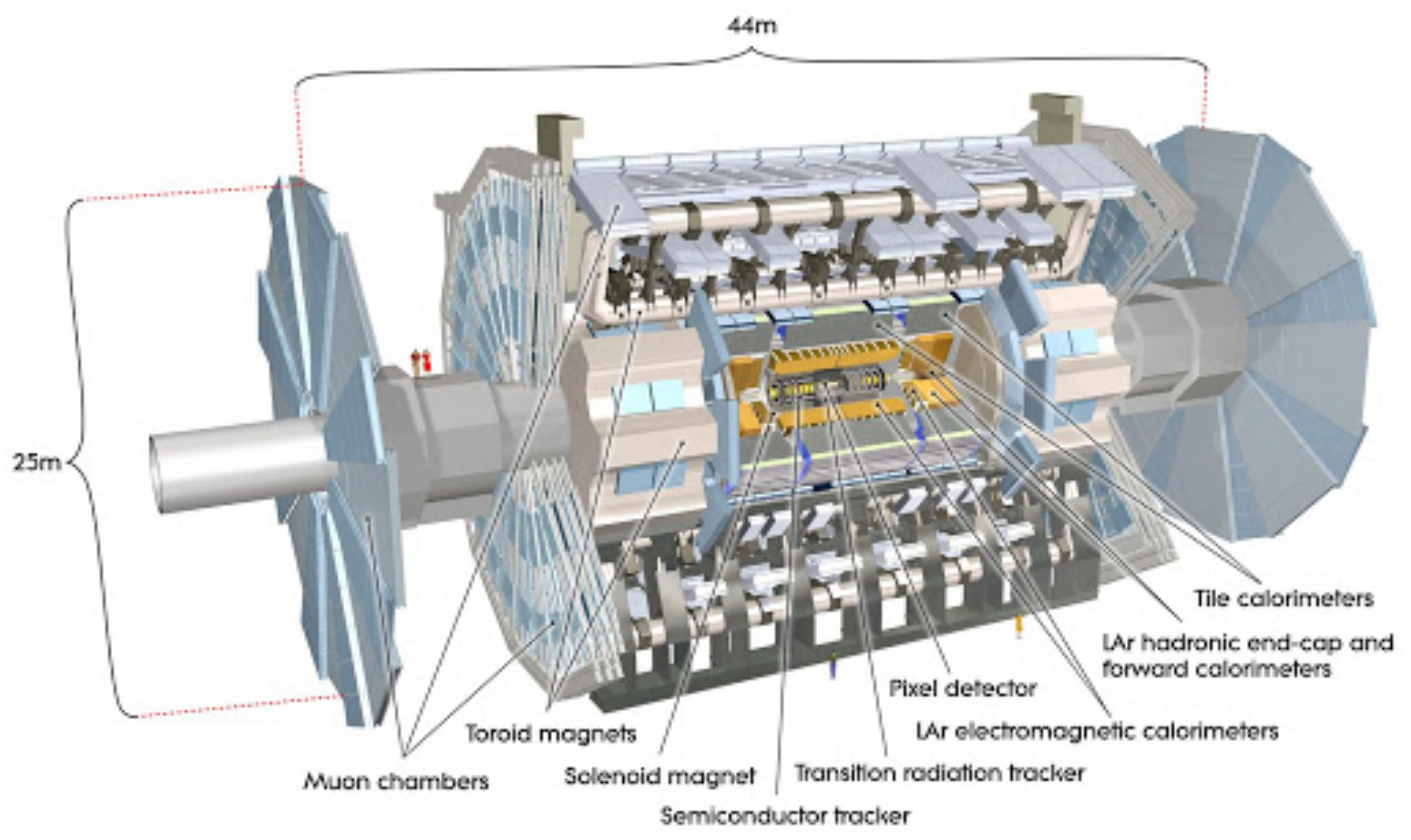
Hadronic Reconstruction Techniques at ATLAS

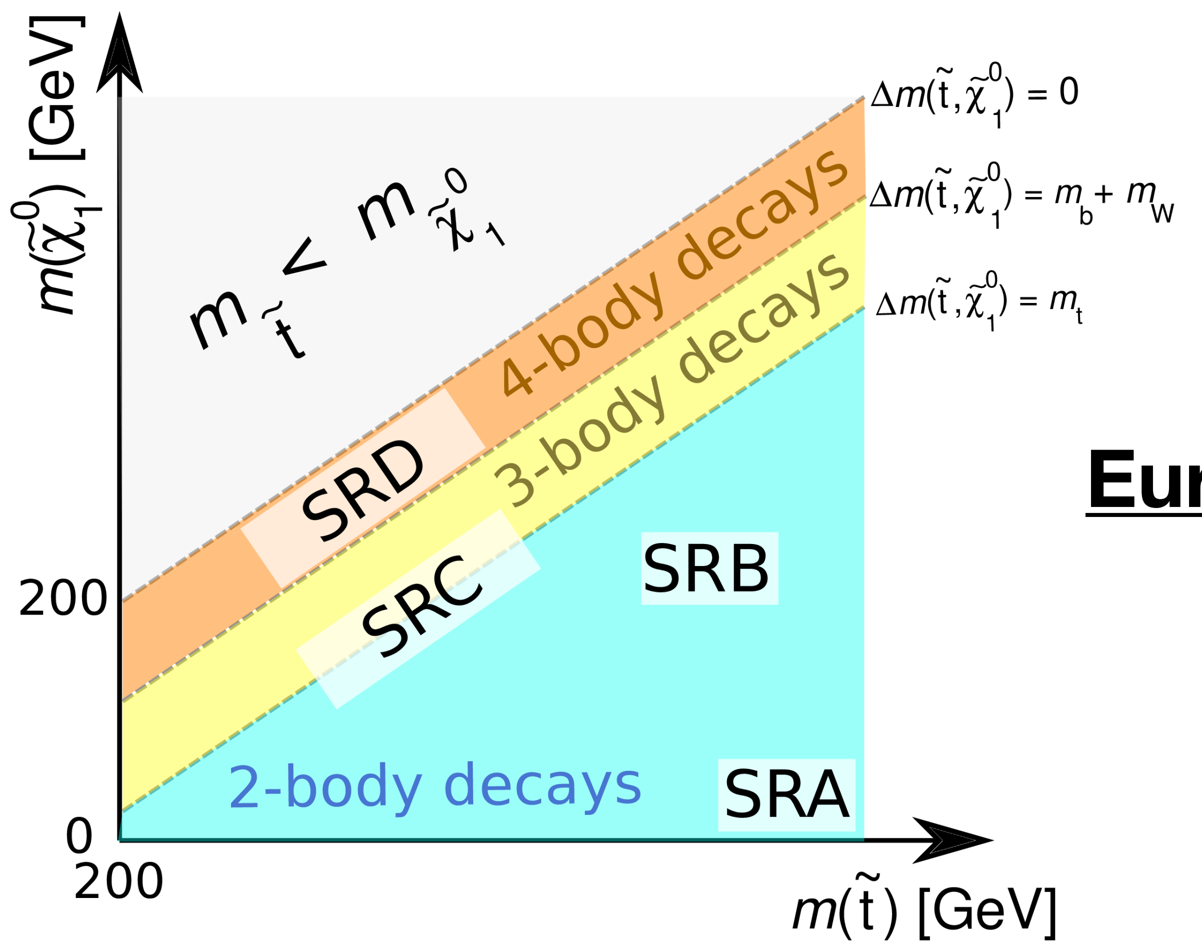
Mark Hodgkinson (on behalf of the ATLAS collaboration)
University of Sheffield
SUSY 2021
m.hodgkinson@sheffield.ac.uk

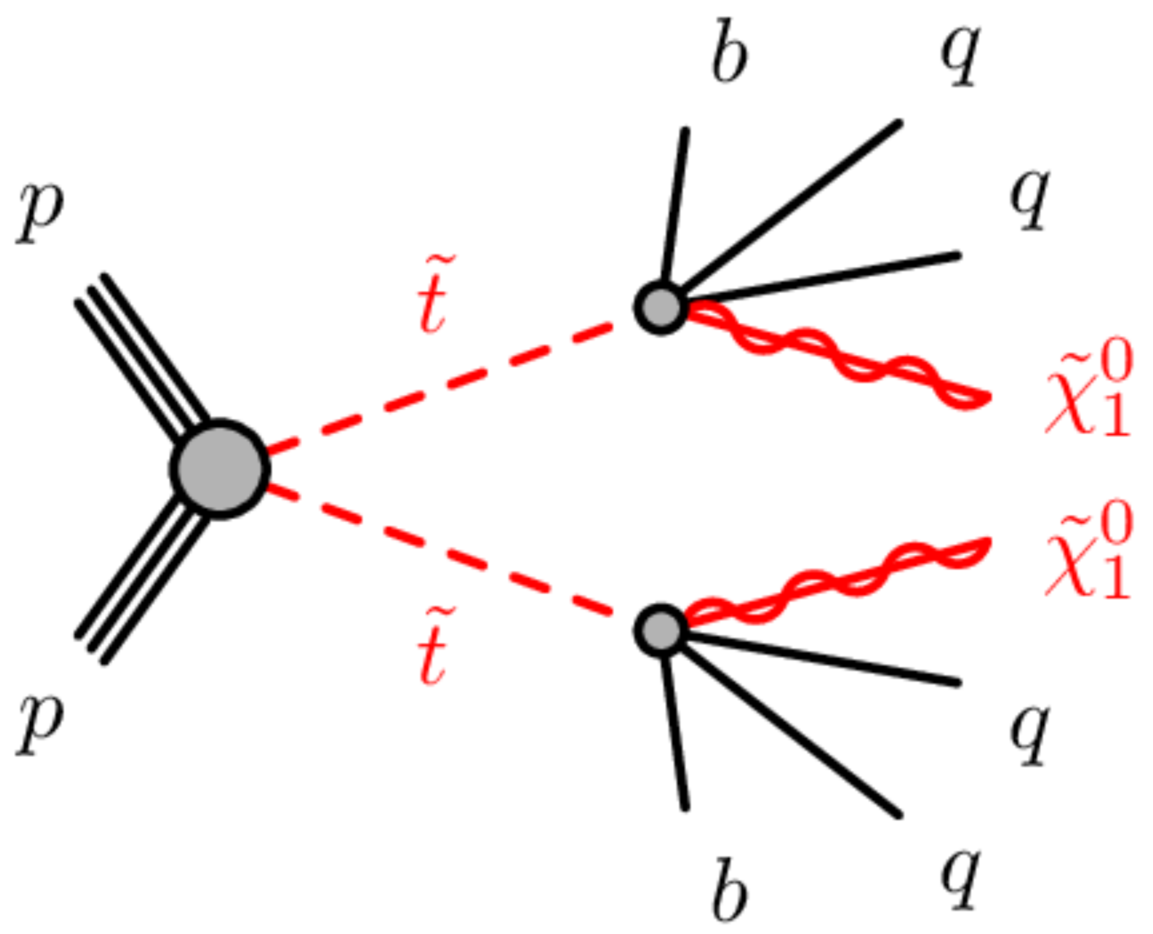
Contents

- Introduction.
- Small radius jets.
- Large radius jets.
- Calorimeter cluster calibration using machine learning.
- Conclusions.

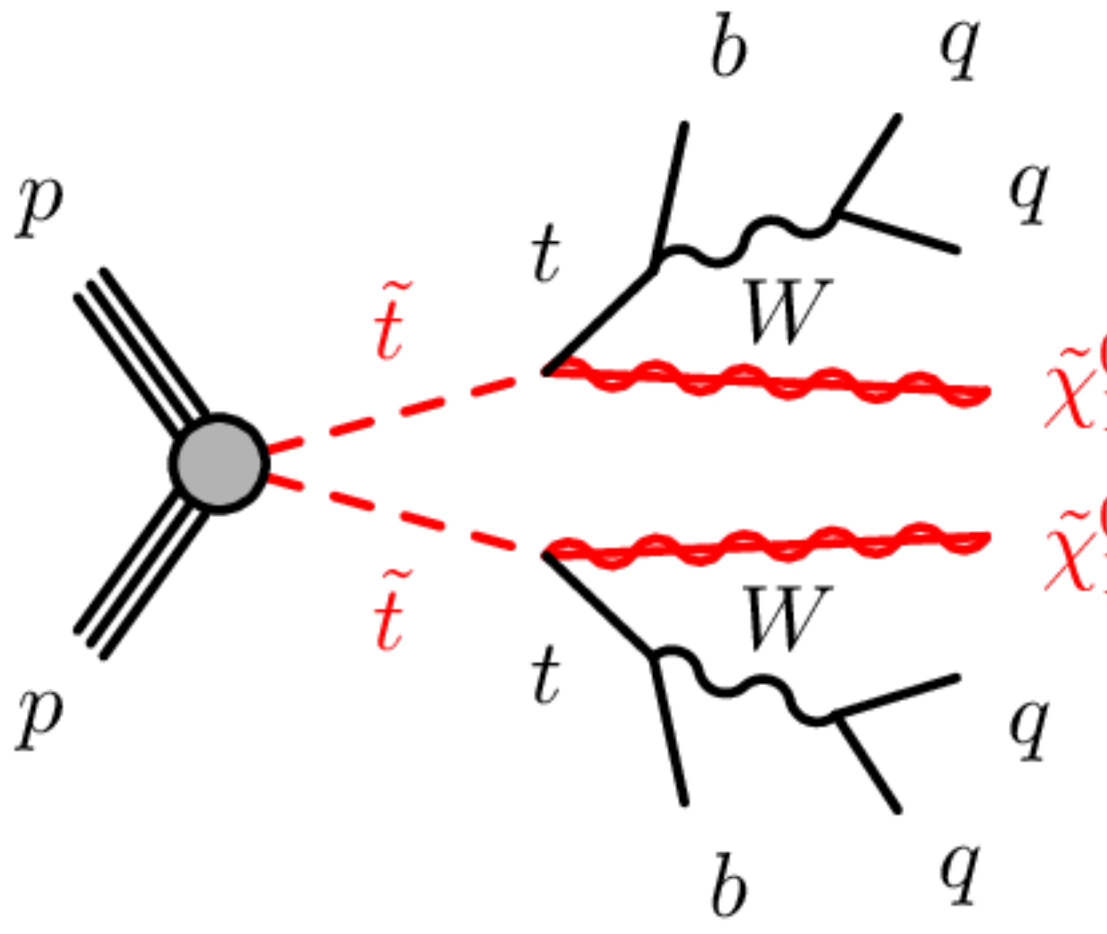
Introduction





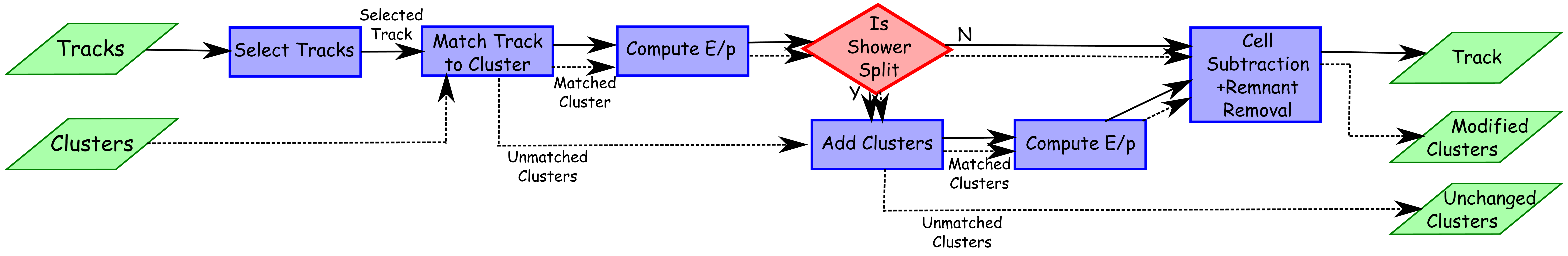


Eur. Phys. J. C 80 (2020) 737



- Jet reconstruction makes use of both Inner Detector tracks and Calorimeter Topoclusters.
  - Relevant for many signatures searched for with ATLAS.
  - e.g Anti-Kt 0.4 jets in direct stop production (top right)
  - e.g Large radius jets to reconstruct boosted systems such as top decay into bottom quark and hadronically decaying W boson (bottom right).
  - Different jet reconstruction techniques relevant in different parts of phase space (middle).
  - Similar arguments apply to many other experimental topologies, which are signatures of production of Supersymmetric particles.

Particle Flow

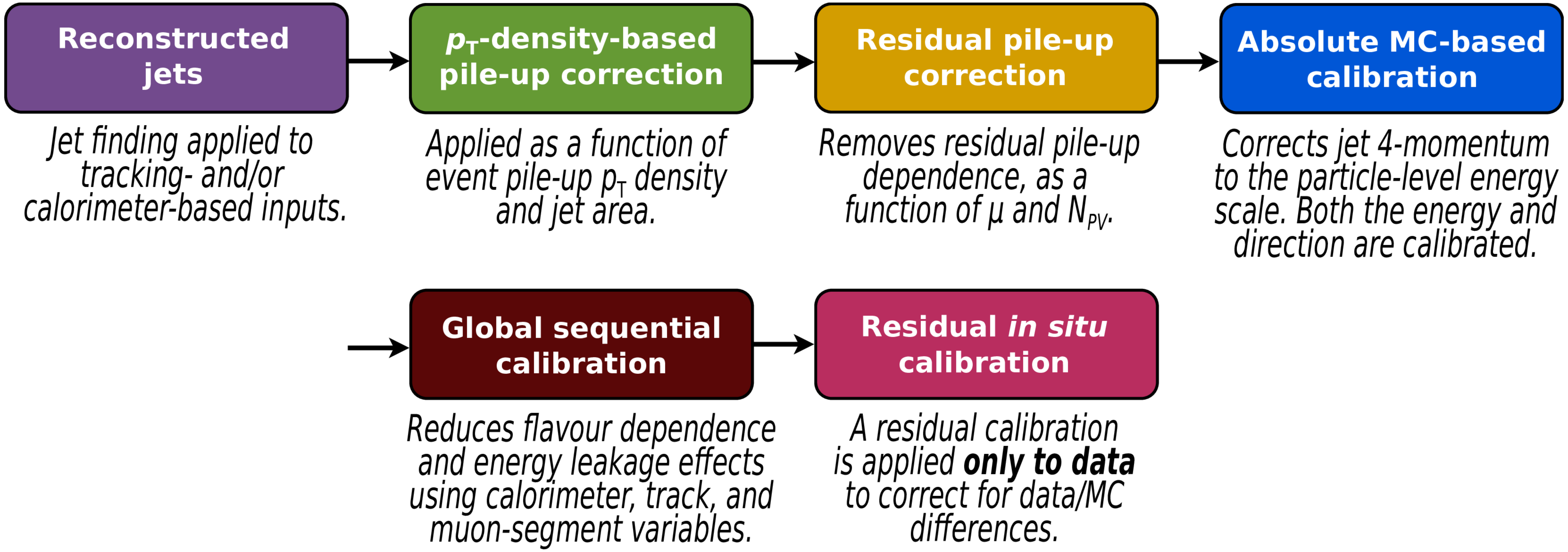


Particle Flow
(Eur. Phys. J. C 77 (2017) 466)

- Starts with Inner Detector tracks and calorimeter topological clusters as input.
- Matching algorithms associate them to each other, and when appropriate subtract out the charged calorimeter shower (based on reference measurements of e/p distributions).

Jet Finding

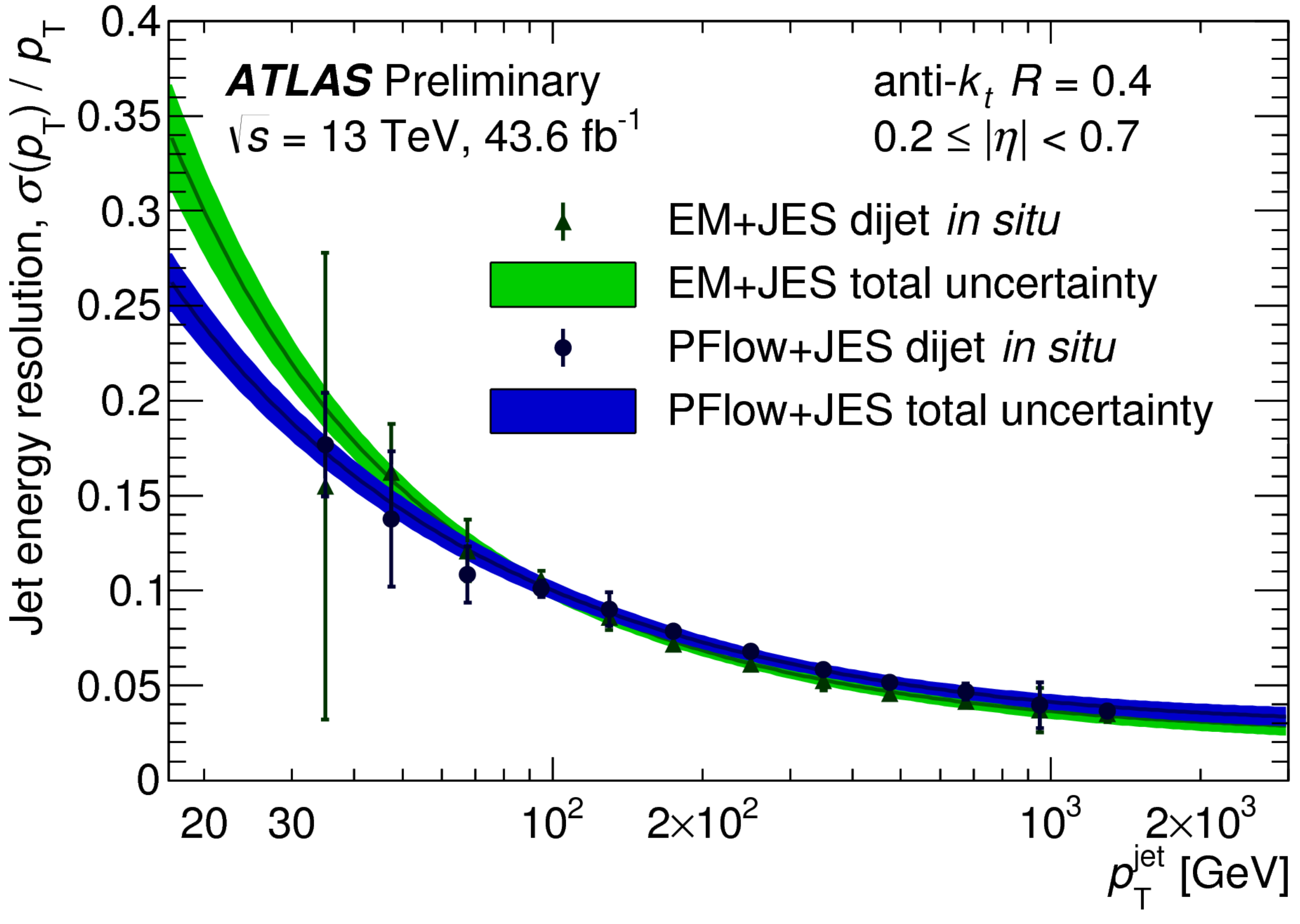
Eur. Phys. J. C 81 (2021) 689



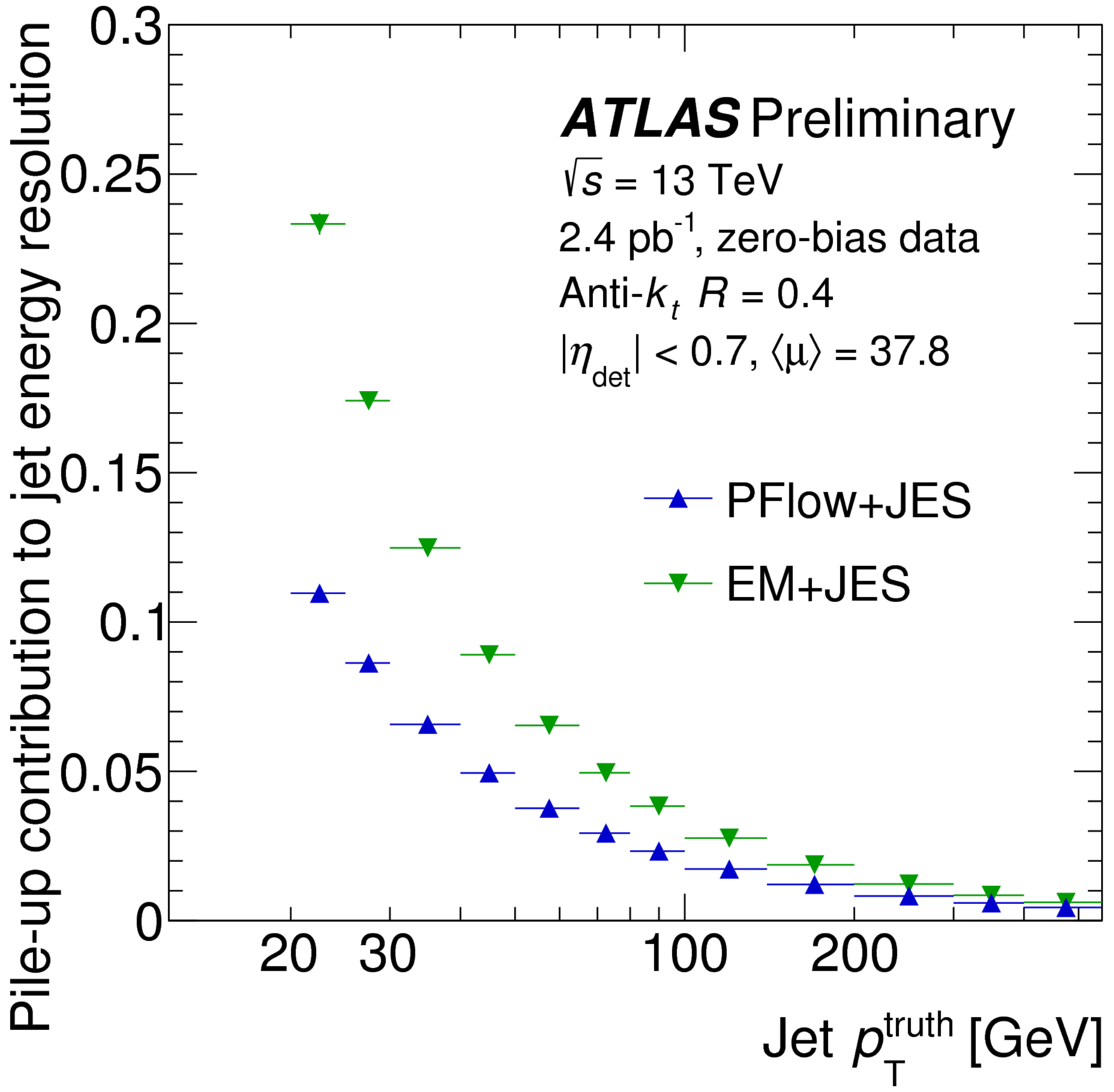
- We use the Anti-kt 4 algorithm to reconstruct jets.
  - Takes as input a set of 4-vectors - could be the Particle Flow objects discussed on previous slide, ID tracks or calorimeter topoclusters.
  - The jet 4-vector that results from this procedure is then calibrated via a set of steps outlined in above diagram - can measure both the Jet Energy Resolution (JER) and the Jet Energy Scale (JES) to quantify the performance.

Small Radius Jet Performance

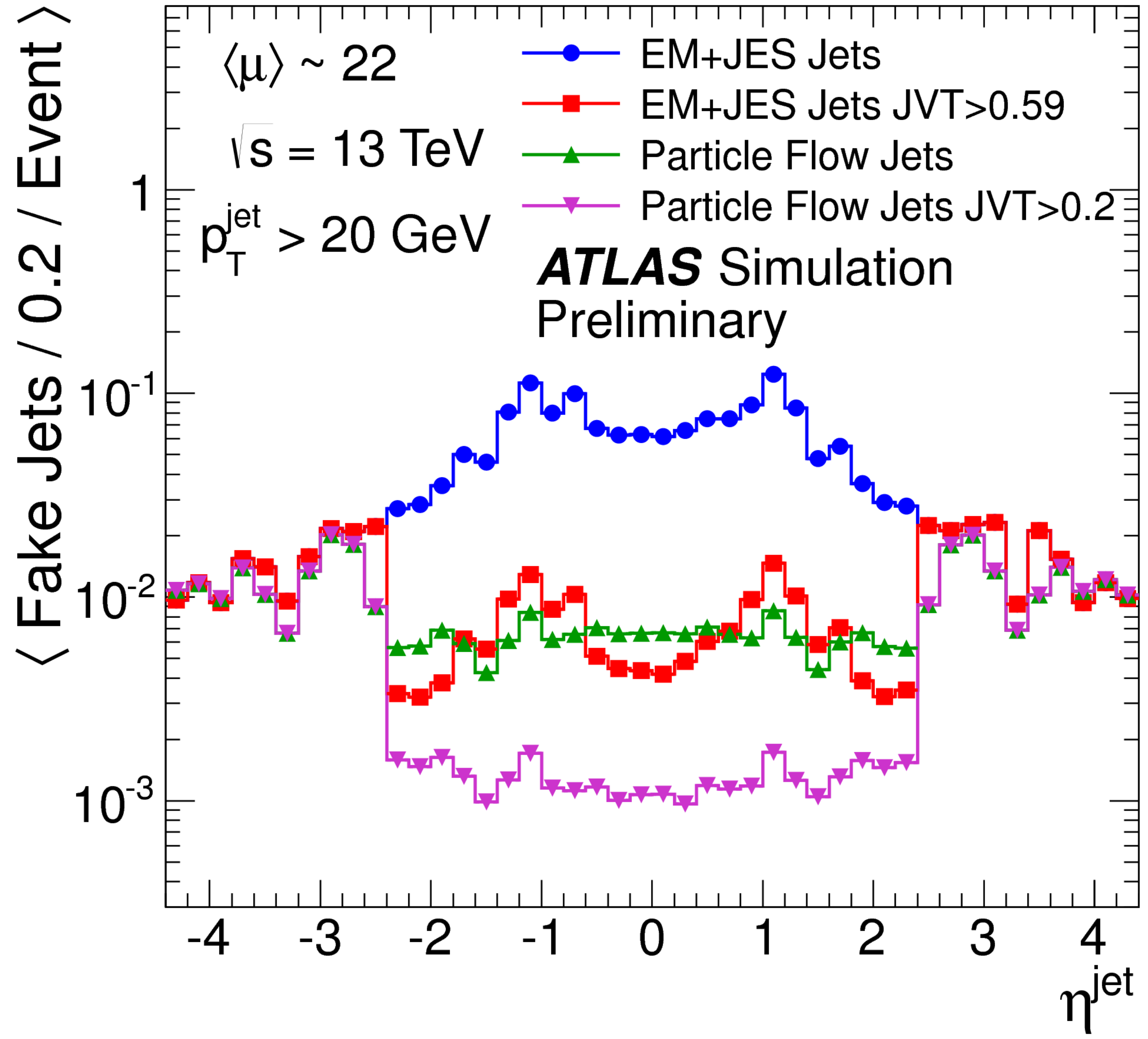
ATLAS-JETM-2018-005



ATLAS-JETM-2019-01



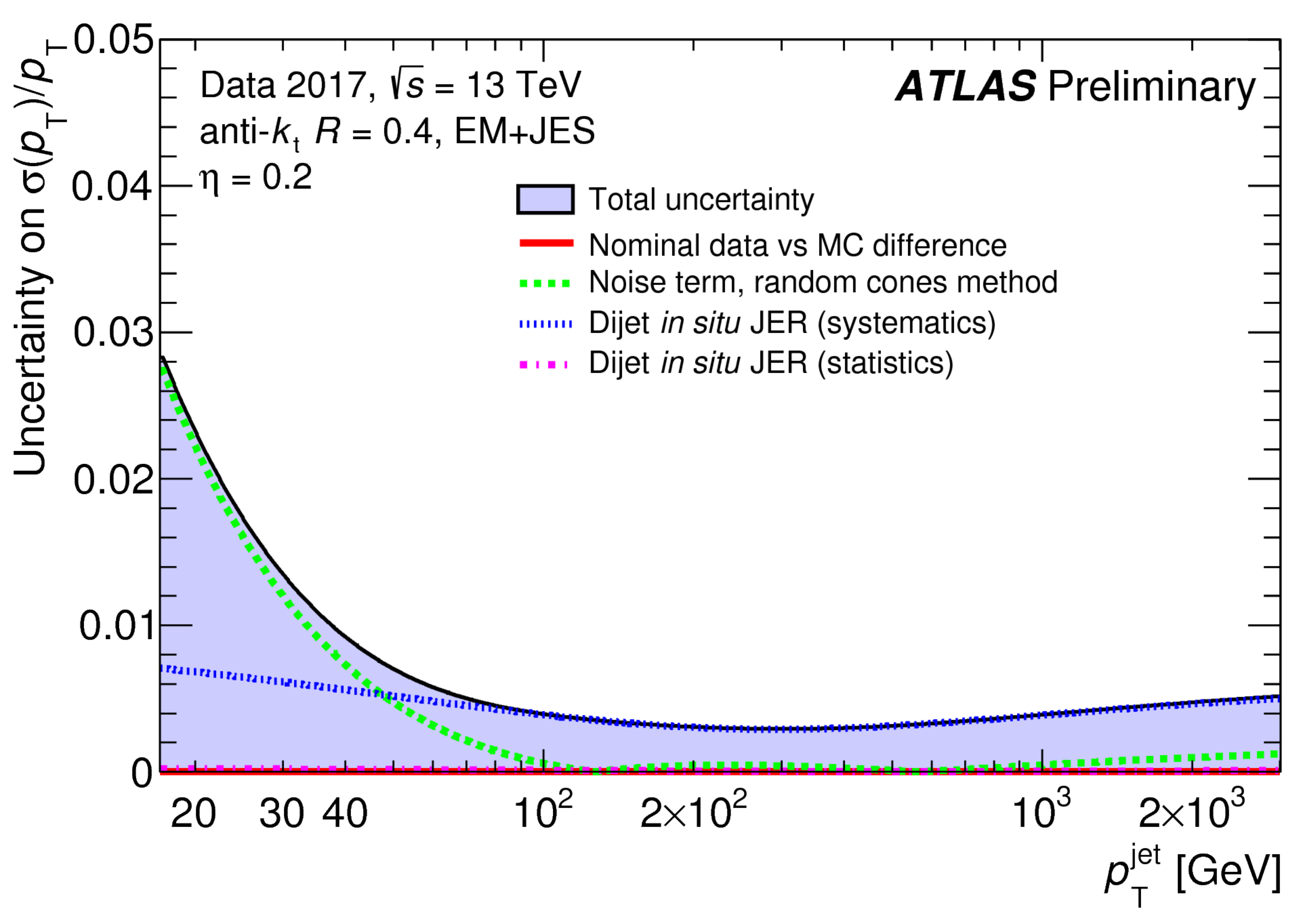
ATLAS-JETM-2017-006

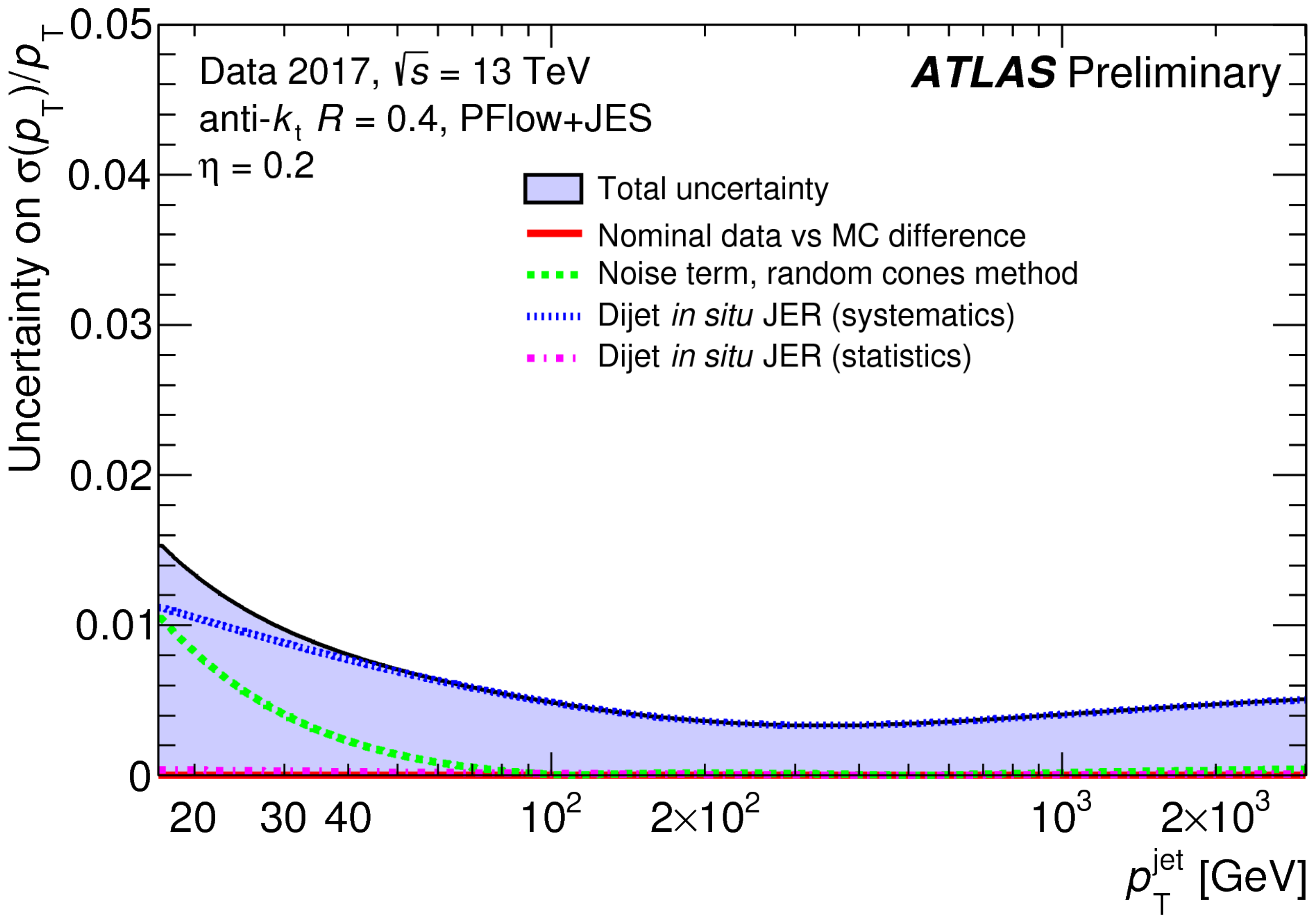


- Improved Particle Flow jet resolution at low P_T (left)
  - Due to smaller contribution to resolution from pileup (middle)
- Fewer Particle Flow pileup jets are reconstructed for the same Hard Scatter efficiency (right)

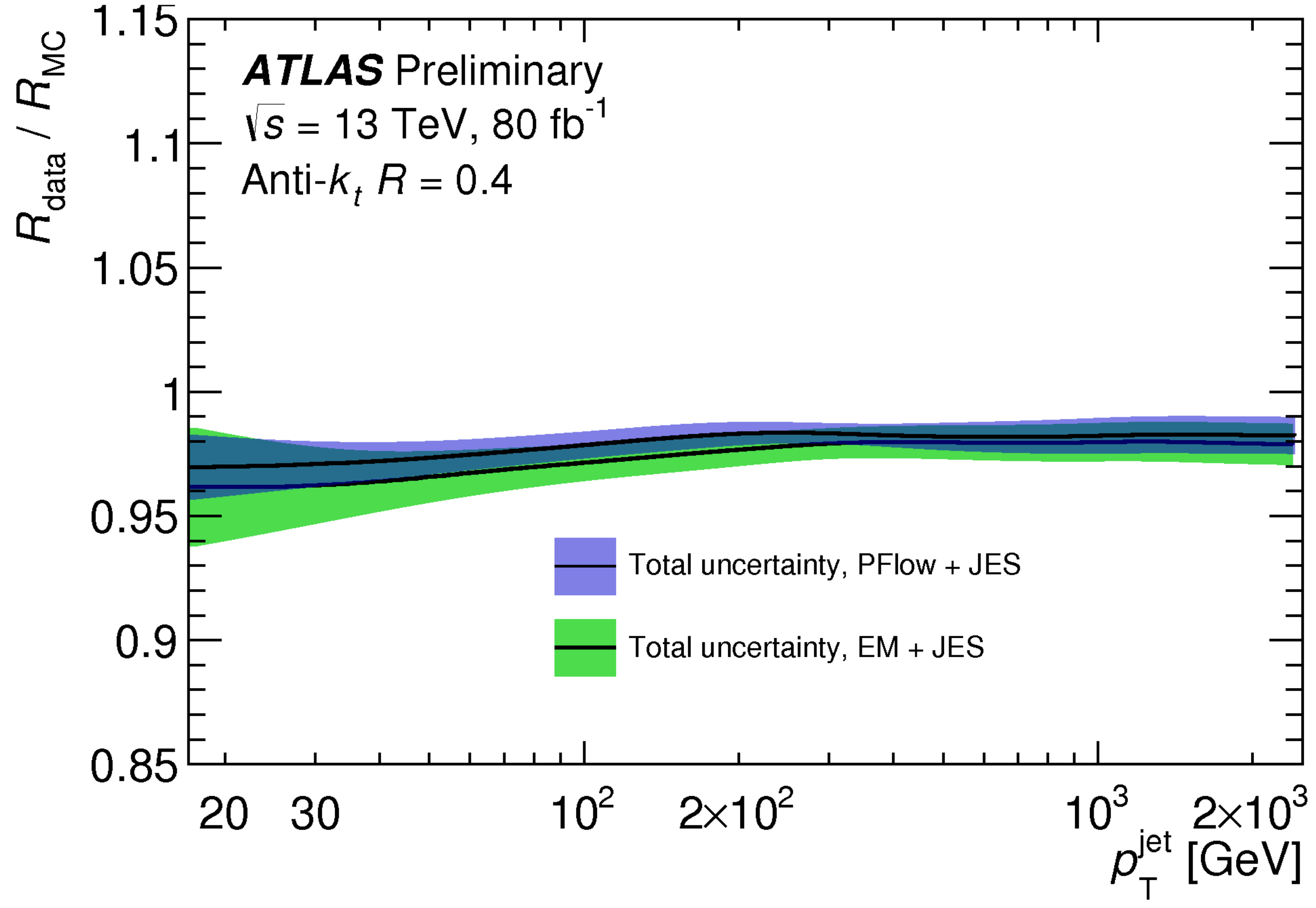
Small Radius Jet Uncertainties

ATLAS-JETM-2018-005



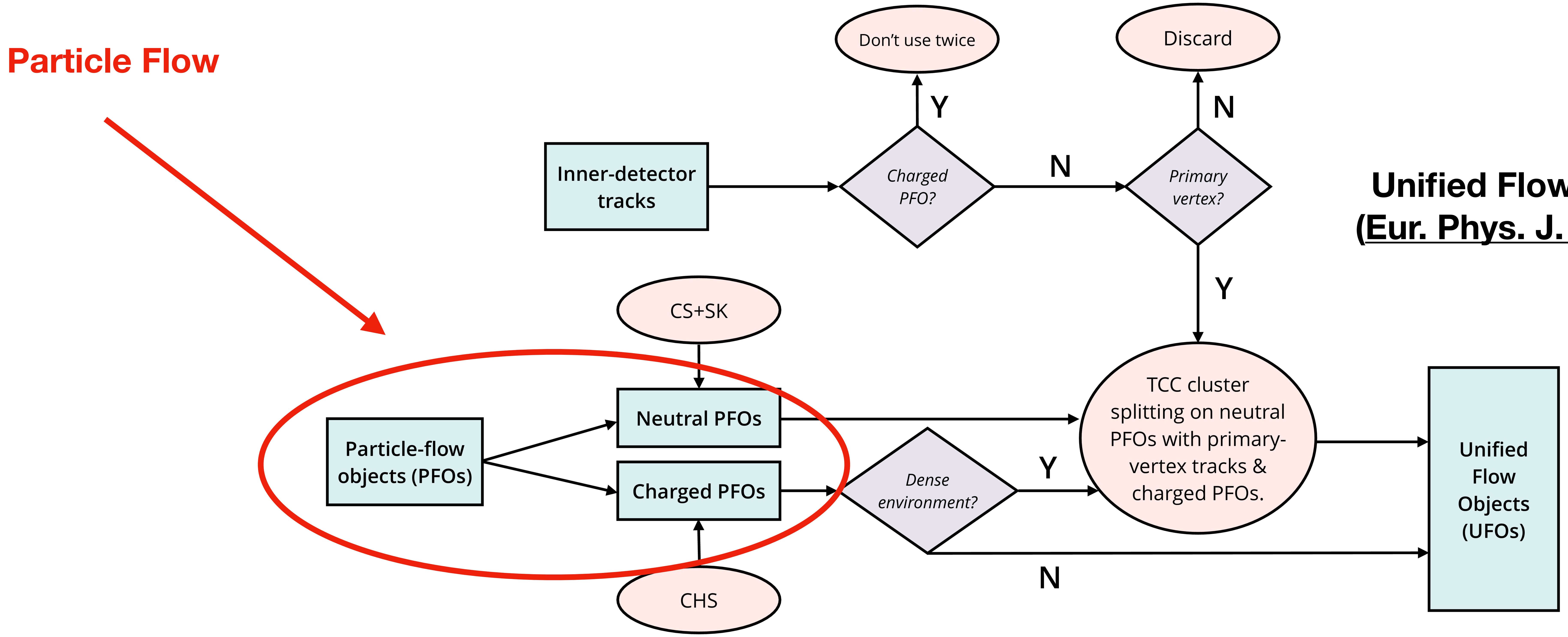


ATLAS-JETM-2018-006



- Particle Flow jets provide similar uncertainties on both noise term and in-situ Jet Energy Resolution (JER) measurement with di-jets (left and middle), except for the lowest P_T bins where Particle Flow uncertainty is smaller.
- Similar overall situation on the Jet Energy Scale (JES) too (right).

Unified Flow Objects (UFO)

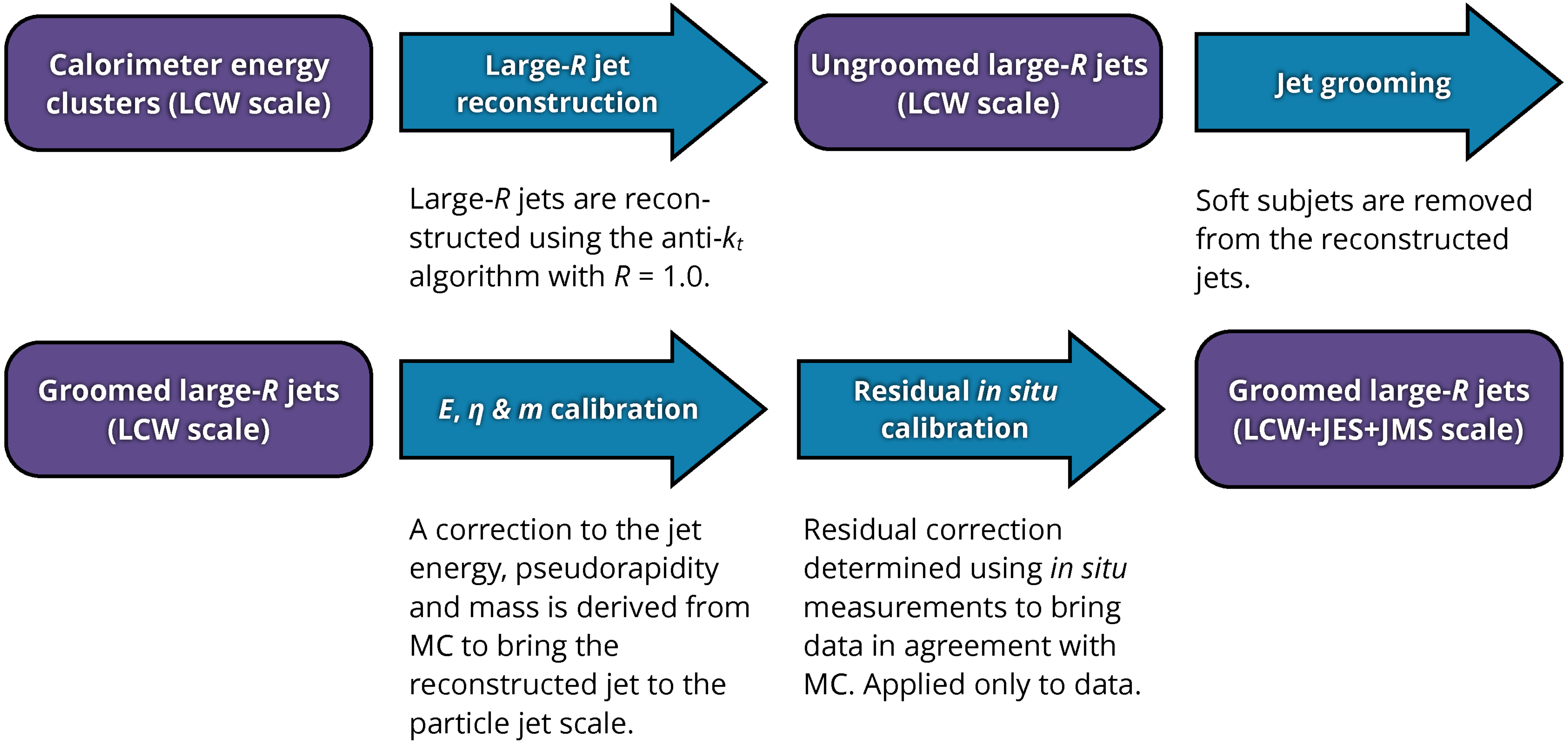


Unified Flow Objects (UFO)
(Eur. Phys. J. C 81 (2021) 334)

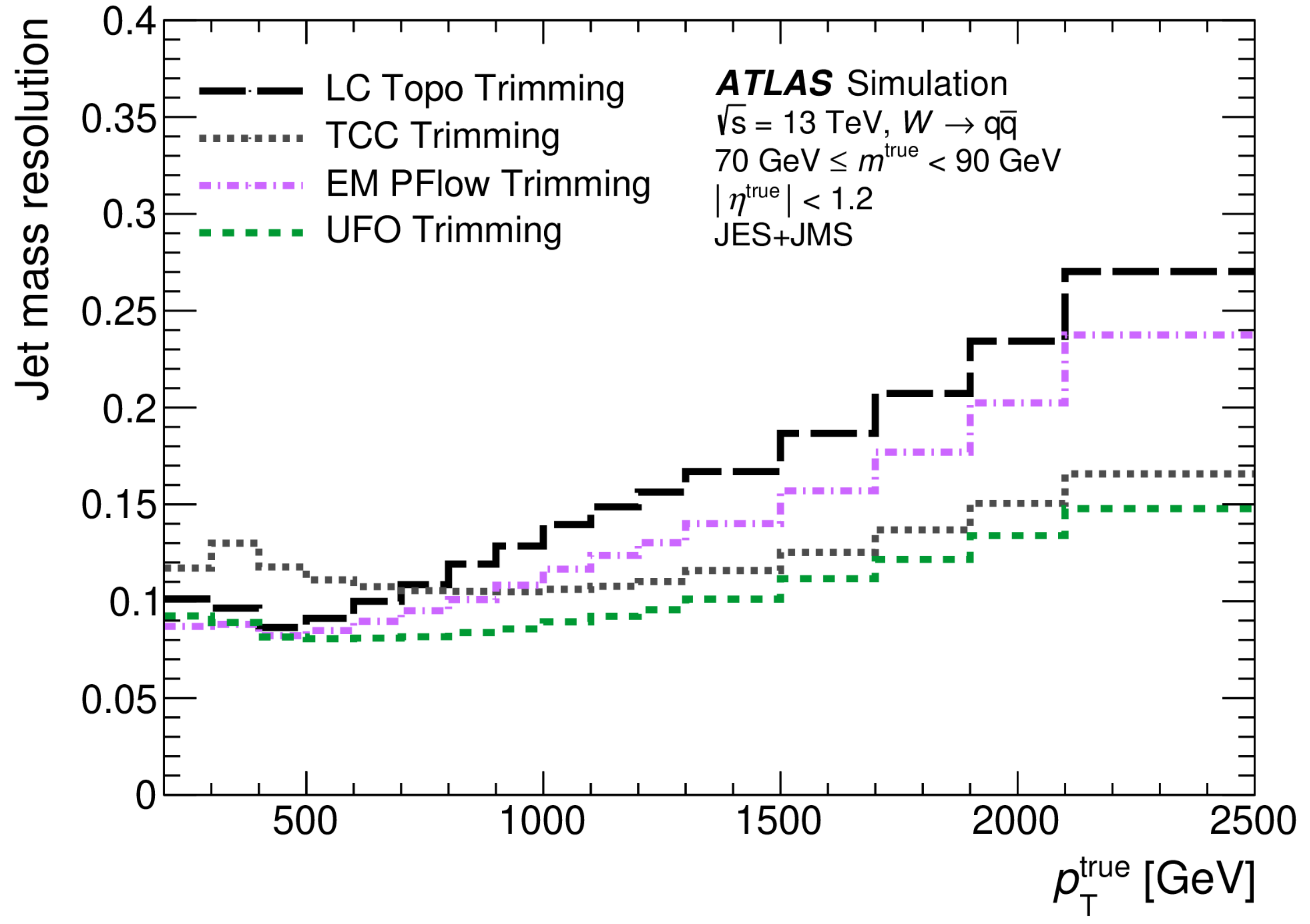
- Particle Flow most relevant for areas of low particle density with charged particles typically having low P_T.
- TrackCaloCluster (TCC) most relevant for areas of high particle density with charged particles typically having higher P_T.
  - TCC matches tracks to topoclusters and uses the ID track angular coordinates and the calorimeter energy measurement.
- UFO combines TCC and Particle Flow to get the best of both worlds - in this scheme TCC matches ID tracks to neutral PFO.
  - Studied in context of large radius jets so far, but can in principle be used for small radius jets.

Large Radius Jet Calibration

Eur. Phys. J C 79 (2019) 135

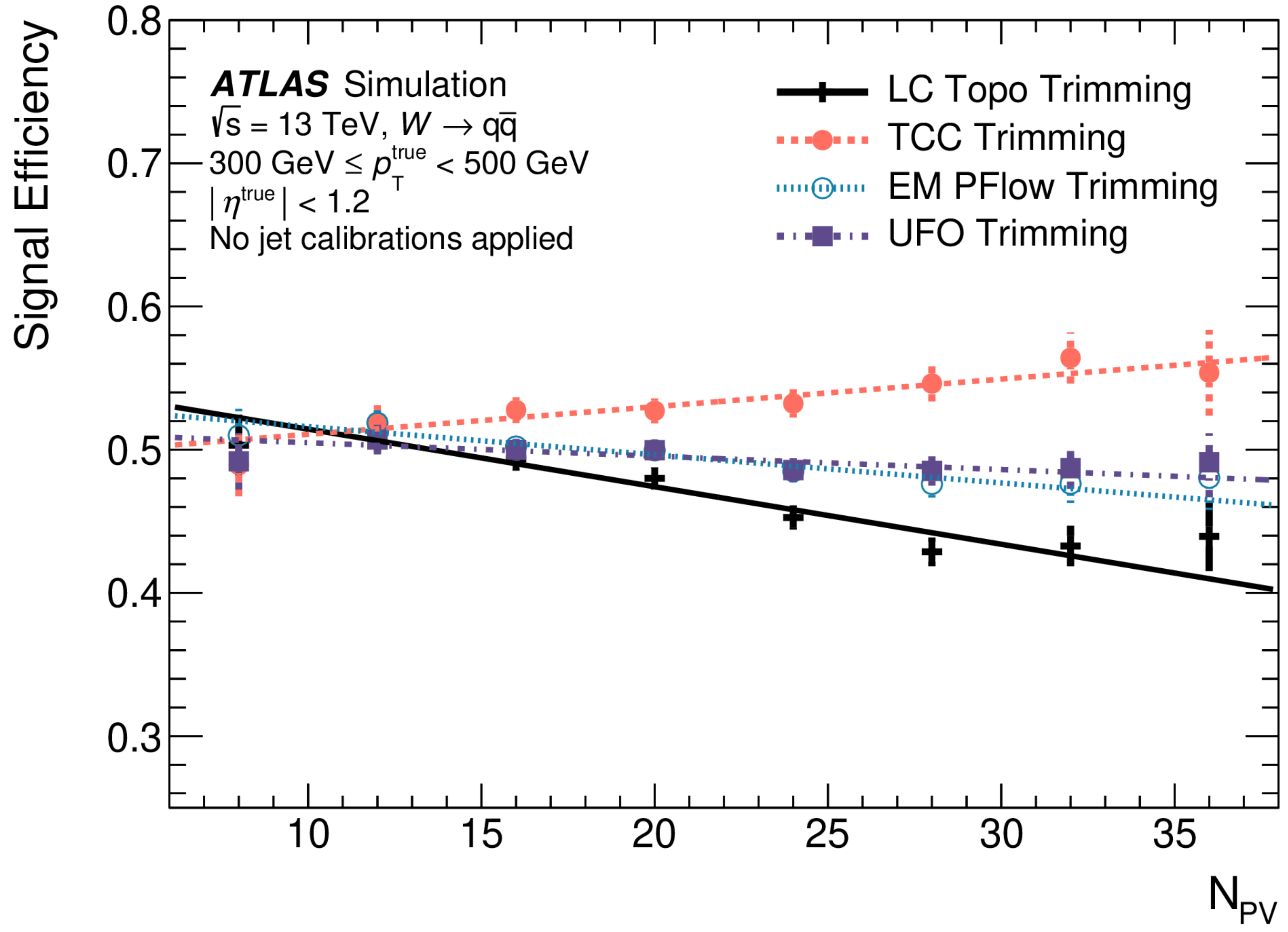


- Takes as input a set of 4-vectors - could be the UFOobjects discussed on previous slide, ID tracks, calorimeter topoclusters or particle flow objects.

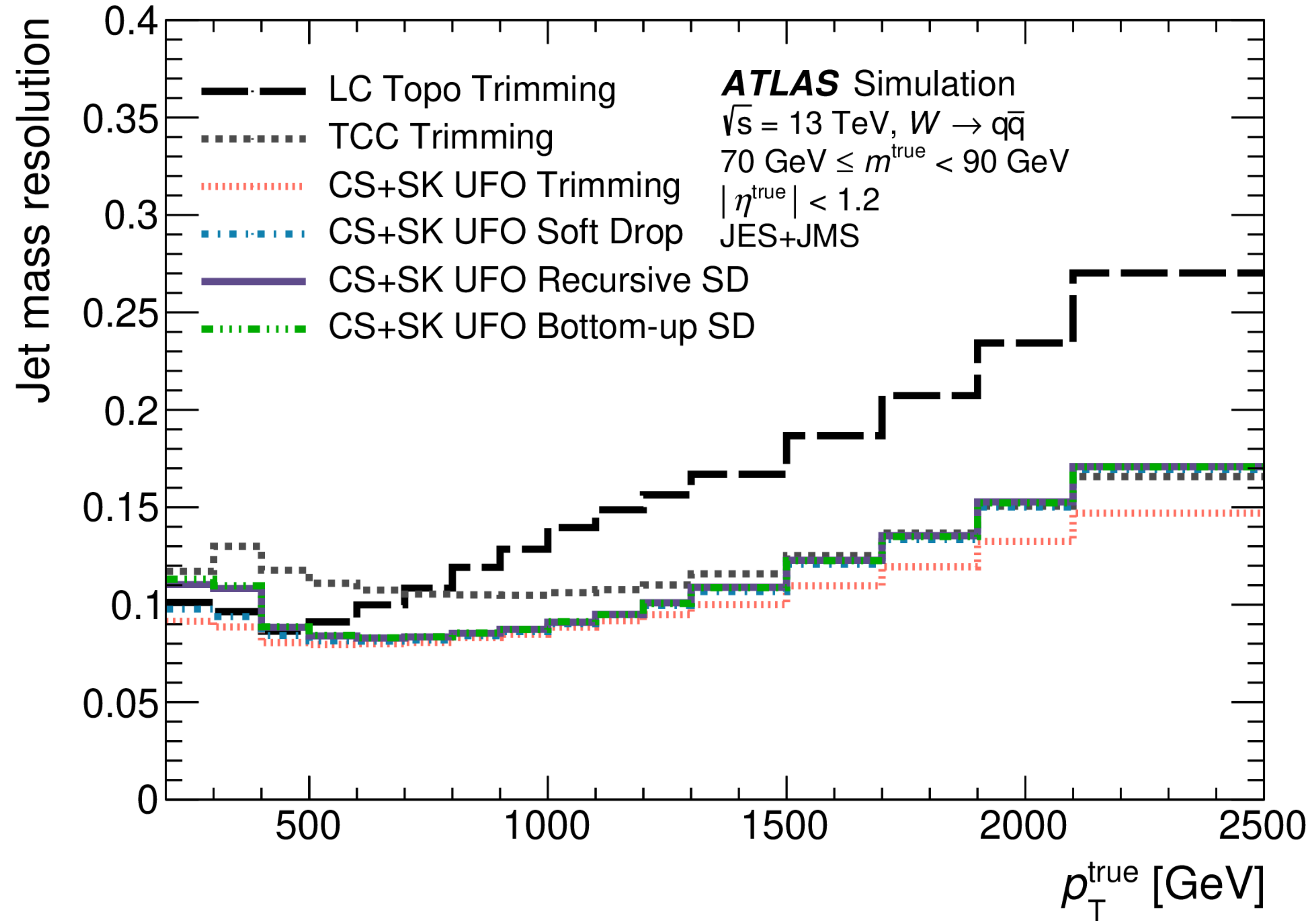


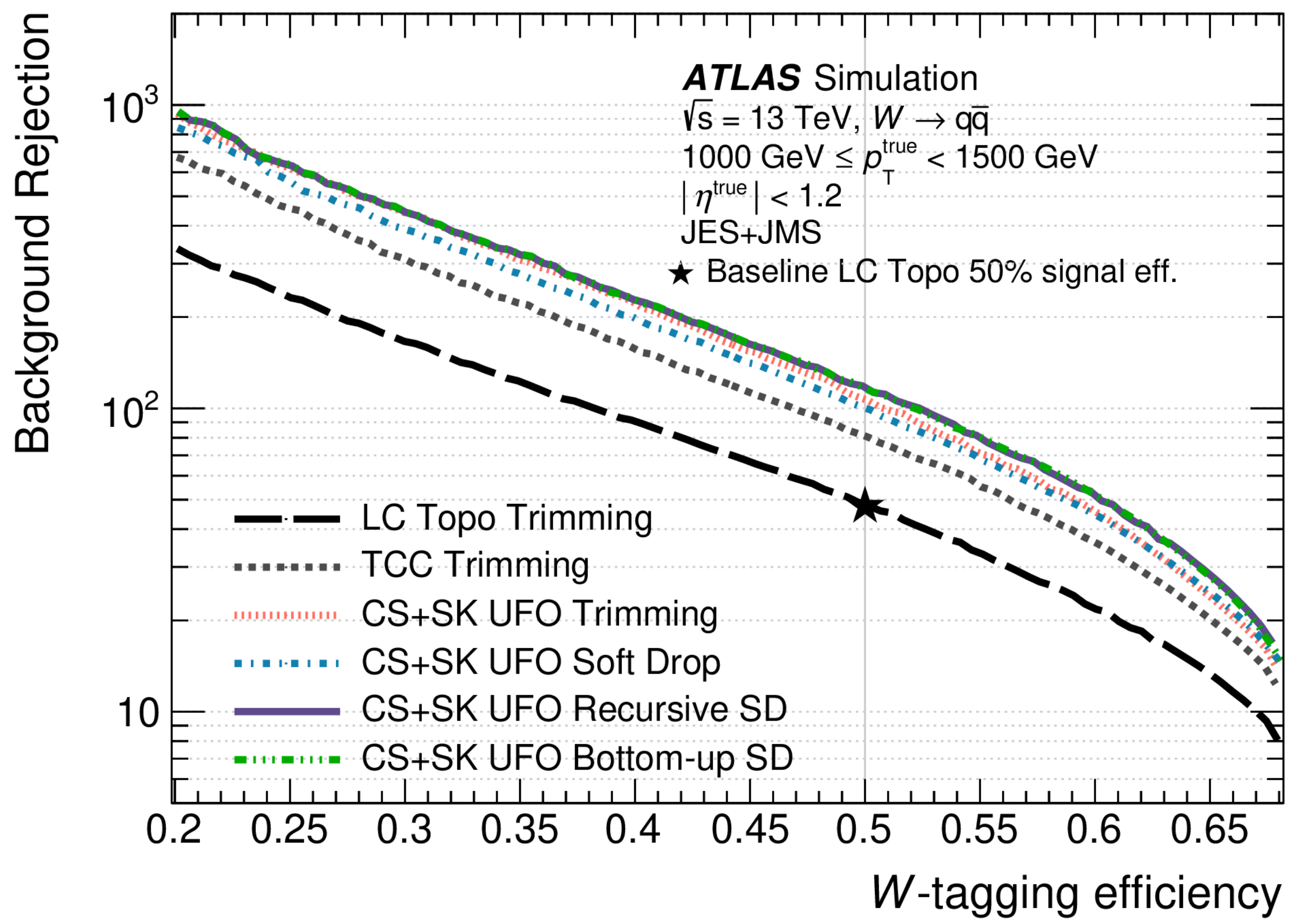
Large Radius Jets

(Eur. Phys. J. C 81 (2021) 334)



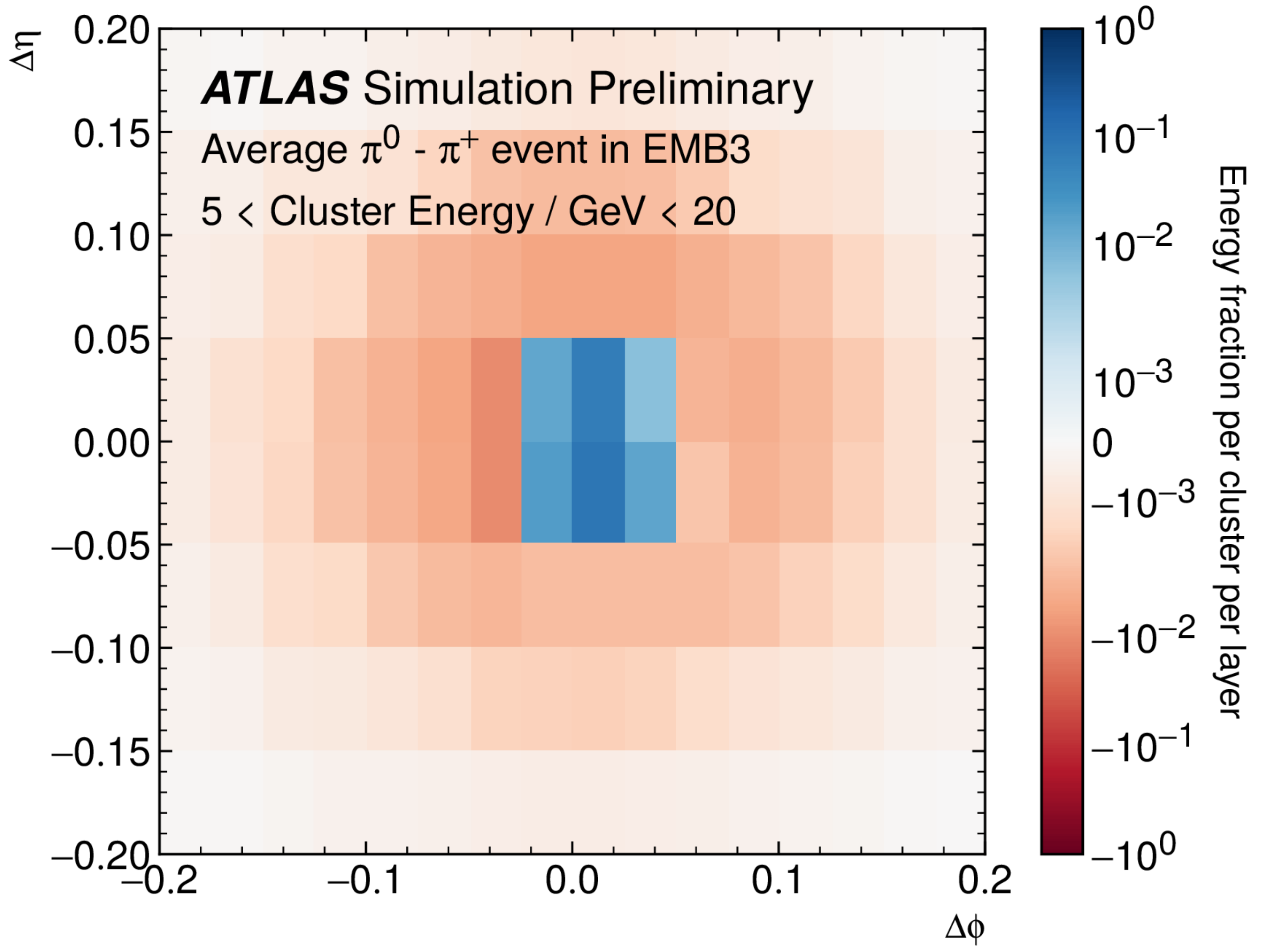
- UFO scheme performs better than TCC, Particle Flow or Topocluster inputs
  - Jet Mass resolution across large P_T range
  - Background rejection across large W-tagging efficiency range
  - Stability across large N_{PV} range

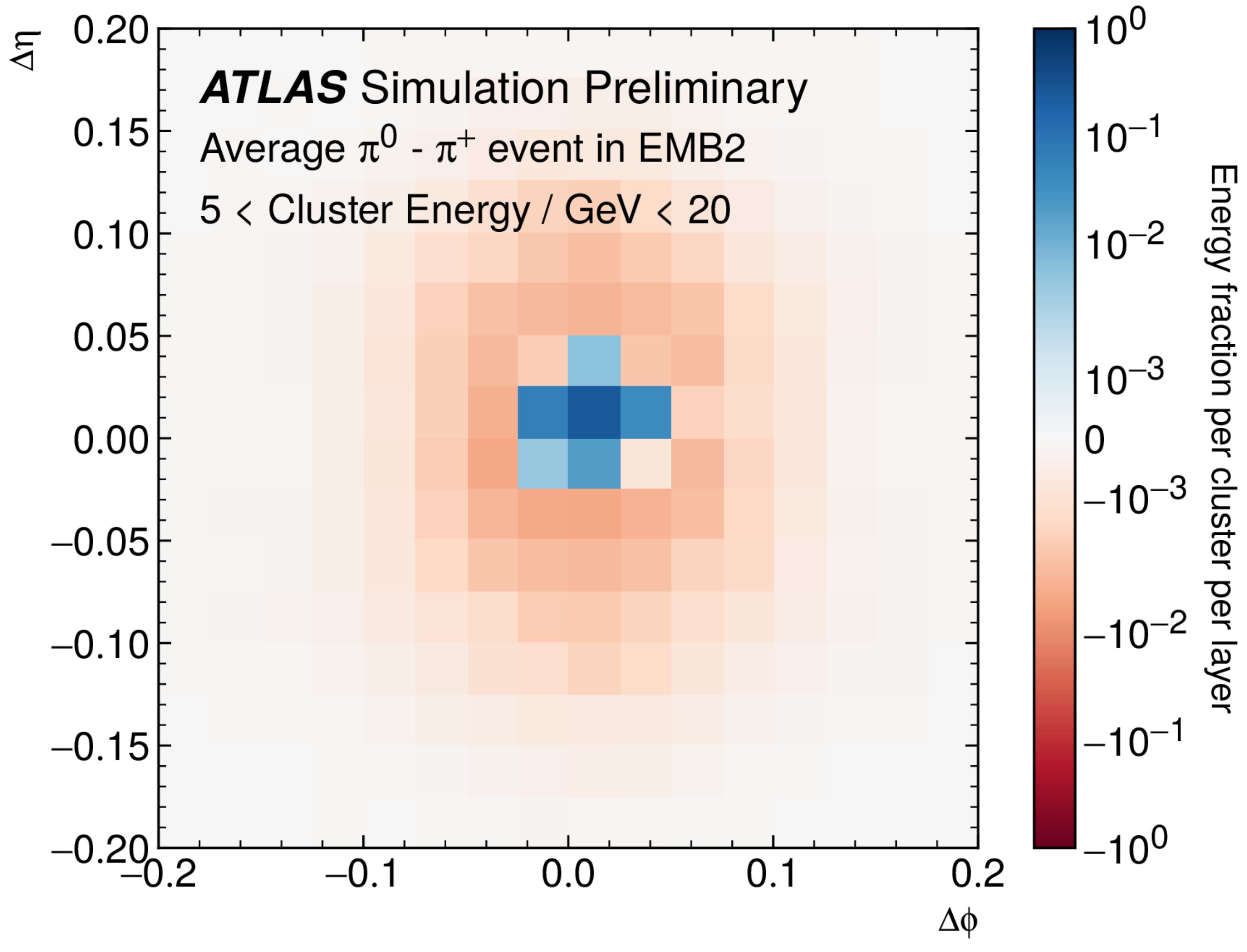




Machine Learning

ATL-PHYS-PUB-2020-018



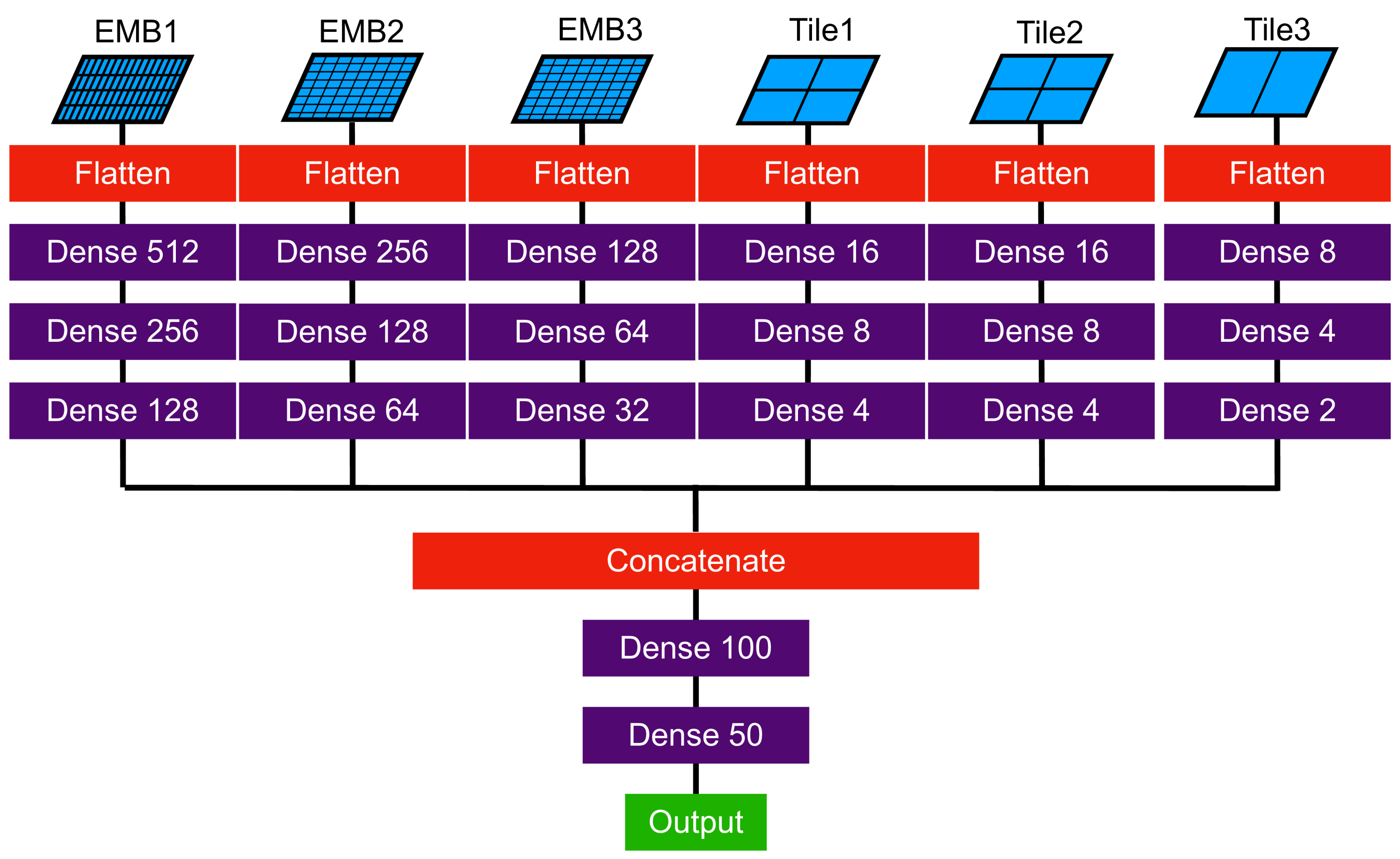


- “LC Topo” (LCW) scheme calibrates individual topoclusters via the Local Hadron Calibration , which is applied to topocluster inputs prior to input to jet finding - has been used for large radius jet finding in ATLAS.
    - Can replace topocluster inputs calibrated to LCW scale with ML calibrated topoclusters.
- Alternative calibration scheme has been studied using Machine Learning (LC)
    - Used samples of isolated charged and neutral pions, without pileup. Calorimeter cluster settings are as used in 2018 data taking conditions.
    - Have considered particles with |eta| < 0.7 (uniform detector layout)

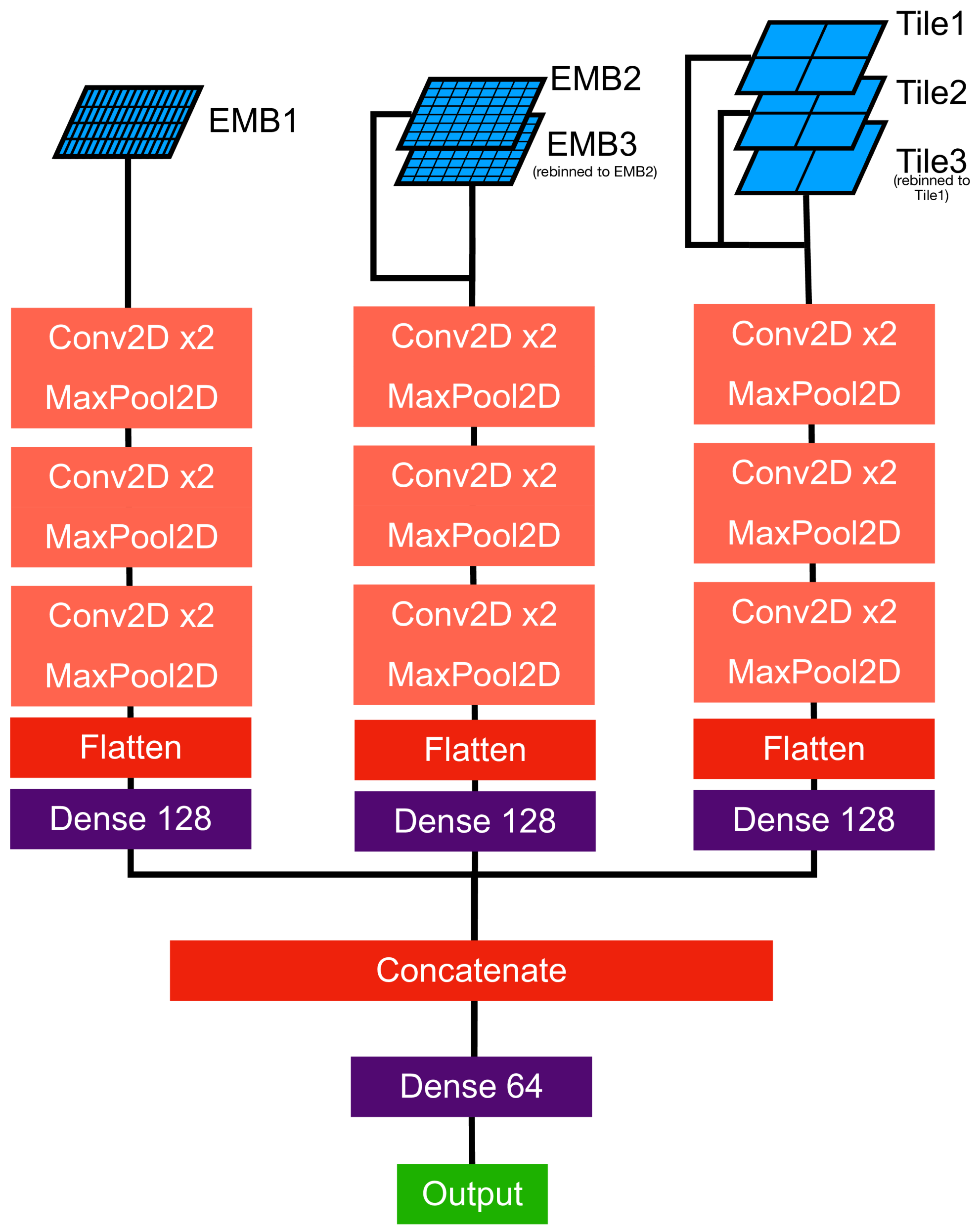
Machine Learning

ATL-PHYS-PUB-2020-018

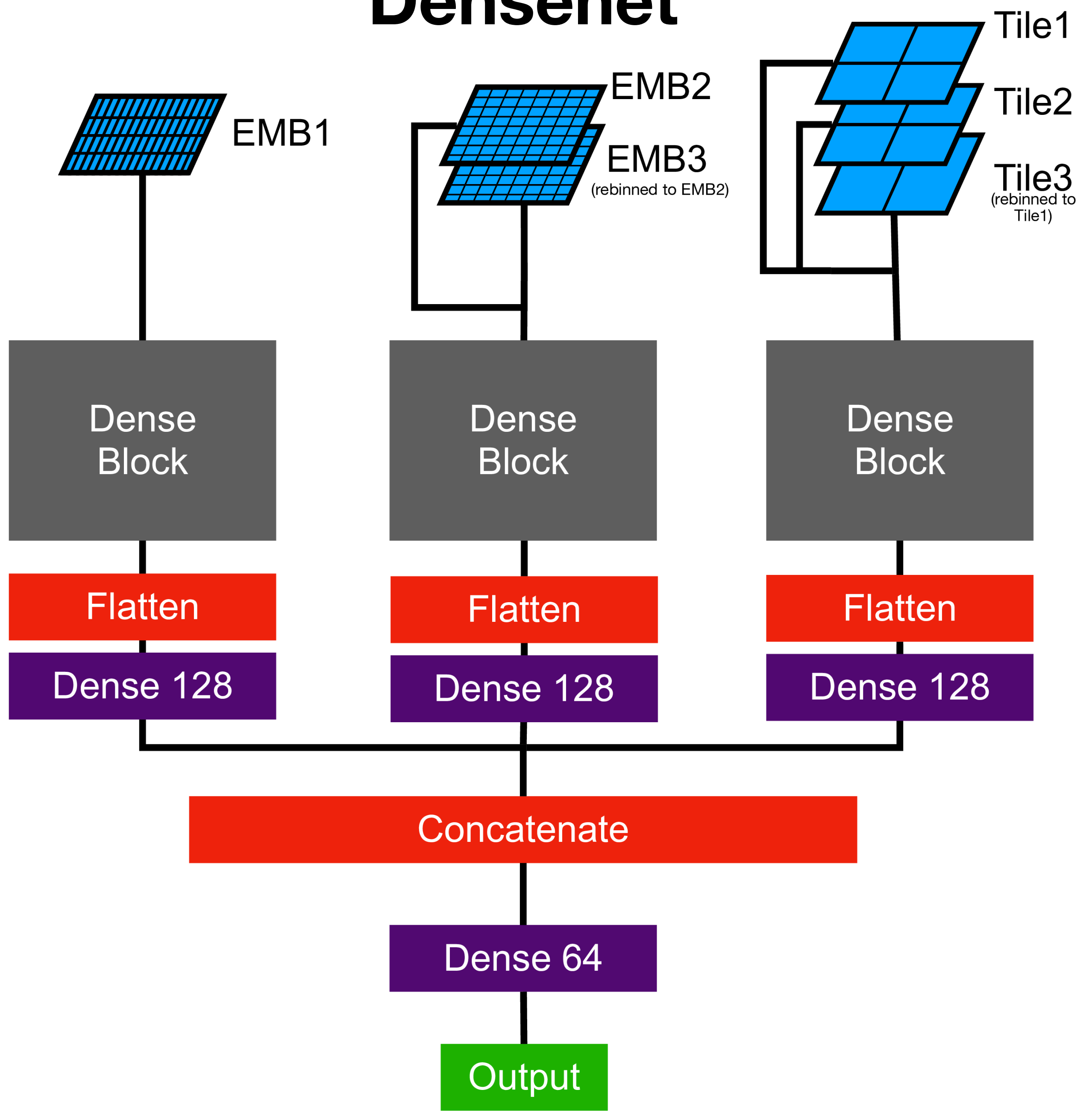
DNN Classifier



CNN Classifier



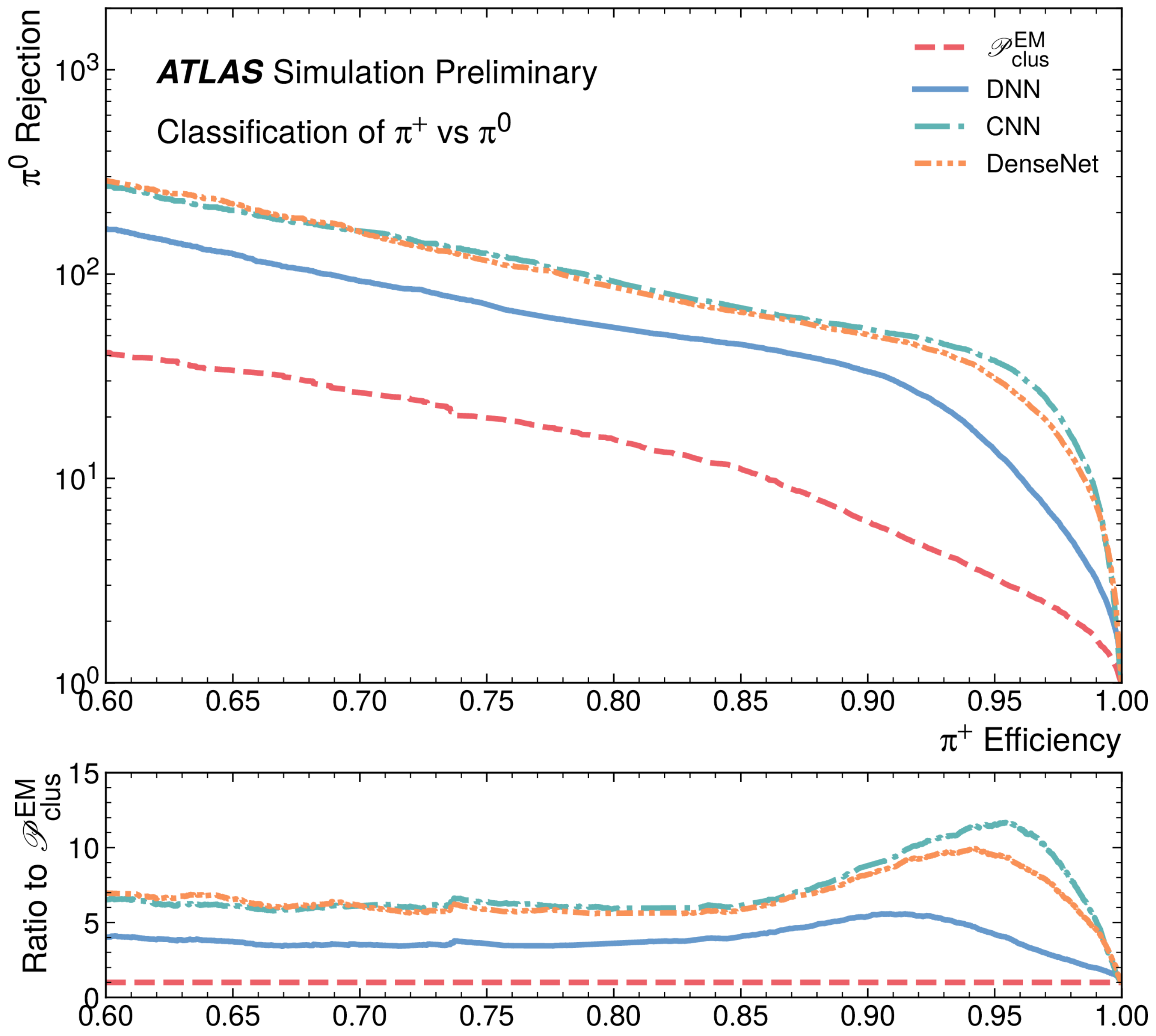
Densenet

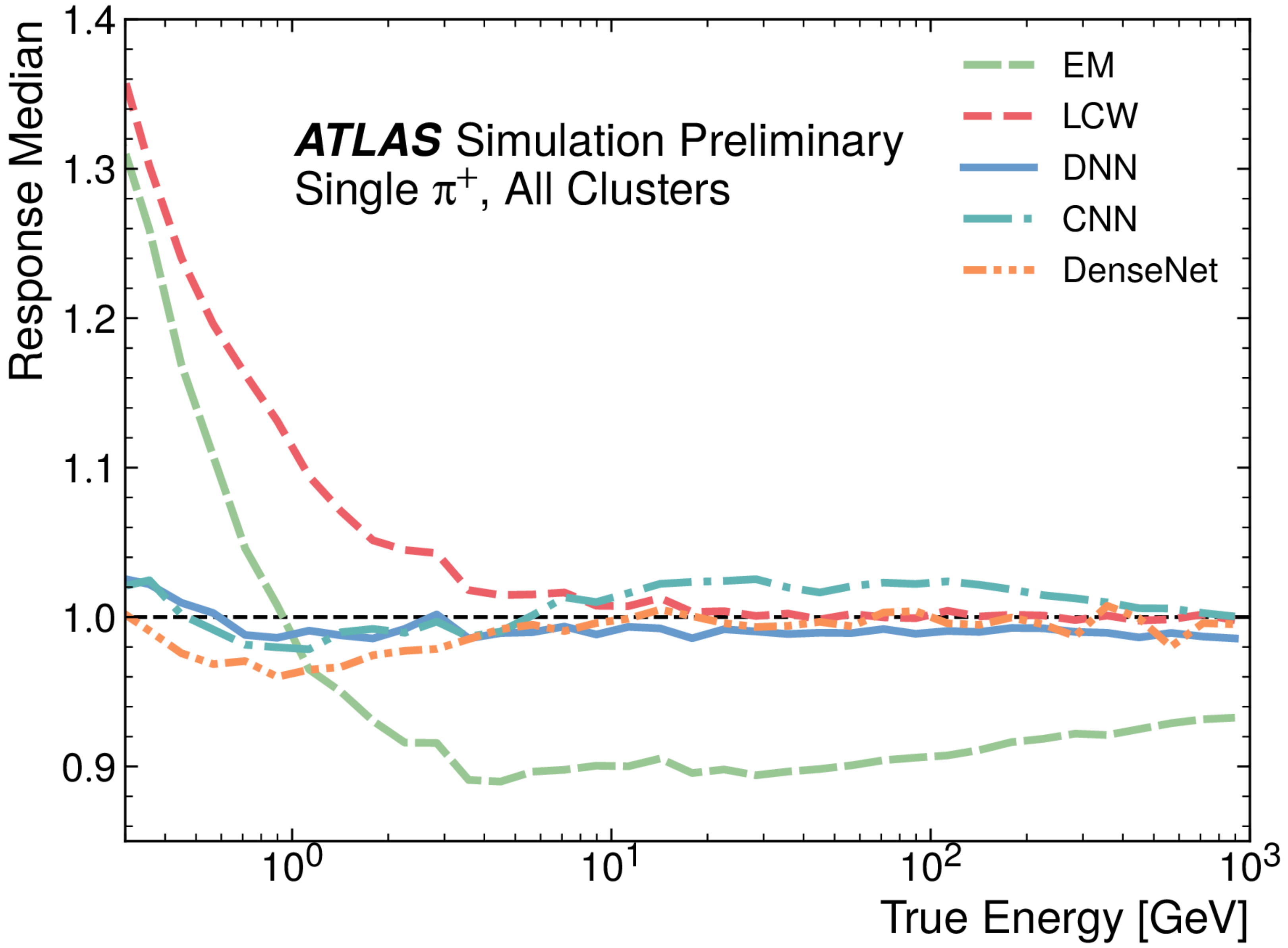


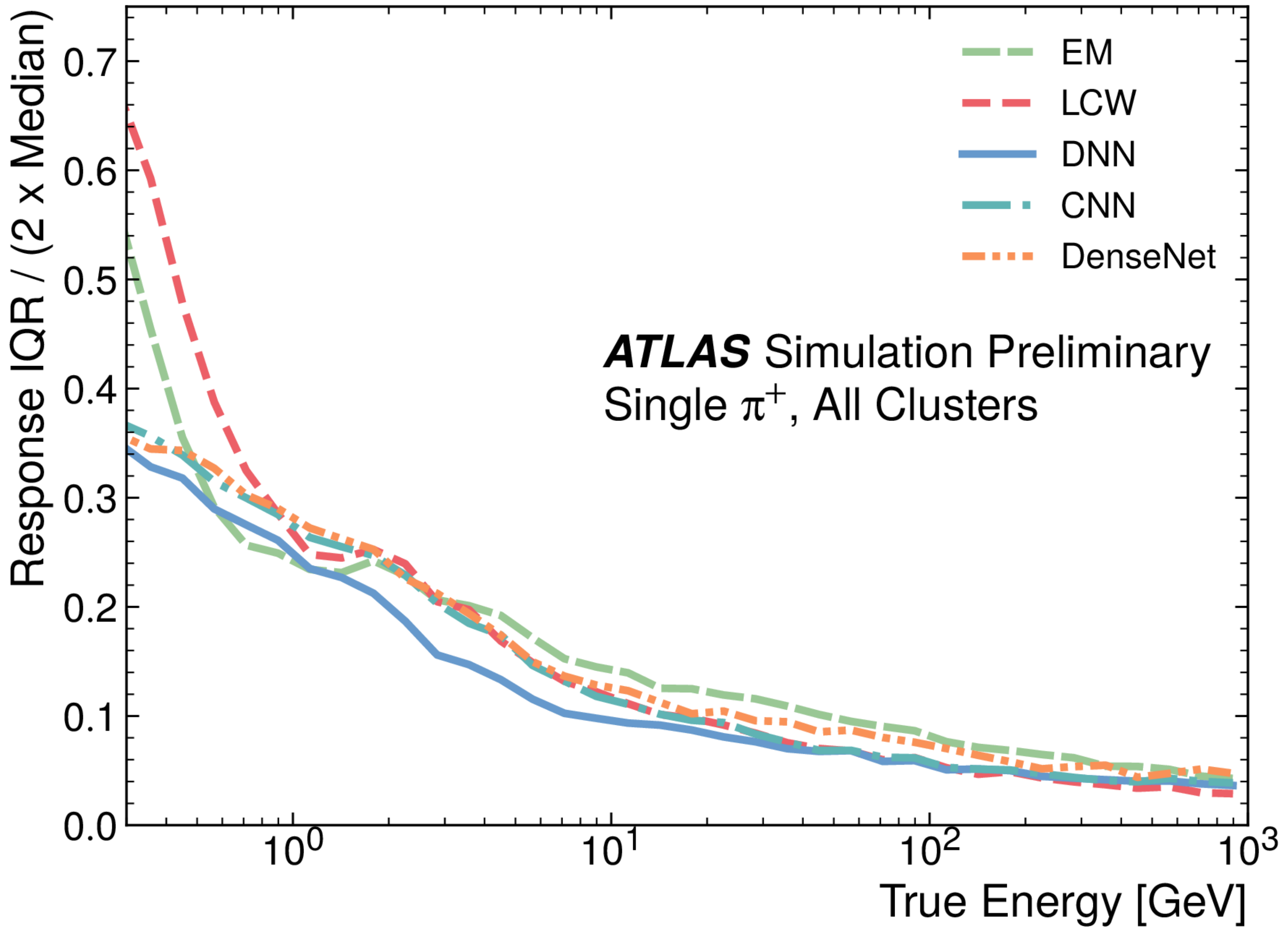
- Deep Neutral Network (DNN), Convolutional Neural Network (CNN) and Densely Connected Convolution Network (DenseNet) have been studied.
- Currently ATLAS LCW scheme uses a Likelihood:
    - Classification step using Likelihood ratio, making use of the cluster energy, eta position, longitudinal depth and average cell energy density.
    - Calibration step deploys calorimeter cell signal weighting which depend on cluster energy and location.
    - The Machine Learning schemes also do both classification and regression.

Machine Learning

ATL-PHYS-PUB-2020-018



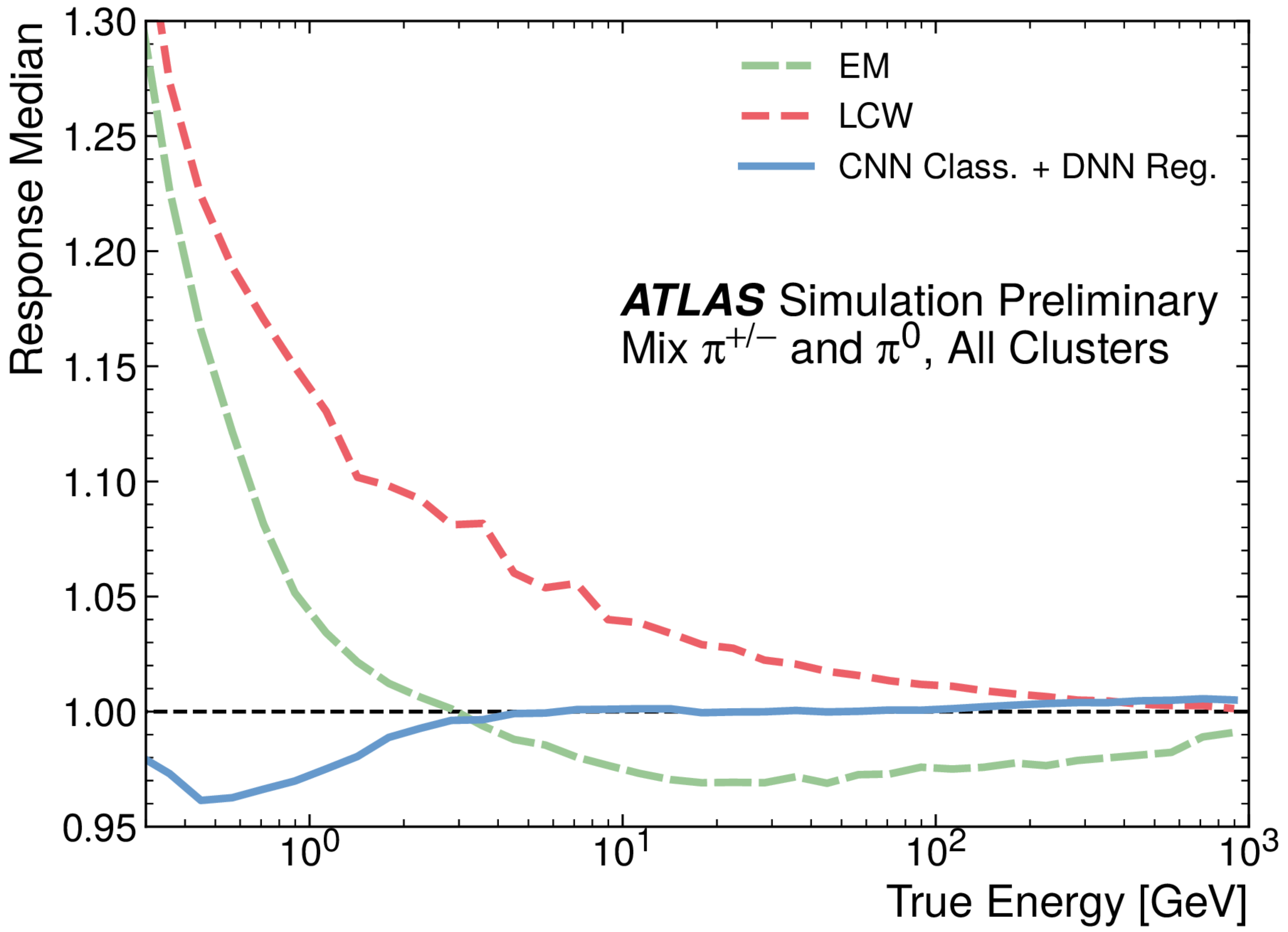


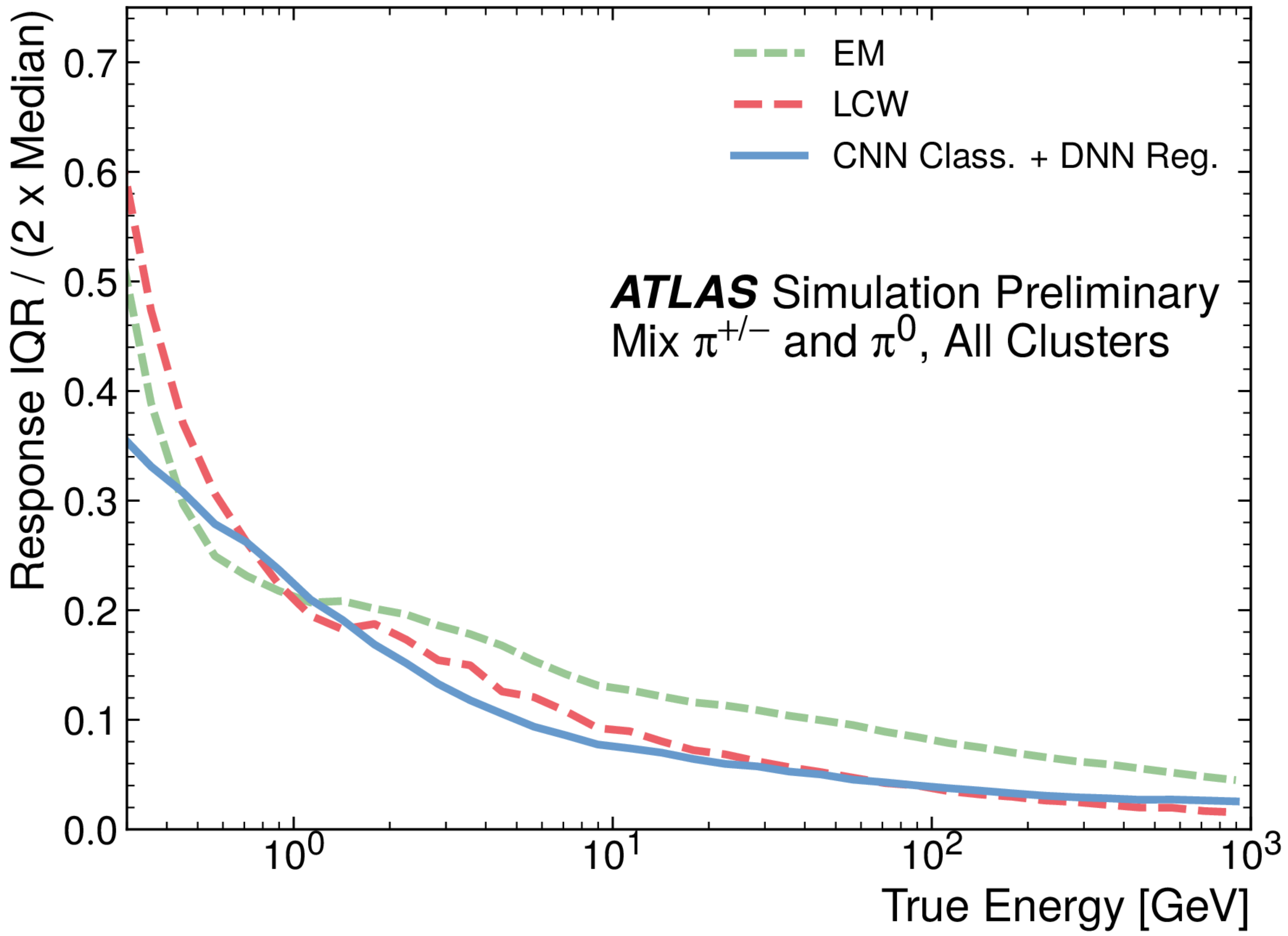


- For the classification problem, shown on the left, all three schemes perform better than the LCW scheme (ρ_{EM}^{clus})
  - DNN not as good as CNN, Densenet.
- For the regression problem, shown in the right two plots, all three schemes perform better than the LCW scheme.
  - DNN gives best resolution and has good linearity.

Machine Learning

ATL-PHYS-PUB-2020-018





- Combined classification and regression test:
  - Compare LCW to combination of CNN Classifier (best) and DNN regression (best)
  - High performance of CNN classifier ensures that the correct energy regression is applied in this mixed particle sample.

Conclusions

- ATLAS reconstruction matches calorimeter clusters and ID tracks in different ways for different environments.
    - Particle Flow and UFO algorithms
- Particle Flow improves jet performance in low P_T regime and leads to similar uncertainties on JES and JER.
- UFO scheme gives best large radius jet mass resolution, tagging efficiency and pileup stability across a wide phase space.
- Machine learning approach to calorimeter cluster calibration improves performance compared to existing LCW scheme.

Extras

Jet Vertex Fraction (JVT)

The quantity corrJVF is a variable similar to JVF, but corrected for the N_{PV} dependence. It is defined as

$$\mathrm{corrJVF} = \frac{\sum_m p_{\mathrm{T},m}^{\mathrm{track}}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T},l}^{\mathrm{track}}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_{\mathrm{T},l}^{\mathrm{track}}(\mathrm{PV}_n)}{(k \cdot n_{\mathrm{track}}^{\mathrm{PU}})}}. \quad (9)$$

Eur. Phys. J. C (2016) 76:581

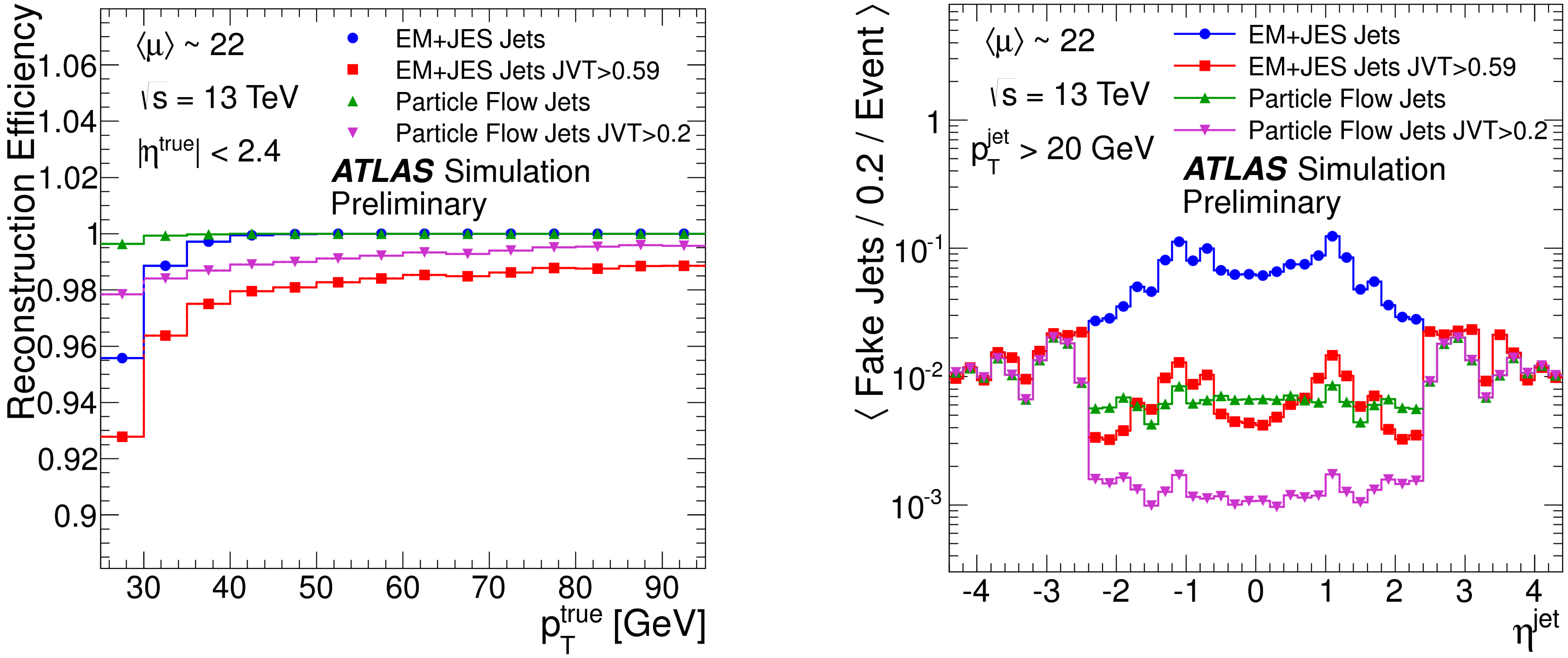
where $\sum_m p_{\mathrm{T},m}^{\mathrm{track}}(\mathrm{PV}_0)$ is the scalar sum of the p_{T} of the tracks that are associated with the jet and originate from the hard-scatter vertex. The term $\sum_{n\geq 1}\sum_l p_{\mathrm{T},l}^{\mathrm{track}}(\mathrm{PV}_n) = p_{\mathrm{T}}^{\mathrm{PU}}$ denotes the scalar sum of the p_{T} of the associated tracks that originate from any of the pile-up interactions.

The variable $R_{p\mathrm{T}}$ is defined as the scalar sum of the p_{T} of the tracks that are associated with the jet and originate from the hard-scatter vertex divided by the fully calibrated jet p_{T}, which includes pile-up subtraction:

$$R_{p\mathrm{T}} = \frac{\sum_k p_{\mathrm{T},k}^{\mathrm{track}}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{\mathrm{jet}}}. \quad (10)$$

A new discriminant called the jet-vertex-tagger (JVT) is constructed using $R_{p\mathrm{T}}$ and corrJVF as a two-dimensional likelihood derived using simulated dijet events and based on a k-nearest neighbour (kNN) algorithm [58]. For each point in the two-dimensional corrJVF–$R_{p\mathrm{T}}$ plane, the relative probability for a jet at that point to be of signal type is computed as the ratio of the number of hard-scatter jets to the number of hard-scatter plus pile-up jets found in a local neighbourhood around the point using a training sample of signal and pile-up jets with $20 < p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$. The local neighbourhood is defined dynamically as the 100 nearest neighbours around the test point using a Euclidean metric in the $R_{p\mathrm{T}}$–corrJVF space, where corrJVF and $R_{p\mathrm{T}}$ are rescaled so that the variables have the same range.

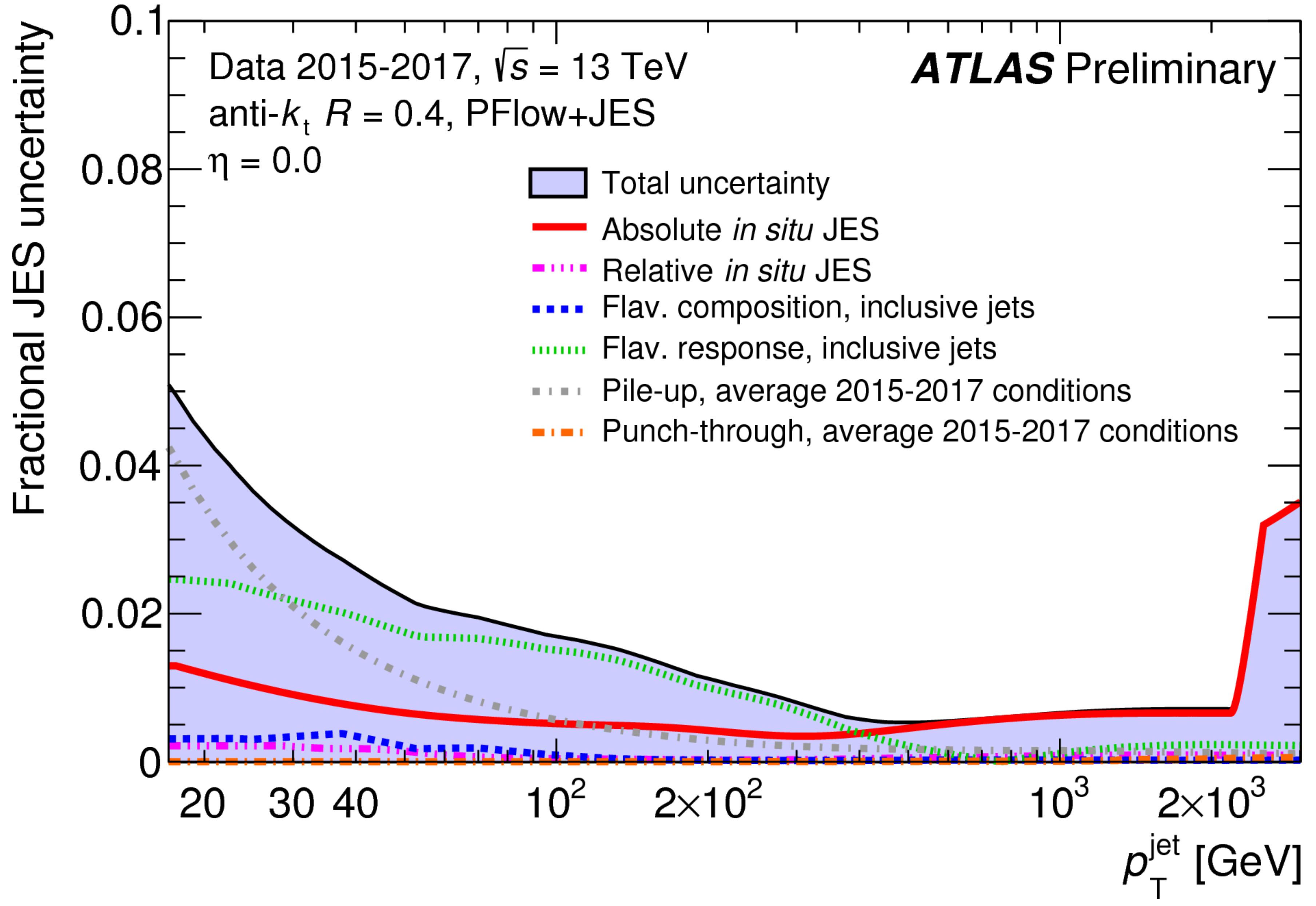
Jet Vertex Fraction (JVT)

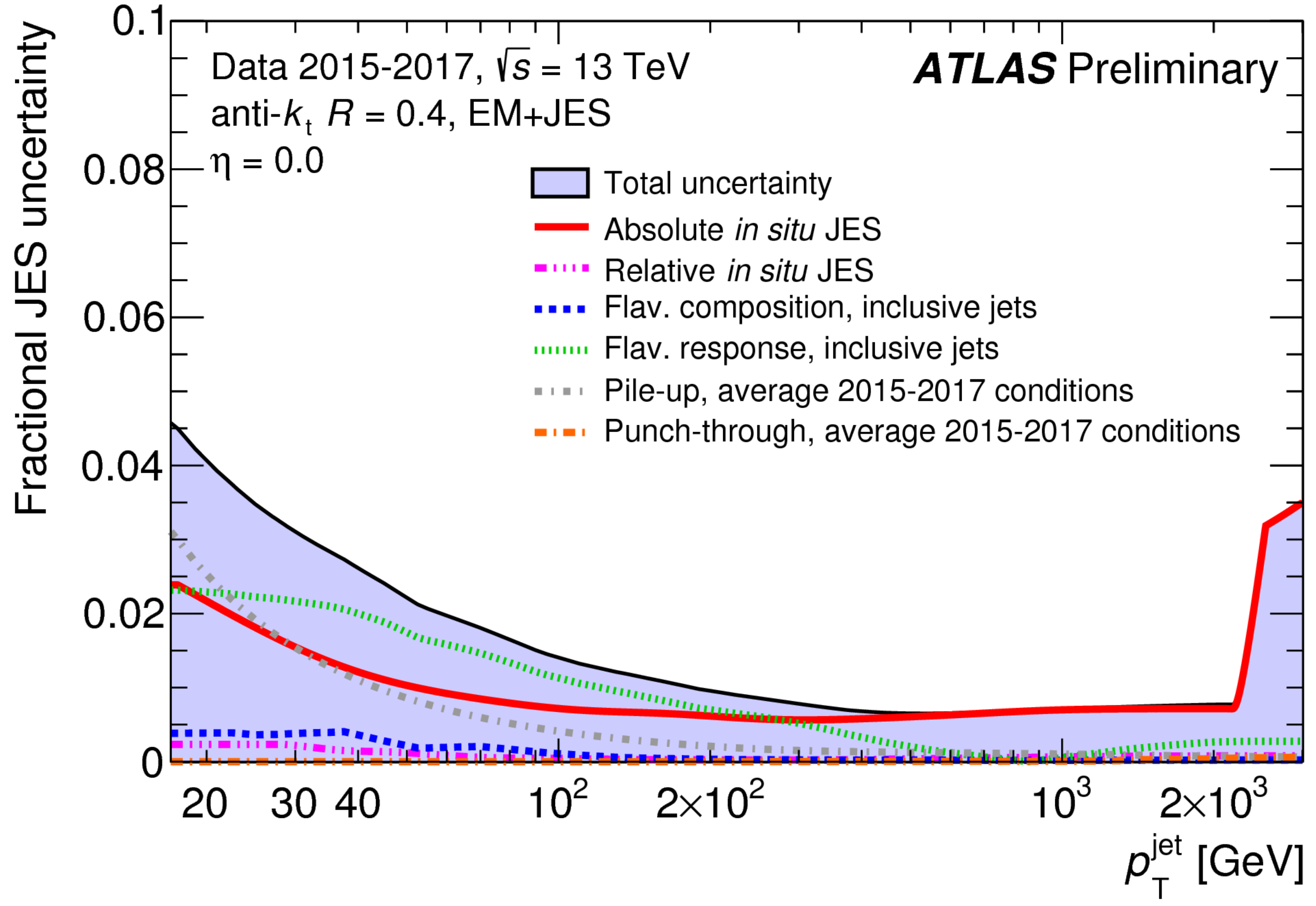


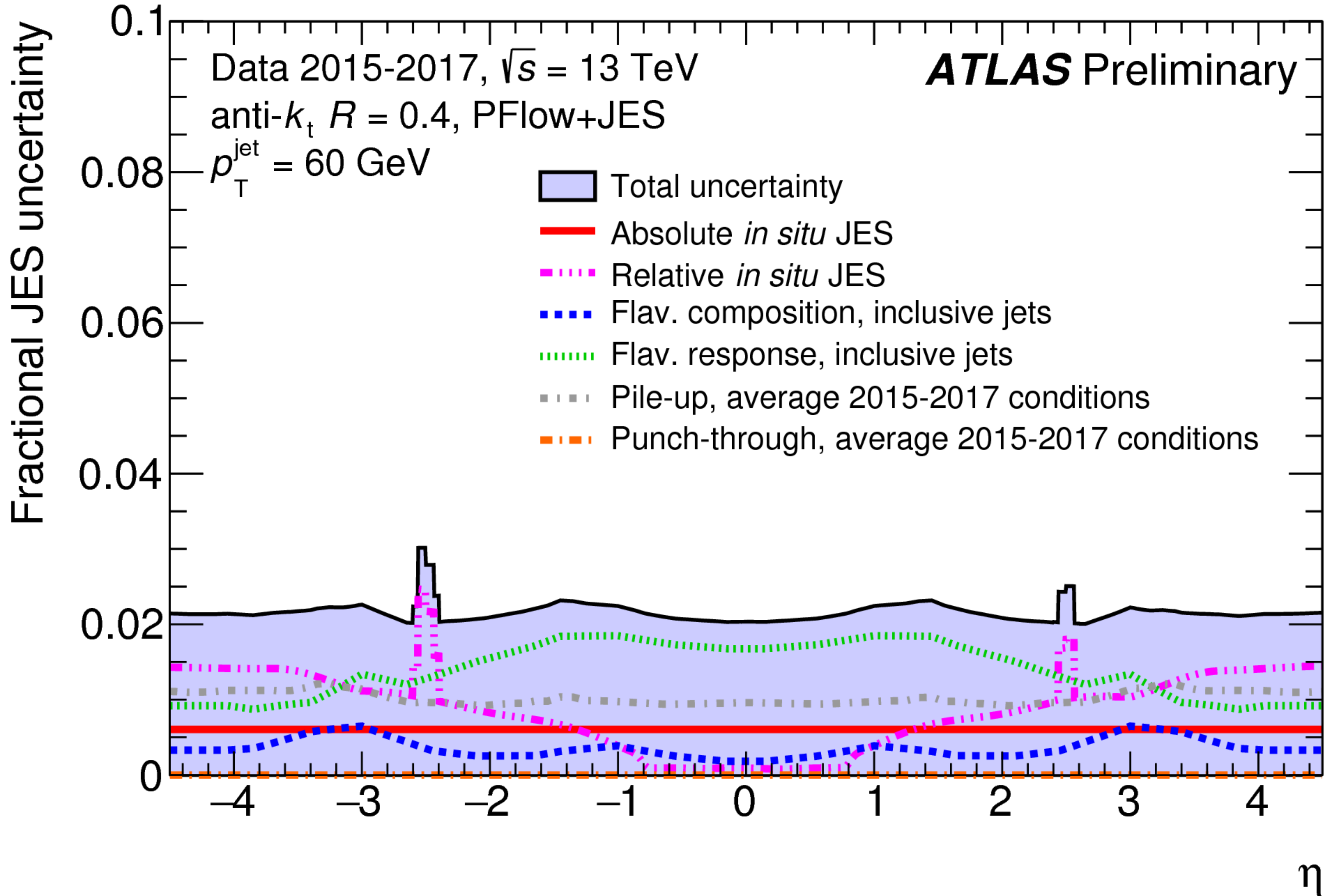
ATLAS-JETM-2017-006

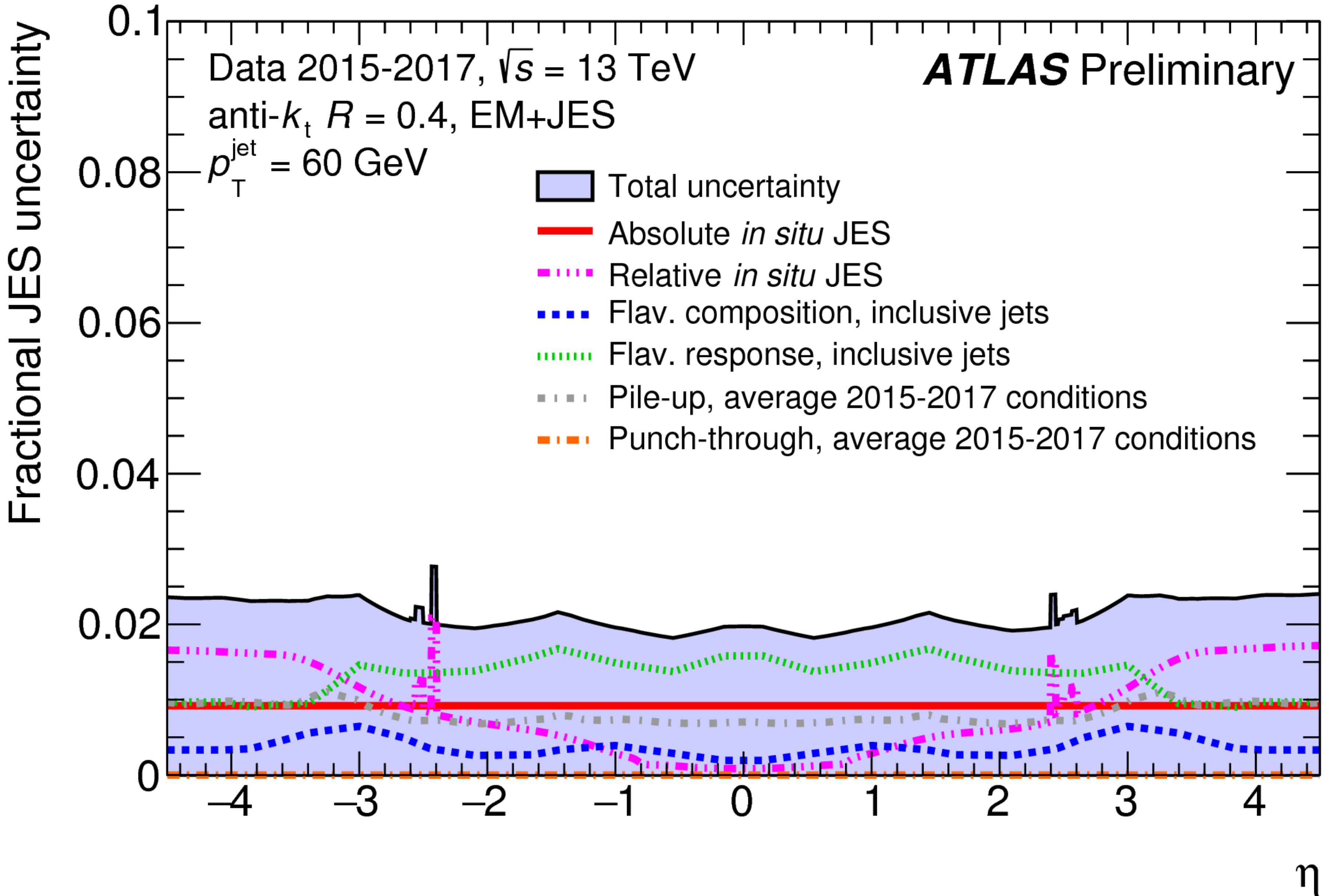
Particle Flow jets have similar hard scatter efficiency to calorimeter jets (left), whilst reconstructing fewer fake jets (right).

JES Uncertainties





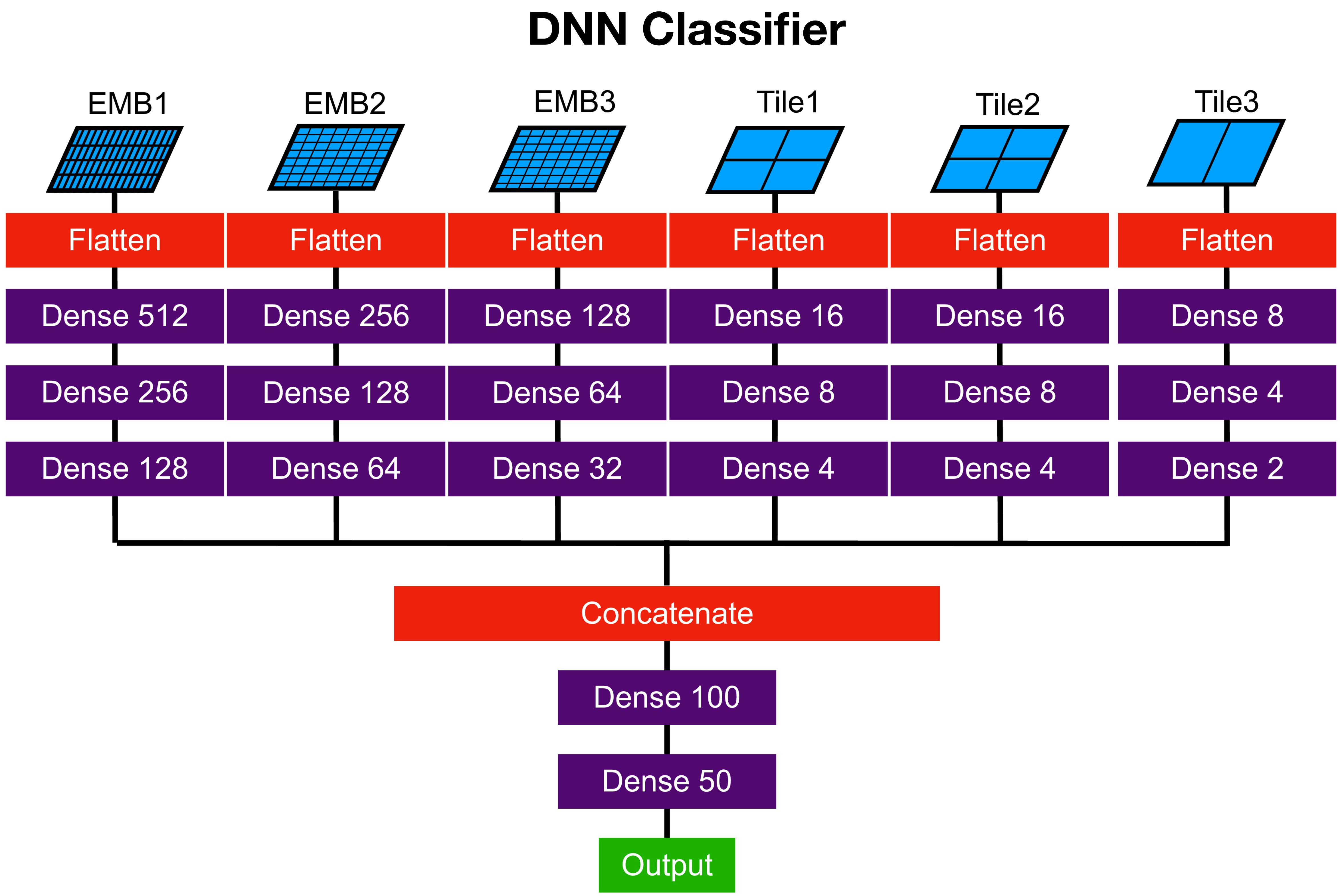


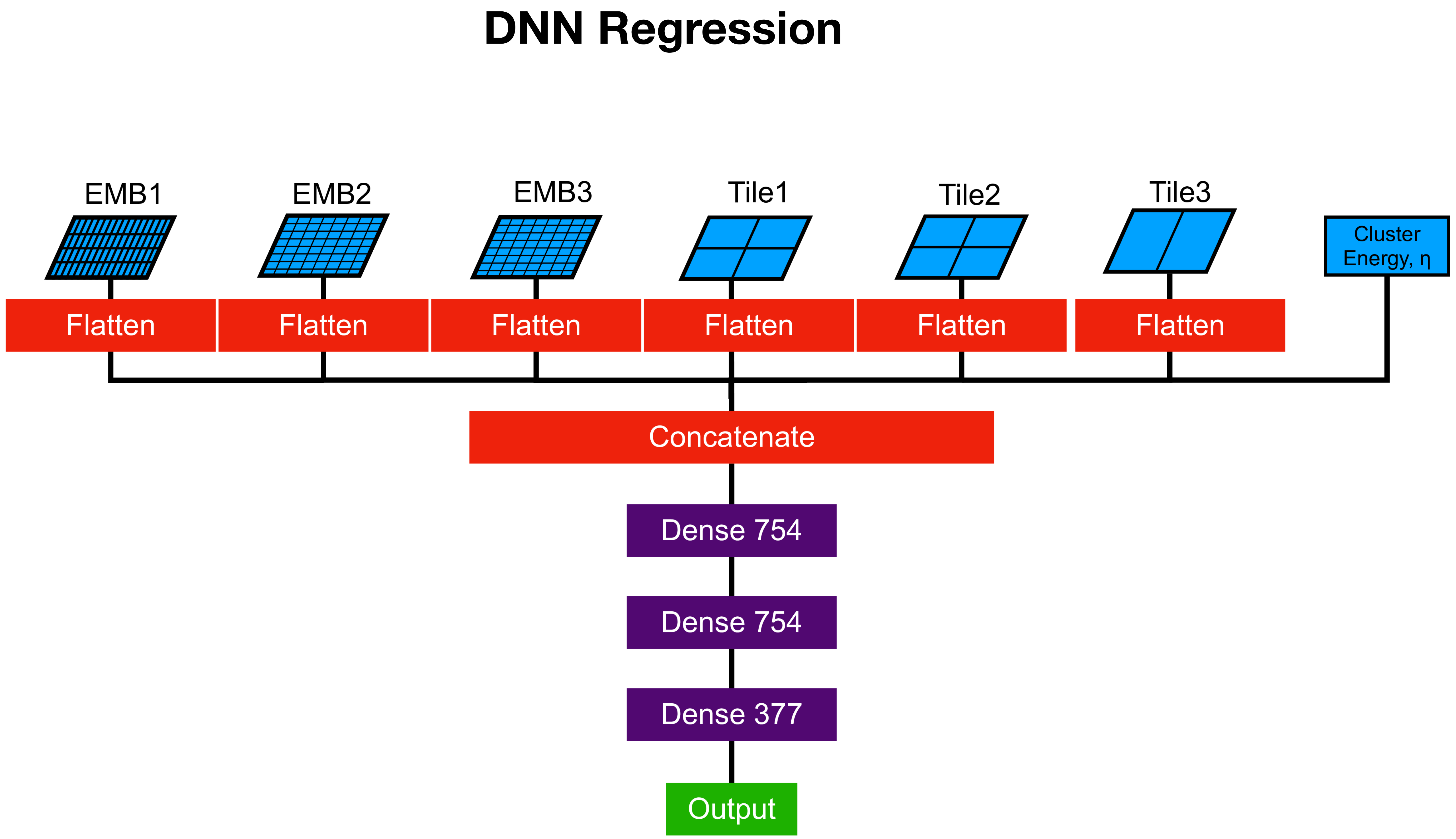


ML

Calorimeter Layer	$\Delta\eta$ Granularity	$\Delta\phi$ Granularity	Interaction Lengths
EMB1	$0.025/8 = 0.003125$	$\pi/32 \approx 0.1$	$\approx 4X_0$
EMB2	0.025	$\pi/128 \approx 0.025$	$\approx 16X_0$
EMB3	0.05	$\pi/128 \approx 0.025$	$\approx 2X_0$
Tile0	0.1	$\pi/32 \approx 0.1$	$\approx 1.5\lambda$
Tile1	0.1	$\pi/32 \approx 0.1$	$\approx 4\lambda$
Tile2	0.2	$\pi/32 \approx 0.1$	$\approx 2\lambda$

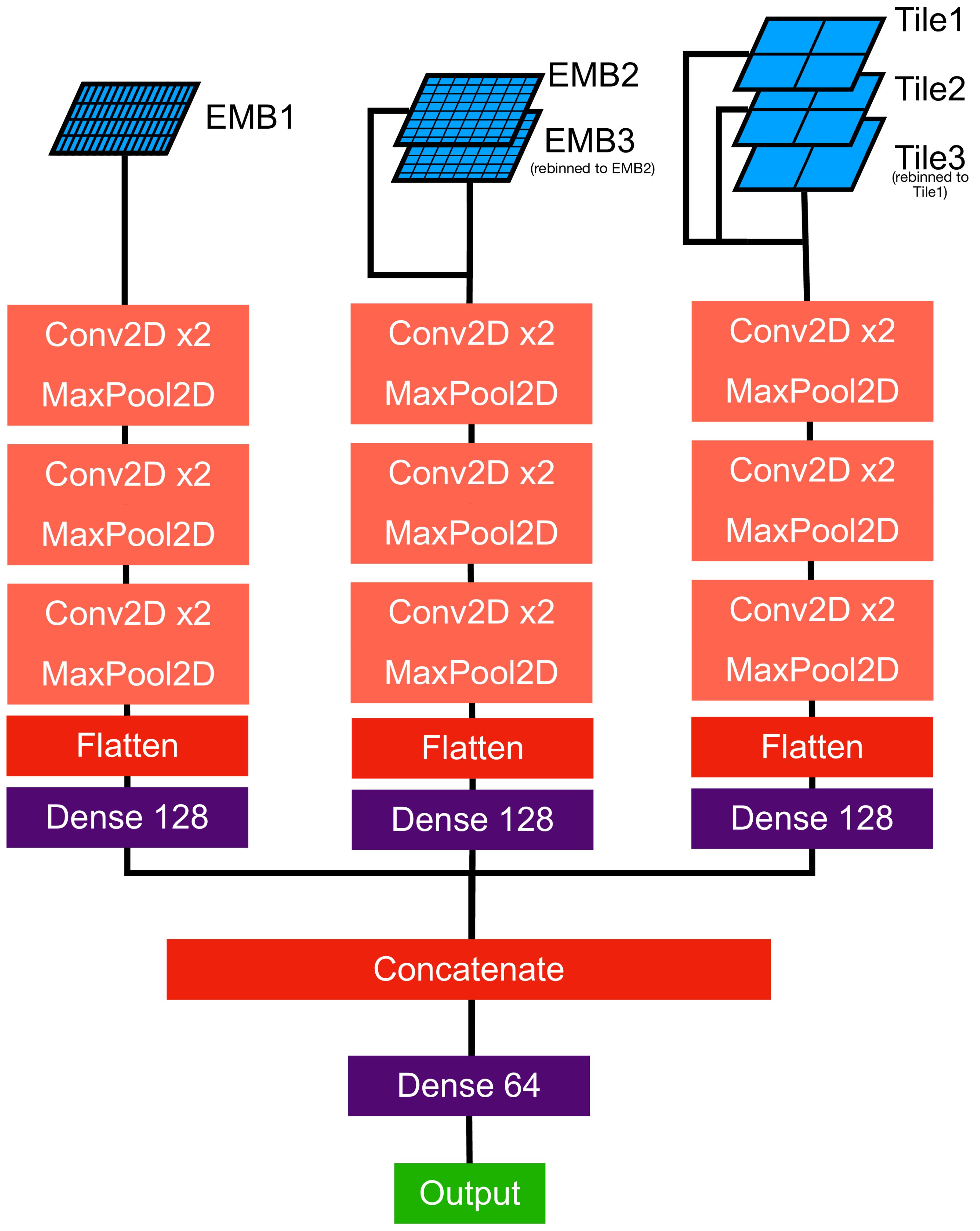
ML


DNN Classifier
EMB1
EMB2
EMB3
Tile1
Tile2
Tile3
Flatten
Flatten
Flatten
Flatten
Flatten
Flatten
Dense 512
Dense 256
Dense 128
Dense 16
Dense 16
Dense 8
Dense 256
Dense 128
Dense 64
Dense 8
Dense 8
Dense 4
Dense 128
Dense 64
Dense 32
Dense 4
Dense 4
Dense 2
Concatenate
Dense 100
Dense 50
Output

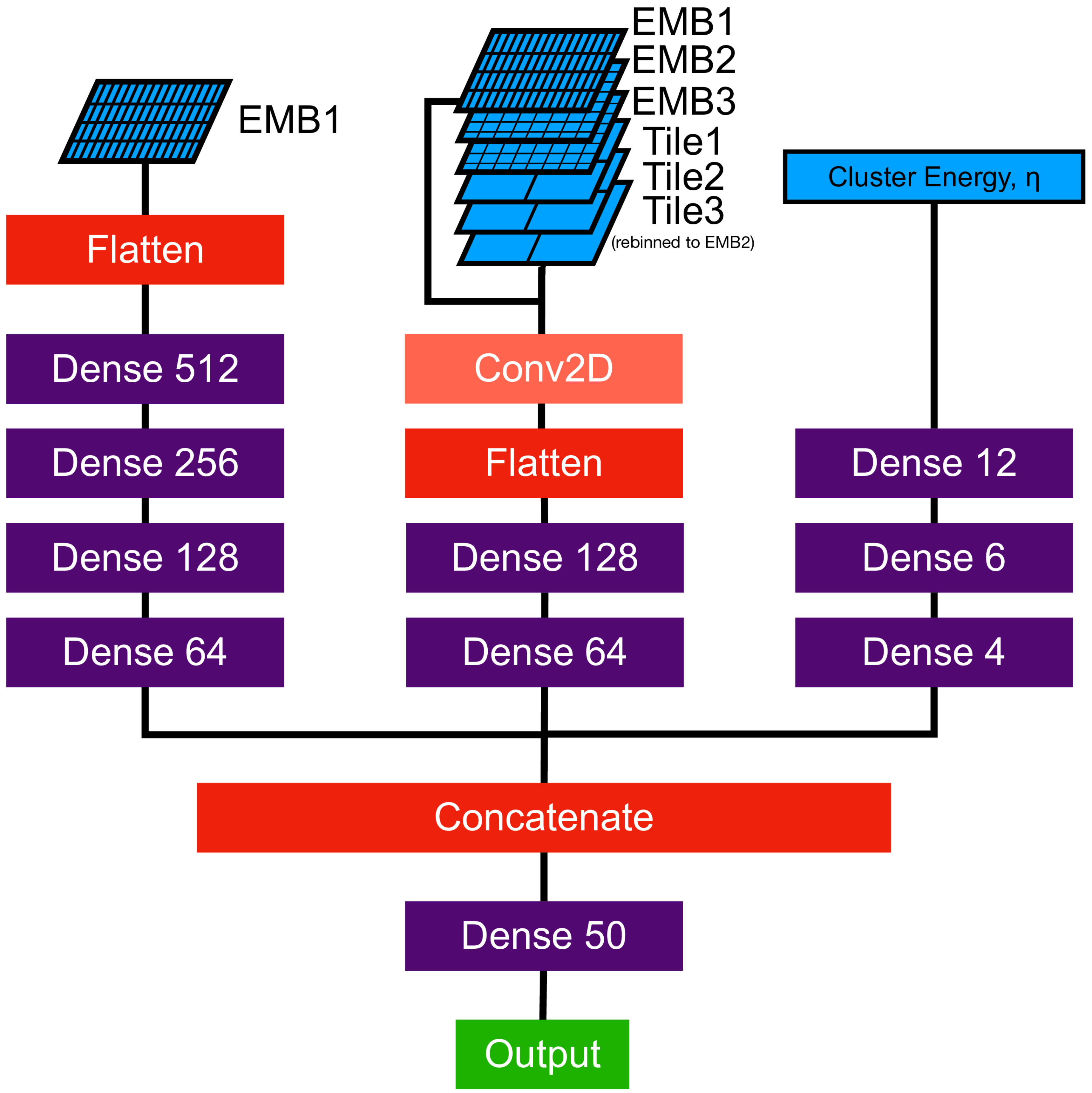

DNN Regression
EMB1
EMB2
EMB3
Tile1
Tile2
Tile3
Cluster Energy, η
Flatten
Flatten
Flatten
Flatten
Flatten
Flatten
Concatenate
Dense 754
Dense 754
Dense 377
Output

ML

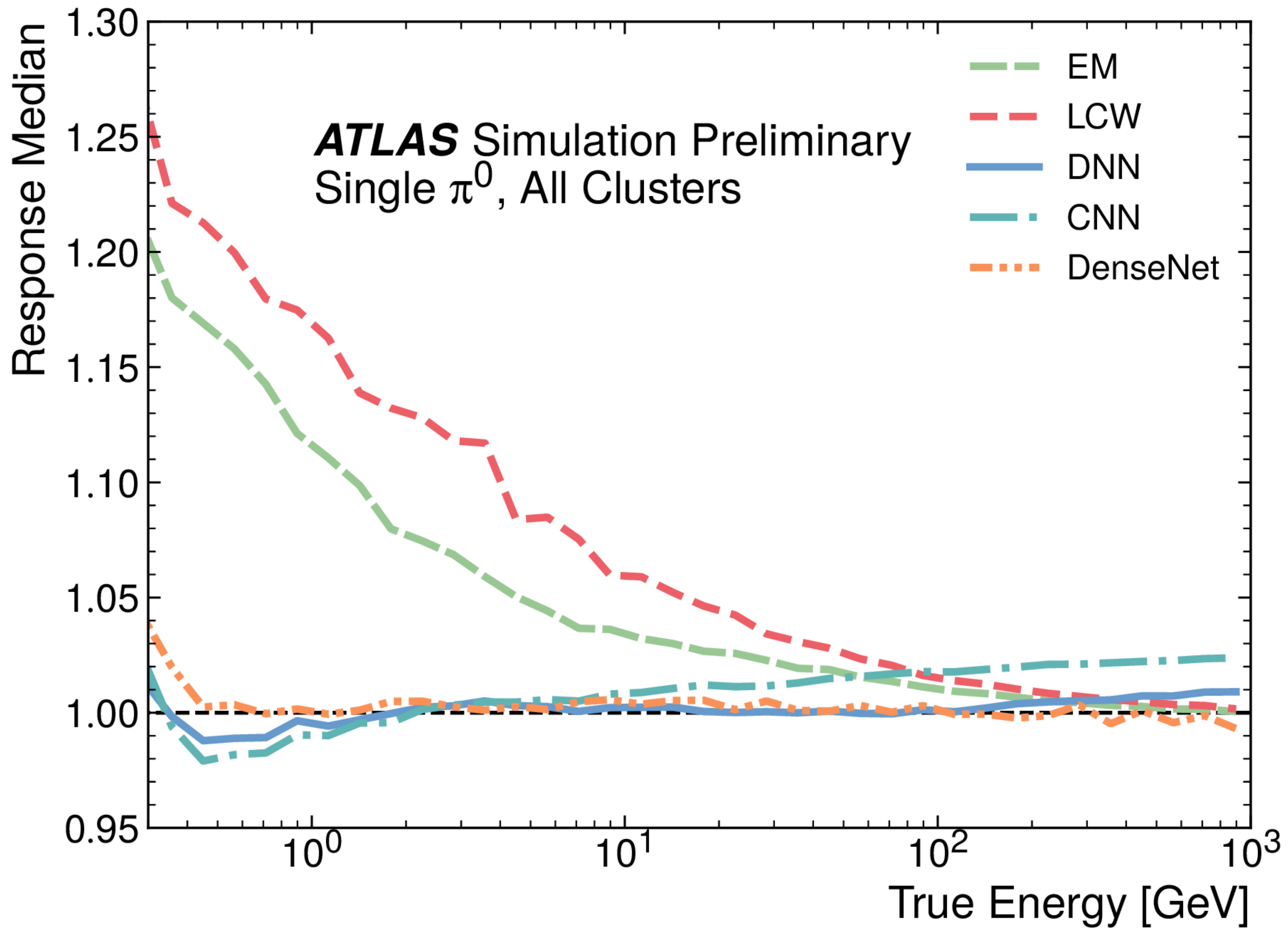
CNN Classifier

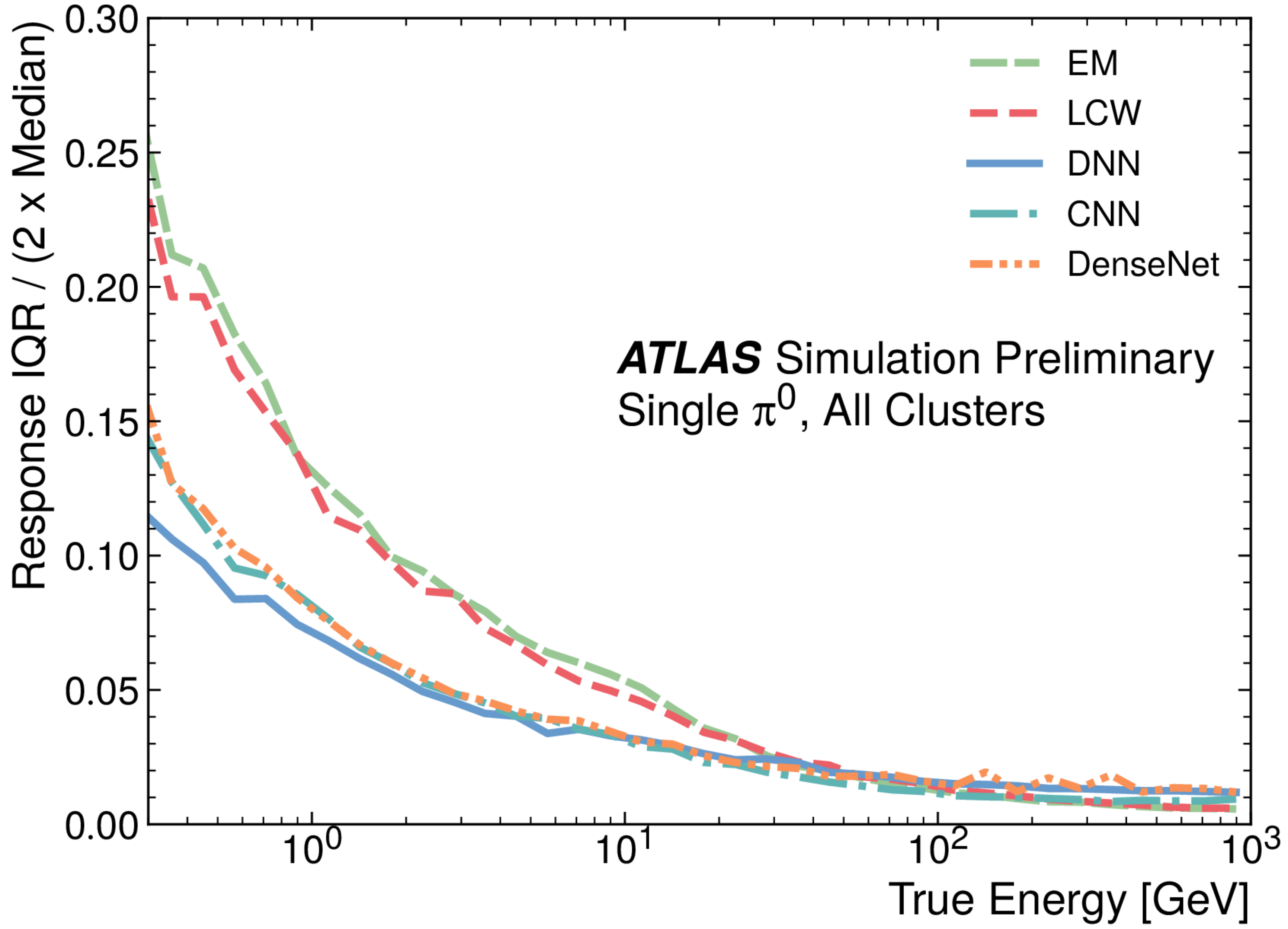


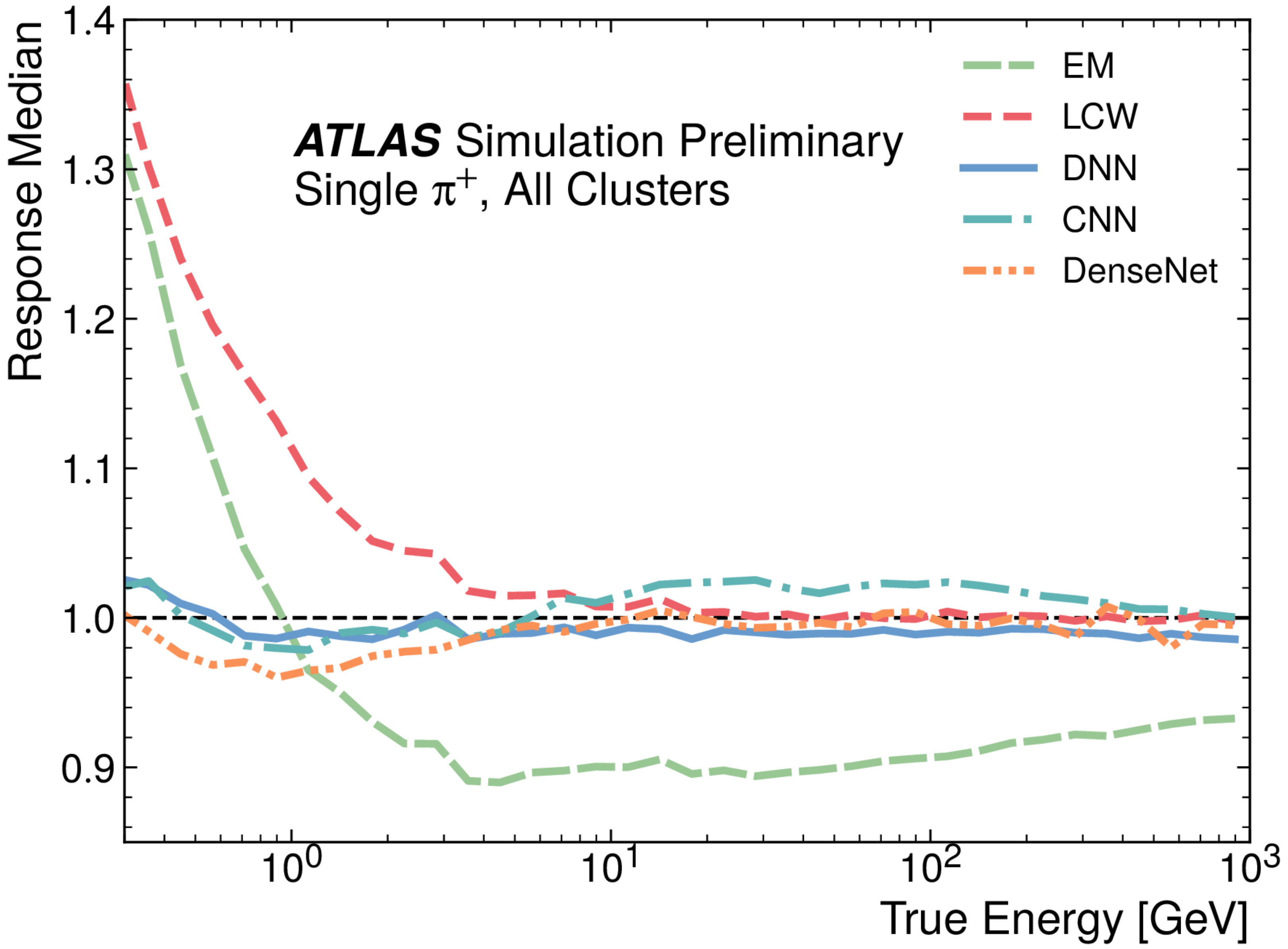
CNN Regression

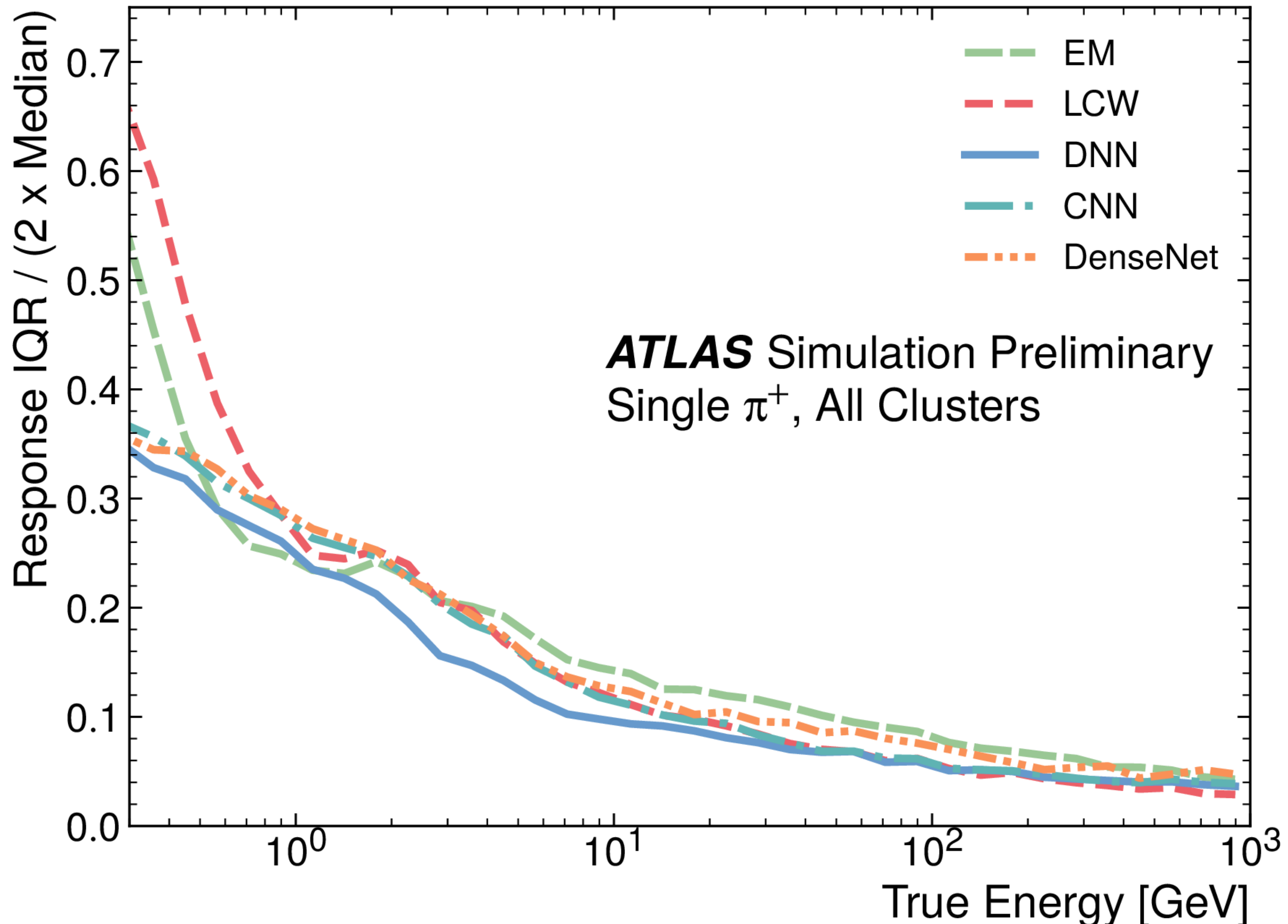


ML


ATLAS Simulation Preliminary
Single π0, All Clusters
EM
LCW
DNN
CNN
DenseNet
Response Median
True Energy [GeV]

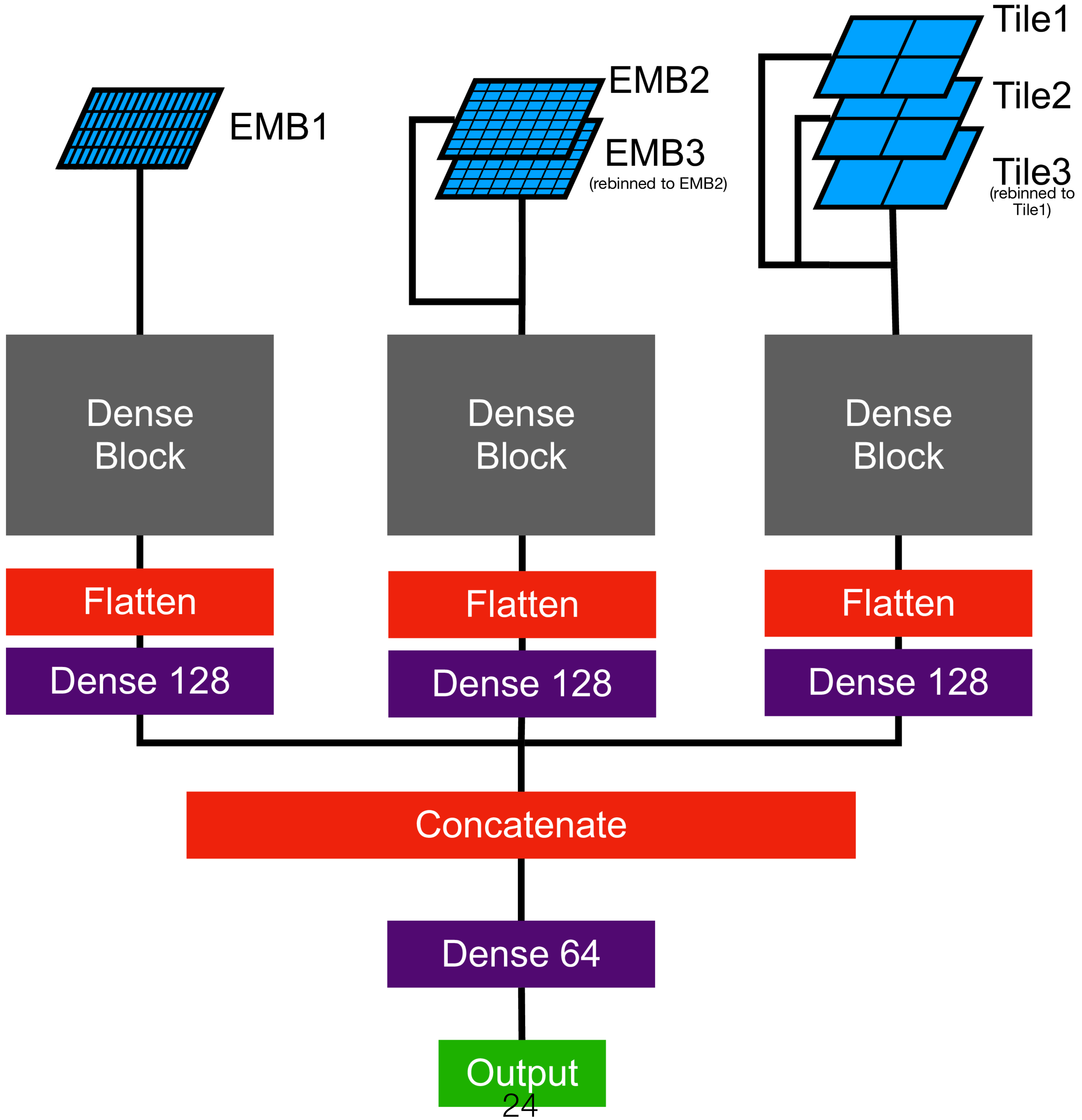

ATLAS Simulation Preliminary
Single π0, All Clusters
EM
LCW
DNN
CNN
DenseNet
Response IQR / (2 x Median)
True Energy [GeV]


ATLAS Simulation Preliminary
Single π+, All Clusters
EM
LCW
DNN
CNN
DenseNet
Response Median
True Energy [GeV]


ATLAS Simulation Preliminary
Single π+, All Clusters
EM
LCW
DNN
CNN
DenseNet
Response IQR / (2 x Median)
True Energy [GeV]

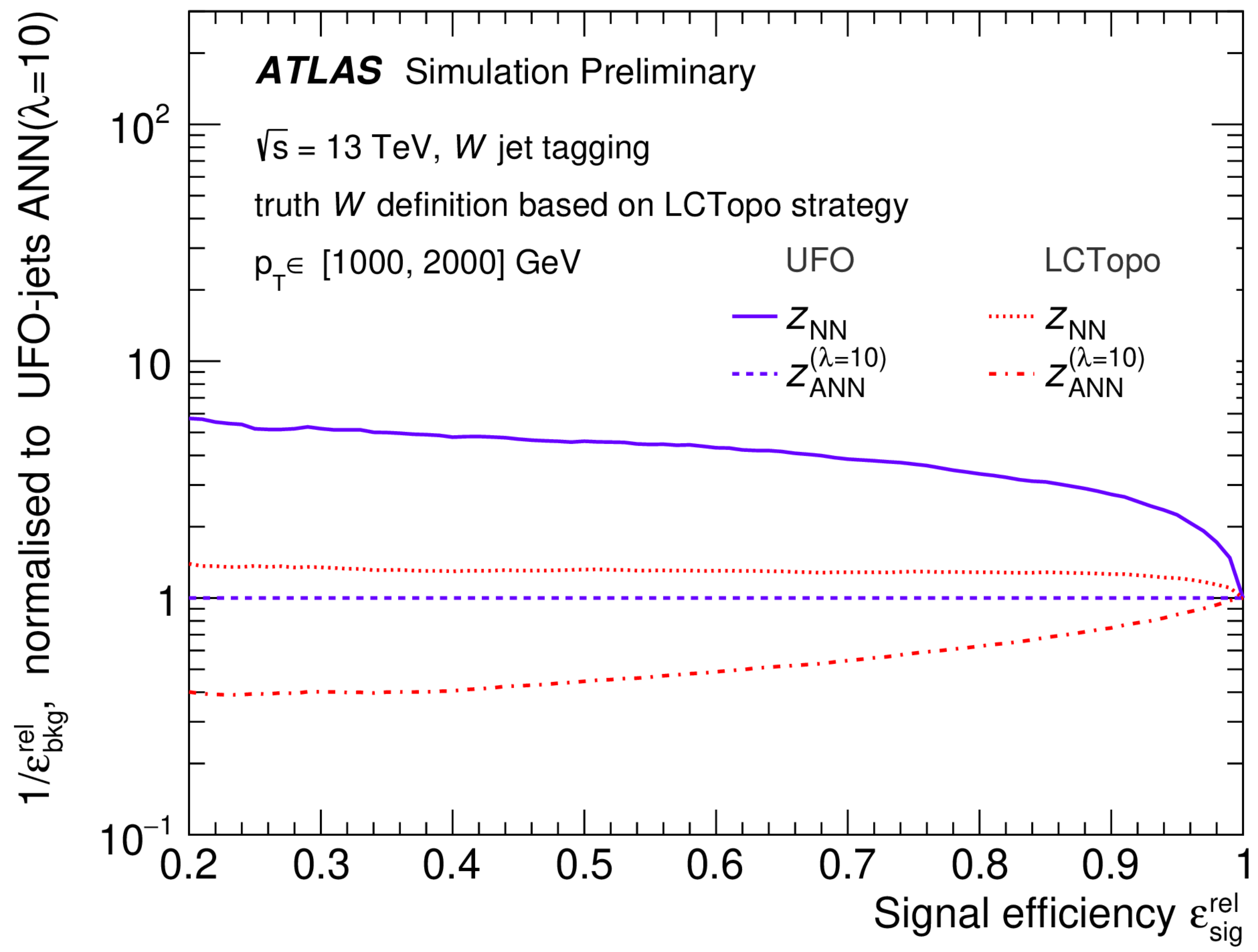
ML

Densenet


EMB1
EMB2
EMB3
(rebinned to EMB2)
Tile1
Tile2
Tile3
(rebinned to Tile1)
Dense Block
Dense Block
Dense Block
Flatten
Flatten
Flatten
Dense 128
Dense 128
Dense 128
Concatenate
Dense 64
Output

ML Taggers

ATL-PHYS-PUB-2021-029



ATL-PHYS-PUB-2021-028