Reconstructing and Calibrating Hadronic Objects with ML Algorithms in ATLAS

ML4Jets 2023

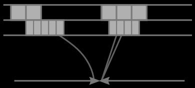
Tobias Fitschen on behalf of the ATLAS Collaboration

06 November 2023

University of Manchester

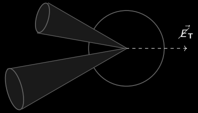
European Research Council
Established by the European Commission

Overview / Publications



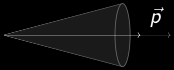
Calorimeter Clusters

07/2020: ML for Pion Identificaton and Energy Calibration
08/2022: Point Cloud DL Methods for Pion Reconstruction
08/2023: Cluster Calibration with ML Techniques
→ Peter Lochs' talk (tomorrow)



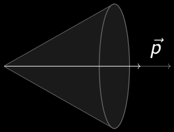
Missing E_T

07/2021: METNet: Combined $\not{E}_{\mathrm{T}}$ Working Point with ML



R=0.4 Jet Calibration

03/2023: New Techniques for Jet Calibration (GNNC)



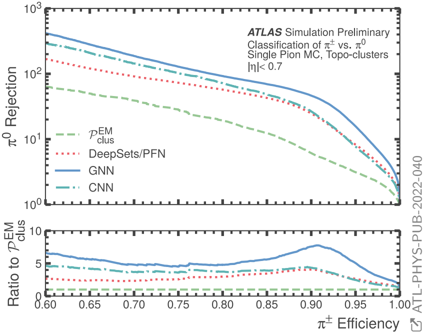
R=1.0 Jet Calibration

New @ ML4Jets: Simultaneous large-R JES+JMS with ML
→ Main part of this talk

π^0 vs $\pi^\pm$ Shower Classification

First step in cluster calibration: Differentiate EM from hadronic clusters
Non-compensating ATLAS calorimeter requires different calibrations for neutral/charged clusters

π^0 vs $\pi^\pm$ classification performance



Point cloud of energy deposits in calorimeter cells

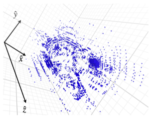
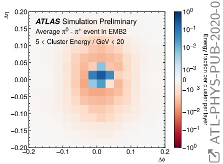


Image: $\pi^0 - \pi^\pm$ difference



Baseline cluster calibration $\mathcal{P}^{EM}_{clus}$:

- Relies on binned EM-scale cluster variables
  - Total cluster energy $E^{EM}_{cluster}$
  - Pseudorapidity η
  - Longitudinal depth λ_{clus}
  - 1st cell energy moment $\langle\rho_{cell}\rangle$
- Combined into likelihood $\mathcal{P}^{EM}_{clus}$

Individual calorimeter cell signals

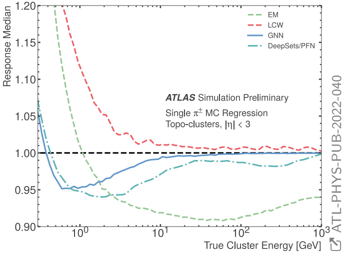
→ As point clouds (**GNN**, **PFN**)

→ Or projected on images (**CNN**)

Observations

- All point cloud methods significantly outperform baseline $\mathcal{P}^{EM}_{clus}$

$\pi^{\pm}$ cluster energy response

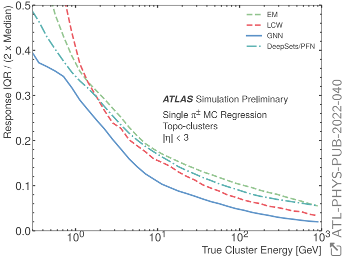


Second step: Energy Calibration Observations

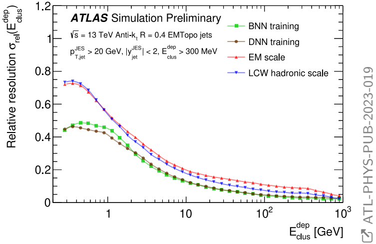
- **GNN** performs best wrt. response and width
- Followed by **Deep Sets**
- New: **Bayesian NN (BNN)**
  - → New results on this in Peter Lochs' talk

Cluster Energy Resolution

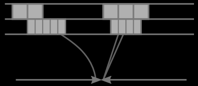
Point Cloud Methods



DNN / BNN

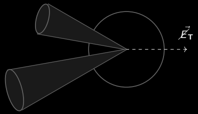


Overview / Publications



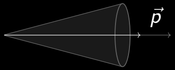
Calorimeter Clusters

07/2020: ML for Pion Identificaton and Energy Calibration
08/2022: Point Cloud DL Methods for Pion Reconstruction
08/2023: Cluster Calibration with ML Techniques
→ Peter Lochs' talk (tomorrow)



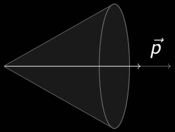
Missing E_T

07/2021: METNet: Combined $\cancel{E}_\mathrm{T}$ Working Point with ML



R=0.4 Jet Calibration

03/2023: New Techniques for Jet Calibration (GNNC)



R=1.0 Jet Calibration

New @ ML4Jets: Simultaneous large-R JES+JMS with ML
→ Main part of this talk

METNet: A combined $p_{\mathrm{T}}^{\mathrm{miss}}$ working point (ATL-PHYS-PUB-2021-025):

- $\not{E}_{\mathrm{T}}$ in ATLAS: Negative sum of calibrated momenta of hard objects (e, μ, τ-jets, γ, jets)
- Plus soft term: Tracks from PV not associate to hard objects
- Different working points (WPs) defined for various pileup conditions
  - E.g. "tight": Higher p_{T} cuts on forward jets
- MetNet: MLP combining $\not{E}_{\mathrm{T}}$ values from different WPs
  - Based on event kinematics and conditions

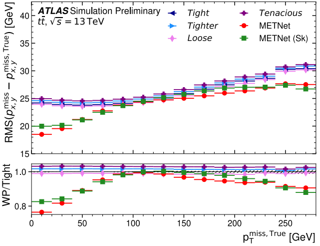
Working Points

Working point	p_{T} [GeV] for jets with: $\lvert\eta\rvert < 2.4$	$2.4 < \lvert\eta\rvert < 4.5$	JVT for jets with $\lvert\eta\rvert < 2.4$	fJVT for jets with $2.5 < \lvert\eta\rvert < 4.5$ and $p_{\mathrm{T}} < 120$ GeV
Loose	> 20	> 20	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	-
Tight	> 20	> 30	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	< 0.4
Tighter	> 20	> 35	> 0.5 for $p_{\mathrm{T}} < 60$ GeV jets	-
Tenacious	> 20	> 35	> 0.91 for $20 < p_{\mathrm{T}} < 40$ GeV jets > 0.59 for $40 < p_{\mathrm{T}} < 60$ GeV jets > 0.11 for $60 < p_{\mathrm{T}} < 120$ GeV jets	< 0.5

ATL-PHYS-PUB-2021-025

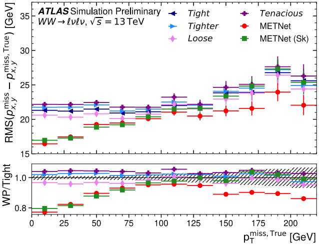
MetNet

Trained among others on $t\bar{t}$



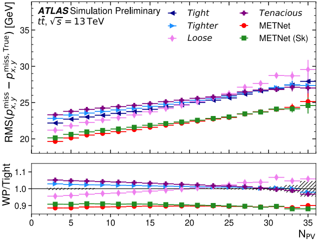
ATL-PHYS-PUB-2021-025

Extrapolates well to $WW \to \ell\nu\ell\nu$



ATL-PHYS-PUB-2021-025

Across different pileup regimes

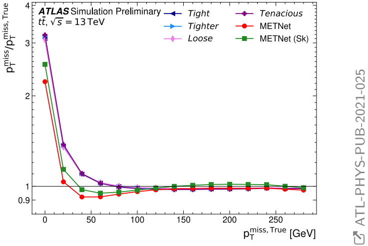


ATL-PHYS-PUB-2021-025

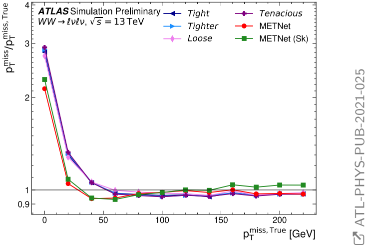
MetNet

- $\not{E}_T$ definition depends on process but MetNet performs best for all
- Significantly better RMS than any WP alone

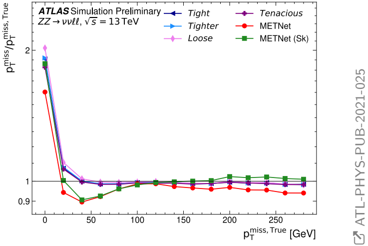
Trained among others on $t\bar{t}$



Extrapolates well to $WW \to \ell\nu\ell\nu$



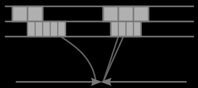
$ZZ \to \nu\nu\ell\ell$



MetNet

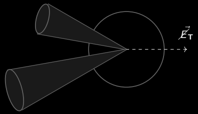
- $\not{E}_T$ definition depends on process but MetNet performs best for all
- Significantly better RMS than any WP alone
- MET Scale generally good but skewed towards 0 for low $\not{E}_T$

Overview / Publications



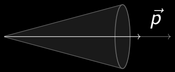
Calorimeter Clusters

07/2020: ML for Pion Identificaton and Energy Calibration
08/2022: Point Cloud DL Methods for Pion Reconstruction
08/2023: Cluster Calibration with ML Techniques
→ Peter Lochs' talk (tomorrow)



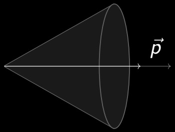
Missing E_T

07/2021: METNet: Combined $\not{E}_T$ Working Point with ML



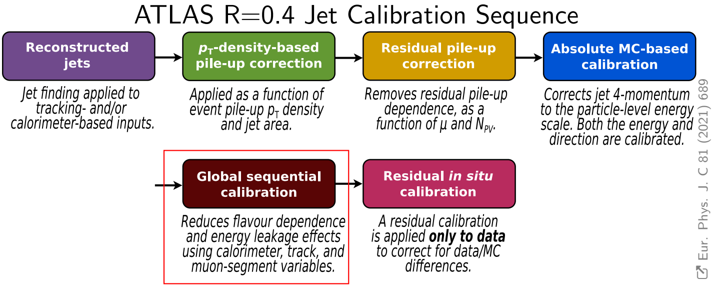
R=0.4 Jet Calibration

03/2023: New Techniques for Jet Calibration (GNNC)

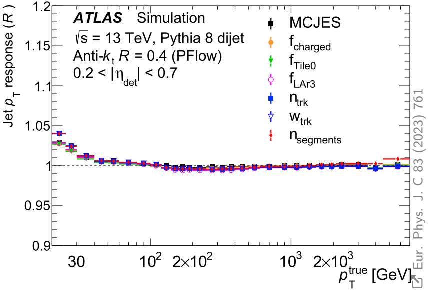


R=1.0 Jet Calibration

New @ ML4Jets: Simultaneous large-R JES+JMS with ML
→ Main part of this talk



p_T response after each GSC step

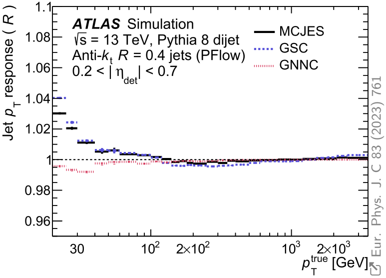


Global Sequential Calibration

- After energy scale calibrated on average, **GSC** corrects for small differences
- E.g. for different jet flavours
- **Sequentially** corrects for each variable

Jet Calibration: GNNC

Response

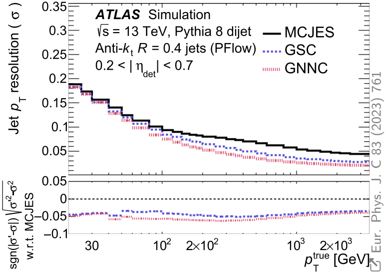


Global NN Calibration (GNNC)

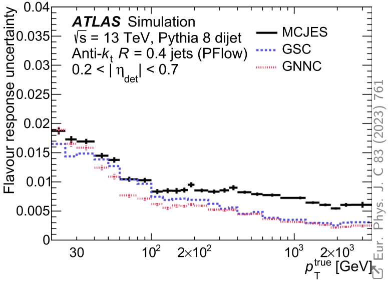
- **GSC** Does not exploit correlations of variables
- New method (**GNNC**) uses MLP trained to predict p_T response

→ Improvement over full p_T range

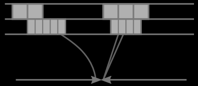
Resolution



Flavour Uncertainty

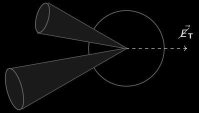


Overview / Publications



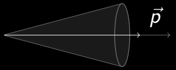
Calorimeter Clusters

07/2020: ML for Pion Identificaton and Energy Calibration
08/2022: Point Cloud DL Methods for Pion Reconstruction
08/2023: Cluster Calibration with ML Techniques
→ Peter Lochs' talk (tomorrow)



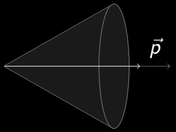
Missing E_{T}

07/2021: METNet: Combined $\not{E}_{\mathrm{T}}$ Working Point with ML



R=0.4 Jet Calibration

03/2023: New Techniques for Jet Calibration (GNNC)

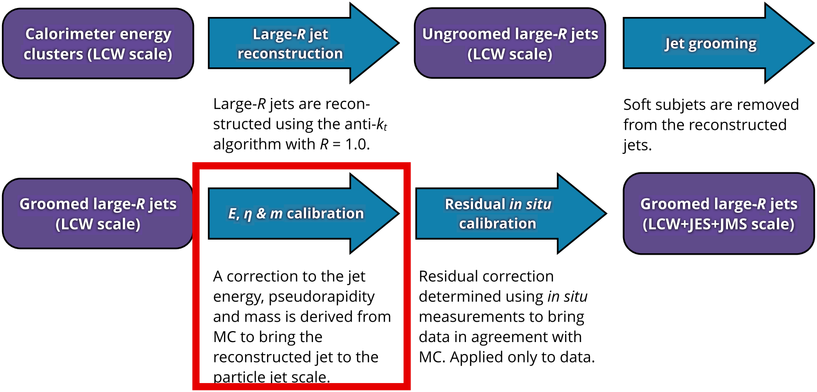


R=1.0 Jet Calibration

New @ ML4Jets: Simultaneous large-R JES+JMS with ML
→ Main part of this talk

ATLAS R=1.0 Jet Calibration

ATLAS R=1.0 Jet Calibration Sequence



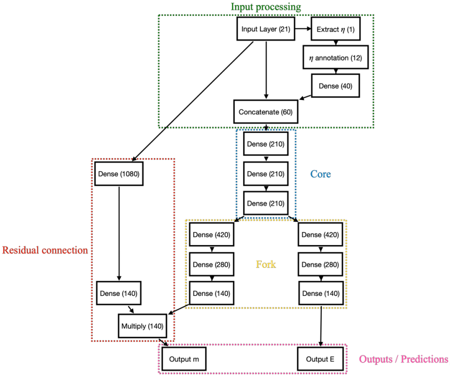
Eur. Phys. J. C(2019) 79:135

New @ ML4Jets: ML variant of **full** MC-based R=1.0 jet calibration

Ref. available soon

Simultaneous Calibration of Jet Energy and Mass using ML

Model Architecture



Ref. available soon

Method:

- Predict responses

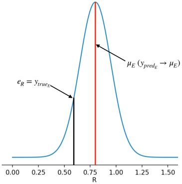
$$R_E = \frac{E_{\text{reco}}}{E_{\text{true}}},\ R_M = \frac{M_{\text{reco}}}{M_{\text{true}}}$$

- Modeled by Gaussians

$$y_{\text{pred}} = (\mu^E, \sigma^E, \mu^m, \sigma^m)_{\text{pred}}$$

$\Rightarrow$ Calibration Factors:

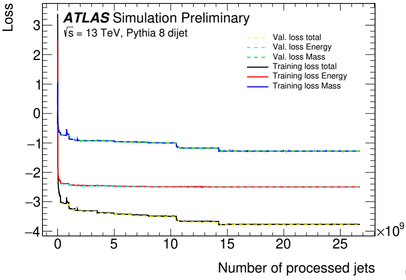
$$E_{\text{calib}} = \frac{E_{\text{reco}}}{\mu^E_{\text{pred}}},\ M_{\text{calib}} = \frac{M_{\text{reco}}}{\mu^M_{\text{pred}}}$$



Mixture density network (MDN) loss:

$$\mathcal{L}_{\text{MDN}} = -\log(P(y_{\text{true}}, y_{\text{pred}})) = \log(\sigma_{\text{pred}}) + \frac{1}{2}\frac{(y_{\text{true}} - \mu_{\text{pred}})^2}{\sigma^2_{\text{pred}}}$$

Training



Ref. available soon

Steps	N°	Number of epochs	Batch size	Loss
Initialisation	1	2	15000	MDNA
	2	2	25000	MDNA
	3	2	35000	MDNA truncated (4.0σ)
	4	2	15000	MDNA truncated (3.5σ)
Common training	5	6	95000	MDNA truncated (3.5σ)
	6	6	95000	MDNA truncated (3.5σ)
	7	6	125000	MDNA truncated (3.2σ)
	8	6	125000	MDNA truncated (3.2σ)
	9	10	155000	MDNA truncated (3.0σ)
	10	15	95000	MDNA truncated (E: 3.0σ, m : 2.0σ)
Exclusive mass training	11	50	95000	MDN truncated (1.0σ)

Training Strategy

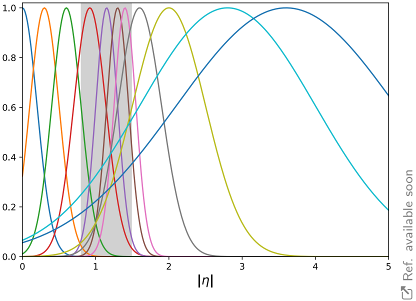
- Multi-stage training process
  - First E & M simultaneous
  - Then only M
- Alternative losses used in some training stages
  - To accommodate for asymmetric response:

Asymmetric MDN:

$$P_{\text{MDNA}}(x) = \begin{cases} 1e^{(x-\mu)^2/2\sigma_1} & \text{if} x < \mu \\ 1e^{(x-\mu)^2/2\sigma_2} & \text{if} x \geq \mu \end{cases}$$

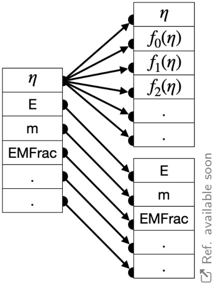
Truncated MDN:

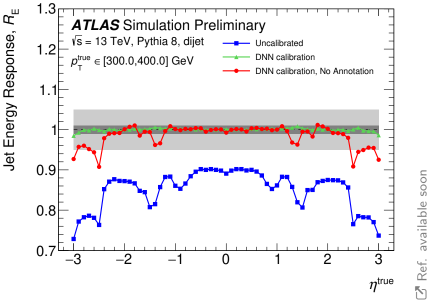
$$P_{\text{trunc}}(x) = \begin{cases} 1e^{(x-\mu)^2/2\sigma} & \text{if} |x < \mu| < N\sigma \\ 0 & \text{otherwise} \end{cases}$$



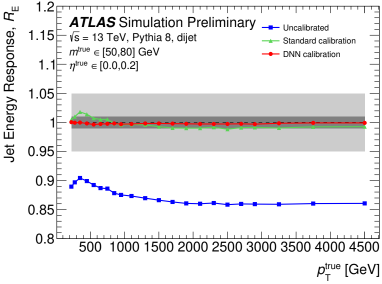
Complex dependence on η

- With sharp changes from bin-to-bin due to detector geometry/instrumentation
- Difficult for DNN to adapt to this
- Annotation strategy
  - $\rightarrow$ Add 12 features that are functions of η
  - $\rightarrow$ Encoding distance to different η regions
- Clear improvement:





Response: E

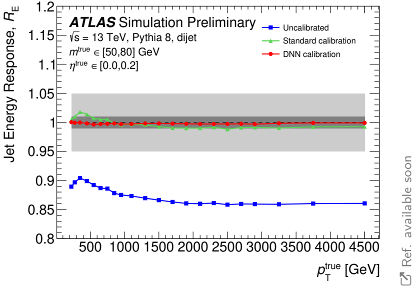


Ref. available soon

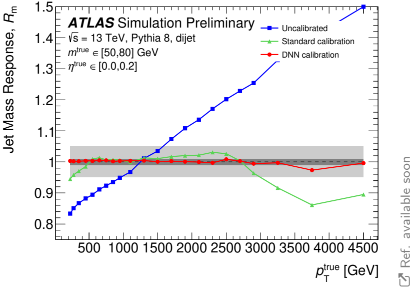
Improvement across the board

- **DNN**: better closure than **standard** calib. in response for E

Response: E



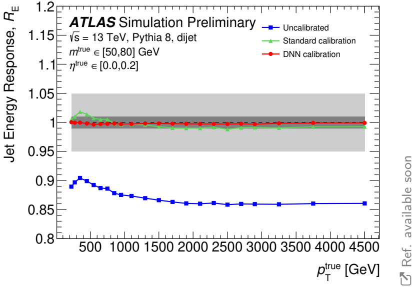
Response: M



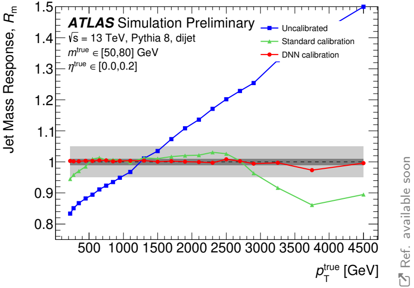
Improvement across the board

- **DNN**: better closure than **standard** calib. in response for E and M

Response: E



Response: M

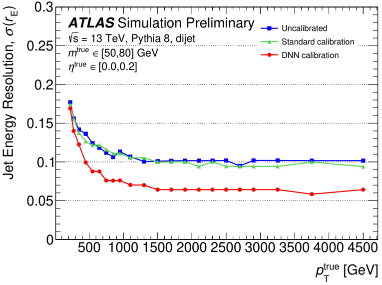


Improvement across the board

- **DNN**: better closure than **standard** calib. in response for E and M
- M response stable even in low and high p_T regime

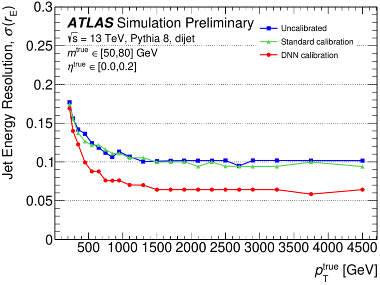
Results: Comparison to Standard Calibration

Resolution: E



Ref. available soon

Resolution: M



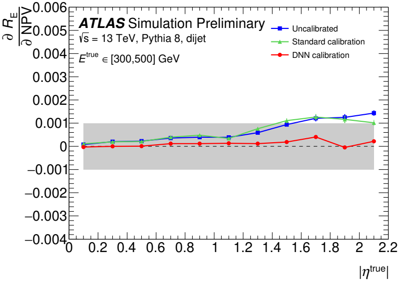
Ref. available soon

Improvement across the board

- **DNN**: better closure than **standard** calib. in response for E and M
- M response stable even in low and high p_T regime
- Resolution drastically improved

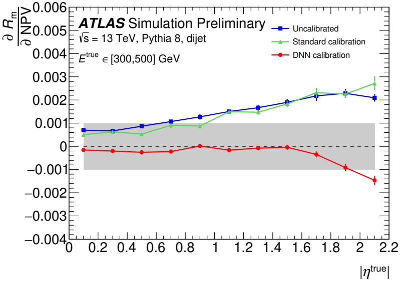
Results: Comparison to Standard Calibration

Pileup Stability: E



Ref. available soon

Pileup Stability: M



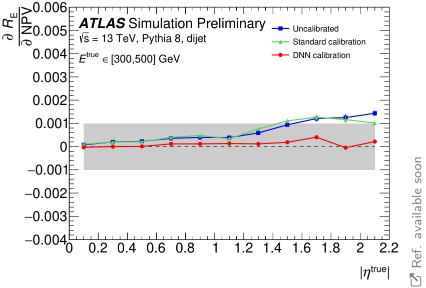
Ref. available soon

Improvement across the board

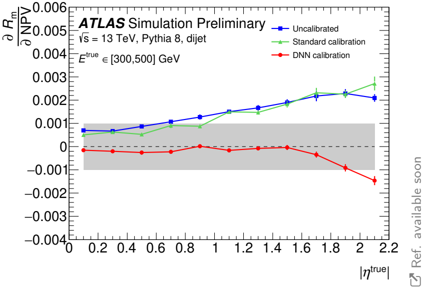
- More stable with respect to pileup

Results: Comparison to Standard Calibration

Pileup Stability: E



Pileup Stability: M



Improvement across the board

- More stable with respect to pileup
- Outperforms standard calibration in many more aspects
  - Modelling: Less dependent on MC generator
  - More stable across different η regions and processes (H, W/Z, top)
  - More consistent across different flavours (q/g)
  - See publication for dedicated studies on these aspects

Summary

Many ML applications for hadronic objects in ATLAS

- Calorimeter cluster classification and energy regression
  - $\rightarrow$ See Peter Loch's talk
- $\not{E}_{\mathrm{T}}$ calibration (MetNet)
- Jet energy scale calibration
  - R=0.4 jets: ML based GSC step (GNNC)
  - R=1.0 jets: Full MC-based calibration (E and M) by single DNN

ML based methods perform best in all domains

- **Important:** Better response & resolution in MC are great, but data/MC agreement & model independence should not be neglected!

Appendix

for new ML-based JES+JMS calibration

Ref. available soon

Samples & Features

Usage	Process type	Generator	Number of jets passing the selections (in millions)
training, validation	QCD dijet	Pythia 8.230	~ 270
validation	W/Z	Pythia 8.230	~ 20
validation	top	Pythia 8.230	~ 15
validation	Higgs	Pythia 8.230	~ 15
validation	QCD dijet	Sherpa 2.2.5	~ 30

Table 1: Simulated samples used for training and validation

Ref. available soon

	Name	Definition
Jet level	Log$E$	$\log(E_{\text{jet}})$ with E_{jet} (energy) in GeV
	Log$M$	$\log(m_{\text{jet}})$ with m_{jet} (mass) in GeV
	η	Jet pseudo-rapidity
Substructure level	groomMRatio	Mass ratio between groomed and ungroomed jets
	Width	$\sum_i p_{\text{T}i} \Delta R(i, \text{jet}) / (\sum_i p_{\text{T}i})$ where ΔR is the angular distance (sum over the jet constituents)
	Split12,Split23	Splitting scales at the 1st and 2nd exclusive k_T declusterings [35]
	C2, D2	Energy correlation ratios [36, 37]
	τ_{21}, τ_{32}	N-Subjettiness ratios using WTA axis [38, 39]
	Qw	Smallest invariant mass among the proto-jets pairs of the last 3 steps of a k_T reclustering sequence
Detector level	EMFrac	Energy fraction deposited in the electromagnetic calorimeter
	EM3Frac	Energy fraction deposited in the third layer of the electromagnetic calorimeter
	Tile0Frac	Energy fraction deposited in the 1st layer of the hadronic calorimeter
	EffNConsts	$(\sum_i E_i)^2 / (\sum_i E_i^2)$ (sum over the jet constituents)
	NeutralFrac	Energy fraction from neutral constituents
	ChargedPTFrac	p_{T} fraction from charged constituents
	ChargedMFrac	Mass fraction from charged constituents
Event level	μ	Mean number of interactions per bunch crossing
	NPV	Number of primary vertices per event

Table 2: The input features of the DNN.

Ref. available soon