.6 ± 0.4
Tighter	Tight	22.2 ± 0.4	-8.6 ± 0.4	21.6 ± 0.4	-7.0 ± 0.4
Tighter	Medium	22.5 ± 0.4	-9.0 ± 0.4	21.4 ± 0.4	-7.0 ± 0.4
Tenacious	Medium	22.2 ± 0.4	-8.7 ± 0.4	-	-

3.4. MC15 and MC16a Samples

Crystal Ball function fitting:

- attempt to find more suitable function for distributions of difference between truth and reco value of MET
- Crystal Ball function seems to be most promising one
- definition (with $t = \frac{x - x_0}{\sigma}$):

$$f(x) = C \cdot e^{-\frac{t^2}{2}} \qquad \text{for} \quad \alpha_{\text{low}} \leq x \leq \alpha_{\text{high}}$$

$$f(x) = C \cdot e^{\frac{\alpha_{\text{low}}^2}{2} - \alpha_{\text{low}} t^2} \qquad \text{for} \quad x \leq \alpha_{\text{low}}$$

$$f(x) = C \cdot e^{\frac{\alpha_{\text{high}}^2}{2} - \alpha_{\text{high}} t^2} \qquad \text{for} \quad \alpha_{\text{high}} \leq x$$

3.4. MC15 and MC16a Sample

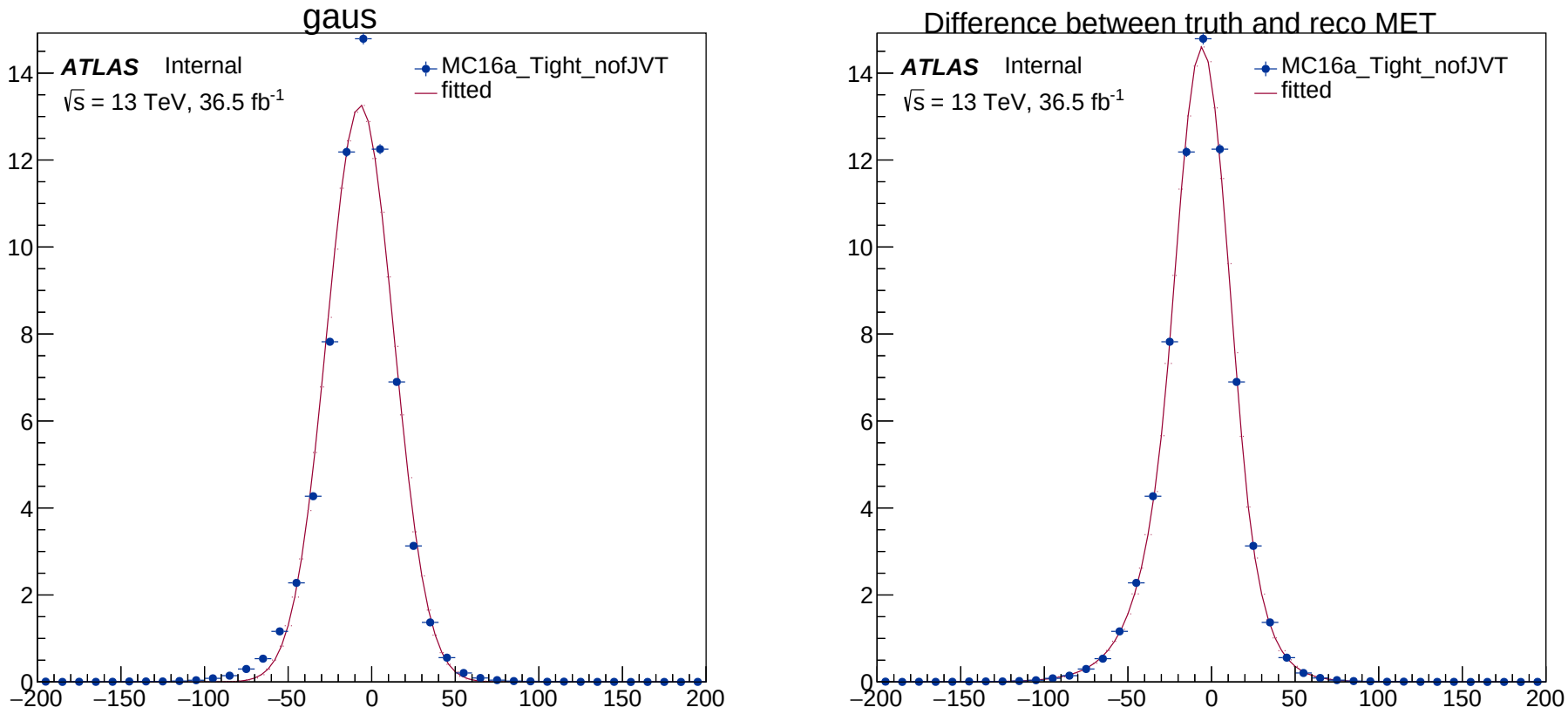


Figure 6: Distribution of difference between truth and reconstructed value of MET for working point Tight (EMTopo, no fJVT) with Gaussian fit (left side) and Crystal Ball function fit (right side) applied

3.4. MC15 and MC16a Samples

- Crystal Ball function gives better fit:
  - Gaussian fitting: $\chi^2/N_{Df} = 48.4$
  - Crystal Ball function fitting: $\chi^2/N_{Df} = 2.6$

- Fitting results (shift and MET resolution) for different working points:

Sample	Jet selection	fJVT	Shift / GeV	Res. / GeV	χ^2/N_{Df}
MC15	Tight (EMTopo)	-	-5.58 ± 0.10	16.60 ± 0.17	2.68
MC16a	Loose (EMTopo)	-	-5.81 ± 0.09	17.12 ± 0.15	3.06
MC16a	Loose (EMPFlow)	-	-5.63 ± 0.10	17.11 ± 0.15	2.20
MC16a	Tight (EMTopo)	-	-5.73 ± 0.10	17.11 ± 0.16	2.61
MC16a	Tight (EMPFlow)	-	-5.62 ± 0.10	17.03 ± 0.15	2.31
MC16a	Tenacious (EMTopo)	-	-5.91 ± 0.10	17.46 ± 0.16	2.79
MC16a	Tenacious (EMPFlow)	-	-5.86 ± 0.10	17.40 ± 0.16	2.33
MC16a	Tight (EMTopo)	Tight	-5.71 ± 0.10	17.10 ± 0.16	2.21
MC16a	Tight (EMPFlow)	Tight	-5.60 ± 0.10	16.98 ± 0.16	2.34

3.4. MC15 and MC16a Samples

- Fitting results (α_{low} and α_{high}) for different working points:

Sample	Jet selection	fJVT	α_{low}	α_{high}	χ^2/N_{Df}
MC15	Tight (EMTopo)	-	-1.098 ± 0.018	1.413 ± 0.028	2.68
MC16a	Loose (EMTopo)	-	-1.102 ± 0.016	1.460 ± 0.029	3.06
MC16a	Loose (EMPFlow)	-	-1.050 ± 0.015	1.378 ± 0.026	2.20
MC16a	Tight (EMTopo)	-	-1.089 ± 0.016	1.441 ± 0.030	2.61
MC16a	Tight (EMPFlow)	-	-1.035 ± 0.015	1.351 ± 0.025	2.31
MC16a	Tenacious (EMTopo)	-	-1.097 ± 0.016	1.447 ± 0.030	2.79
MC16a	Tenacious (EMPFlow)	-	-1.032 ± 0.015	1.357 ± 0.025	2.33
MC16a	Tight (EMTopo)	Tight	-1.089 ± 0.016	1.444 ± 0.029	2.21
MC16a	Tight (EMPFlow)	Tight	-1.037 ± 0.016	1.346 ± 0.025	2.34

3.5. Pile-up Dependence

- MET resolution value expected to increase with growing pile-up influence
- pile-up dependence investigated by use of sample MC16a (and MC15 as reference) and Crystal Ball function fits
- data seperated into following pile-up bins:

Index of pile-up bin	Range
1	$0 \leq N_{\text{Pileup}} \leq 17$
2	$18 \leq N_{\text{Pileup}} \leq 24$
3	$25 \leq N_{\text{Pileup}} \leq 31$
4	$32 \leq N_{\text{Pileup}}$

3.5. Pile-up Dependence

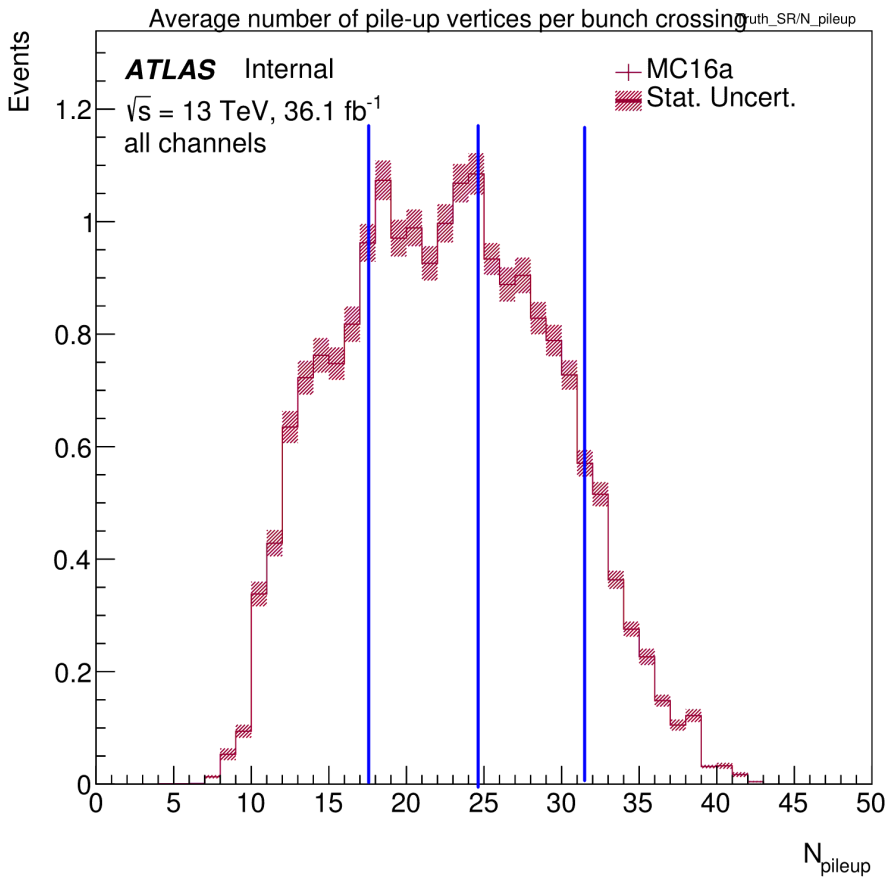


Figure 7: Visualization of the pile-up bin division of the data contained in sample MC16a: the three vertical blue lines symbolize the limits of the bins

3.5. Pile-up Dependence

- results from two releases are compatible
- MET resolution value increases indeed with growing pile-up

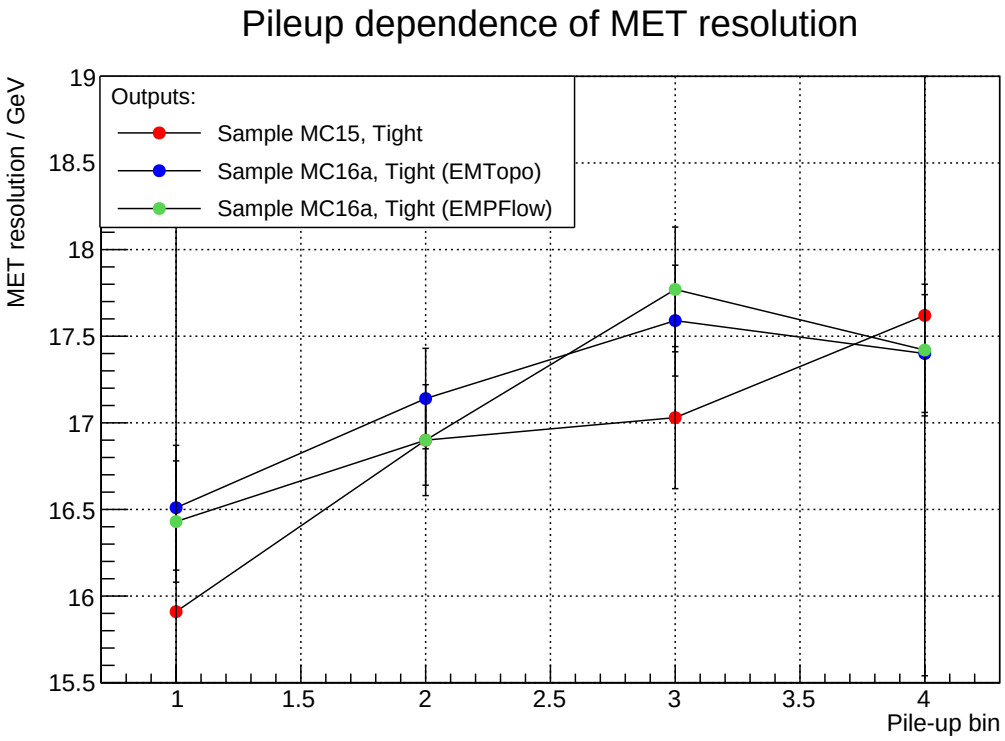


Figure 8: MET resolutions for different PU bins obtained by use of Tight for MC15 sample and MC16a sample (EMTopo/EMPFlow)

3.5. Pile-up Dependence

- less tight jet selection working points seem to be more appropriate for low pile-up

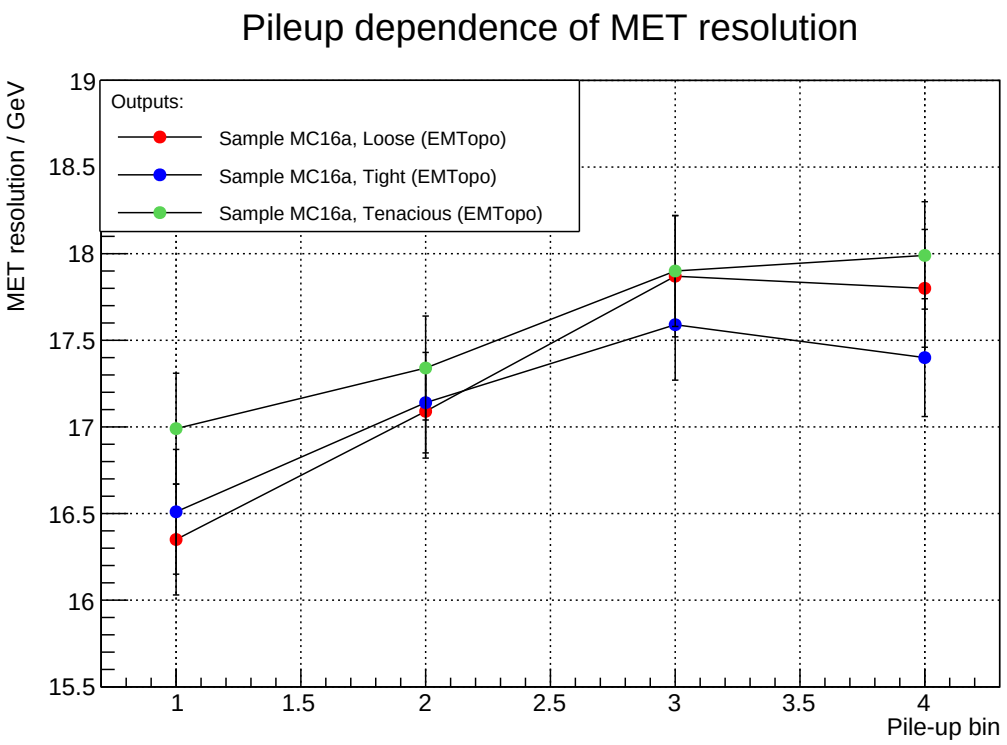


Figure 9: MET resolutions for different PU bins obtained by use of Loose, Tight and Tenacious with MC16a sample (EMTopo)

3.5. Pile-up Dependence

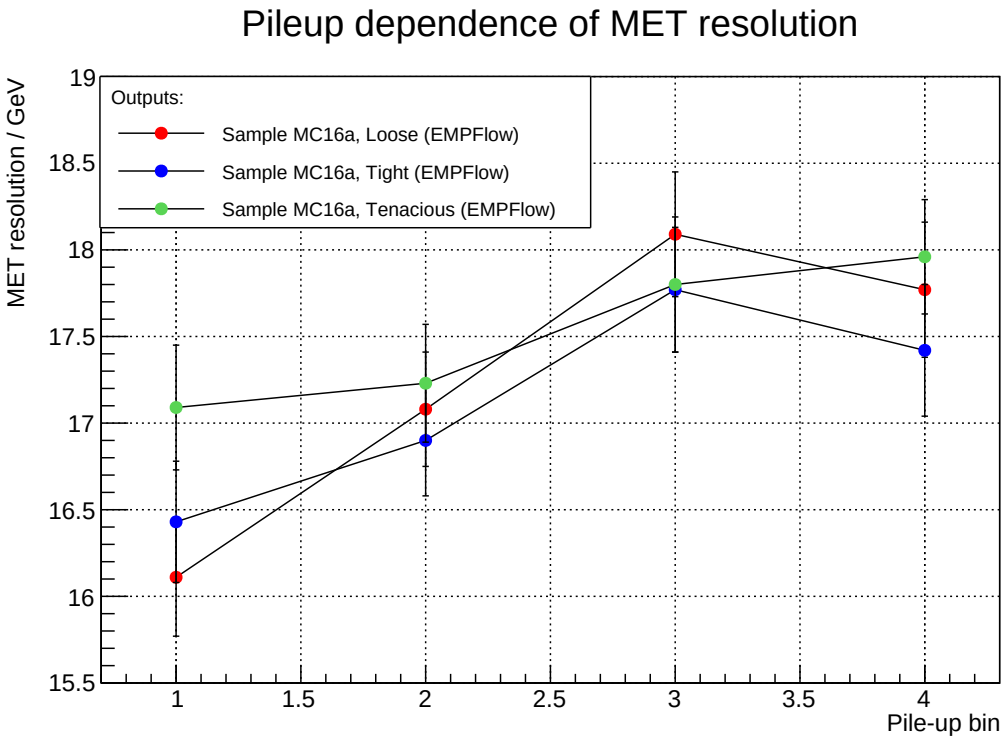


Figure 10: MET resolutions for different PU bins obtained by use of Loose, Tight and Tenacious with MC16a sample (EMPFlow)

4. Summary/Outlook

- distribution of difference between truth and reconstructed MET: Crystal Ball function found as better fitting function than Gaussian distribution
- results obtained with Release 21 are compatible with results obtained with Release 20
- fitting results indicate strong similarities between variations, only minor influence of application of EMPFlow jet reconstruction or fJVT algorithm
- pile-up dependence also very similar for different variations
- no variation can be identified as most appropriate one
- possible future investigations:
  - regard larger amount of data to reduce relative statistical uncertainties
  - considering different phase spaces, different decay modes

Bibliography

[1] ATLAS Collaboration, *Study of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*,.

[2] "Event cross section in a computer generated image of the atlas detector." https://cds.cern.ch/record/1096081.

[3] J. M. Ramos, *WZ diboson measurements with the ATLAS experiment at the LHC and Performance of resistive Micromegas in view of HL-LHC applications*.
PhD thesis, Université Paris Diderot, 2013.

[4] "Introduction: General production and performance public results of the atlas data preparation group." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/DataPrepGenPublicResults.
Accessed: 2019-06-04.

[5] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC pp collisions*,.

[6] ATLAS Collaboration, M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D96** no. 7, (2017) 072002, arXiv:1703.09665 [hep-ex].

[7] ATLAS Collaboration, G. Aad et al., *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. **C76** no. 11, (2016) 581, arXiv:1510.03823 [hep-ex].

[8] ATLAS Collaboration, M. Aaboud et al., *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. **C77** no. 9, (2017) 580, arXiv:1705.02211 [hep-ex]. [Erratum: Eur. Phys. J.C77,no.10,712(2017)].

[9] ATLAS Collaboration, M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. **C77** no. 7, (2017) 466, arXiv:1703.10485 [hep-ex].

[10] ATLAS Collaboration, G. Aad et al., *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector*, Eur. Phys. J. **C77** no. 4, (2017) 241, arXiv:1609.09324 [hep-ex].

BACK-UP

Jet Vertex Tagging (JVT)

- JVT discriminant for jet i and regarded primary vertex j:

$$\mathrm{JVT}_i(\mathrm{jet}_i, \mathrm{PV}_j) = \frac{\sum_m p_T(\mathrm{track}_m^{\mathrm{jet}_i}, \mathrm{PV}_j)}{\sum_n \sum_l p_T(\mathrm{track}_l^{\mathrm{jet}_i}, \mathrm{PV}_n)}$$

  - value for probability that jet is originated in regarded vertex
  - used to remove pile-up jets

- modified version used with multivariate methods in Release 21

Foward Jet Vertex Tagging (fJVT)

- for pile-up vertex i MET reconstructed on basis of central region detector signals
  - should be associated with one jet (JVT < 0.14)
- forward jet should be QCD pile-up jet: $\Delta R_{pT} > 0.2$
- fJVT discriminant for pile-up vertex i:

$$\mathrm{fJVT}_i = \frac{\vec{p}_{\mathrm{T,i}}^{\,\mathrm{miss}} \cdot \vec{p}_{\mathrm{T}}^{\,\mathrm{fj}}}{|\vec{p}_{\mathrm{T}}^{\,\mathrm{fj}}|^2}$$

  - normalized projection of the MET on transversal momentum of forward jet
  - value for probability that jet is originated in that vertex

Foward Jet Vertex Tagging (fJVT)

- phase space definitions (table from [1]):

Variable	Total	Fiducial inclusive	Fiducial VBS
Lepton $\lvert\eta\rvert$	—	< 2.5	< 2.5
p_T of ℓ_Z, p_T of ℓ_W [GeV]	—	$> 15, > 20$	$> 15, > 20$
m_Z range [GeV]	$66 - 116$	$\lvert m_Z - m_Z^{PDG}\rvert < 10$	$\lvert m_Z - m_Z^{PDG}\rvert < 10$
m_T^W [GeV]	—	> 30	> 30
$\Delta R(\ell_Z^-, \ell_Z^+)$, $\Delta R(\ell_Z, \ell_W)$	—	$> 0.2, > 0.3$	$> 0.2, > 0.3$
p_T two leading jets [GeV]	—	—	> 40
$\lvert\eta_j\rvert$ two leading jets	—	—	< 4.5
Jet multiplicity	—	—	≥ 2
$\eta_{j1} \cdot \eta_{j1}$	—	—	< 0
m_{jj} [GeV]	—	—	> 500
$\Delta R(j, \ell)$	—	—	> 0.3
$N_{b-\text{quark}}$	—	—	$= 0$