Recent Improvements in FastCaloGAN in ATLAS fast calorimeter simulation

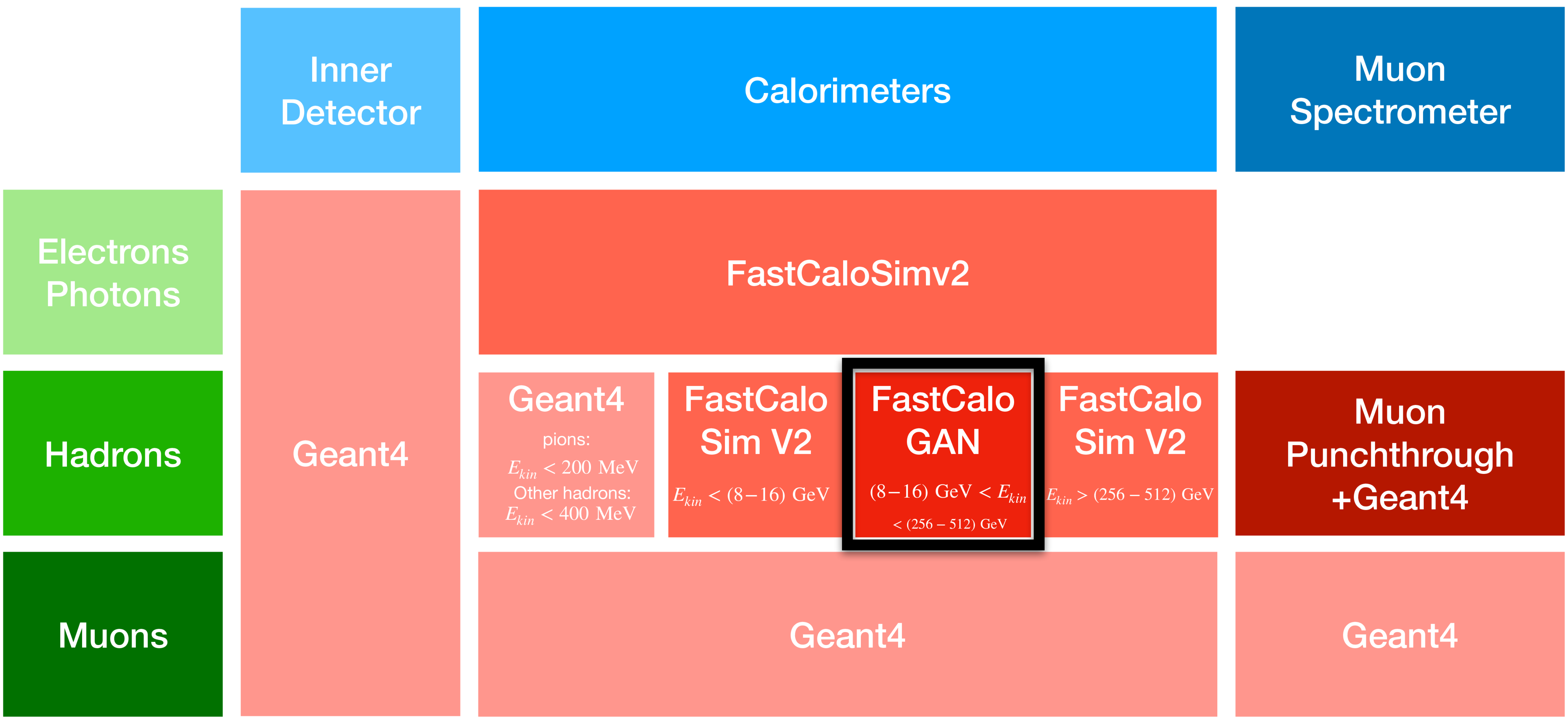
Rui Zhang
On behalf of the ATLAS collaboration
University of Wisconsin-Madison, Wisconsin

HSF Detector Simulation WG

27.11.2023

ATLAS fast calorimeter simulation

- FastCaloGAN V2 is a machine learning based model used as part of the ATLAS fast calorimeter simulation (AF3)



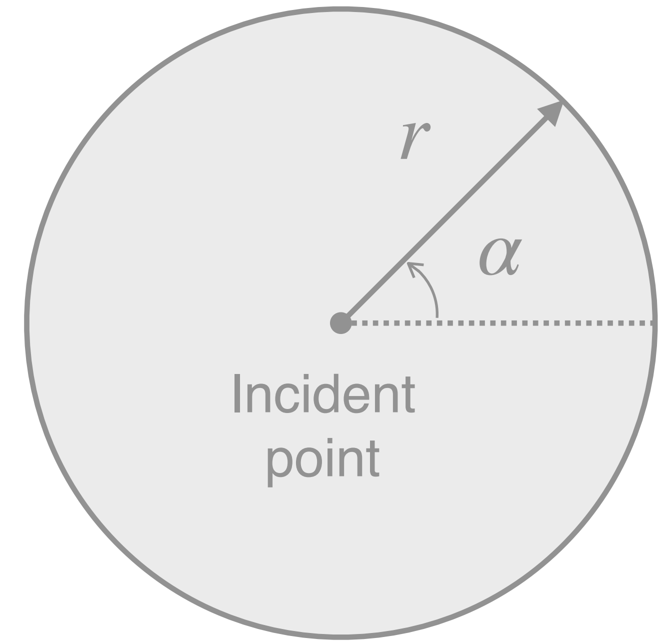
COMPUT SOFTW BIG SCI 6, 7 (2022)

FastCaloGAN V1 strategy

- Simulate photons, electrons and pions to represent all particles interacting in calorimeters
  - All hadrons share the pion parametrisation with a correction for the mass
- Divide calorimeter in 100 slices in |η|
  - In each slice we studied 17 energy points from 64 MeV to 4 TeV (in powers of two)
- For each particle/energy/|η| point a Geant4 sample (10k) is produced at the calorimeter surface
  - Noise and other imperfections are removed to parametrise on "perfect" calorimeter showers
- Custom voxels as input to avoid handling the complex and non homogeneous calorimeter structure
- In simulation, hits are generated by sampling the area of each voxel

Voxelisation

- Hits are transformed from ATLAS (x,y,z) coordinates to cylindrical (r, α, R) coordinates
  - R is not use, as hits are grouped in layers
- GAN cannot be trained on hits, so they are grouped in areas in the (r, α) plane in each layer
  - each volume in the (r, α, layer) space defines a voxel

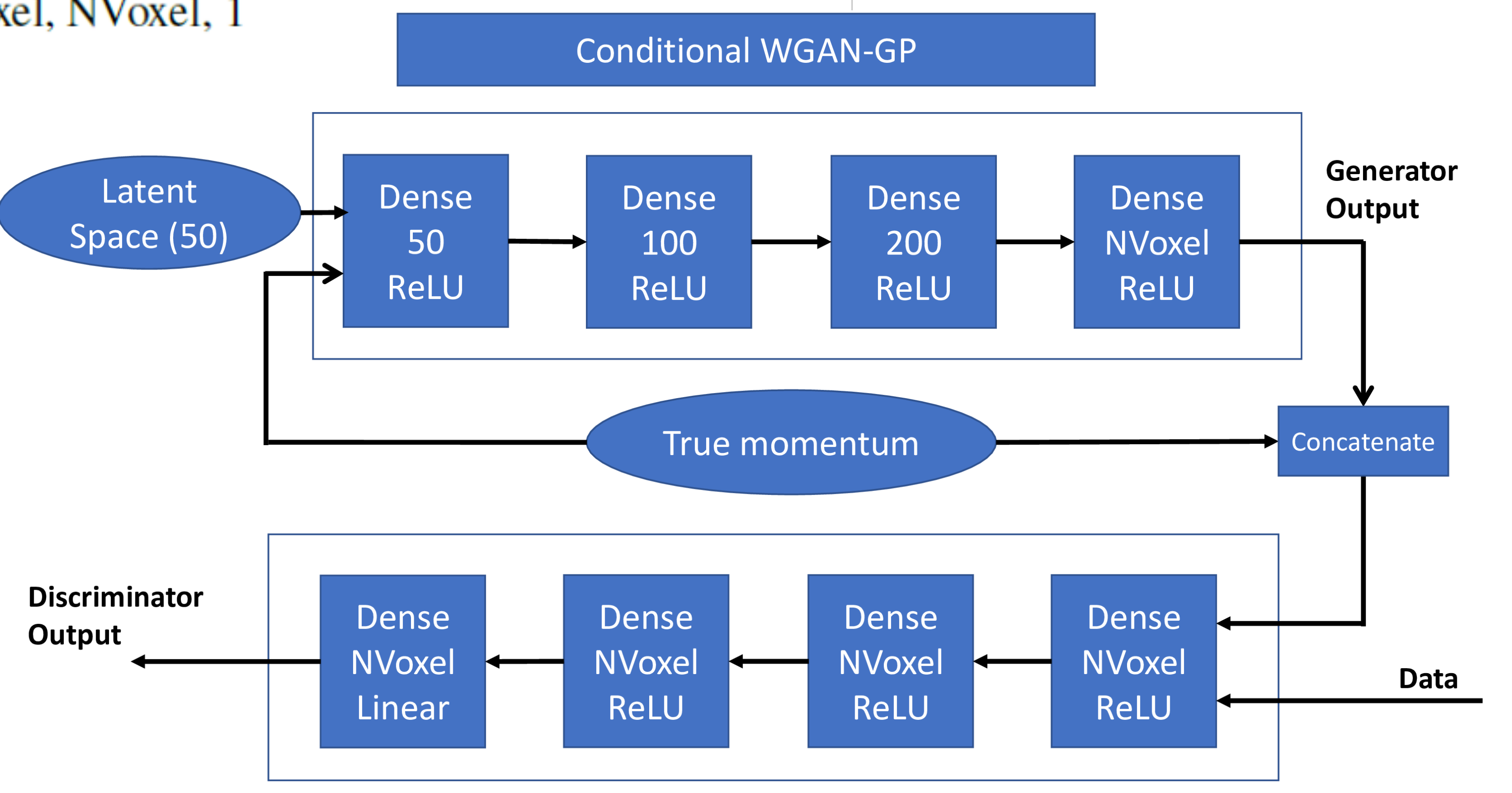


- Voxels are optimised to
  - Have enough energy to avoid large event-by-event fluctuations
  - Binning in α only for layers with a large energy deposits (i.e. EMB1 and EMB2)
  - Contain all the energy in a shower
  - Reproduce the Energy Centroids in η and ϕ, $EC_\eta = \frac{\sum_i E_i \eta_i}{E}$

FastCaloGAN V1

- A similar structure is used for all GANs
  - 300 GANs are trained in total

G	50 (Input latent Space), 50, 100, 200, NVoxel (pid and η dependent)
D	NVoxel, NVoxel, NVoxel, NVoxel, 1
Activation function	ReLU (in all layers)
Optimiser	Adam [21]
Learning Rate	10^{-4}
β	0.5
Batchsize	128
Training ratio (D/G)	5
Gradient penalty λ	10



- AtlFast3 that us deployed in Run 2 is tuned for Run 2; Geant4 is updated in Run 3
  - The shape of electromagnetic and hadronic showers are different

FastCaloGAN V2

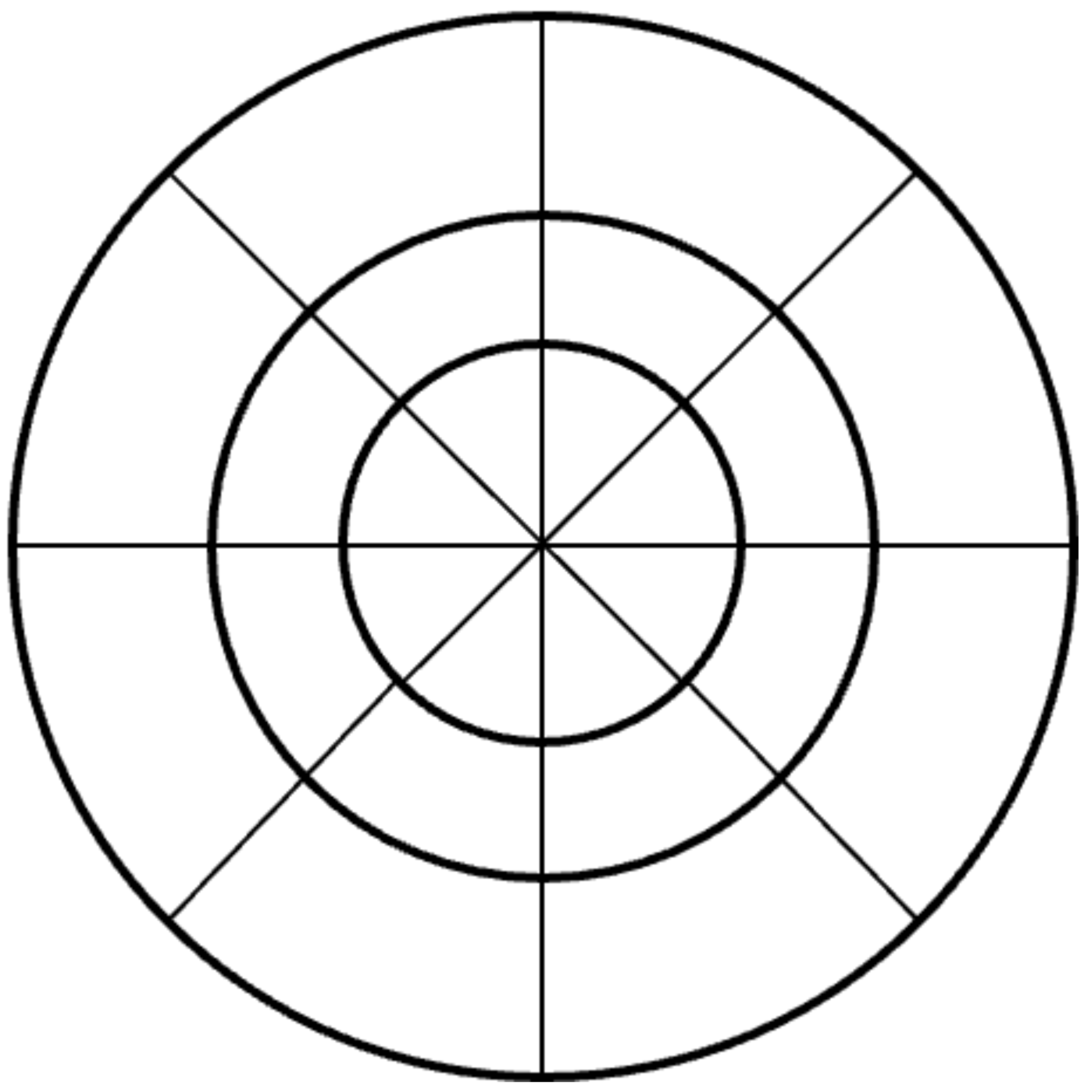
- FastCaloGAN V2 has developed
- New TensorFlow provide more stable and faster training
  - Pre-Train 6h pion, 16h e/gamma x 10 eta regions
  - Production 3h pions, 8h e/gamma x 100 eta slices
  - Full detector training time ~2 weeks on CERN HTCondor GPUs
    - V1 required 3 months, now we can easily retrain!
- More granular voxelisation for a more accurate voxel-to-cell energy assignment
  - This is further improved by exploiting energy-independent lateral shower profile
- Separate training for low- and high-energy electron and photon showers; the lowest kinetic energy reach to 64 MeV from 256 MeV in V1
- Change training strategy to two-step training
  - Divide into regions based on continuous $|\eta|$ and train a single $|\eta|$ for an extended period in each region
  - Train with other $|\eta|$ slices, starting from the best trained model obtained in the first step

FastCaloGAN V2 (cont'd)

- Architecture and hyperparameters optimised for each particle, energy range, and |η|
  - Bigger networks (due to larger input dimensions)
  - Swish activation for e/γ barrel and endcap (relu works better in pions and forward)
  - High batchsize, tuned D/G ratio, lambda
  - New conditional labels $\hat{E} = \log \frac{E_{\text{kin}}}{E_{\text{min}}} \Big/ \log \frac{E_{\text{max}}}{E_{\text{min}}}$
- Improved voxel-to-cell energy assignment exploiting energy-independent lateral shower profile
- Training use all energy samples since the first iteration, rather than starting from a single range and adding more energies after some iterations
- Introduced proton FastCaloGANs for baryon simulations, instead of using the pion in V1

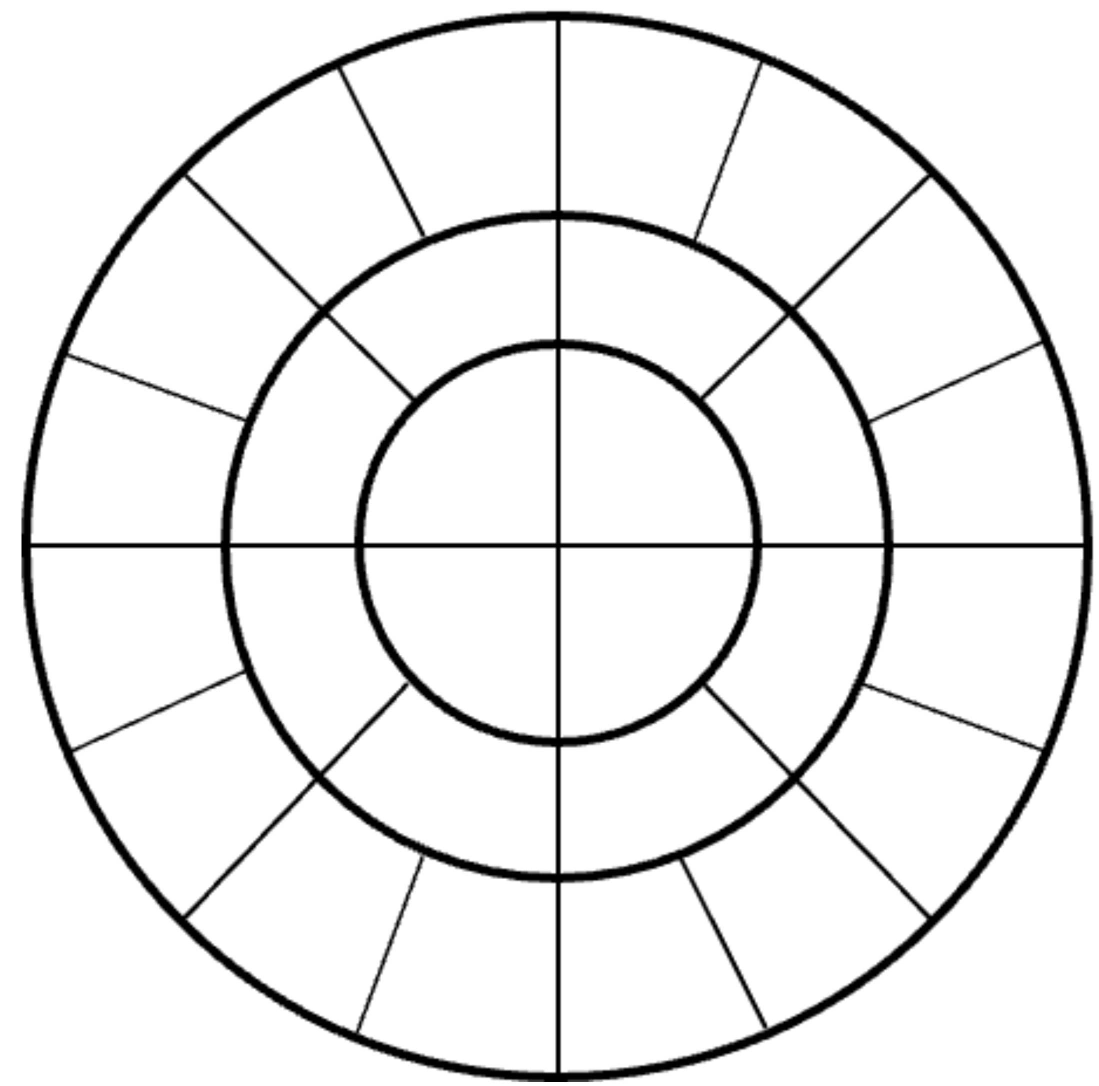
New voxelisation: more bins

V1



Better to parametrise the shower shape

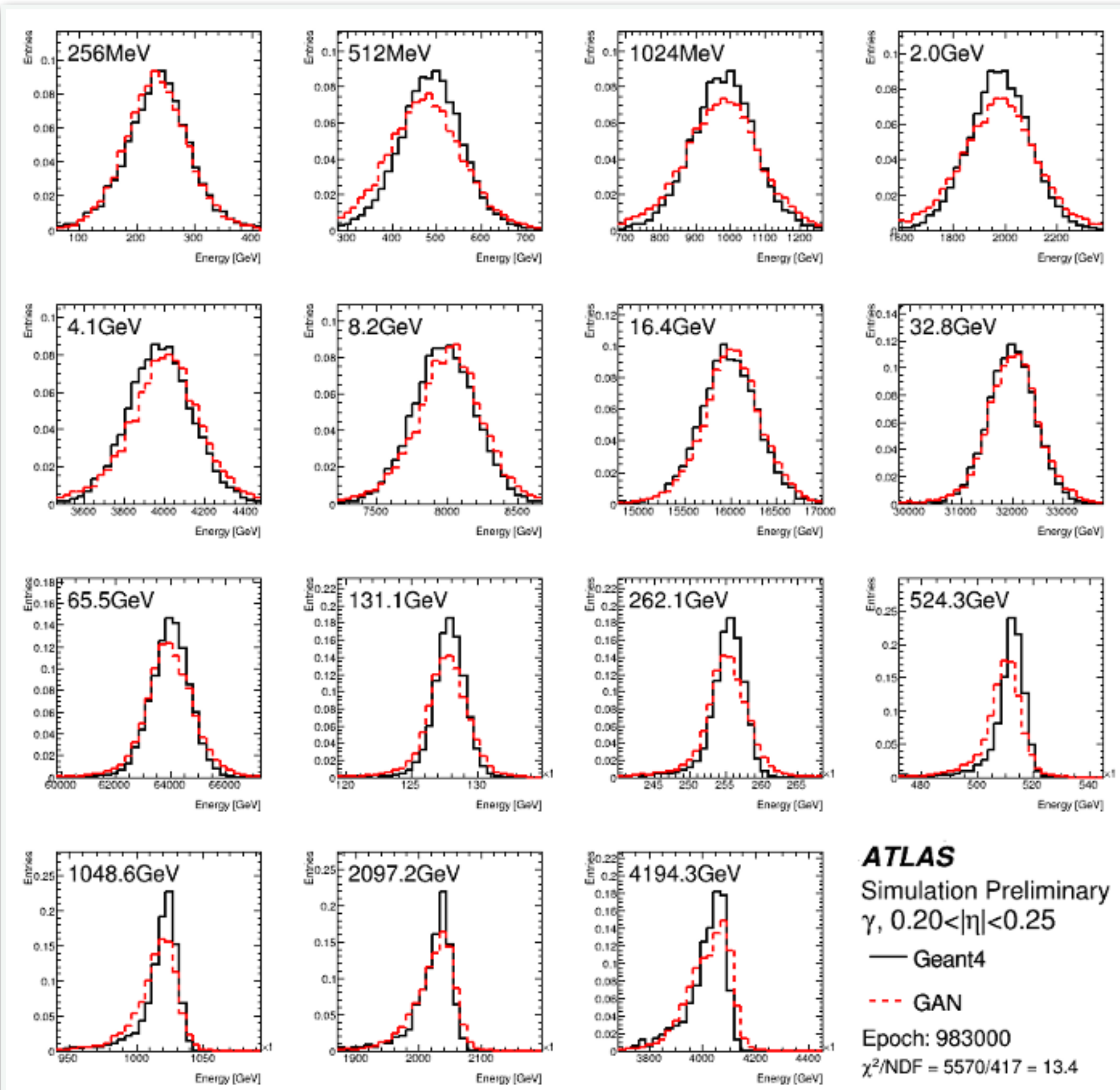
V2



More information to learn
but need higher memory

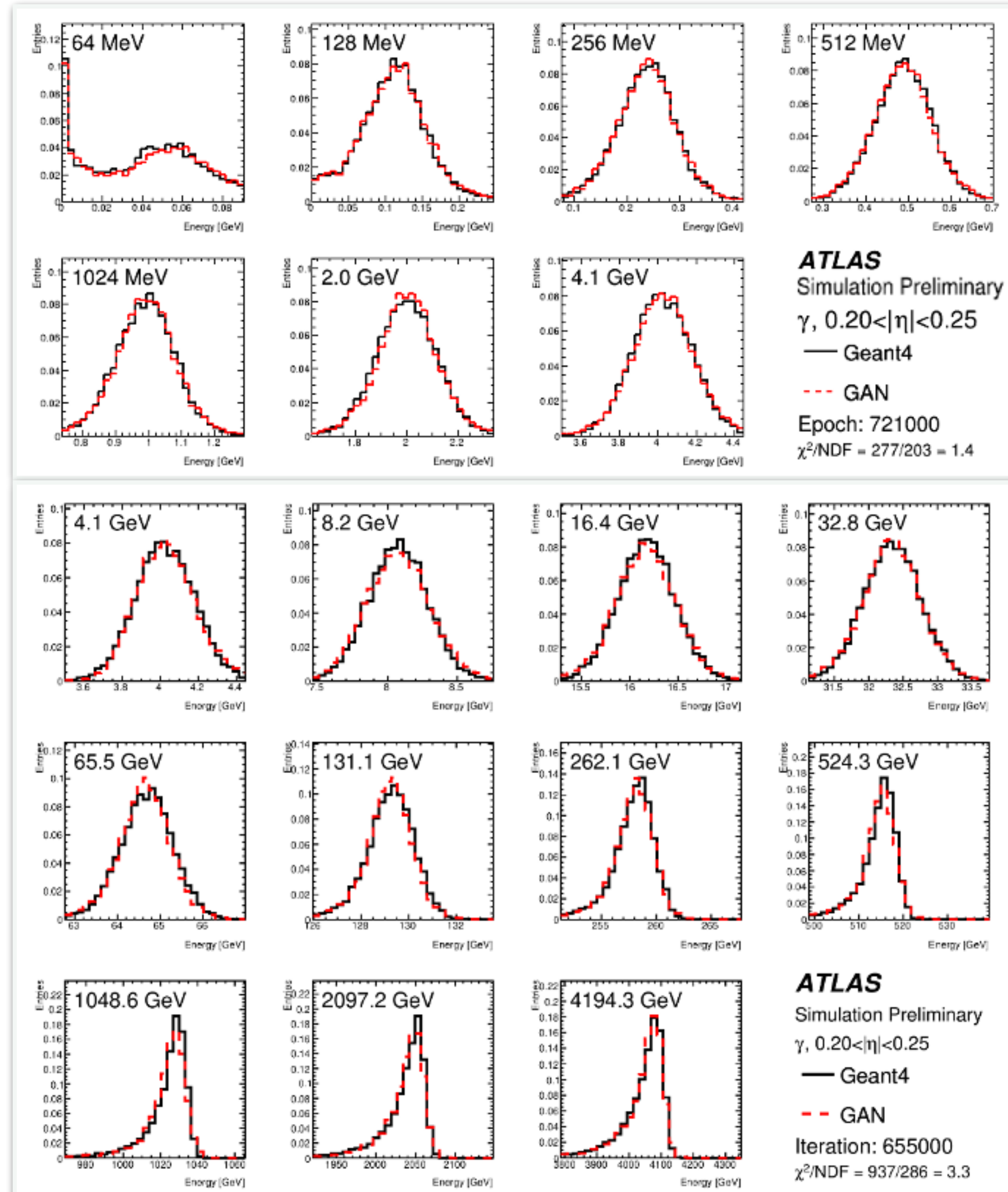
Photon FastCaloGAN

V1



$\chi 2$/NDF = 12.3

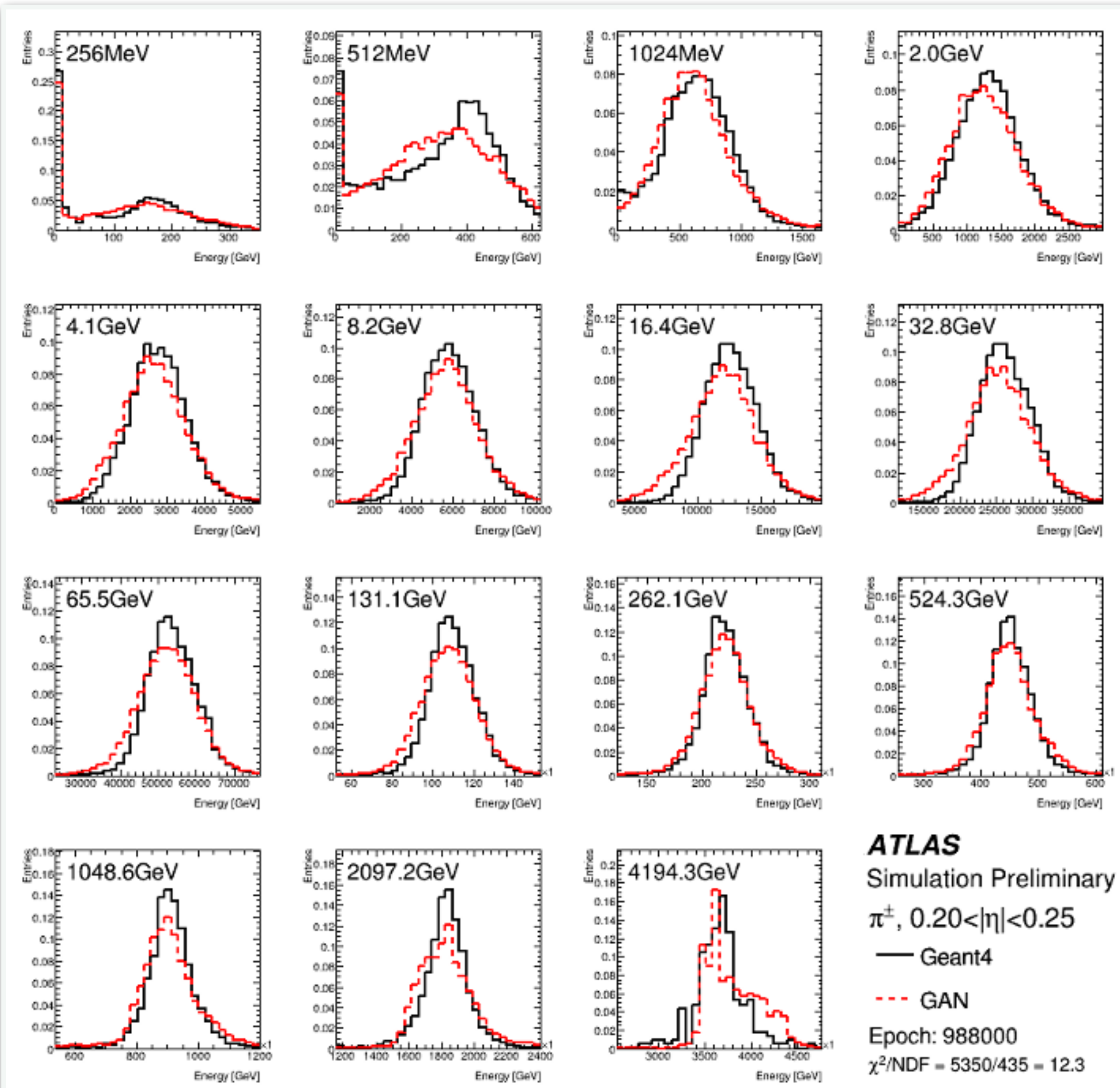
V2



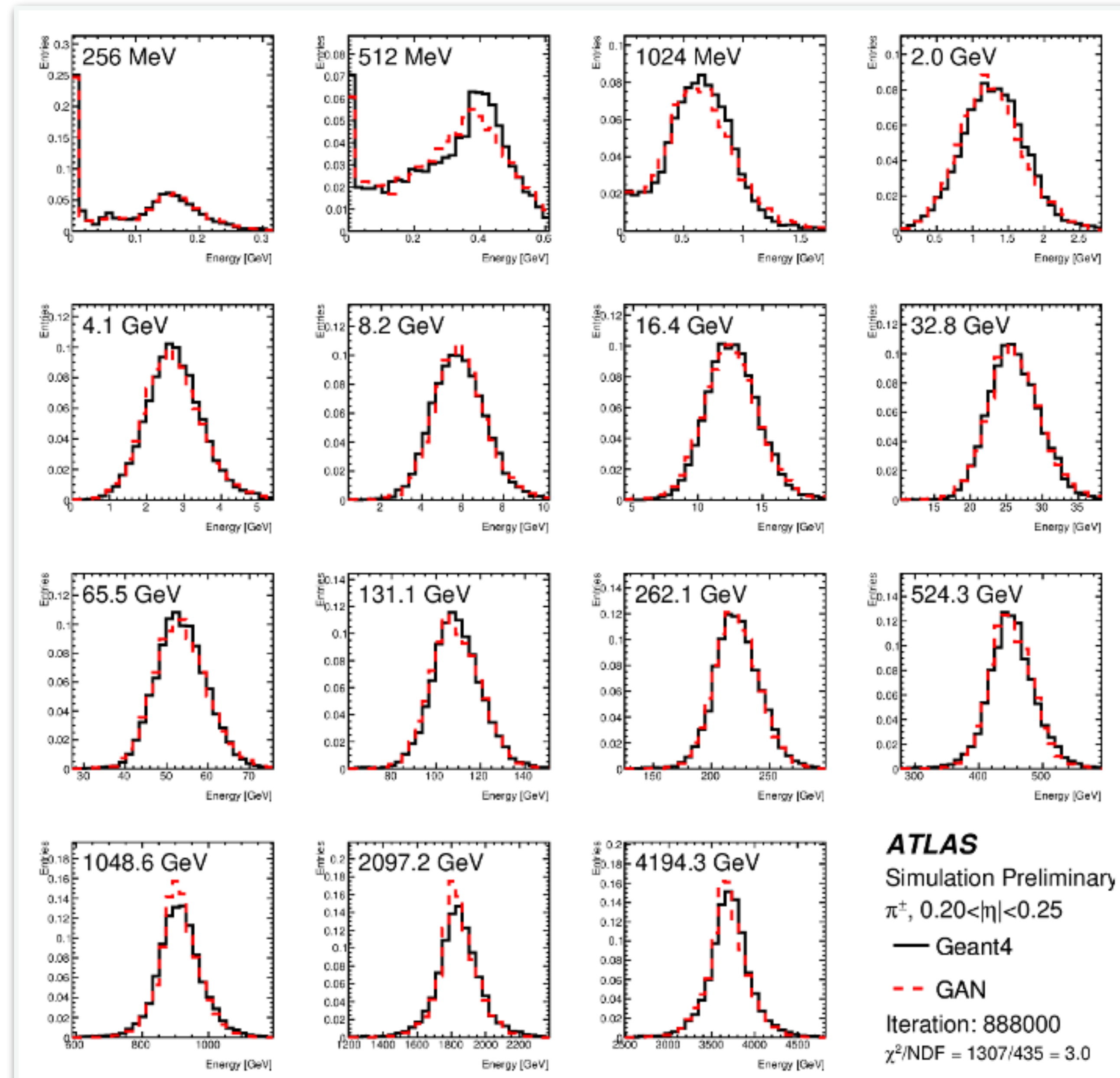
$\chi 2$/NDF = 2.5

Pion FastCaloGAN

V1



V2



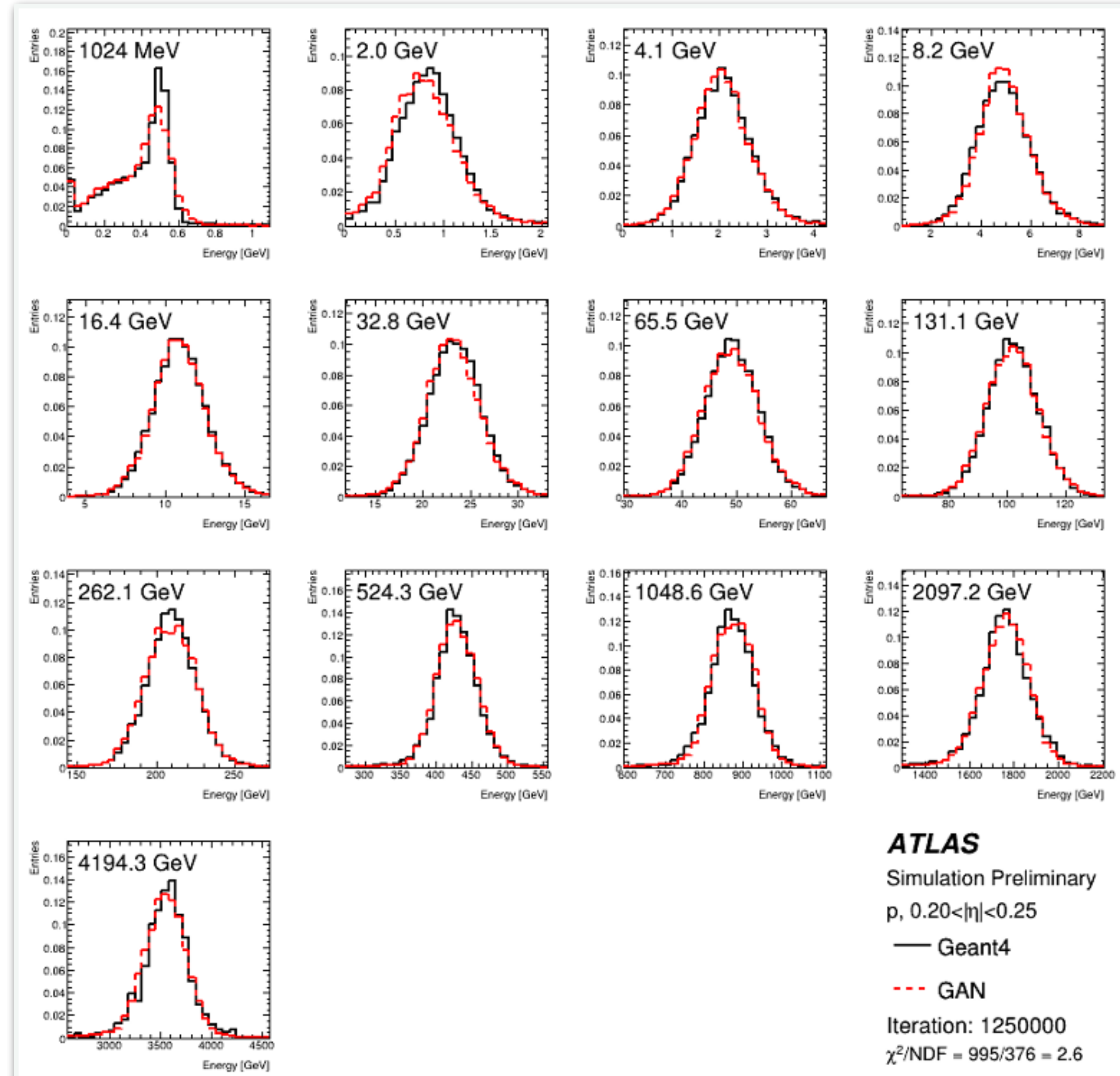
χ2/NDF = 12.3

χ2/NDF = 3.0

Proton FastCaloGAN

- Fewer energies are used as they are produced by momentum
- The same voxelisation and hyperparameter as pion
- Good performance is achieved

V2



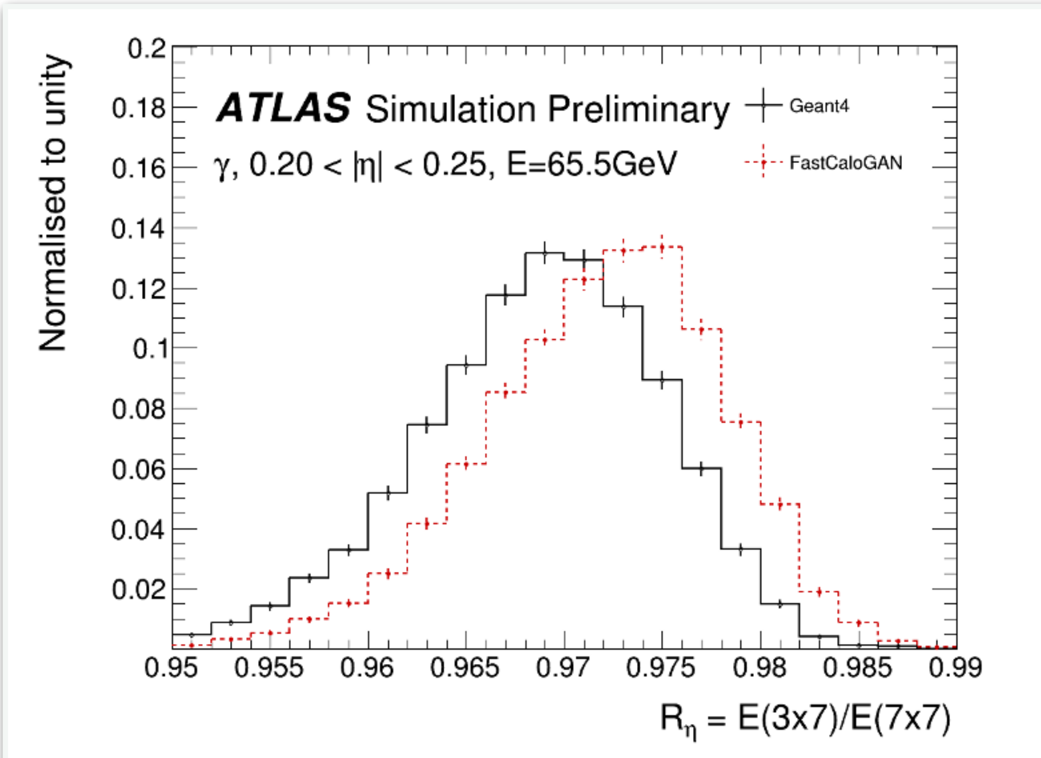
$\chi 2/NDF = 2.6$

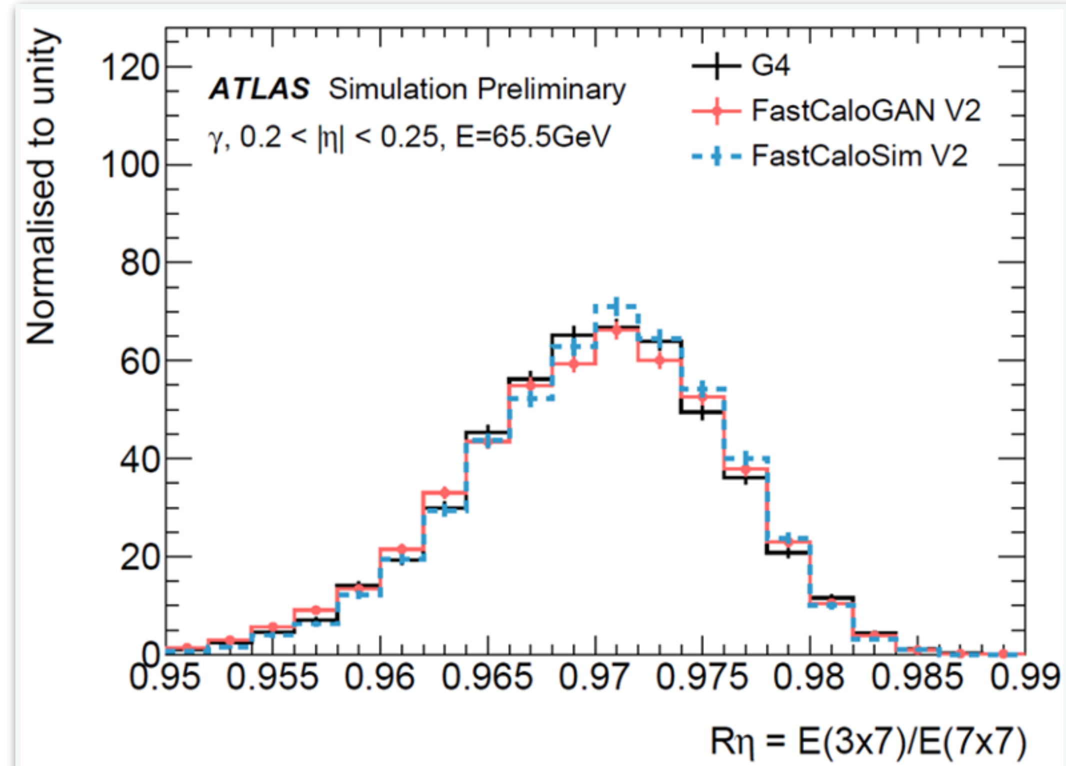
FastCaloGAN performance in single particles

V1

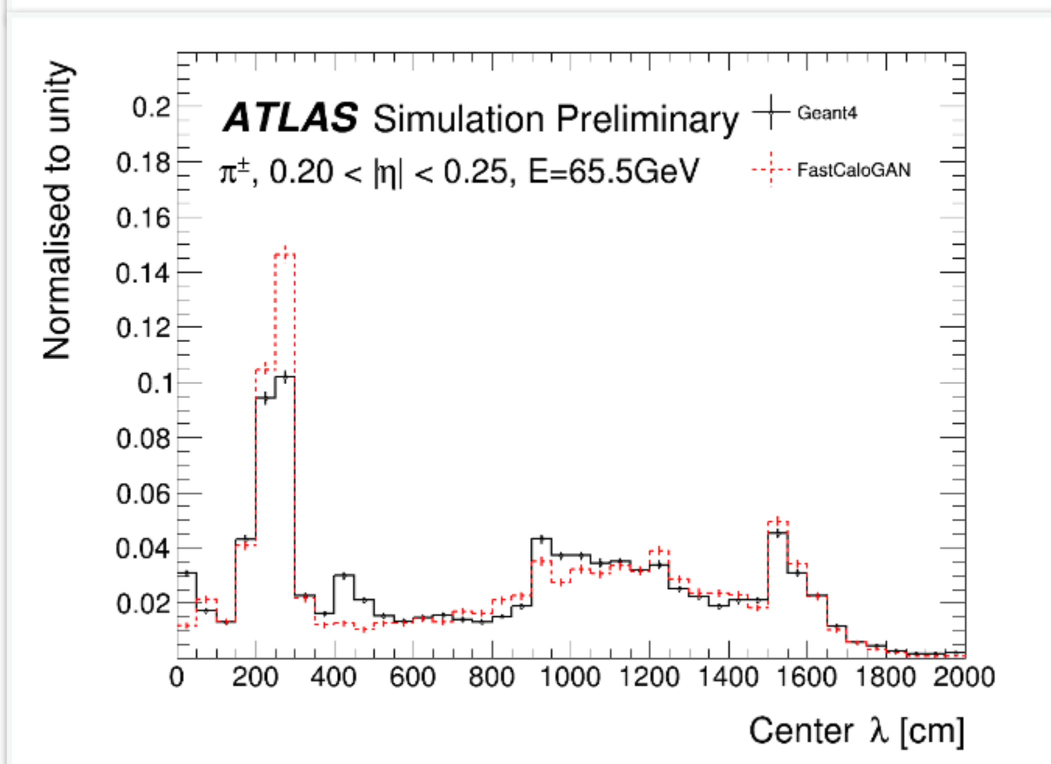
V2

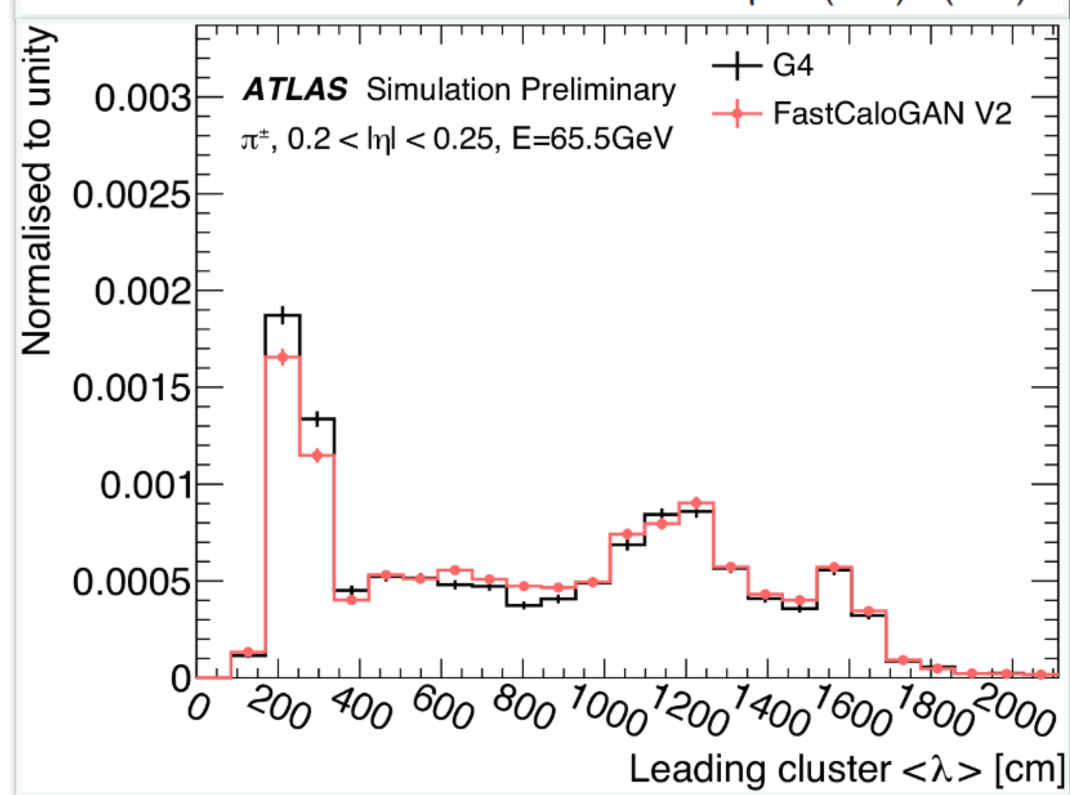
γ





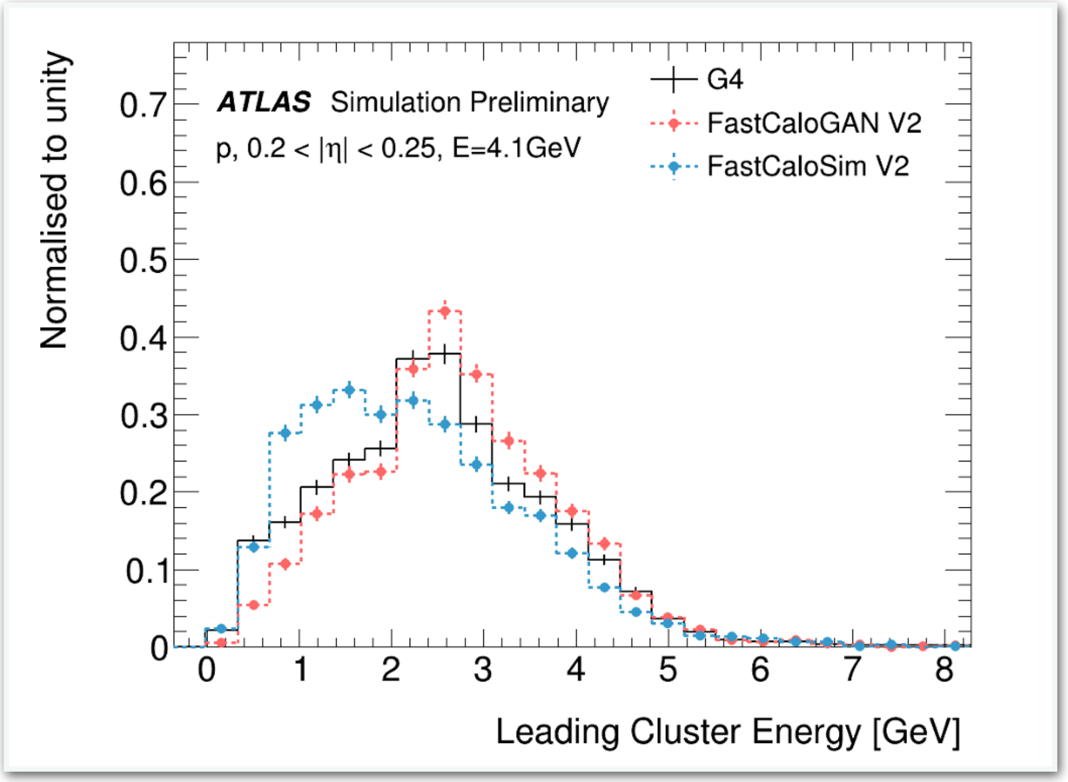
π





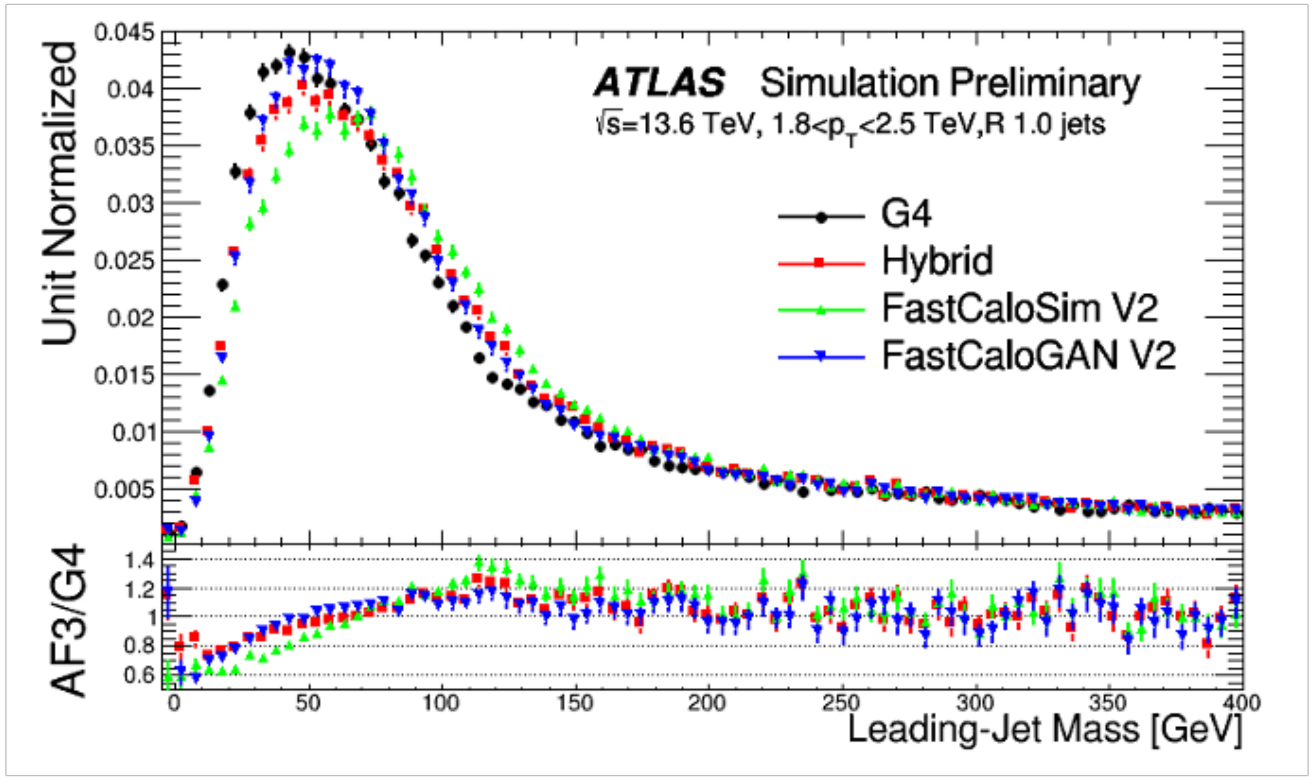
FastCaloGAN performance in single particles

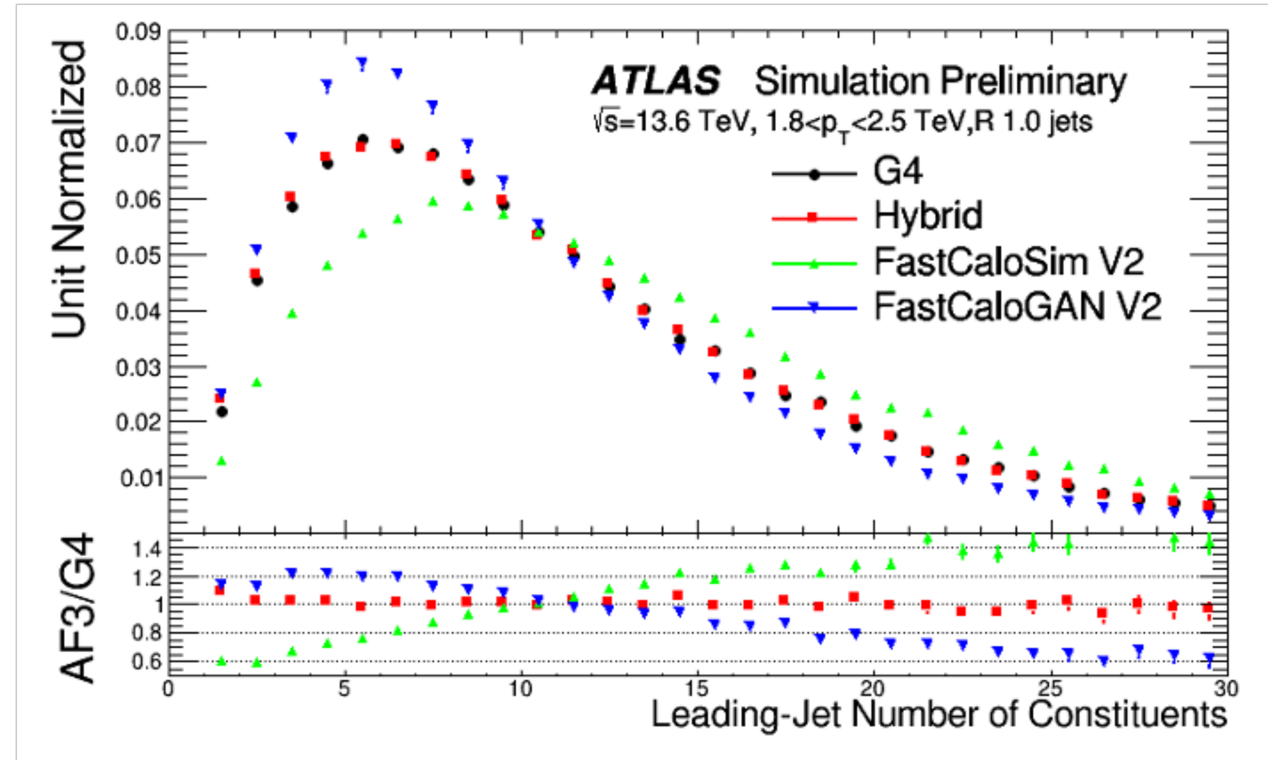
V2

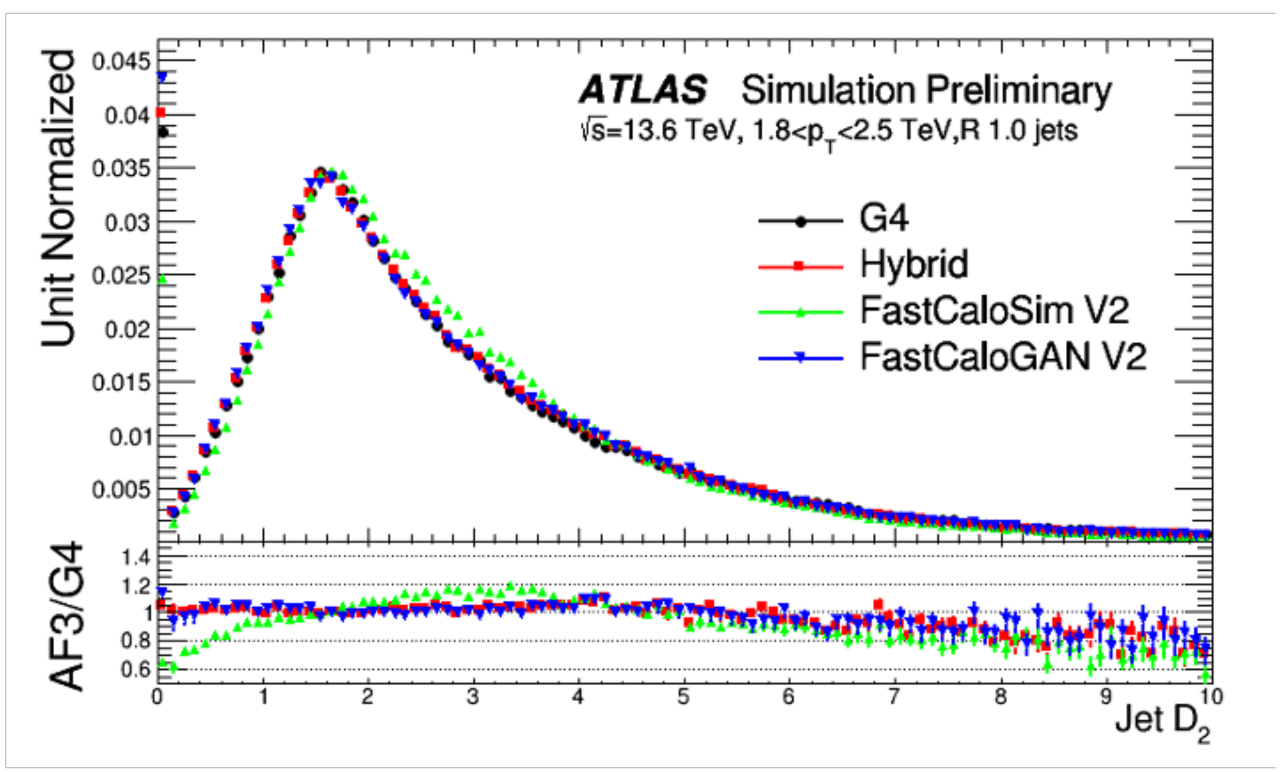


p

Jet performance in AF3







- Better performed achieved by combining FastCaloGAN and FastCaloSim in different energy and |η|

Simulation in Athena

- The inference is done in LWTNN but we are working on an ONNX implementation to reduce the memory
  - More and bigger networks than in V1 resulted in a tenfold increase in memory requirement
  - Reducing memory will allow to add GANs for other hadrons that are currently simulated as “pions+energy correction”
  - Speed at inference should also increase but we are not limited by inference time
- Improved voxel-to-cell energy assignment; hits are no longer generated in a grid, instead:
  - alpha direction is randomised
  - Along the radial direction hits are generated according to a pdf taken from input GEANT4 samples

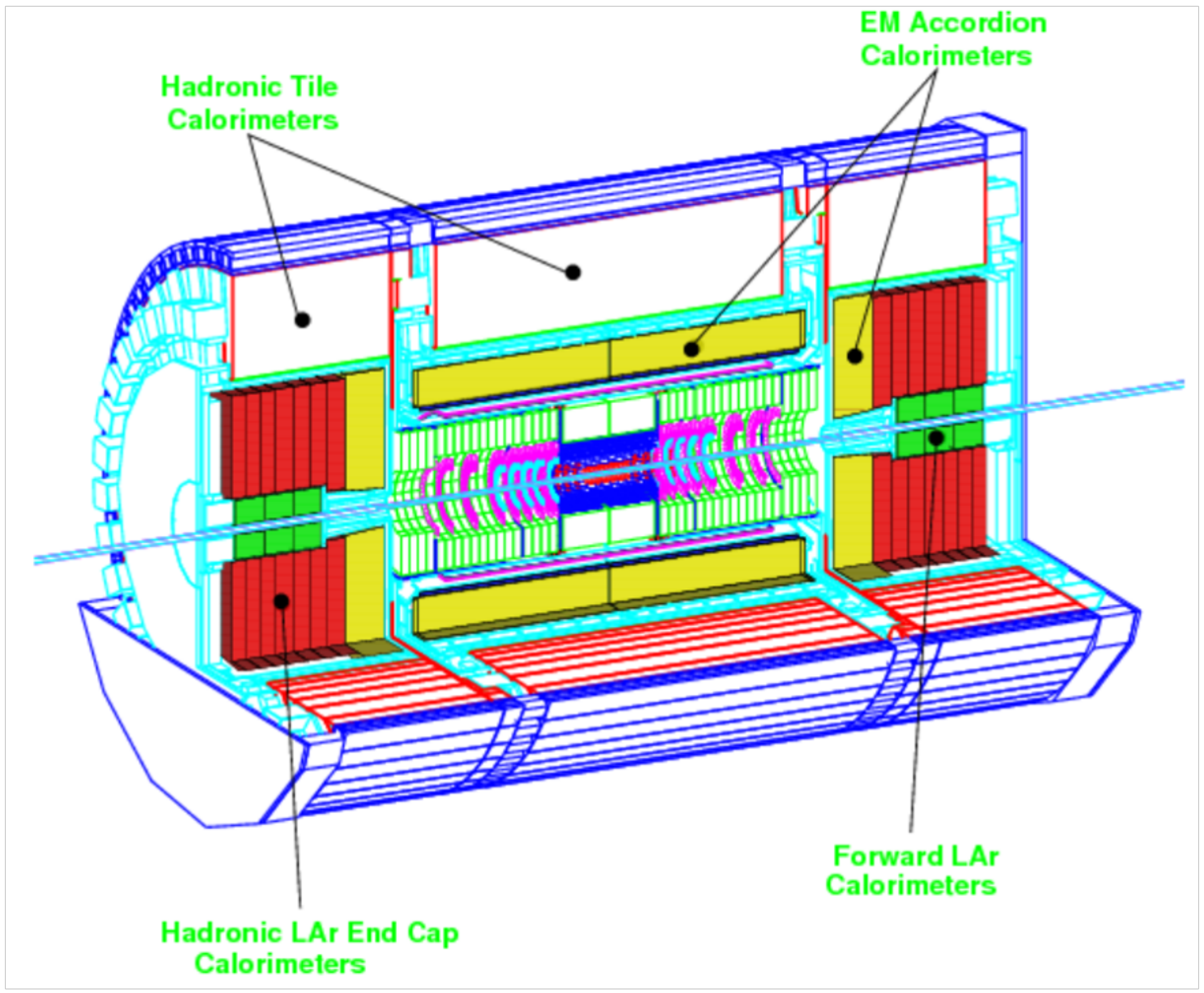
Conclusion

- AtlFast3 is a significant improvement w.r.t. AtlFast2 in reproducing key variables used in analysis
  - This is crucial to allow a wider use of fast simulation required to match the collected luminosity in Run3
- FastCaloGAN is a critical ingredient in the mix of tools that allows to simulate the jet sub-structure with high accuracy
- FastCaloGAN V2 was developed for Run3 and further improves the accuracy of the simulation
- There is still room for improvement and we will continue to explore new models to further improve the simulation in ATLAS
  - Public datasets used for the #calochallenge includes examples from ATLAS dataset and we see promising models that can achieve good results (see Anna's talk)
  - It is also time to think about whether the current input format is optimal
  - Adapting available good models to ATLAS is very welcome

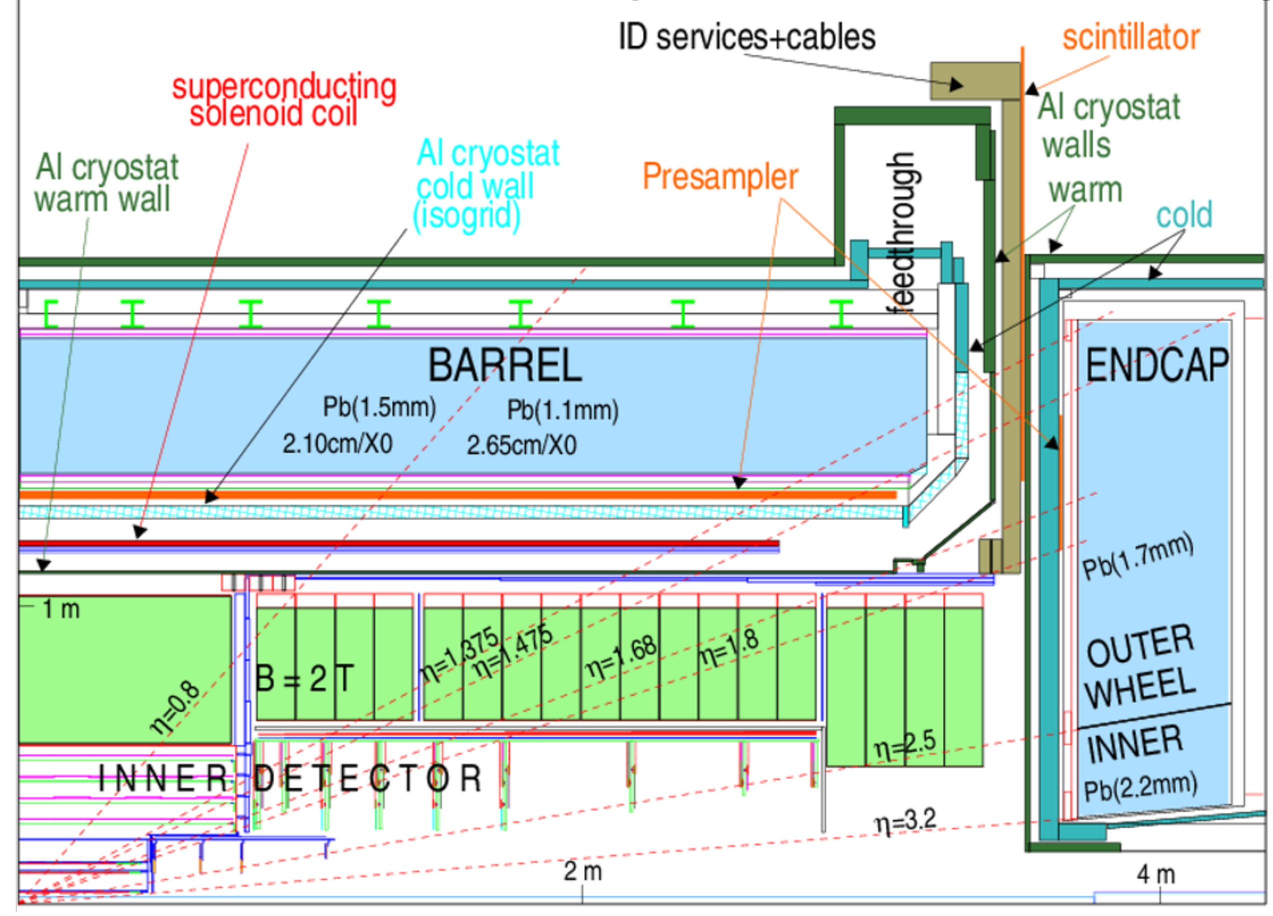
Backup

The ATLAS calorimeters

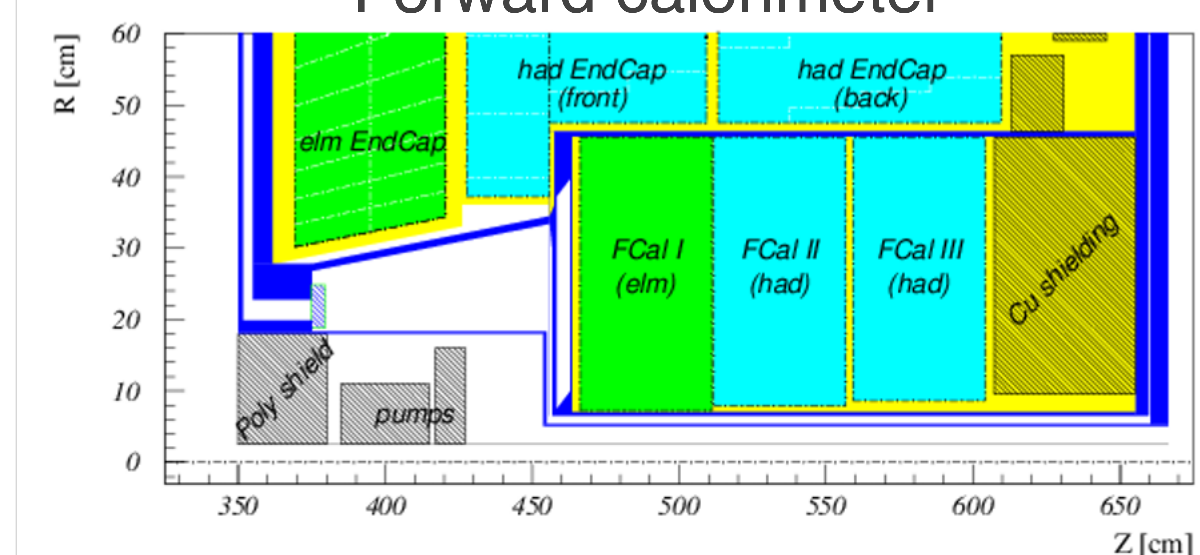
Complex geometry



EM calo: each line represent a discontinuity

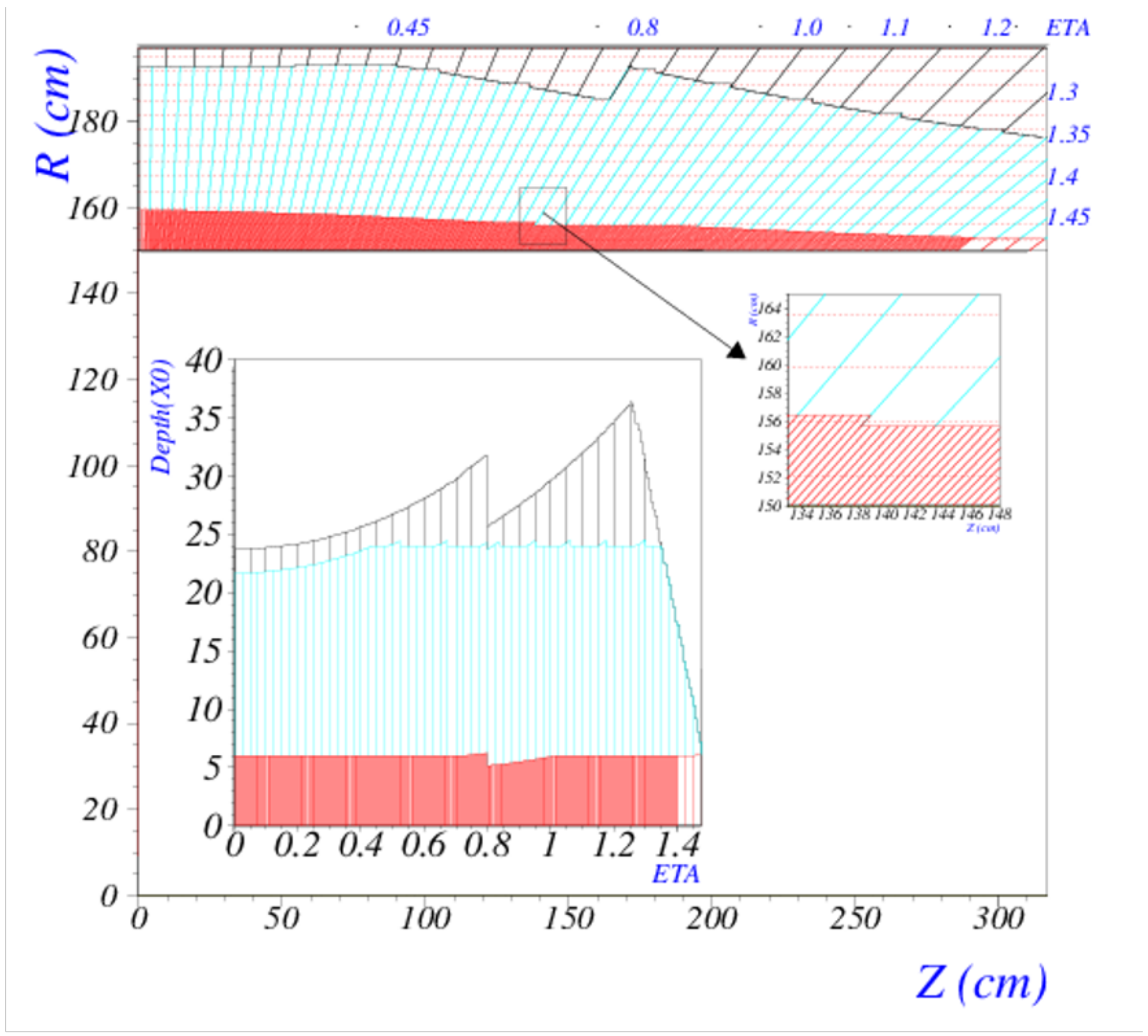


Forward calorimeter



The ATLAS calorimeters

EM calorimeter: barrel (4 layers: pre-sampler, EMB1, EMB2, EMB3)



Hadronic calorimeter: transition region between barrel and endcap