East Bay BOOST 2023

The application of neural networks for the calibration of topological cell clusters in the ATLAS calorimeters

Principal ATLAS calorimeter signal: topological cell clusters

Eur.Phys.J.C 77 (2017) 490

Topo-clusters – 3-dimensional clusters of topologically connected cell signals collected by following spatial signal significance patterns
Massless pseudo-particle representation with basic signal E^{EM}_{clus}, rapidity y^{EM}_{clus} and azimuth φ^{EM}_{clus} at electromagnetic (EM) energy scale
Cluster moments comprise shapes, signal-to-noise, compactness, location, internal time structures – sensitive to signal source

MC Simulations pp @ $\sqrt{s} = 13$ TeV

Calorimeter Jet: $p_T^{jet} > 20$ GeV, $|y_{jet}| < 2$

TopoCluster $E^{dep}_{clus} < 300$ MeV

TopoCluster $E^{dep}_{clus} > 300$ MeV

Randomly sampled for training, testing & validation, performance

Local topo-cluster calibration with Monte Carlo simulations

Standard (LCW) local hadronic calibration

Eur.Phys.J.C 77 (2017) 490

EM-likelihood | hadronic calibration | outside energy losses

Topo-cluster Formation → Cluster Classification → Calibration → Out-of-cluster Corrections → Dead Material Corrections

energy deposited in cluster (from MC)

$$E^{had}_{clus} = \sum_{\text{cells } i \text{ in cluster}} w^{geo}_{cell,i} \cdot \left(\mathcal{P}^{EM}_{clus} + (1 - \mathcal{P}^{EM}_{clus}) \cdot w^{cal}_{cell,i}\right) \times E_{cell,i} \xrightarrow{\text{measures}} E^{dep}_{clus}$$

$\mathcal{P}^{EM}_{clus}$ & w^{cal}_{cell} from *look-up tables* employing cluster moments & cell signal features

1. **Classification** – look up likelihood for cluster being of EM nature $0 \leq \mathcal{P}^{EM}_{clus} \leq 1$
2. **Hadronic calibration** – apply cell signal weights to correct hadronic signals to EM scale ($e/h > 1 \xRightarrow{\text{calibrate to}} e/h = 1$)
3. **Out-of-cluster corrections** – correct for energy lost due to cells missed in cluster formation
4. **Dead material corrections** – correct for energy losses in inactive material in the proximity of the topo-cluster

Machine-learning-based (ML-based) local calibration

(proof of concept study)

Topo-cluster Formation → Calibration →

learned calibration

$$\underbrace{E^{ML}_{clus}}_{\hat{=}\, E^{had}_{clus}} = \mathcal{C}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus}) \cdot \underbrace{E^{EM}_{clus}}_{\text{basic cluster signal (EM scale)}} \xrightarrow{\text{measures}} E^{dep}_{clus}$$

energy deposited in cluster (from MC)

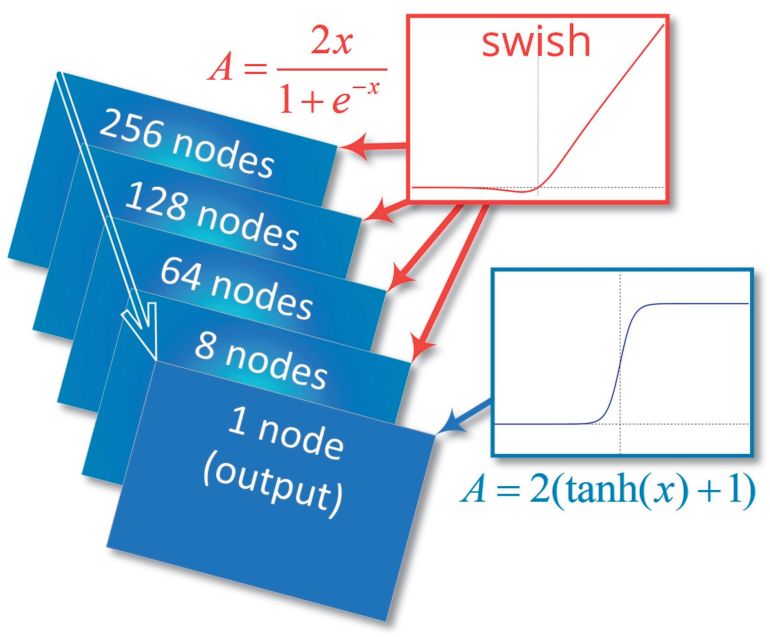
$$\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus}) \xrightarrow{\text{predicts}} \mathcal{R}^{EM}_{clus} = \frac{E^{EM}_{clus}}{E^{dep}_{clus}} \xrightarrow{\text{yields}} \mathcal{C}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus}) = \frac{1}{\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})}$$

1. **No explicit classification** – variables used to determine $\mathcal{P}^{EM}_{clus}$ added to feature set $\mathfrak{O}^{ML}_{clus}$
2. **Hadronic calibration** – apply scale factor to cluster signal E^{EM}_{clus} to measure E^{dep}_{clus}

Calibration from response predictions $\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})$ obtained by regression fits $\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus}) \mapsto \mathcal{R}^{EM}_{clus}$ employing feature set $\mathfrak{O}^{ML}_{clus}$

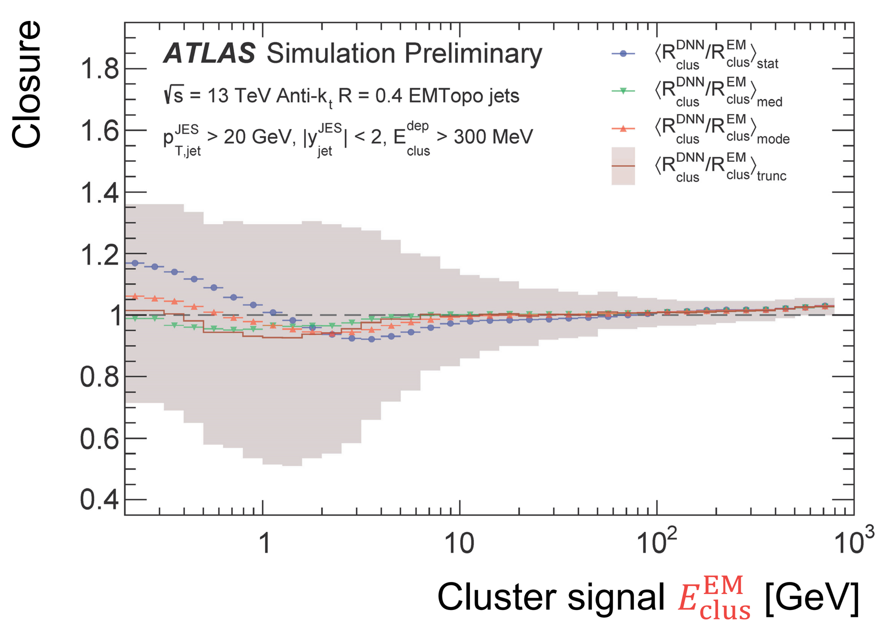
Networks

Dense Neural Net (DNN)

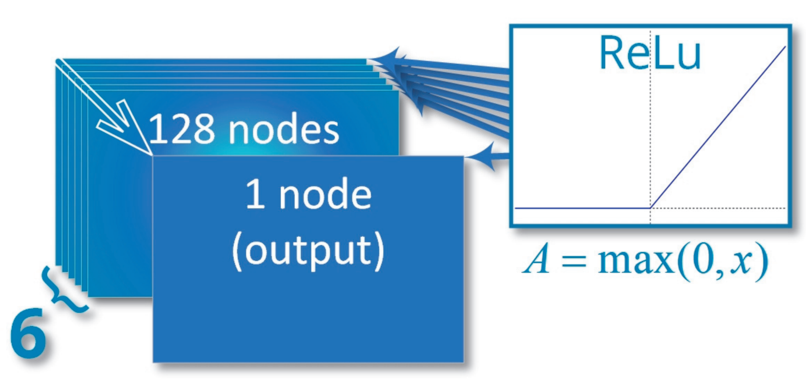


Loss function: $h = 0.1$, $\alpha = 0.05$ (1st pass), $\alpha = 0$ (2nd pass):

$$\mathcal{L}_{LGK} = -\frac{1}{\sqrt{2\pi}h} \exp\left[\frac{1}{2h}\left(\frac{\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})}{\mathcal{R}^{EM}_{clus}} - 1\right)^2\right] + \alpha\left|\frac{\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})}{\mathcal{R}^{EM}_{clus}} - 1\right|$$



Bayesian Neural Net (BNN)



Negative log-likelihood loss function with regularization by (reverse) Kullback-Leibler (KL) divergence D_{KL} *models* $\mathcal{R}^{EM}_{clus}$ *distribution*

$$D_{KL}\left[p\left(\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})\right), q(\mathcal{R}^{EM}_{clus})\right] = \left\langle \log \frac{p(\mathcal{R}^{EM}_{clus})}{q\left(\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})\right)} \right\rangle_{q(\mathcal{R}^{EM}_{clus})}$$

$$\mathcal{L}_{BNN} \sim -\log\left(p_{train} = \sum_{i=1}^{6} \mathcal{G}_i(\mathcal{R}^{EM}_{clus}, \mathfrak{O}^{ML}_{clus}; \mu_i, \sigma_i)\right) + D_{KL}\left[p\left(\mathcal{R}^{ML}_{clus}(\mathfrak{O}^{ML}_{clus})\right), q(\mathcal{R}^{EM}_{clus})\right]$$

↪ uncertainties (1) for model and (2) from training statistics → under further investigation

Common Network Feature Set $\mathfrak{O}^{ML}_{clus}$

$$\mathfrak{O}^{ML}_{clus} = \{E^{EM}_{clus}, y^{EM}_{clus}, t_{clus}, \mathrm{Var}_{clus}(t_{cell}), \zeta^{EM}_{clus}, f_{emc}, |\vec{c}_{clus}|, \lambda_{clus}, \langle \mathfrak{m}^2_{long}\rangle, \langle \mathfrak{m}^2_{lat}\rangle, \rho_{clus}, p_T D, f_{iso}, \mu, N_{PV}\}$$

Kinematics and signal reconstruction

Eur.Phys.J.C 77 (2017) 490

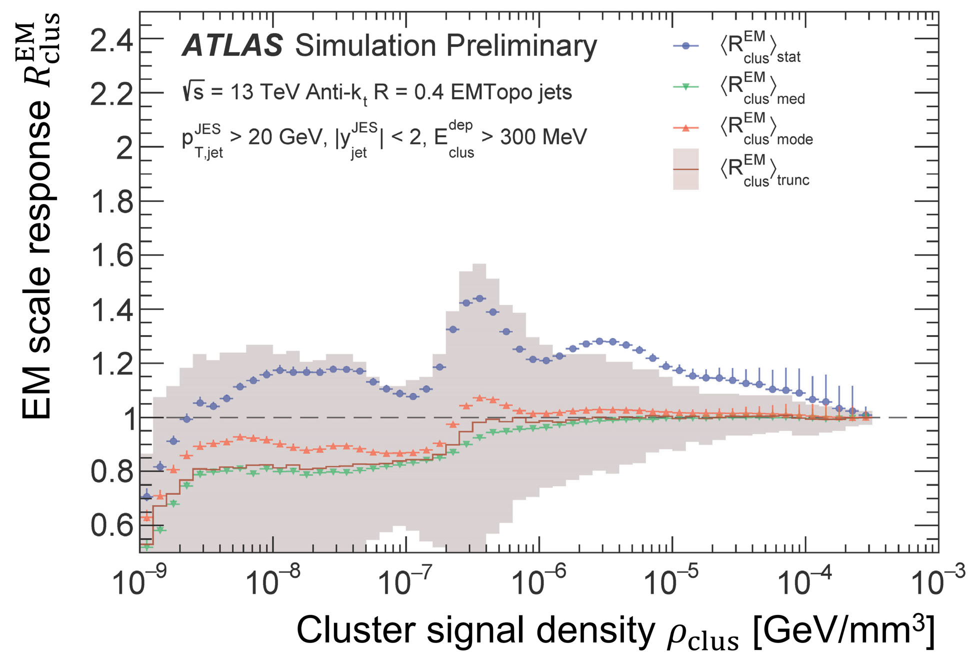
- Regional signal characteristics: E^{EM}_{clus}, y^{EM}_{clus} – at EM scale
- Cluster signal time t_{clus}; internal time structure $\mathrm{Var}_{clus}(t_{cell}) = \langle t^2_{cell}\rangle - t^2_{clus}$
- Significance – $\zeta^{EM}_{clus} = E^{EM}_{clus}/\sigma^{EM}_{clus}$ (noise σ^{EM}_{clus} includes pile-up fluctuations for LHC Run 2)

Shower development

- Sharing, location – EM calorimeter signal fraction f_{emc}, distance to vertex $|\vec{c}_{clus}|$, depth in calorimeter λ_{clus}
- Compactness – spatial energy dispersion along $\langle \mathfrak{m}^2_{long}\rangle$, perpendicular $\langle \mathfrak{m}^2_{lat}\rangle$ to shower axis; signal density $\rho_{clus} = \langle \rho_{cell}\rangle$; core density measure $p_T D = \sqrt{\sum_{E_{cell}>0} E^2_{cell} / \sum_{E_{cell}>0} E_{cell}}$

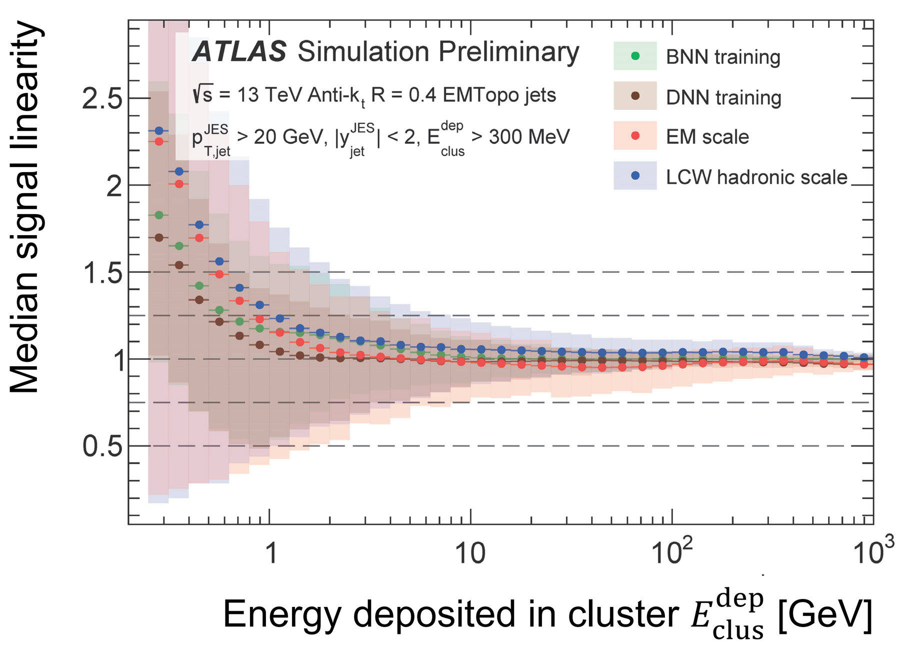
Event environment

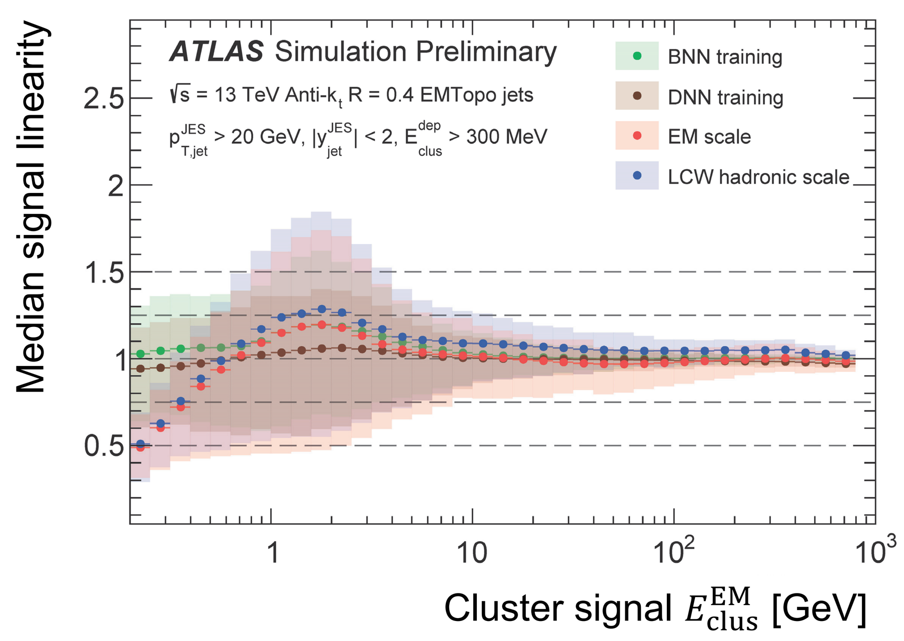
- Topology/neighboring signals – cluster isolation f_{iso}
- Pile-up – out-of-time ↦ number of pile-up interactions/bunch crossing μ; in-time ↦ number of reconstructed primary vertices N_{PV}

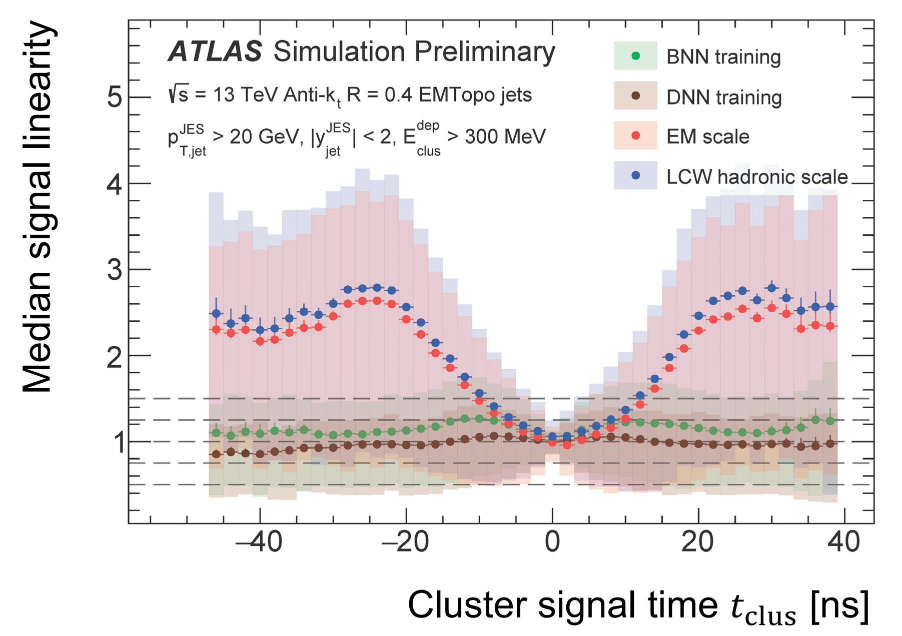


Performance Comparisons

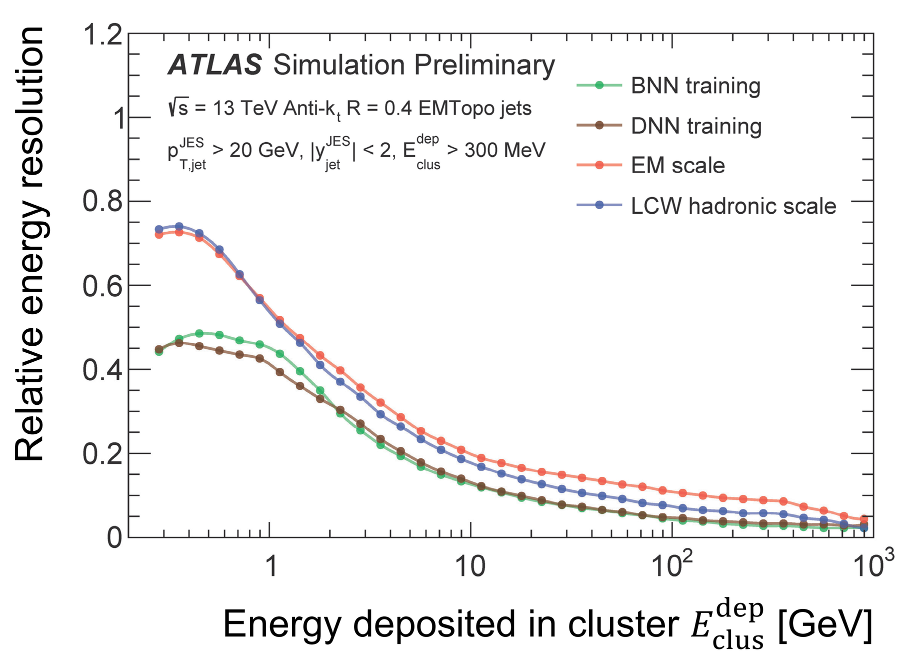
Signal linearity

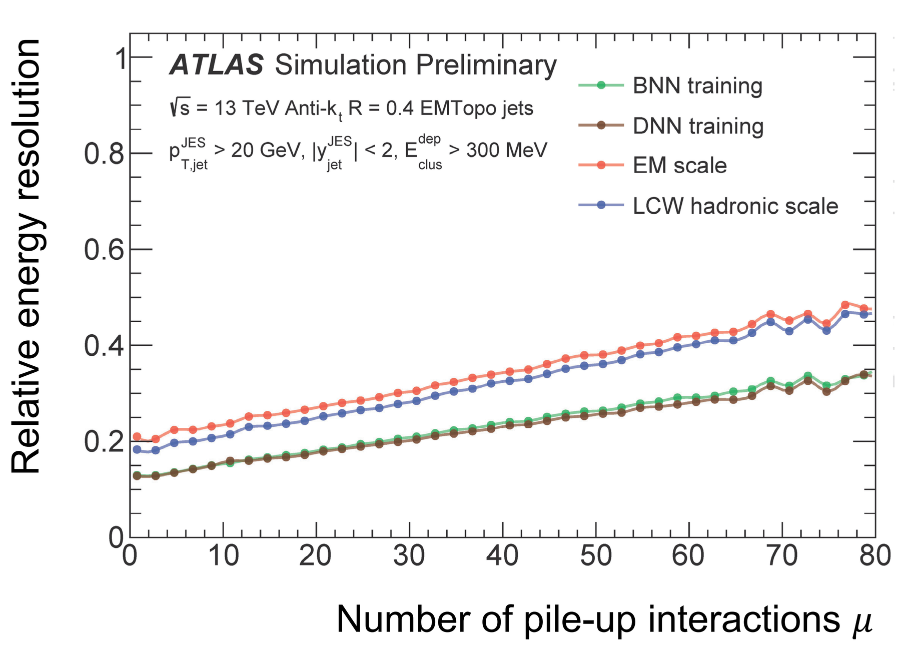


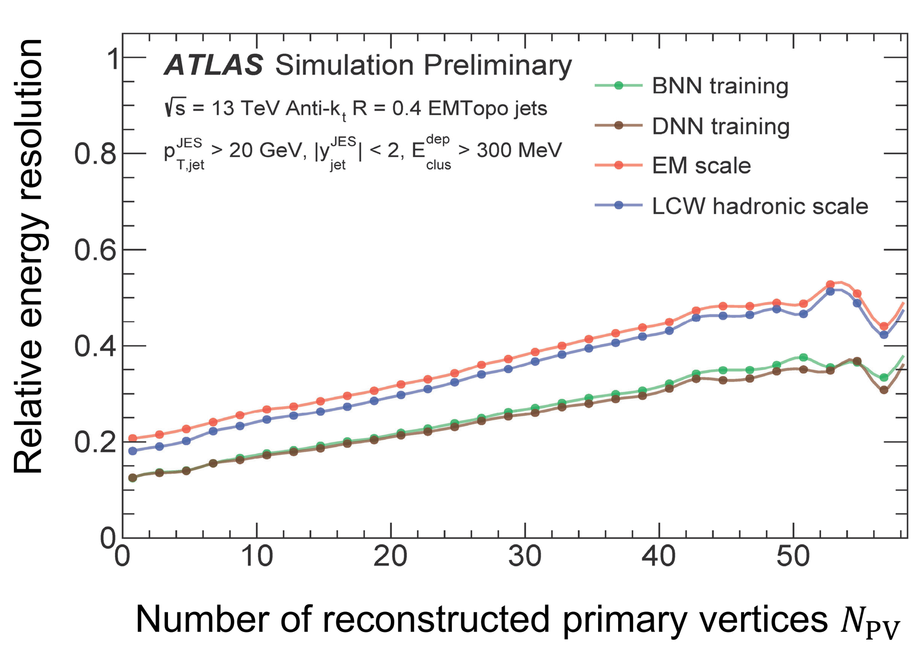




Relative local energy resolution







Peter Loch, for the ATLAS collaboration

ATL-PHYS-PUB-2023-019

ATLAS EXPERIMENT