New approaches using machine learning for fast shower simulation in ATLAS

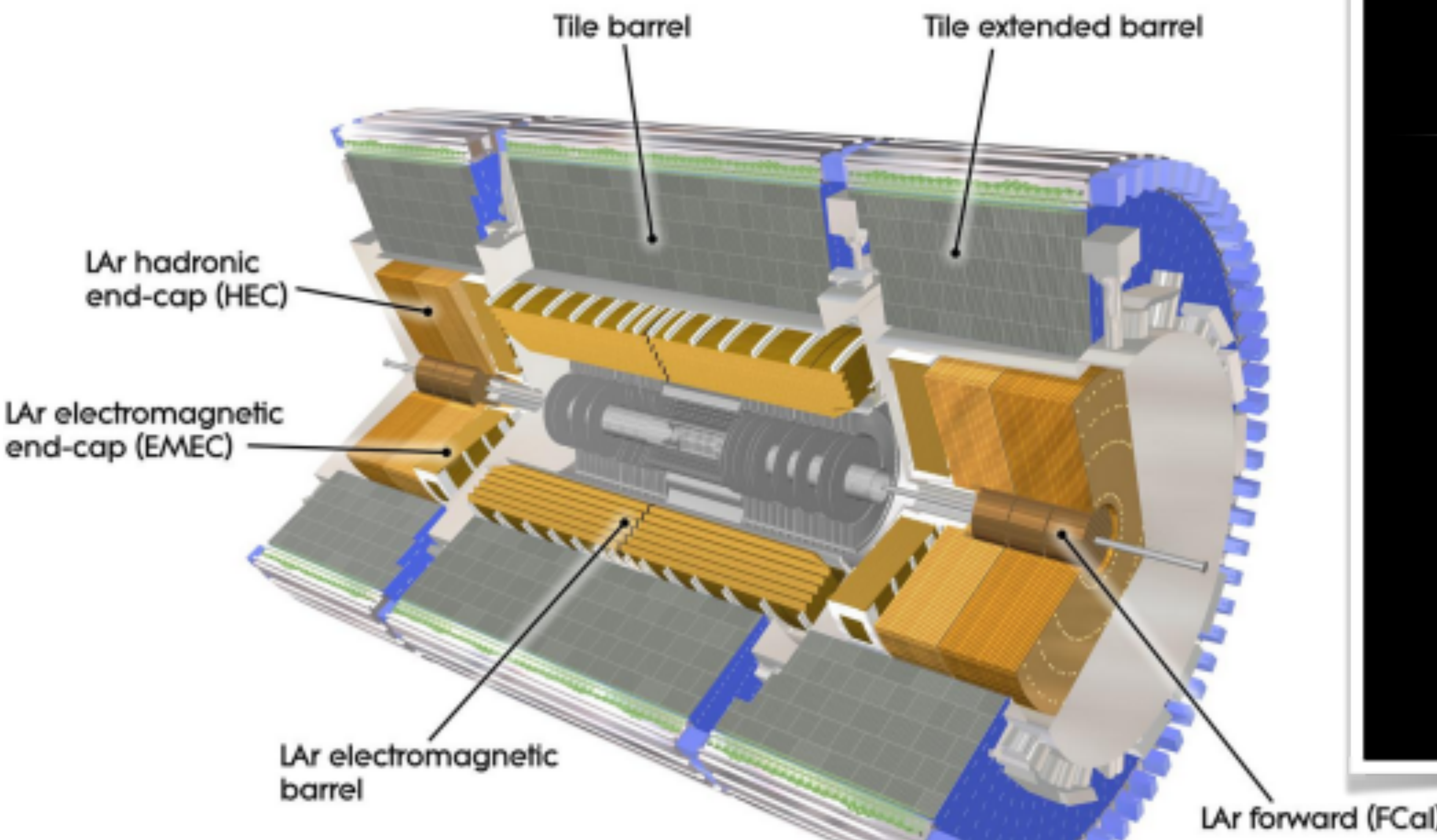
Hasib Ahmed

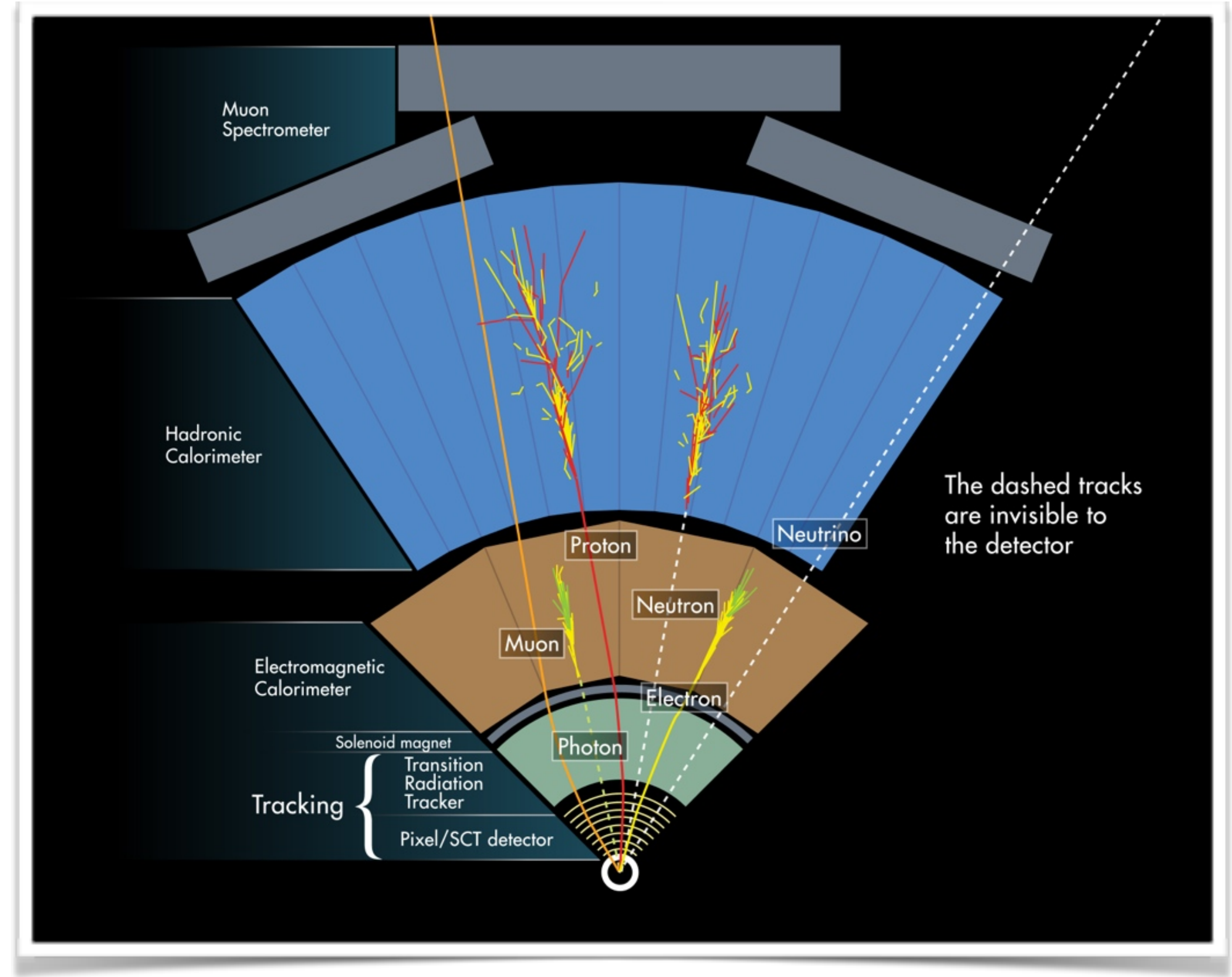
On behalf of the ATLAS Collaboration

ICHEP, 2018

The ATLAS Calorimeter and shower generation

Sampling calorimeter covering $|\eta| < 4.9$





Total readout channels: ~190 k

Number of layers: 24

System	EM Barrel	EM EC	Hadronic EC	FCAL	Tile
#Channels	110k	64k	5.6k	3.5k	9.8k

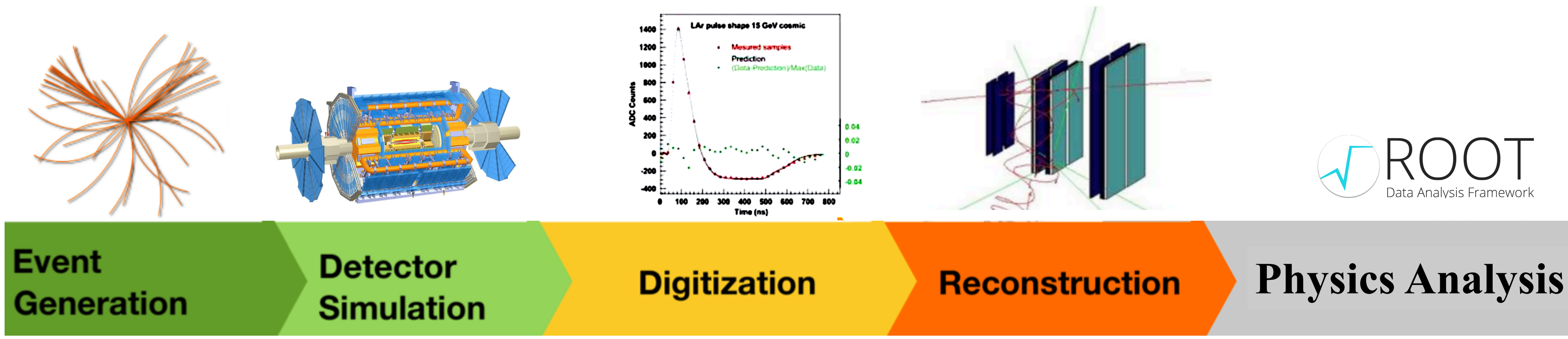
Electromagnetic (EM) Cal:

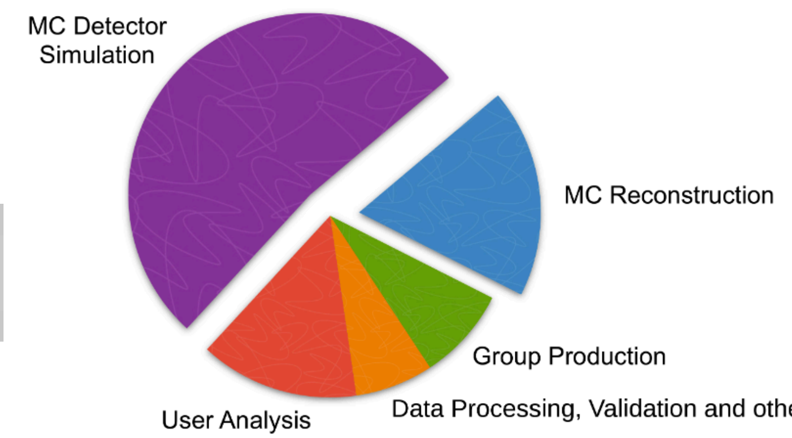
- Liquid Argon (active)
- Pb/Cu/Tungsten (absorber)

Hadronic/Tile Cal:

- Scintillating tiles (active)
- Steel (absorber)

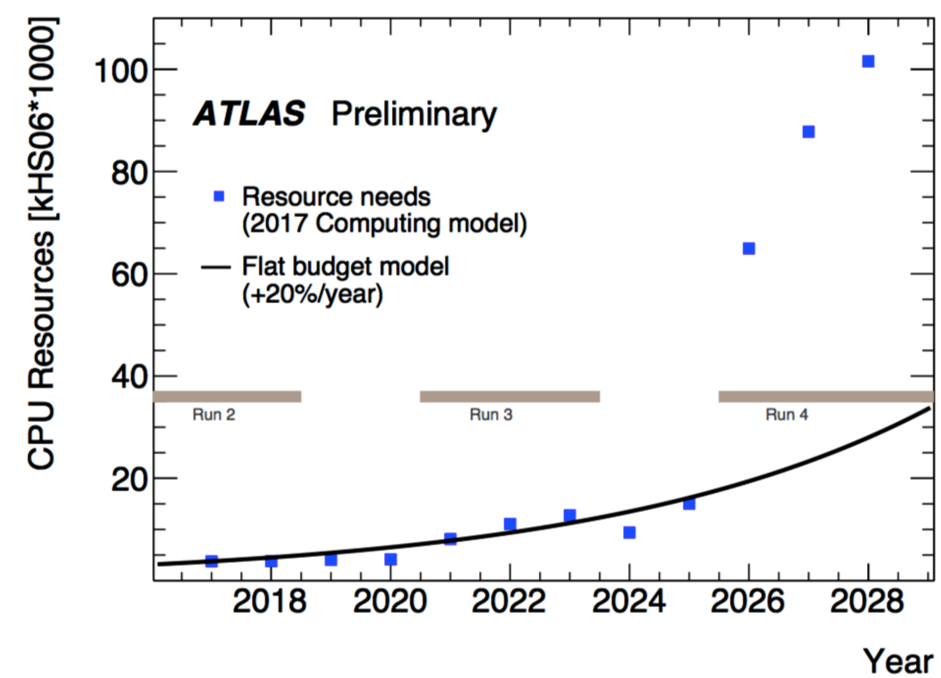
Need for fast shower simulation: Monte Carlo Production





Successful Physics program in ATLAS depends on the availability of high statistics Monte Carlo simulated events

Geant4 requires significant resources with ~75% spent in shower simulation i.e. Calorimeter simulation

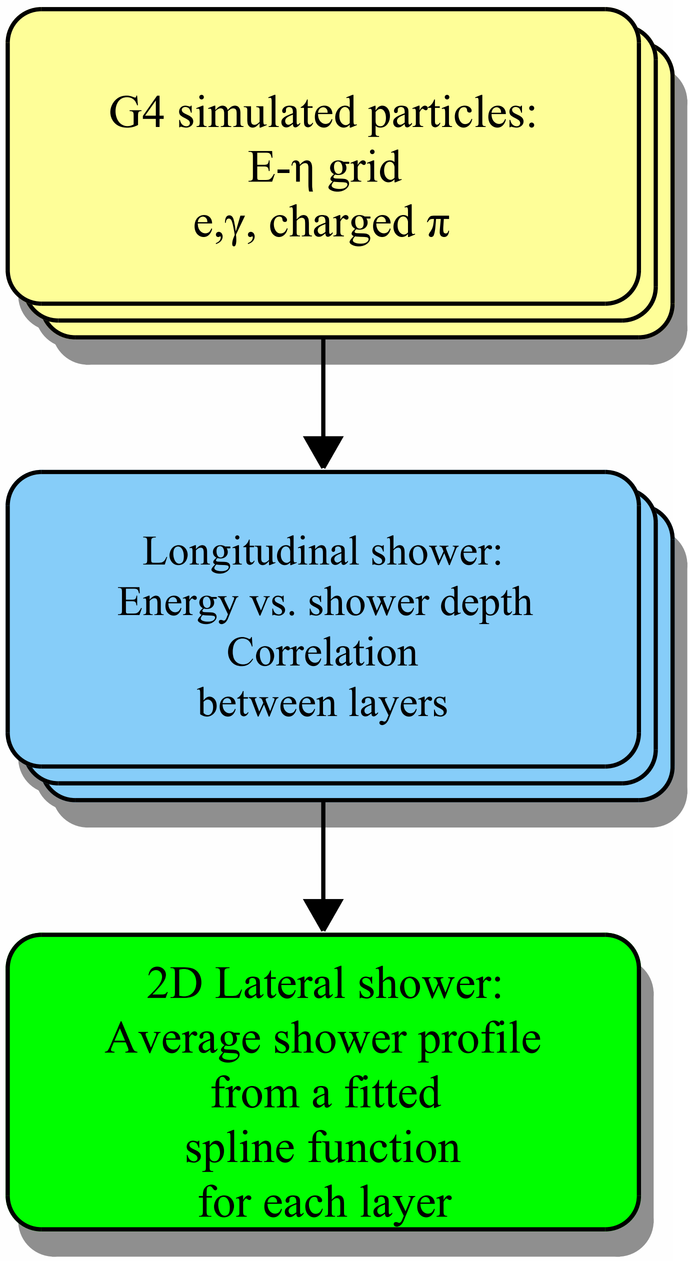


The increased pileup at HL-LHC will also increase the CPU requirement for the same number of hard scattered events

Imperative to develop fast shower simulations compared to Geant4

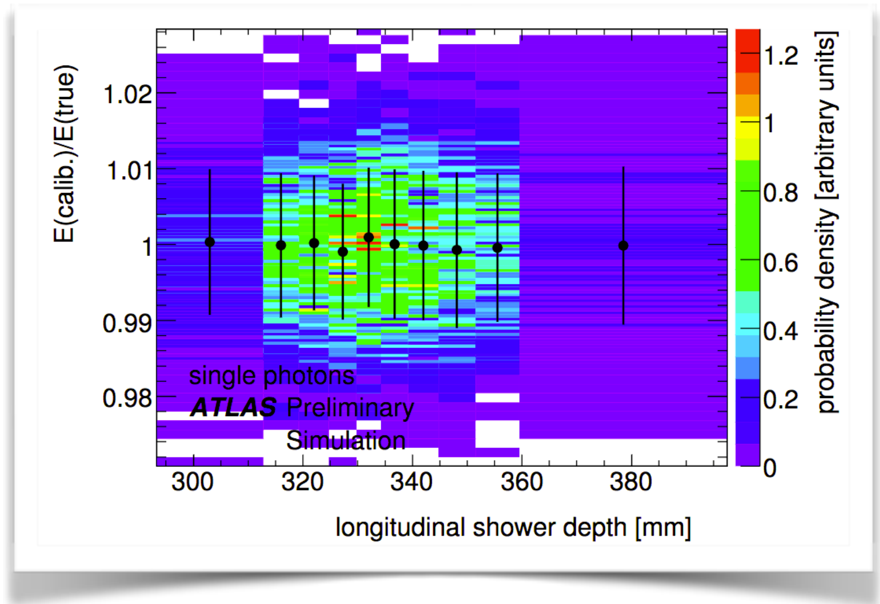
The Current Fast Simulation: *FastCaloSimV1*

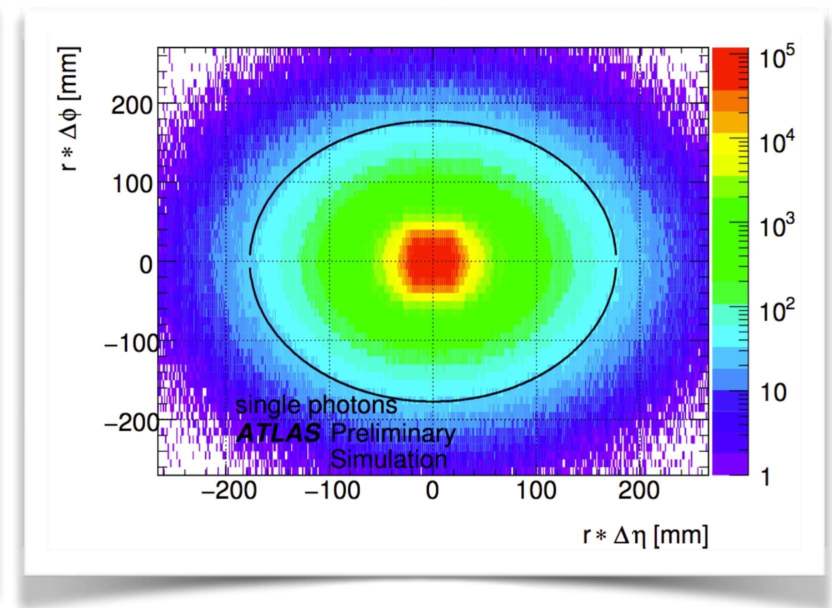
Fast simulation utilizes parametrized calorimeter response



e,γ for EM interaction

$\pi^{\pm}$ for hadronic interaction



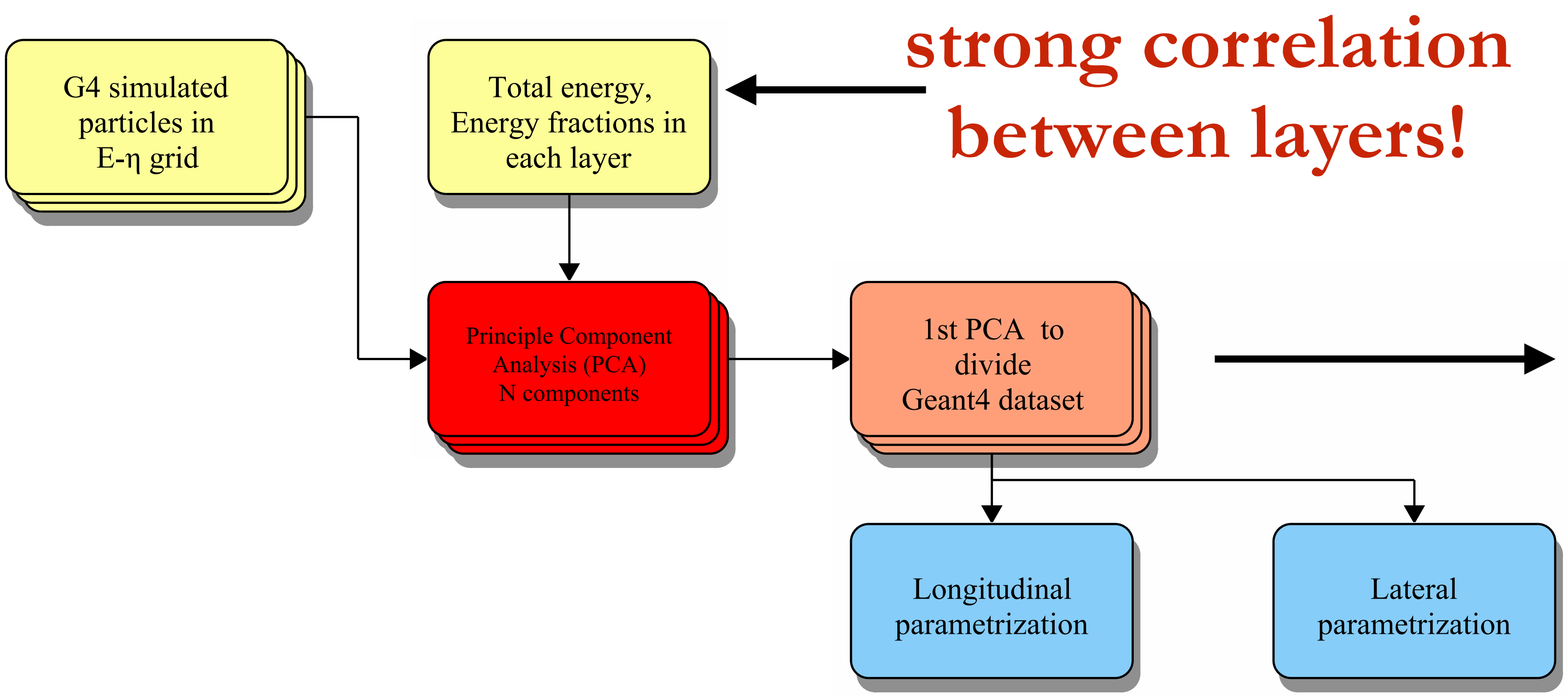


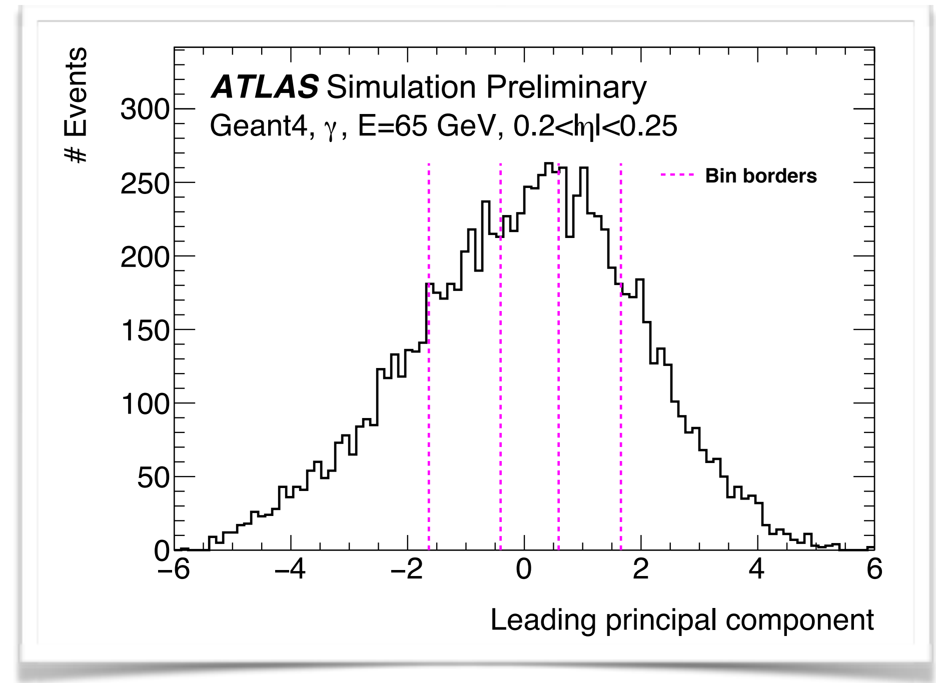
Poor modeling of some physics variables i.e. jet substructure

Forward Calorimeter (FCal) not implemented

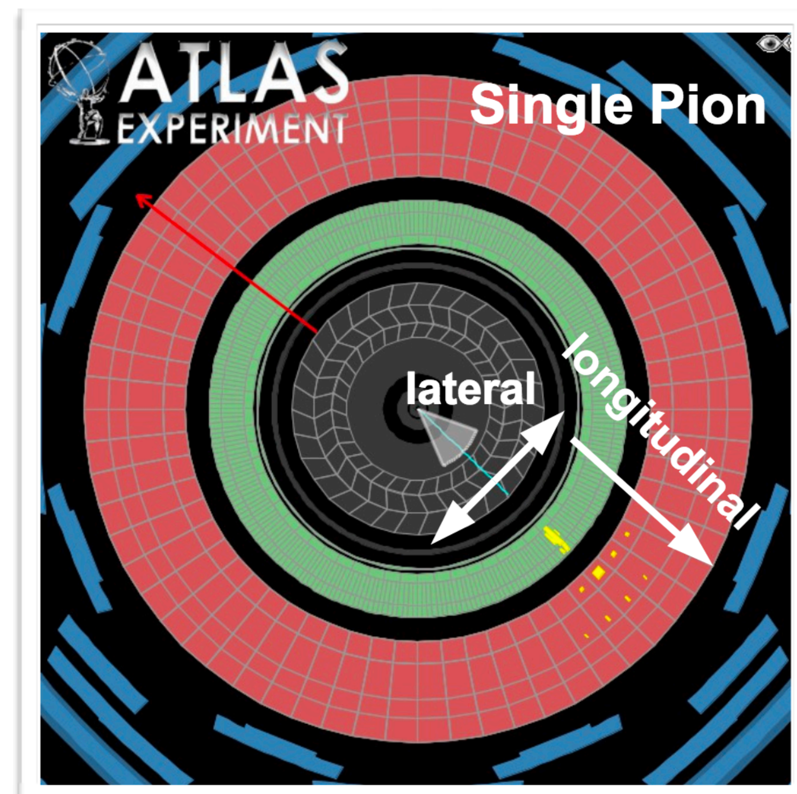
New approaches of fast simulation: *FastCaloSimV2*

Parametrization based approach following *FastCaloSimV1*



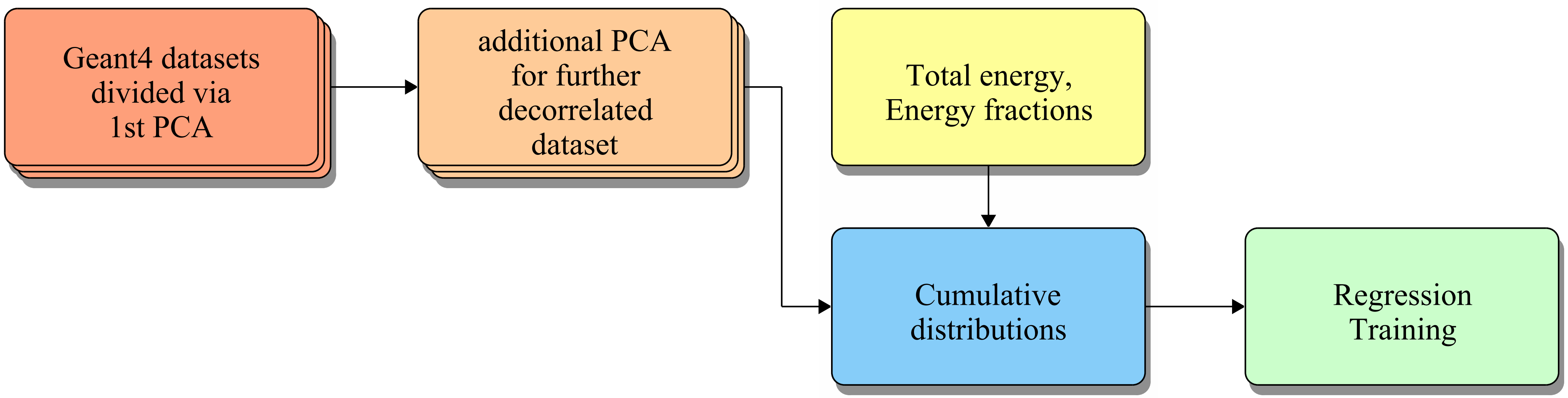


- PCA transformation to decorrelate energy deposit in each layer
- Leading PCA component is used to divide the Geant4 dataset into subsets
- Each subset represents shower with similar feature
- Longitudinal and lateral parametrization for each subset



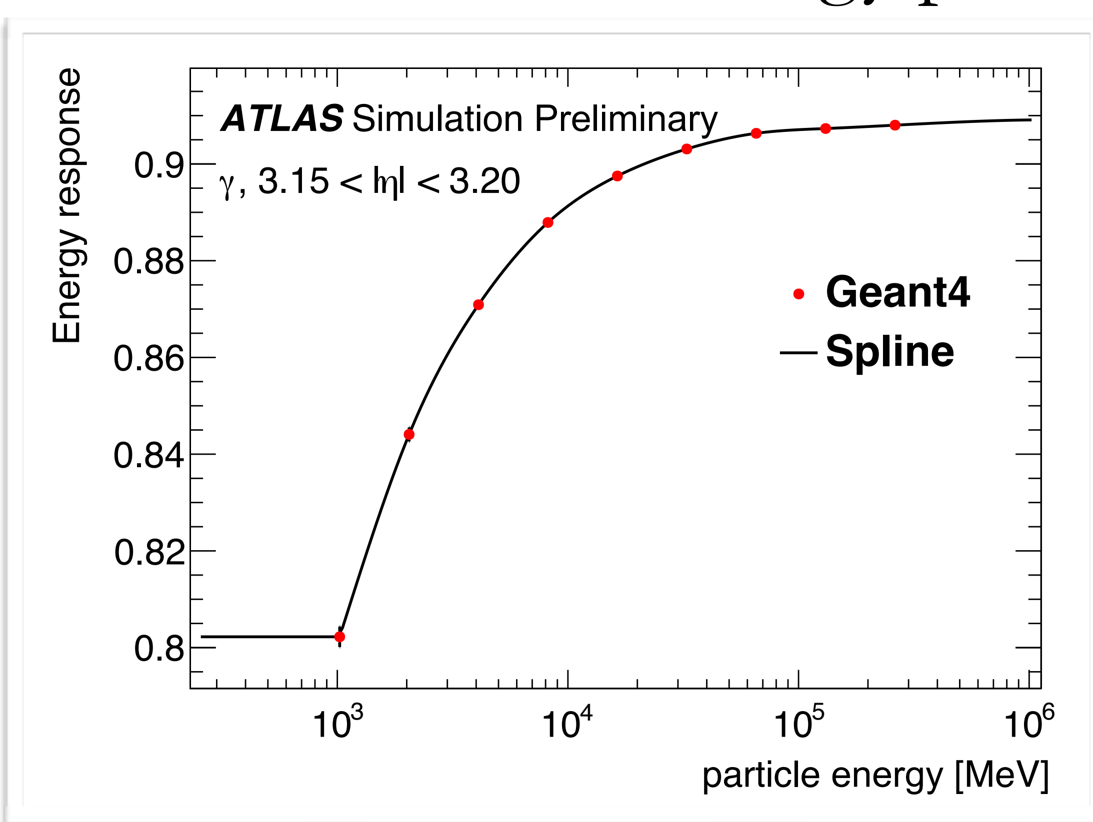
Longitudinal Shower Parametrization

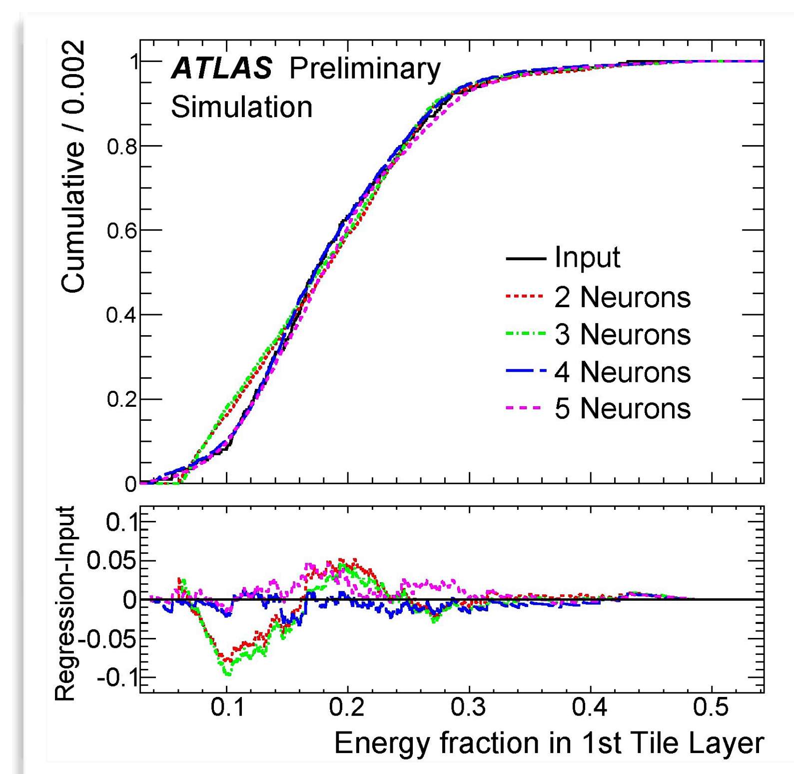
Additional PCA transformation to further decorrelation



Multi-layer perceptron (MLP) for regression of energy cumulants

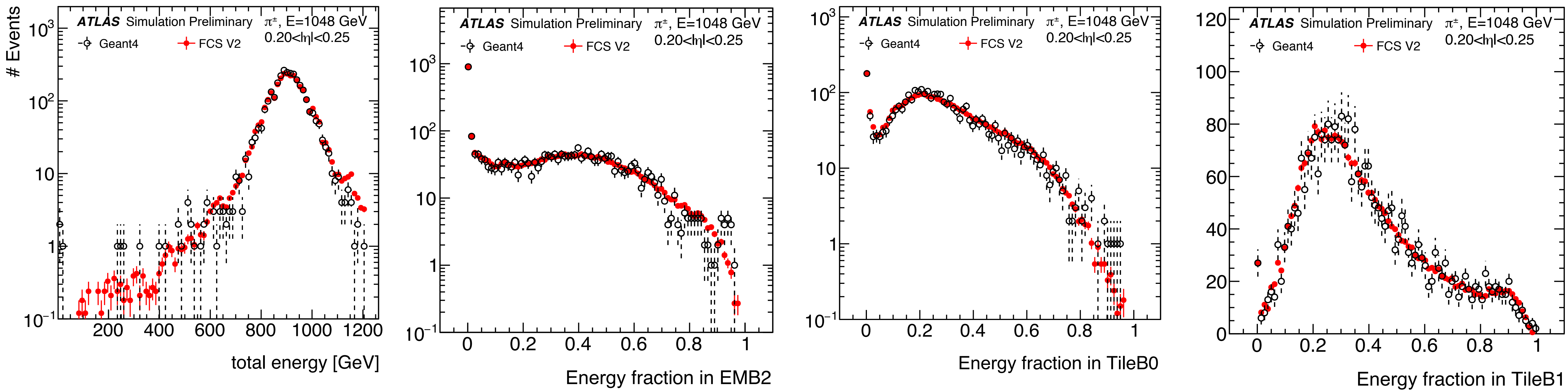
Parametrization of discrete energy points, spline function for interpolation

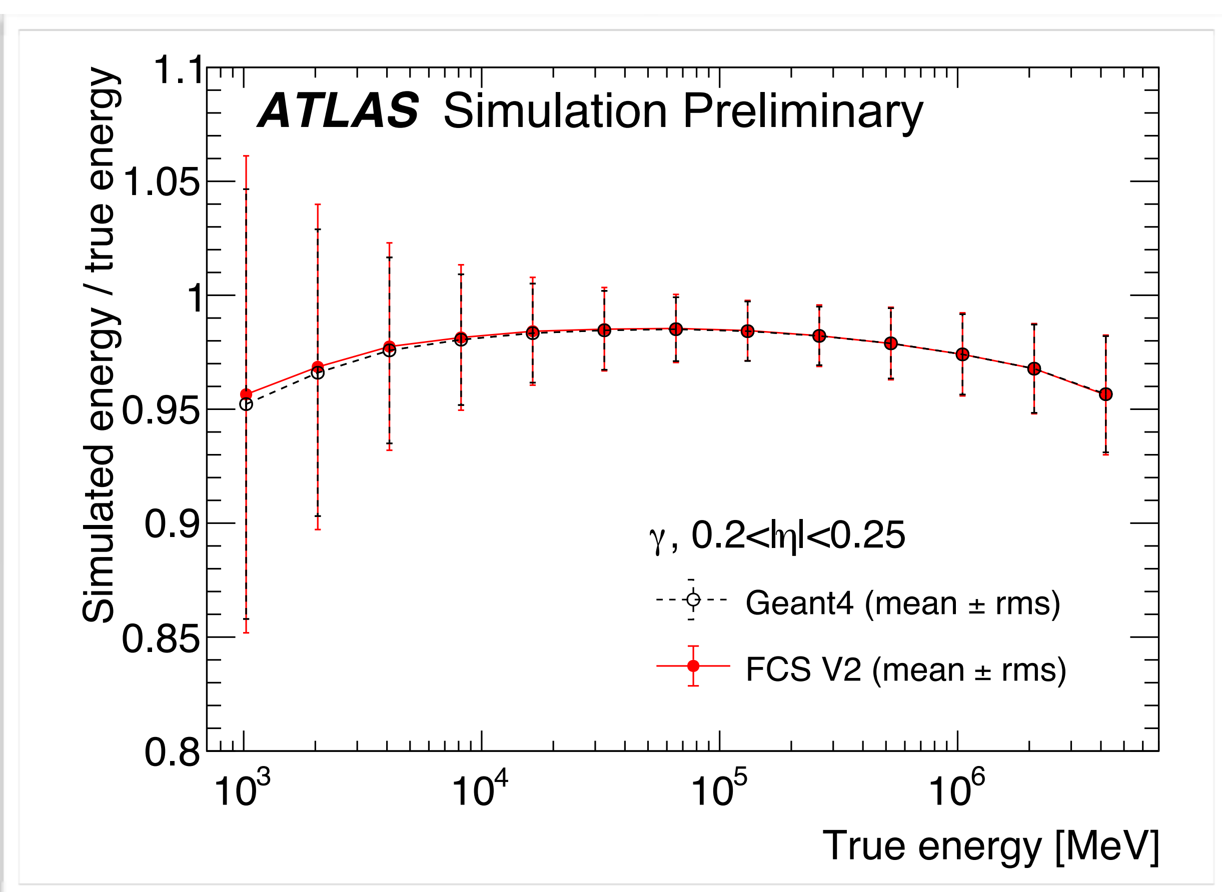


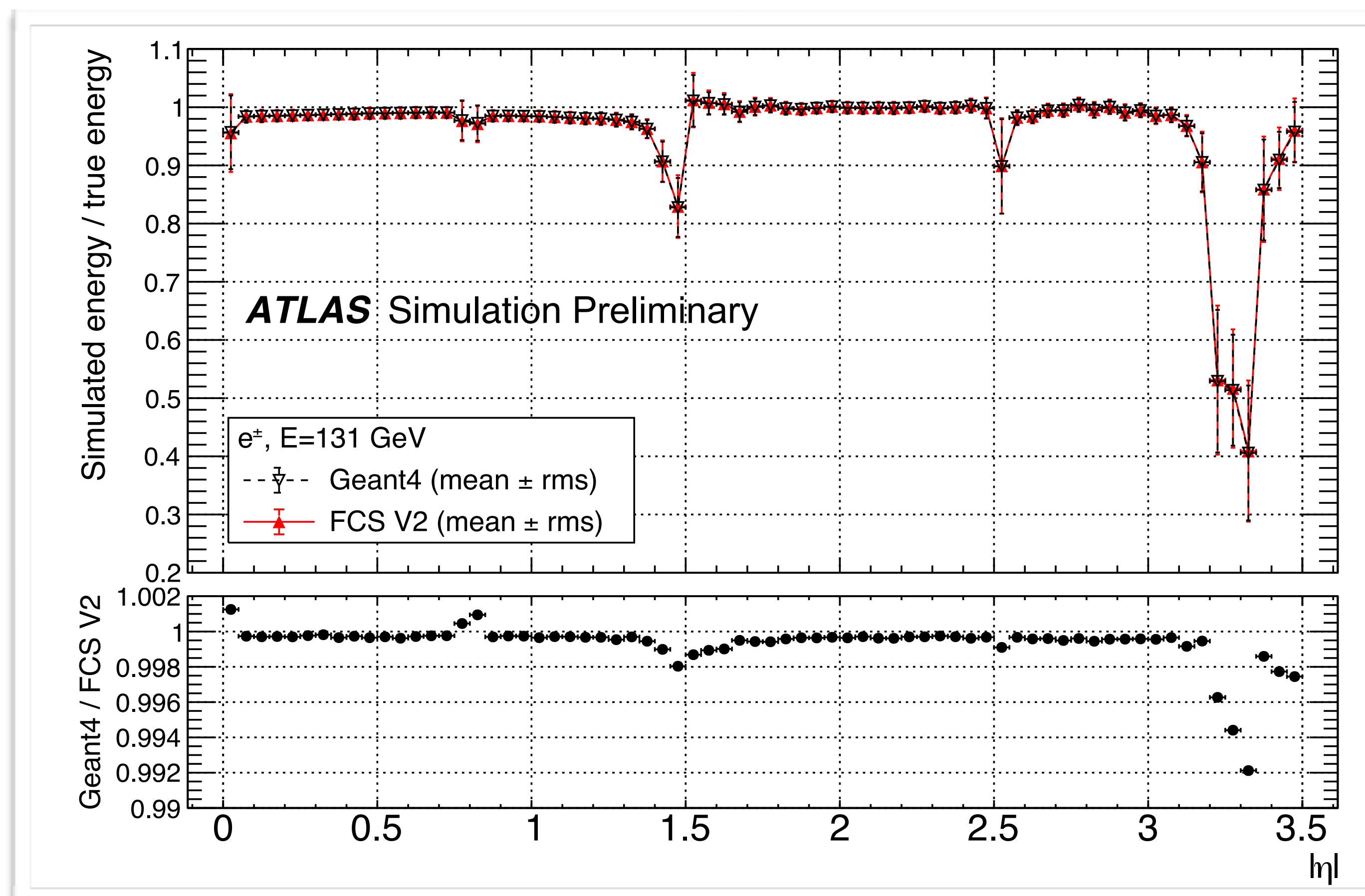


Longitudinal Shower Parametrization

FastCaloSimV2



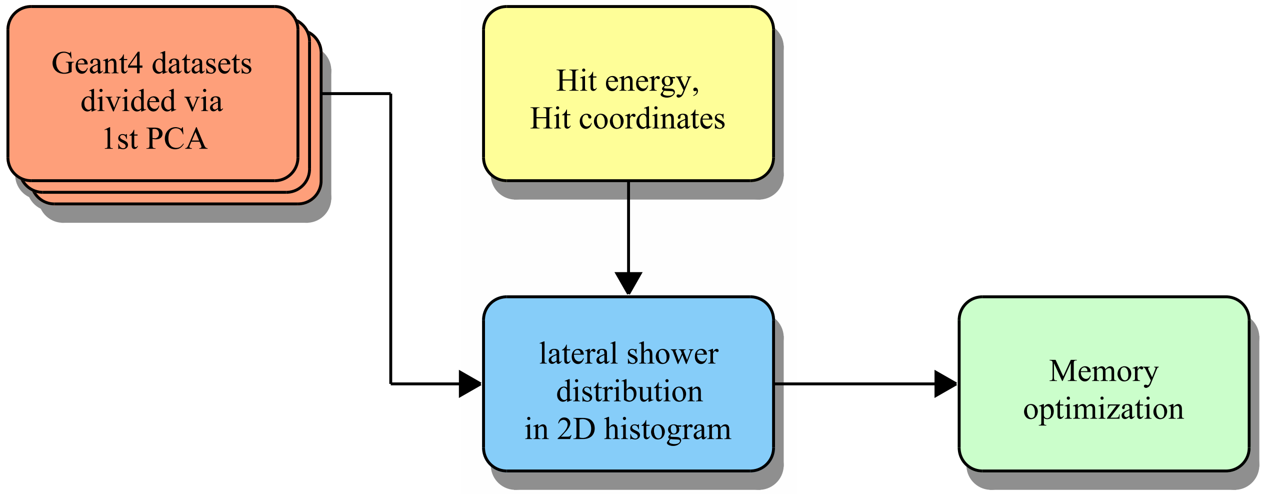




Excellent agreement across various energy/eta regions!

Lateral Shower Parametrization

Lateral shower parametrization performed in each layer and PCA divided dataset

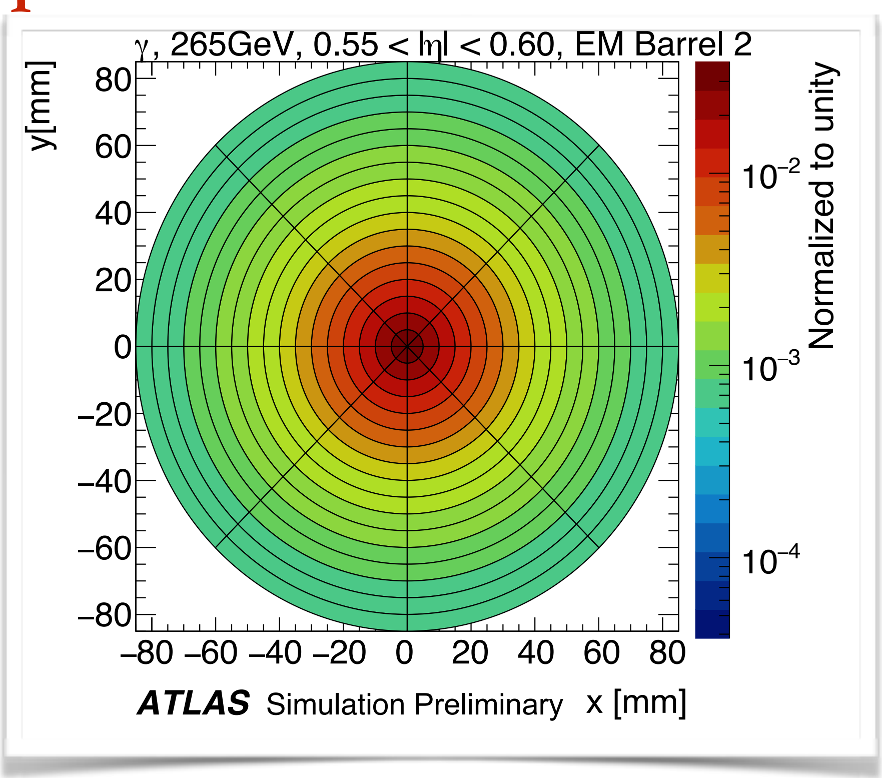


Memory optimization:
- Utilize φ symmetry of the shower
- Use smart rebin/spline in the radial direction

Sample random hit positions from the histogram for simulation

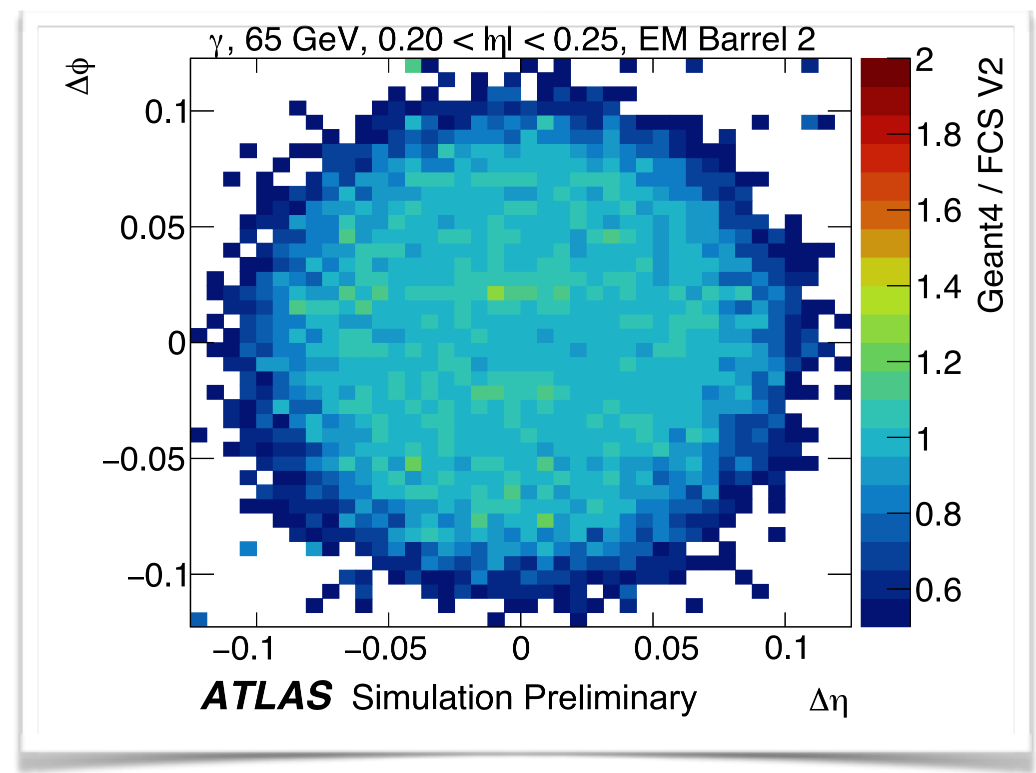
Simulated showers generated from parametrization and compared to Geant4

parametrized shower



simulation

simulated shower

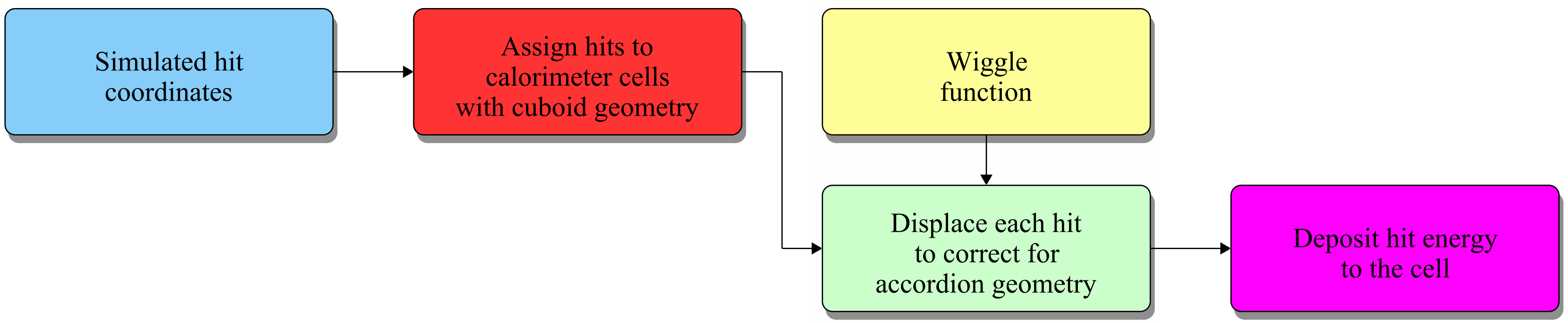


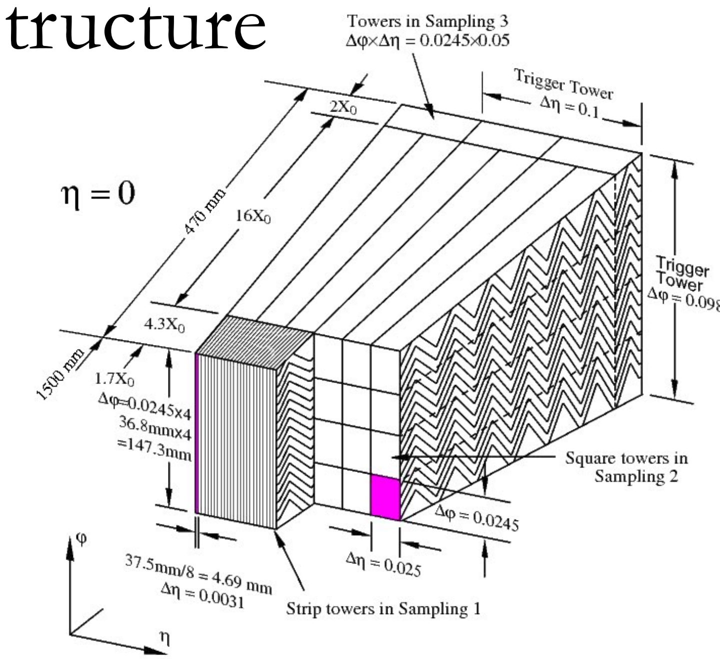
Lateral Shower: Assign hits to cell

FastCaloSimV2

Simulated hits assigned to cells assuming simplified cuboid geometry

Electromagnetic calorimeter have accordion structure

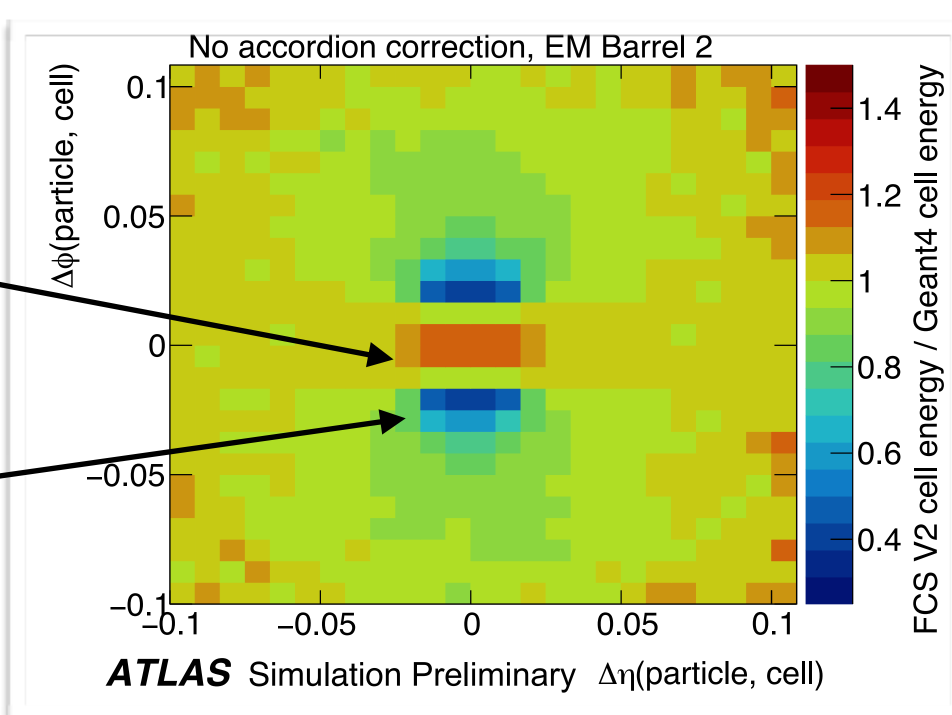


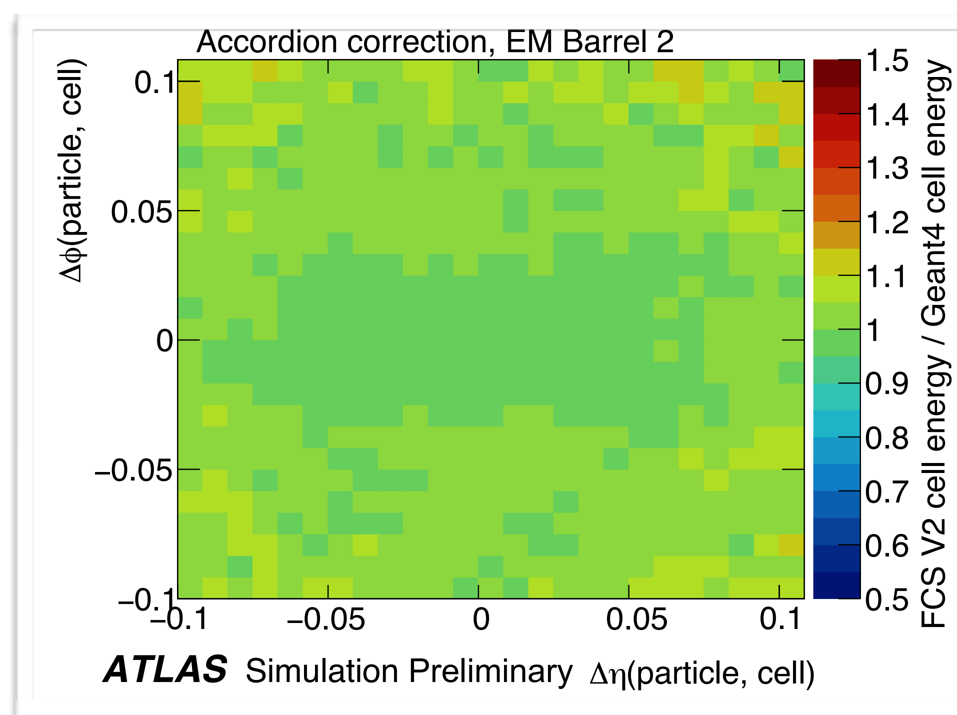


Wiggle the hit position based on a function describing the probability that a hit belongs to a neighboring cells

too much energy

not enough energy

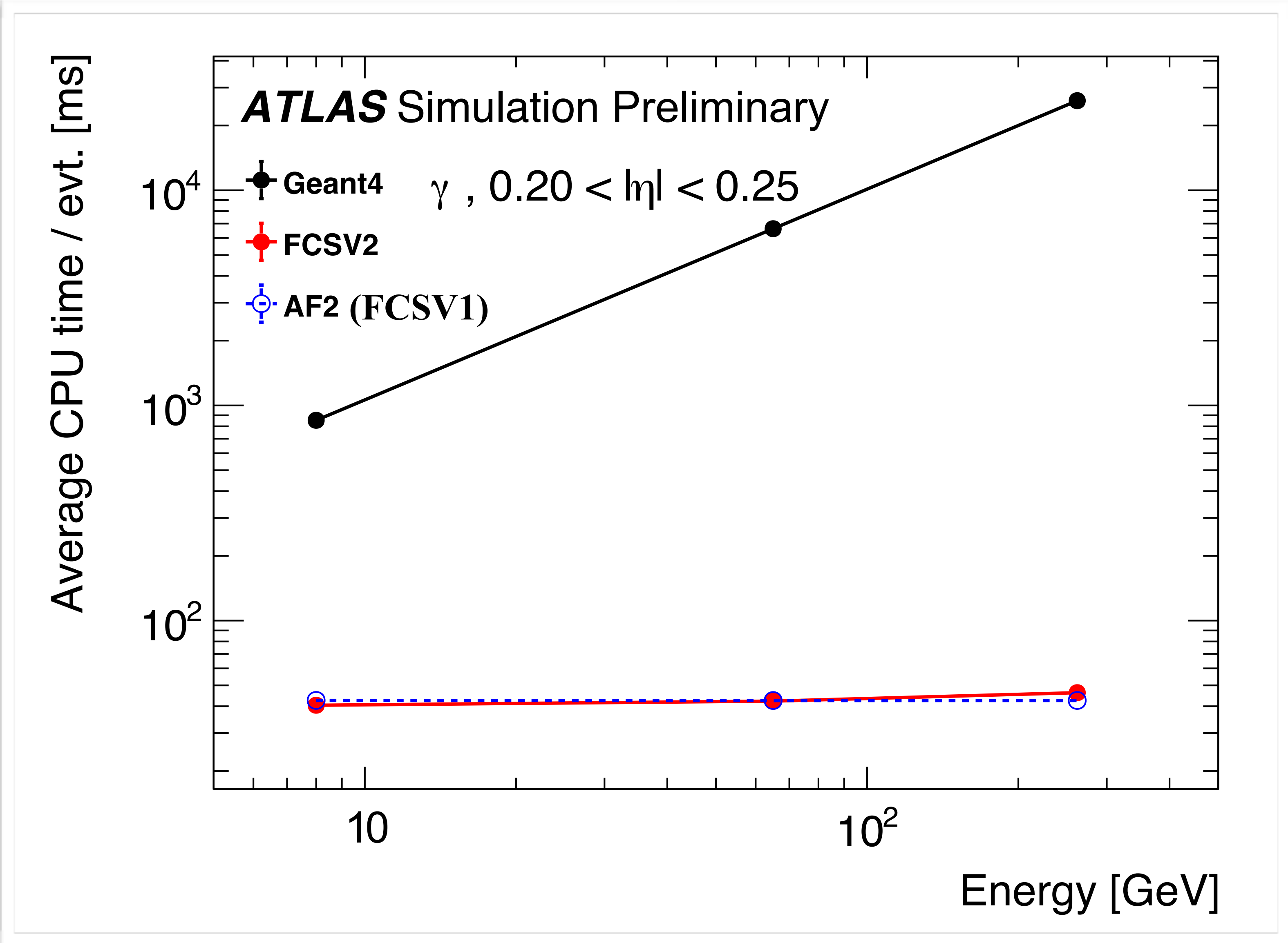




FastCaloSimV2 Performance: CPU

Single particle simulation compared to Geant4 and FastCaloSimV1

Particles are generated on the calorimeter surface



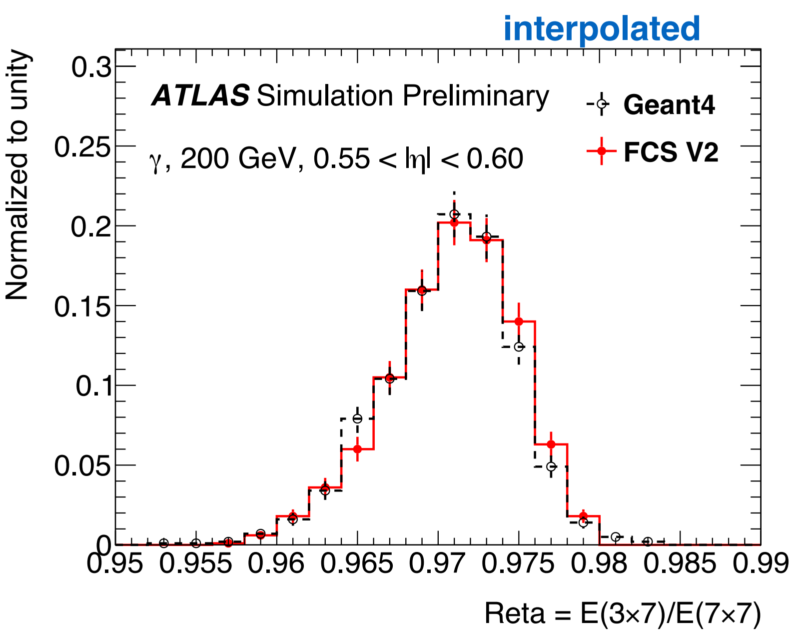
A factor of ~10-25 times faster than Geant4

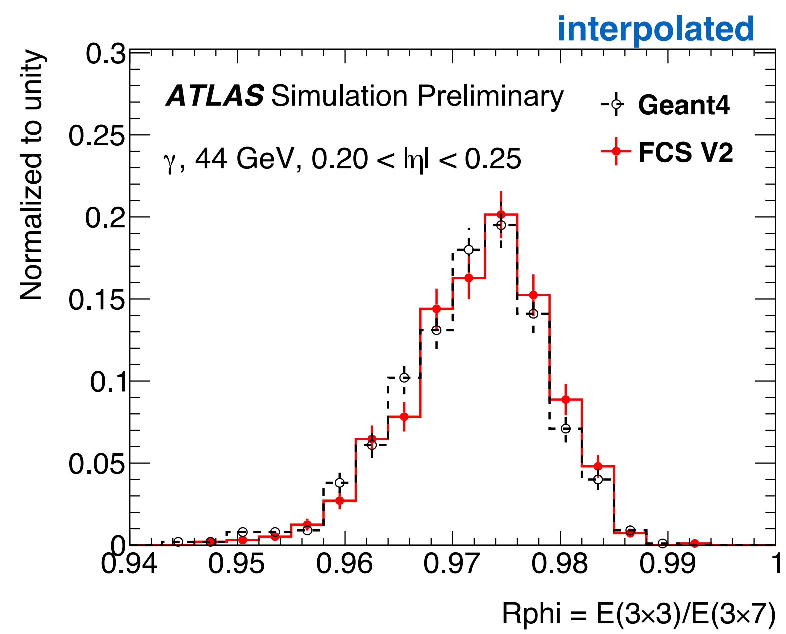
Similar performance as FastCaloSimV1 (AF2) out of the box!

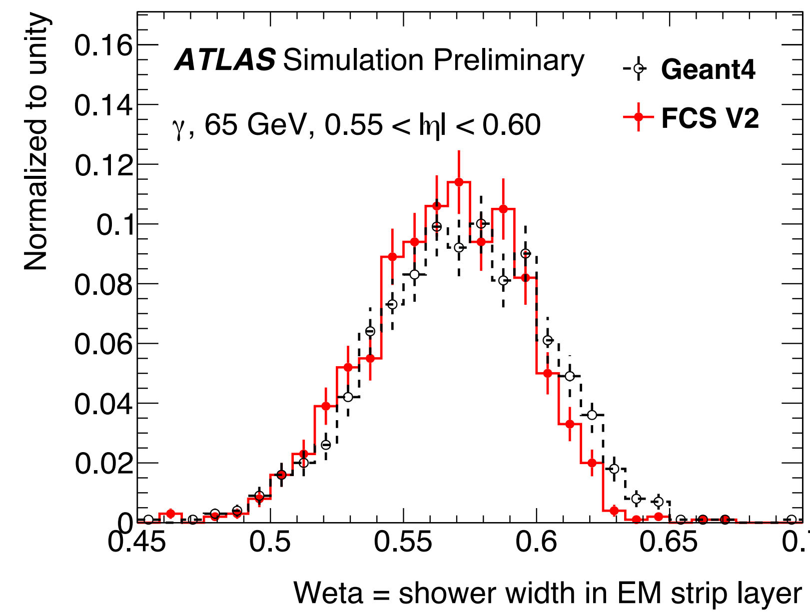
FastCaloSimV2 Performance

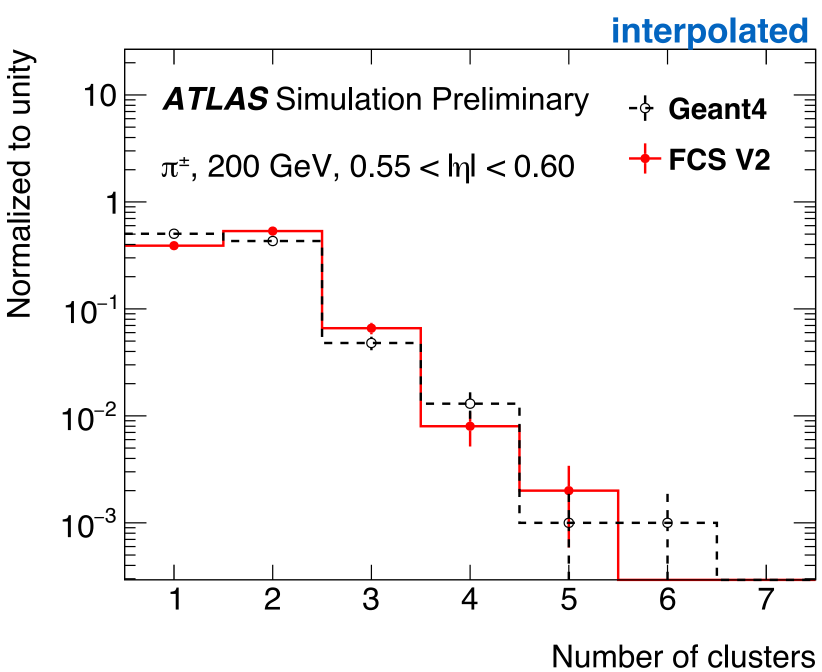
Single particle events are simulated with FastCaloSimV2

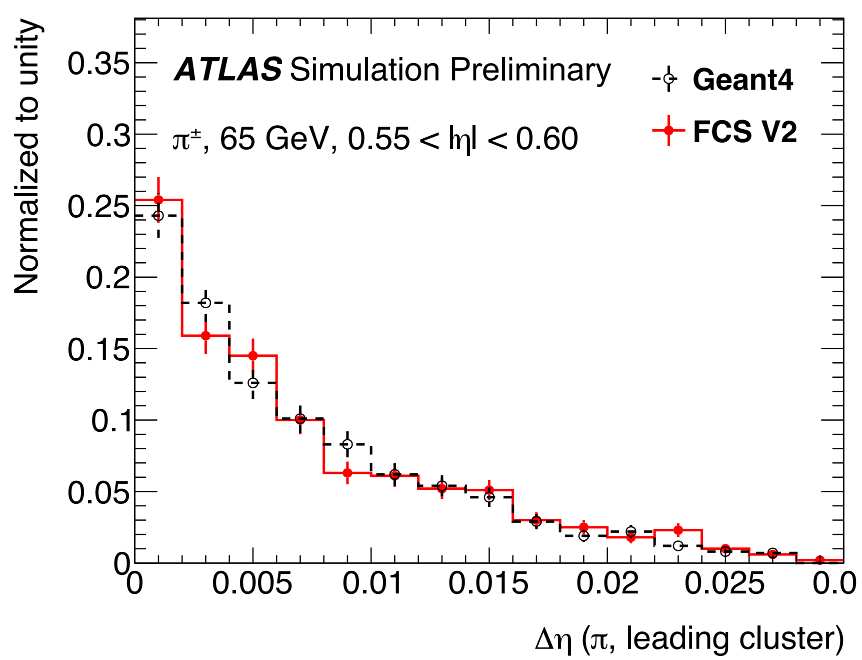
Various EM shower and cluster variables are calculated and compared to Geant4

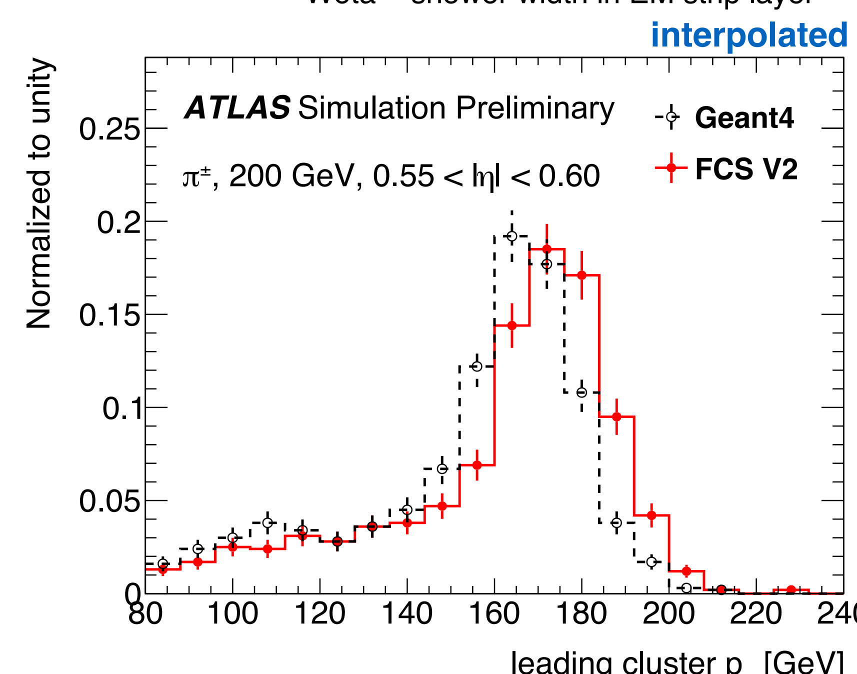






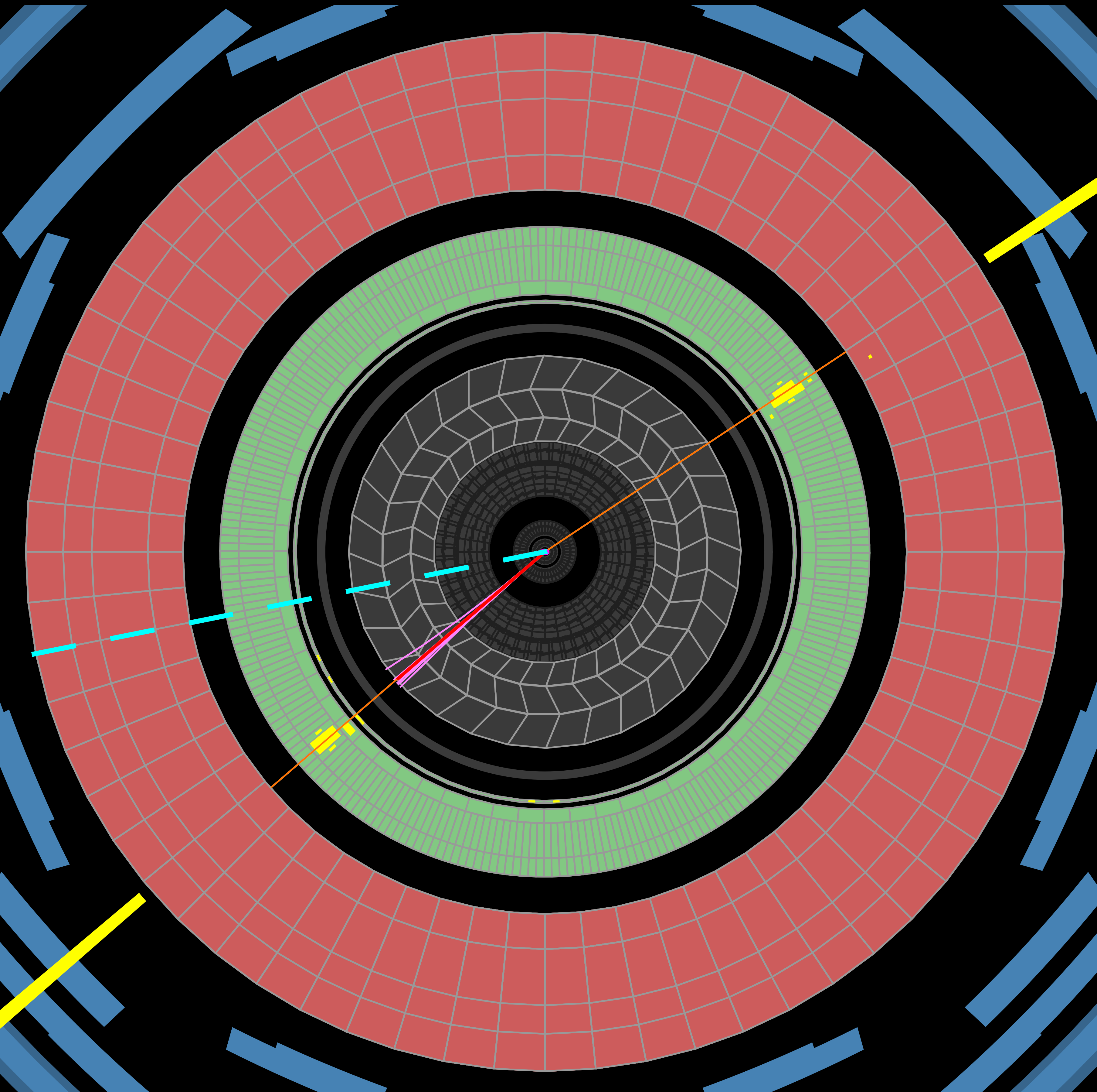






Good agreement over various regions without any corrections applied!

Event Display of a FCS V2 simulated H →γγ event



ATLAS EXPERIMENT

FCS V2, H -> γγ MC

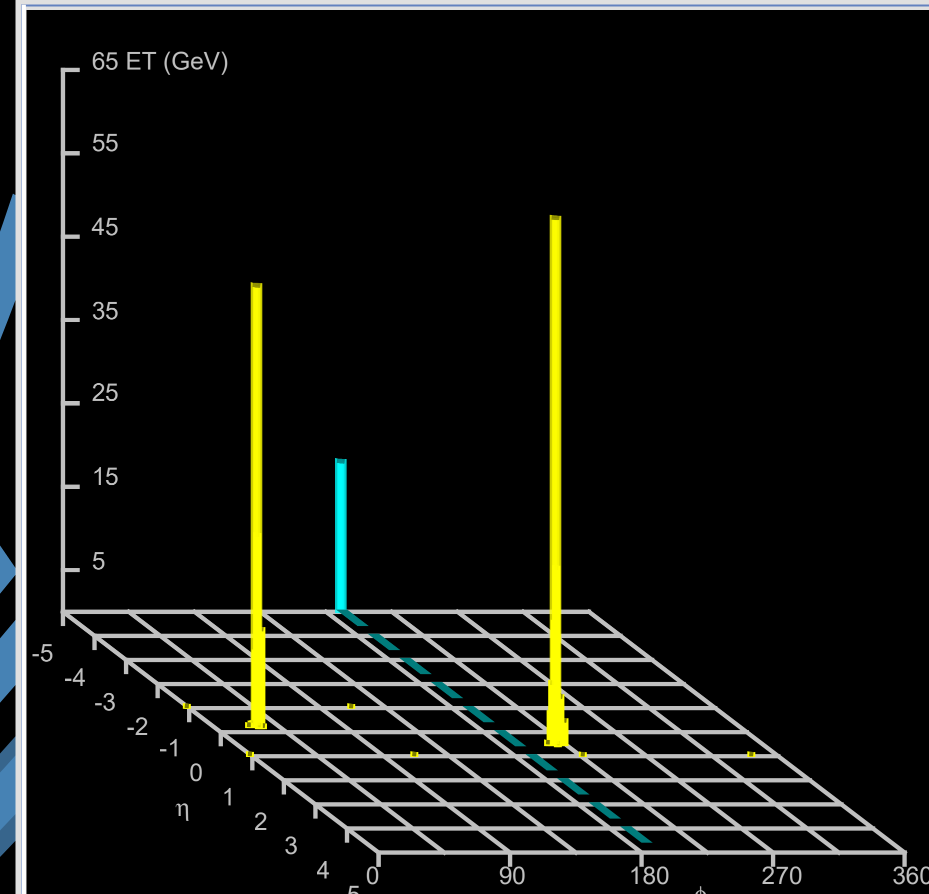
Reconstructed photon

Reconstructed track

MET

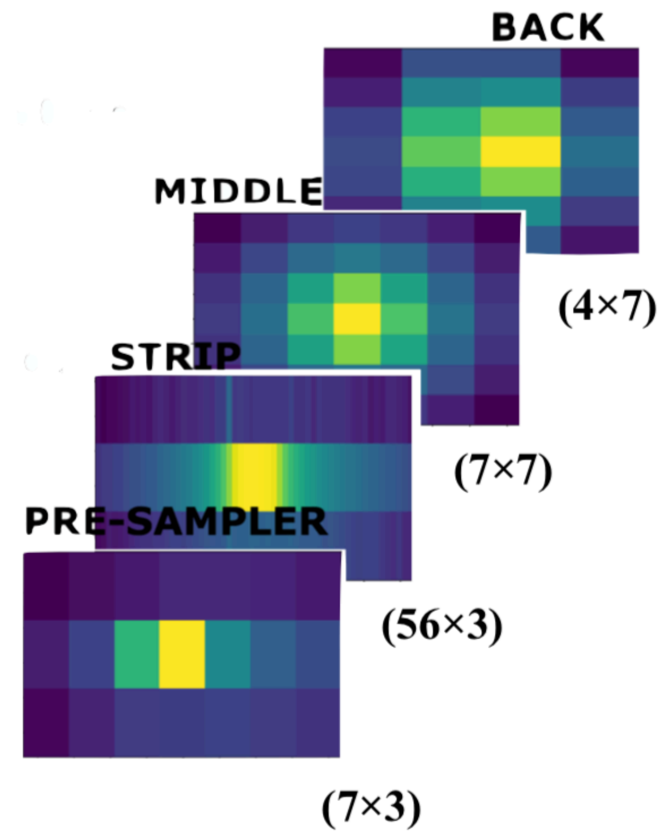
Simulated charged particle

Simulated neutral particle

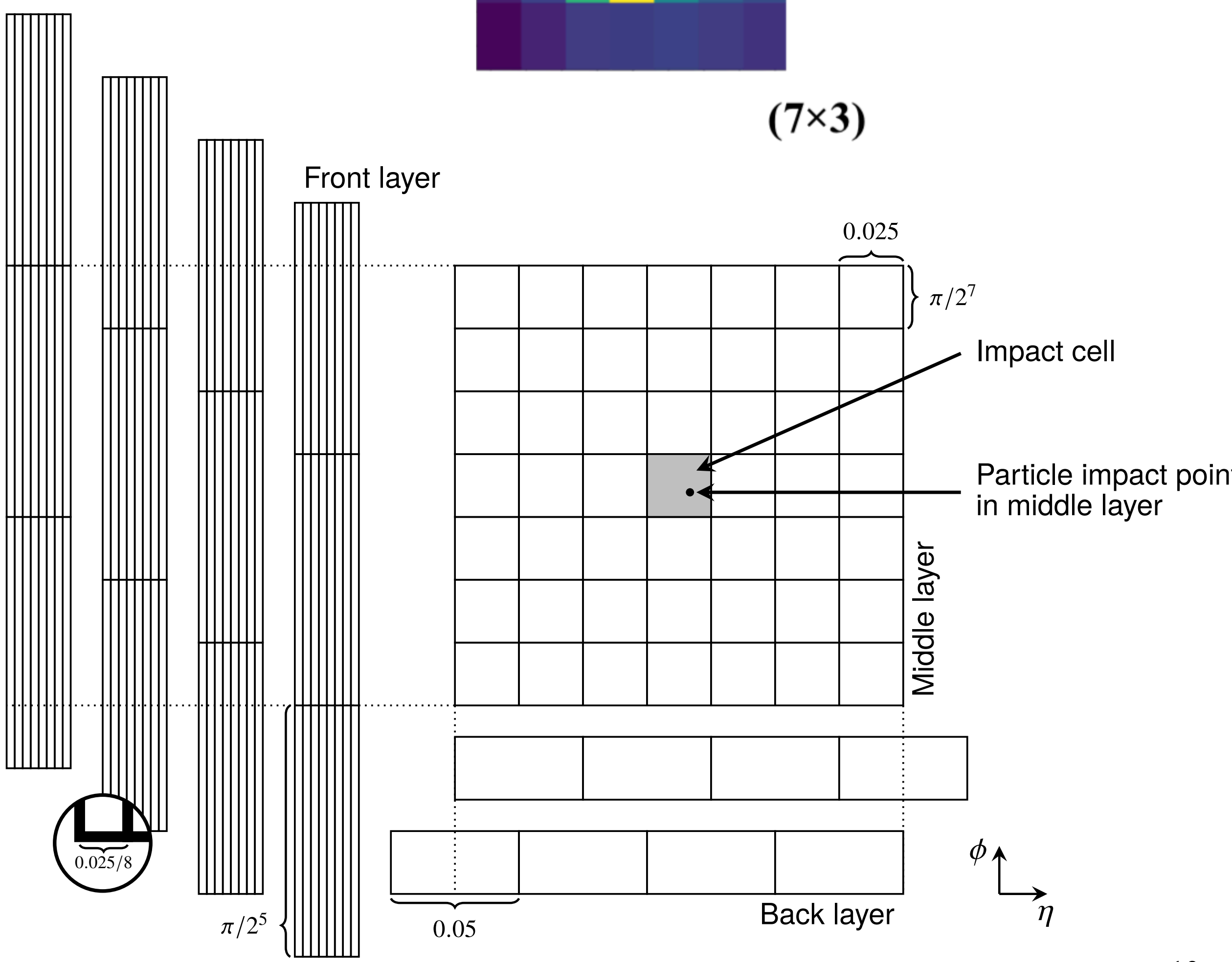


Deep generative networks to generate EM showers

Networks investigated:
Variational Auto Encoder (VAE)
Generative Adversarial Network (GAN)

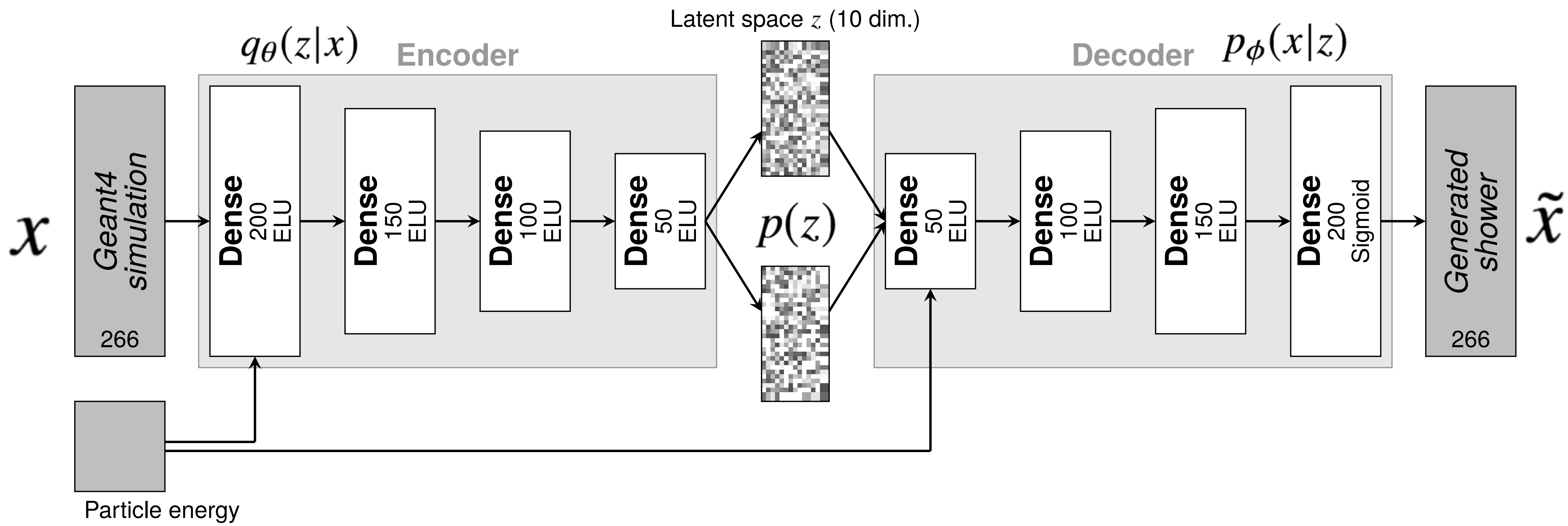


- Only photons in EM calorimeter (< 1% leakage to hadronic calorimeter)
- Energies [1, 260] logarithmically spaced
- Pseudo rapidity $0.20 < |\eta| < 0.25$
- The energy deposits are voxalized into rectangular shapes
- A total of 266 cells are considered for energy deposits
- The networks are trained with energies normalized to the energy of the incident particle



Variational Auto Encoder

Unsupervised deep learning with variational Bayesian method



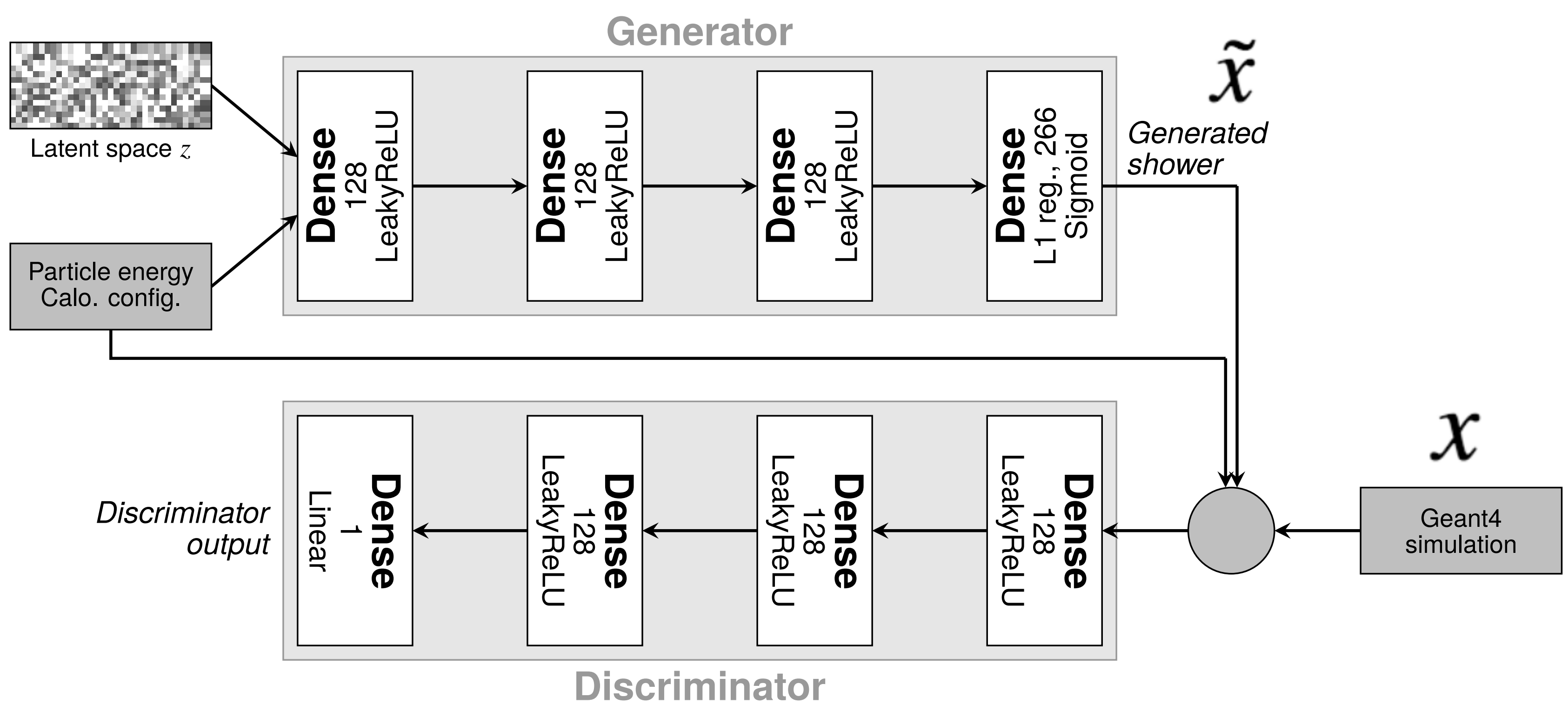
Encoder and decoder used together to maximize the negative log likelihood of the loss function

$$\mathcal{L}_{\mathrm{VAE}} = -\underbrace{w_{\mathrm{reco}} E_{z\sim q_\theta(z|x)}[\log p_\phi(x|z)]}_{\text{reconstruction loss}} + \underbrace{w_{\mathrm{KL}} KL(q_\theta(z|x)||p(z))}_{\text{regularizer}}$$

$$+\underbrace{w_{E_{\mathrm{tot}}} L_{E_{\mathrm{tot}}}(x,\tilde{x})}_{\text{total energy}} + \sum_i^M \underbrace{w_i L_{E_{\mathrm{i}}}(x,\tilde{x})}_{\text{energy fraction}}$$

Generative Adversarial Network

Generative network with a feedback from a Discriminator network



Improve the robustness of training by calculating Wasserstein loss with a two sided gradient penalty

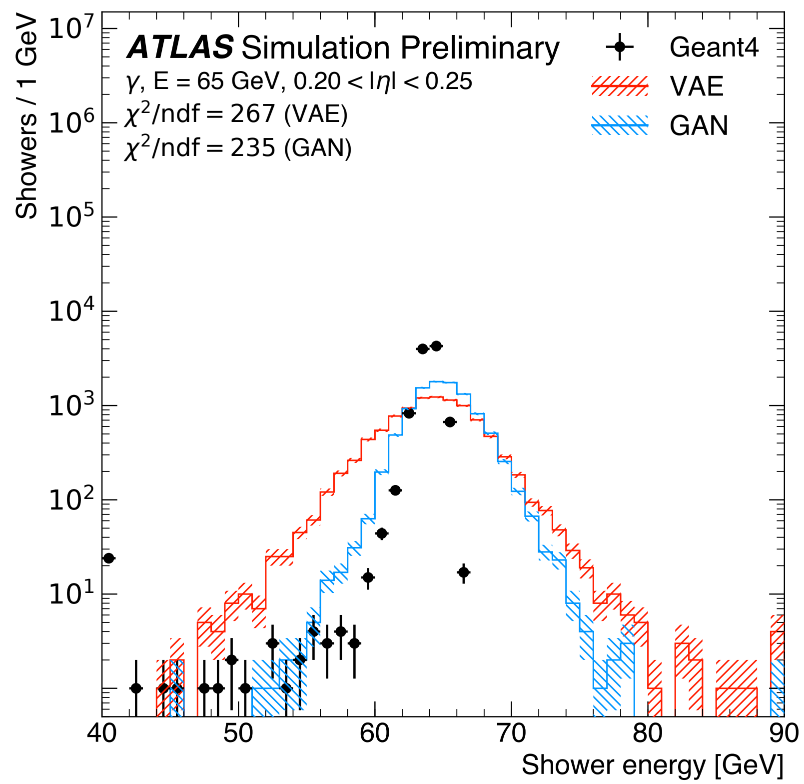
$$L_{\text{GAN}} = \underset{\tilde{x} \sim p_{\text{gen}}}{E}[D(\tilde{x})] - \underset{x \sim p_{\text{Geant4}}}{E}[D(x)] + \lambda \underset{\hat{x} \sim p_{\hat{x}}}{E}[(||\Delta_{\hat{x}} D(\hat{x})||_2 - 1)^2].$$

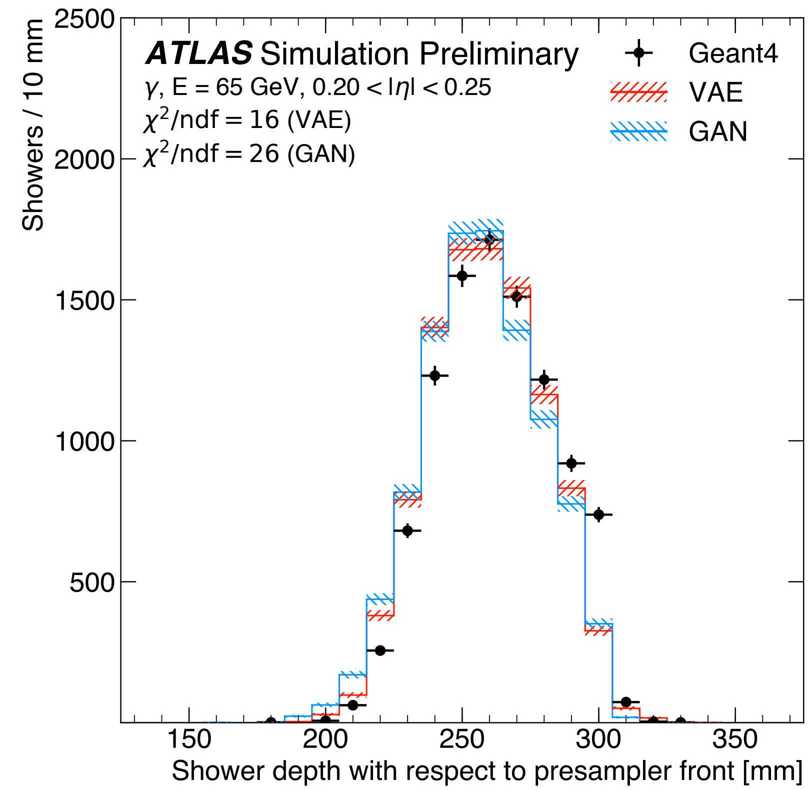
ability to identify generated shower correctly

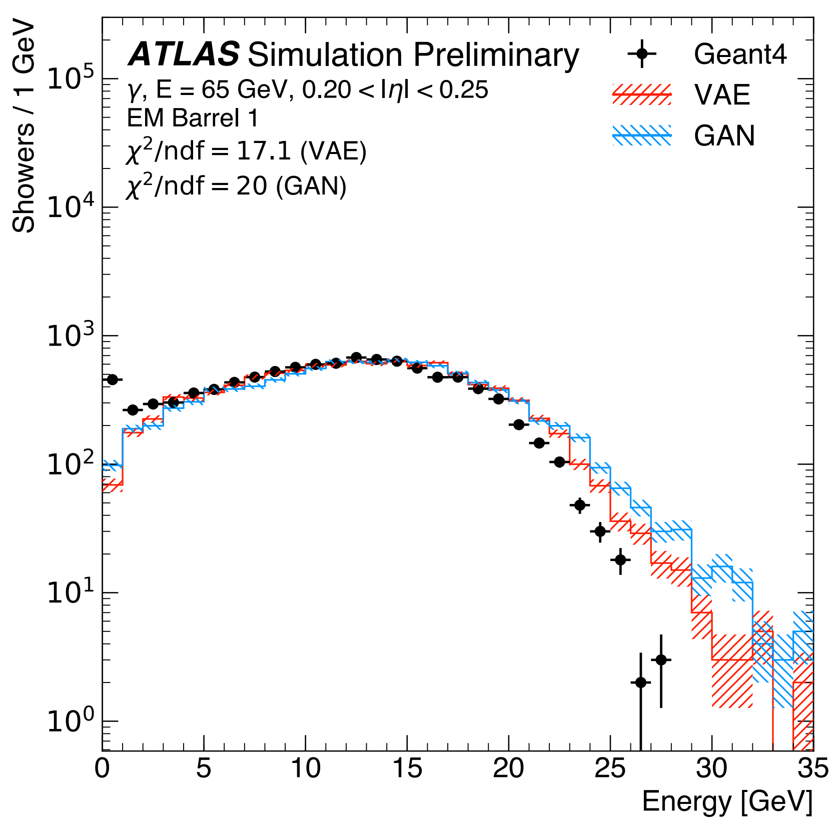
ability to identify Geant4 shower correctly

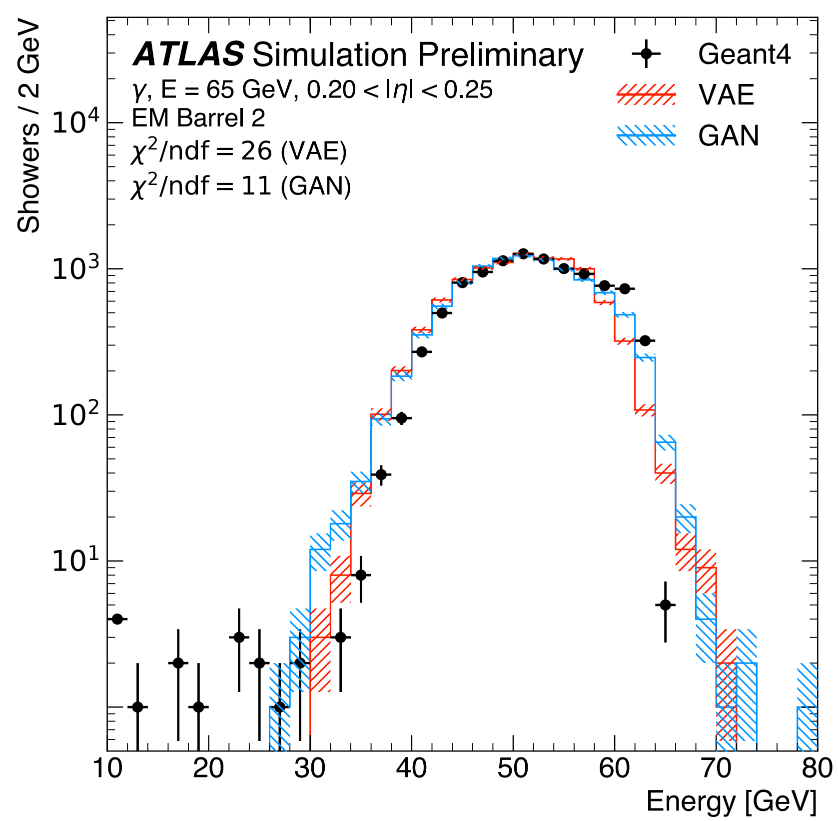
penalizes by calculating Wasserstein loss

Performance in Longitudinal Shower



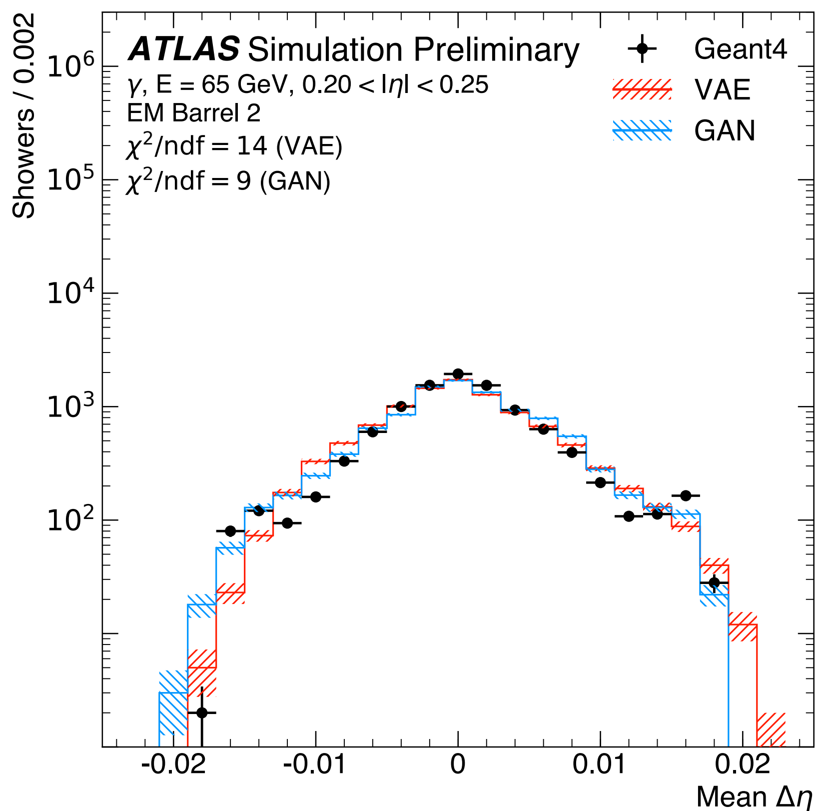


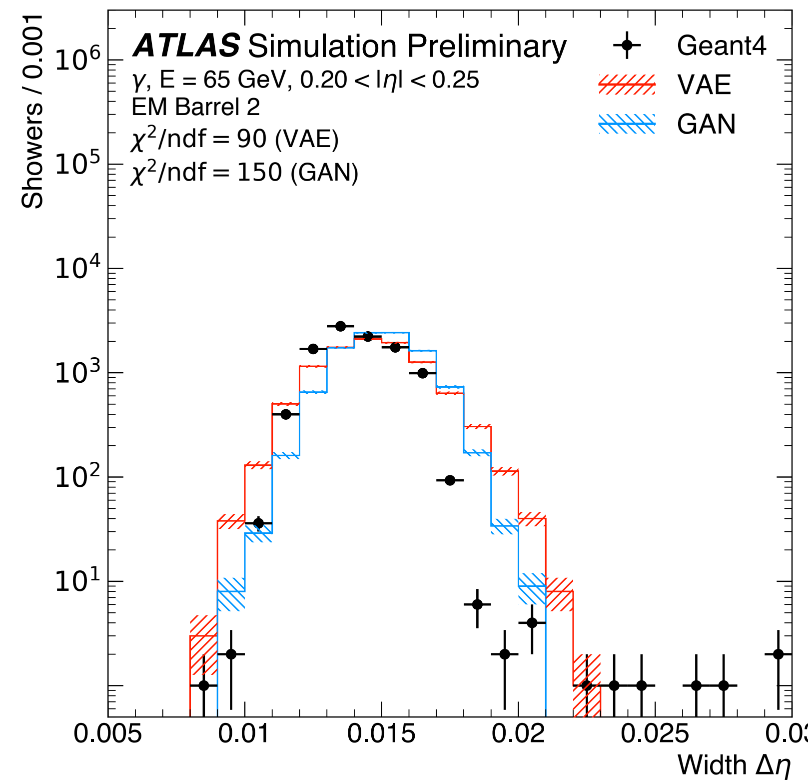


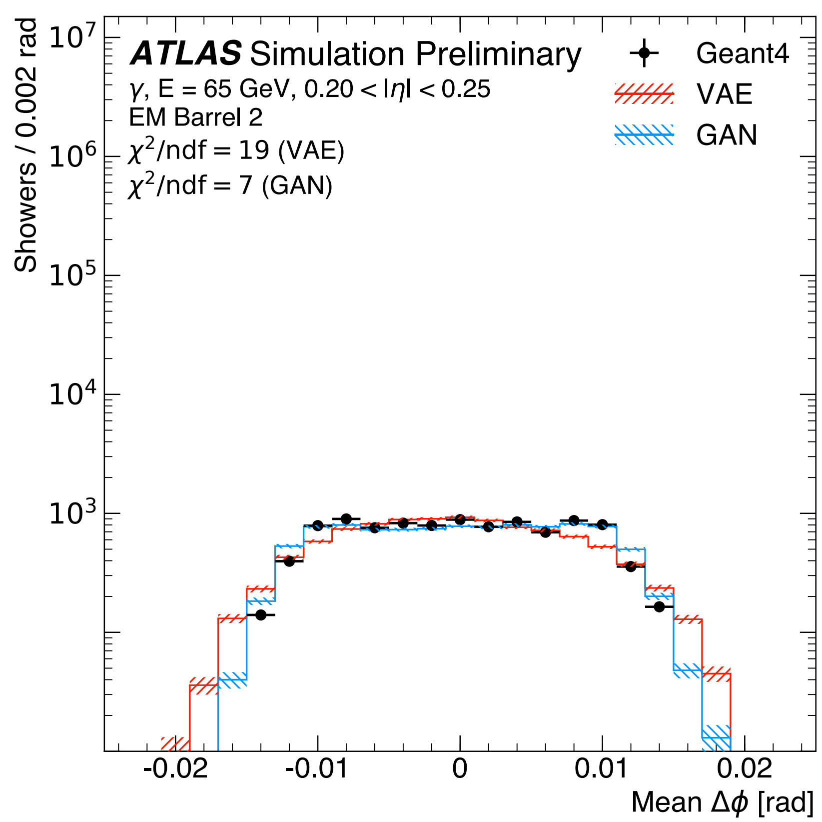


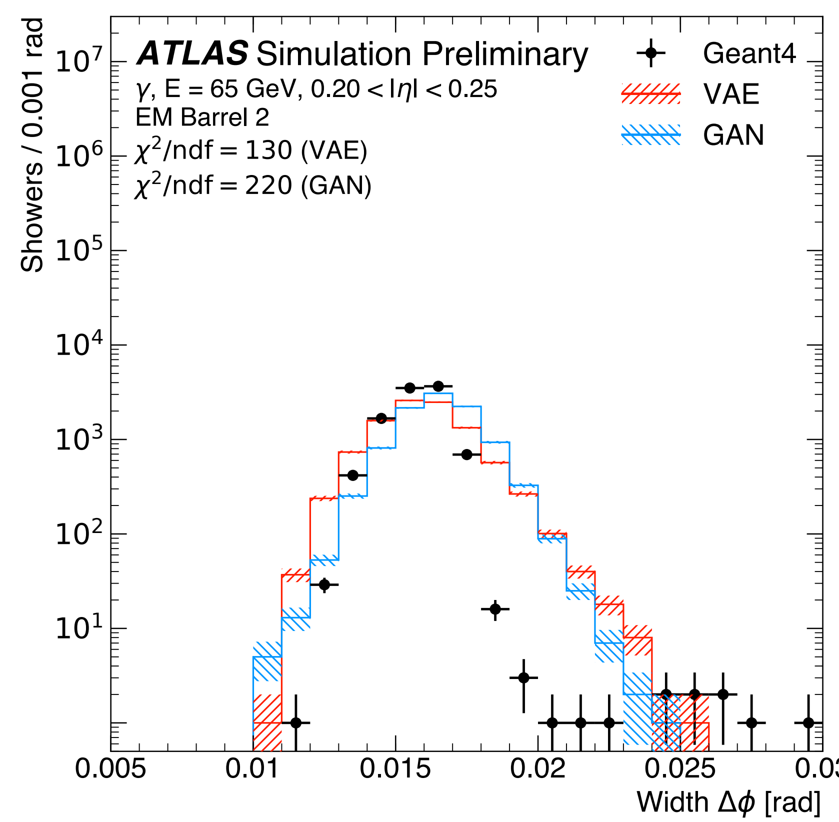
ATLAS EXPERIMENT

Performance in Lateral Shower









Summary

Fast shower simulation is essential for ATLAS physics program

FastCaloSimV1 does not describe collision data adequately to be used in precision measurements

Several approaches of fast simulation is under active development

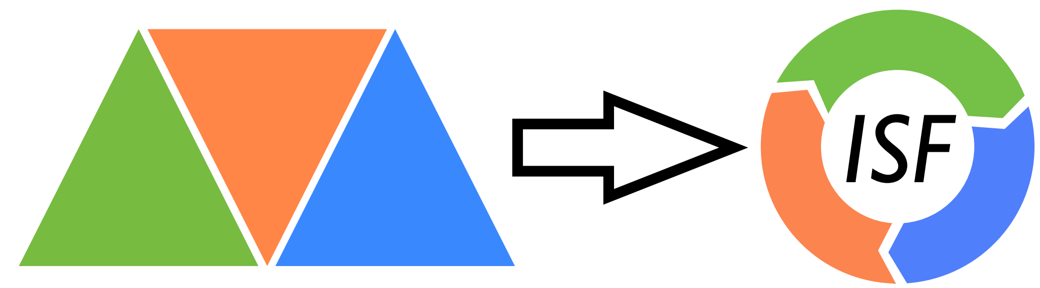
FastCaloSimV2 shows good agreement with Geant4 and is expected to be in production soon

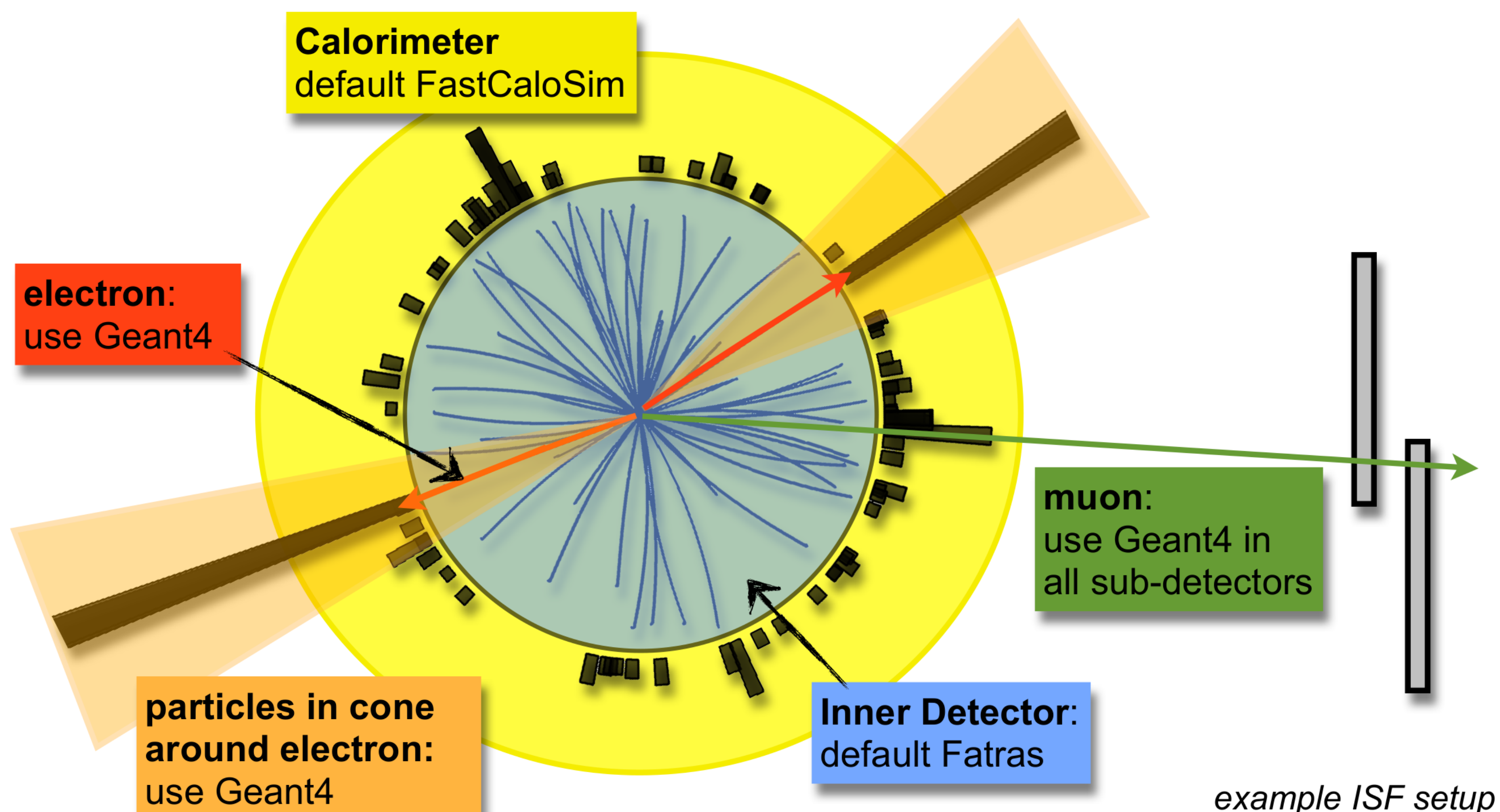
DNNCaloSim have shown promising results as the first application of generative models and continue the development towards achieving required accuracy for physics analyses

BONUS

ISF - Integrated Simulation Framework

- Combines different simulation approaches in ATLAS into one framework
  - ➡ Output format is always the same independent c simulation chosen
  - ➡ Configuration is done a one central place and standardized
  - ➡ Fast and full simulation setup can be mixed and used alongside
- Compatible with multithreading and multiprocessing





example ISF setup

Calorimeter fast simulation can be combined with full simulation of Inner Detector/Muon Systems based on physics requirements

Energy Parametrization

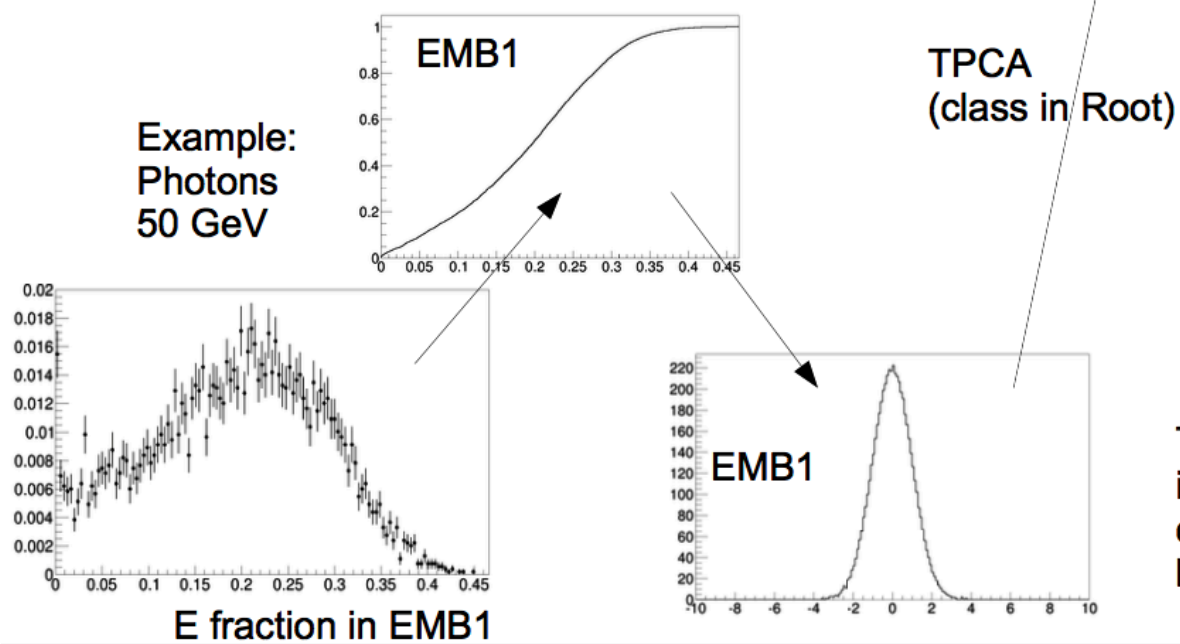
1st PCA chain:

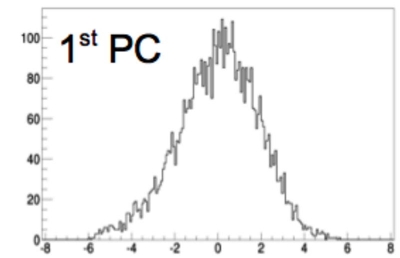
G4 Inputs: Energy fractions f, Total energy → N inputs → **Cumulative distributions** → Inverse error function → Gaussians → **PCA** → **PCA output data** N components

N outputs

Example:
Photons
50 GeV

TPCA
(class in Root)





The first principle component is that eigenvector of the covariance matrix with the largest eigenvalue (variance)

Bin 5

Bin 1

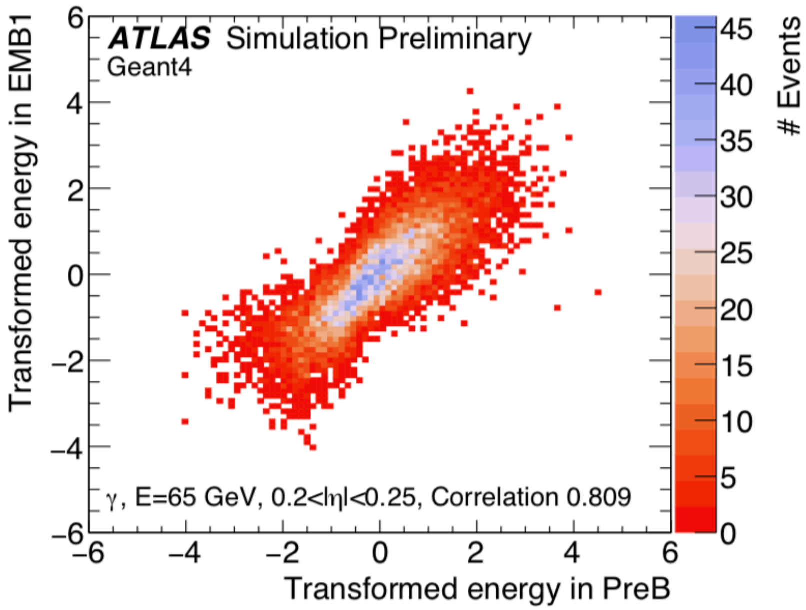
G4 Inputs: Energy fractions f, Total energy → N inputs → **Cumulative distributions** → Inverse error function → Gaussians → **PCA** → **PCA output data** N components

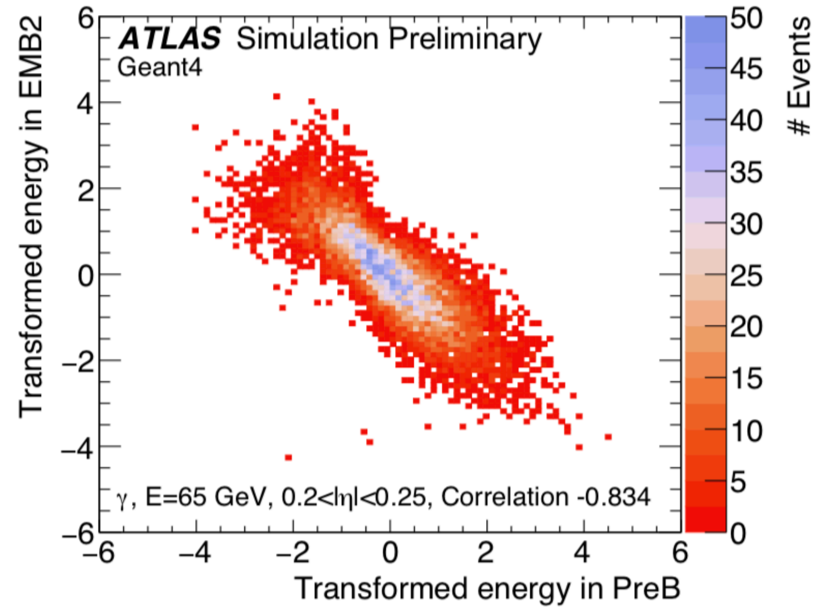
During simulation, this chain is performed back-wards:

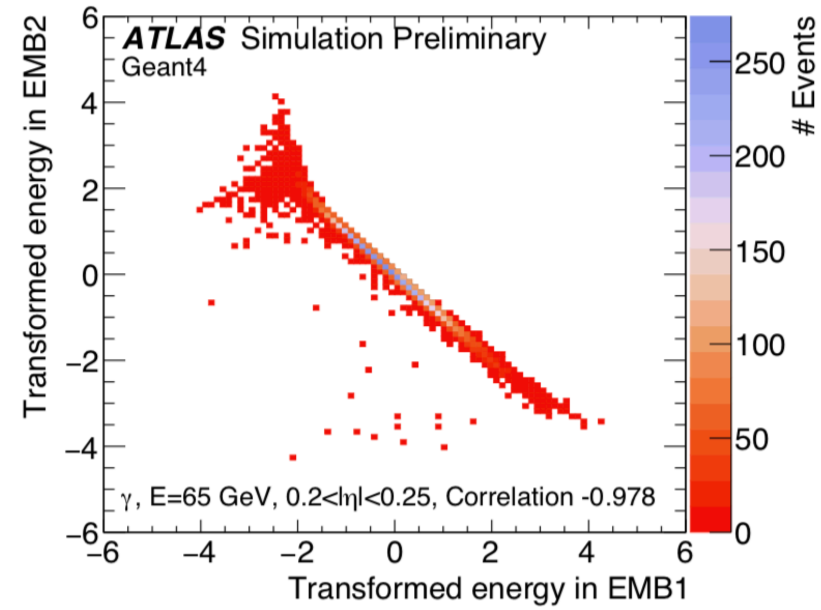
Gaussian random numbers → **Inverse PCA** → Inverse PCA output (Gaussians) → Error function → Uniform numbers [0,1] → **Inverse Regression Or Histogram** → **Simulated Inputs**

PCA correlations

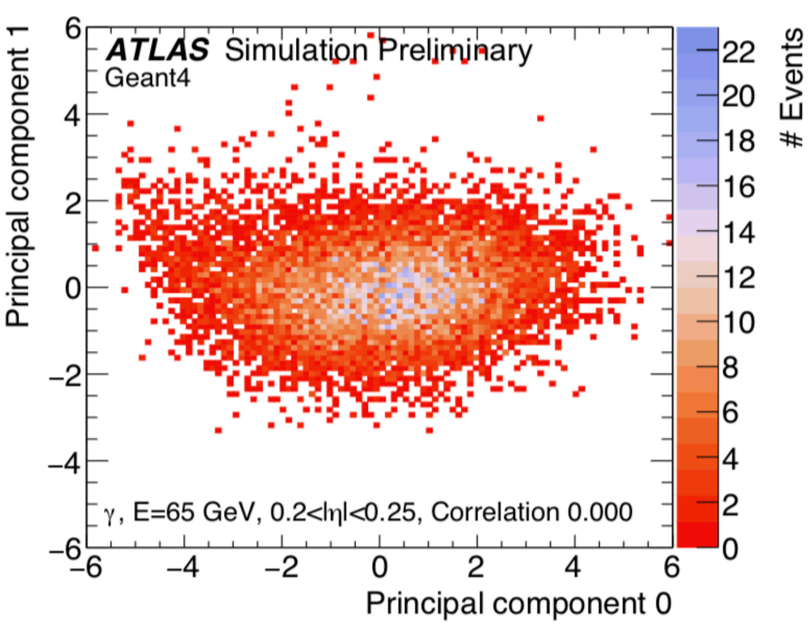
Before PCA transformation

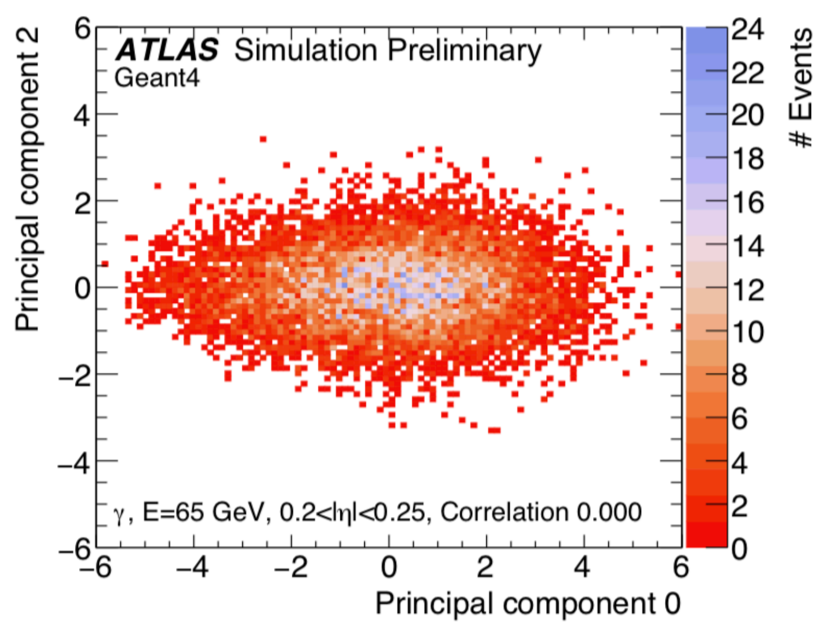


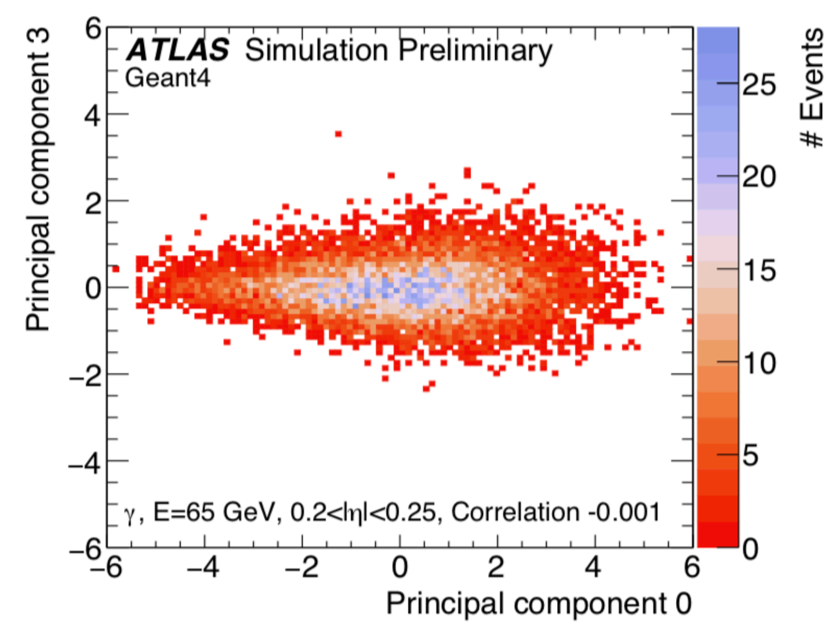




After PCA transformation







Lateral Shower Simulation

FastCaloSimV2

Randomly sample hit position from the 2D histograms

Number of hits sampled in each layer for a given energy

Determine the number of hits such that the statistical fluctuation corresponds to the stochastic term of energy resolution of each layer:

$$\frac{\Delta E}{E} = \frac{\alpha}{\sqrt{E}} \oplus \beta \oplus \frac{\gamma}{E}$$

The position of each hit in global coordinates is calculated using a numeric solution

Sufficient to describe fluctuations is electromagnetic showers

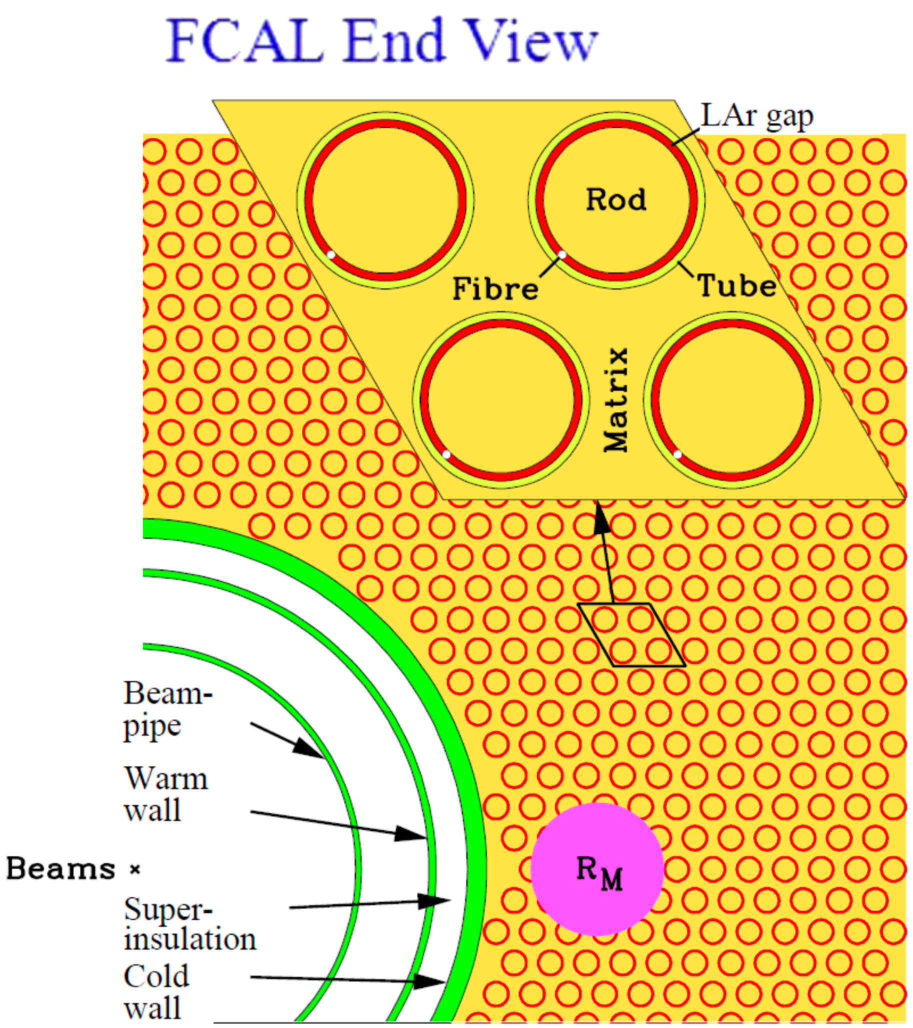
Forward Calorimeter Geometry

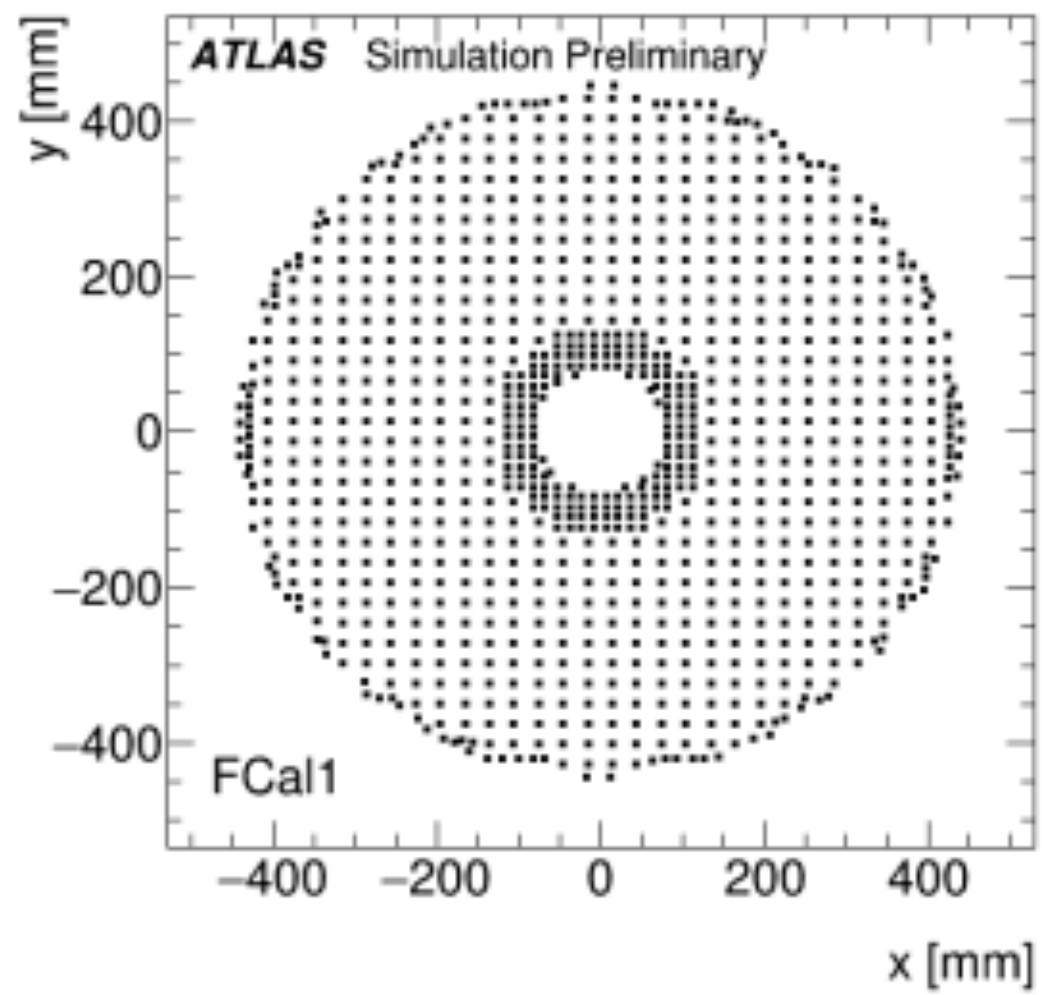
FastCaloSimV2

Cylindrical anodes are arranged in a rhombus-like formation for the forward calorimeters (FCal)

Significantly different geometry compared to cuboid barrel layers

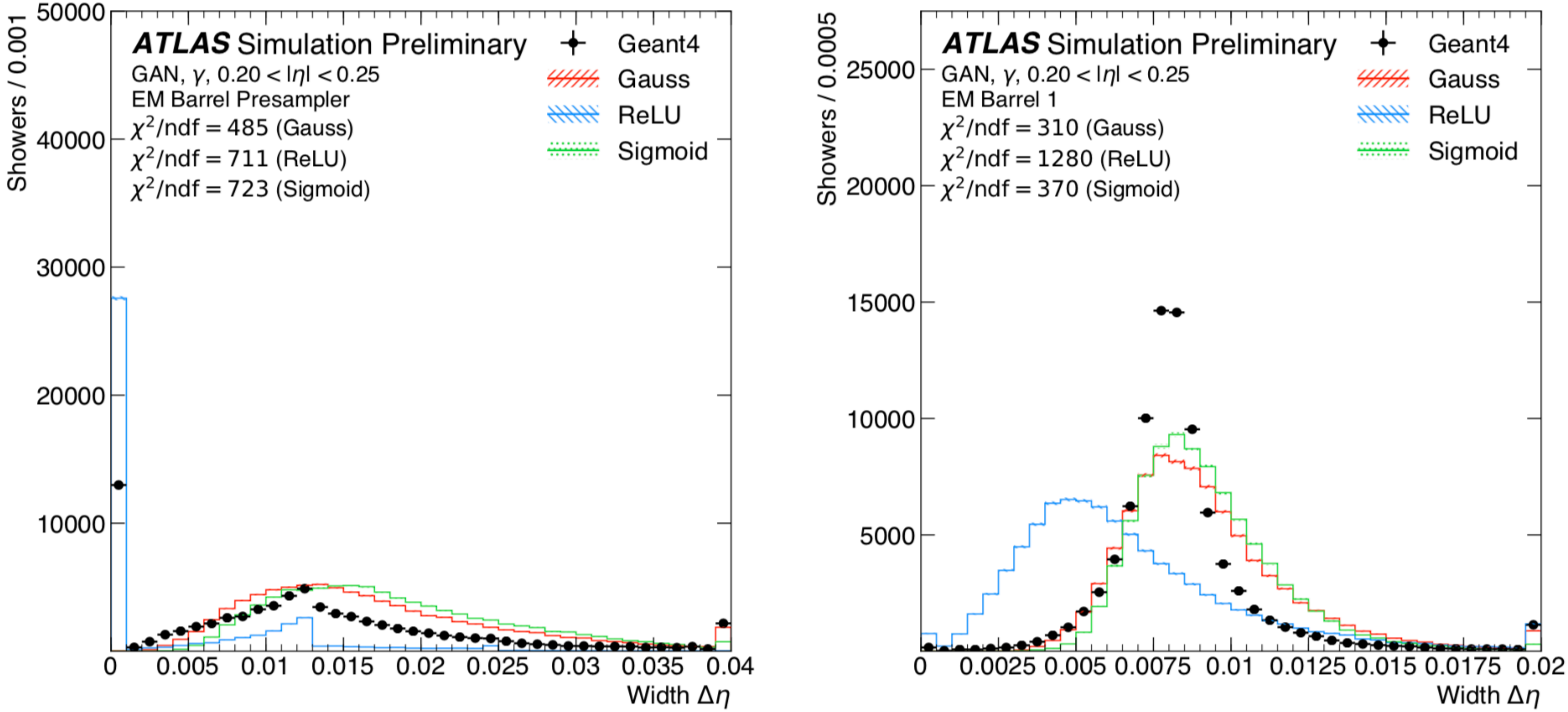
Correct geometry is implemented in the FastCaloSimV2





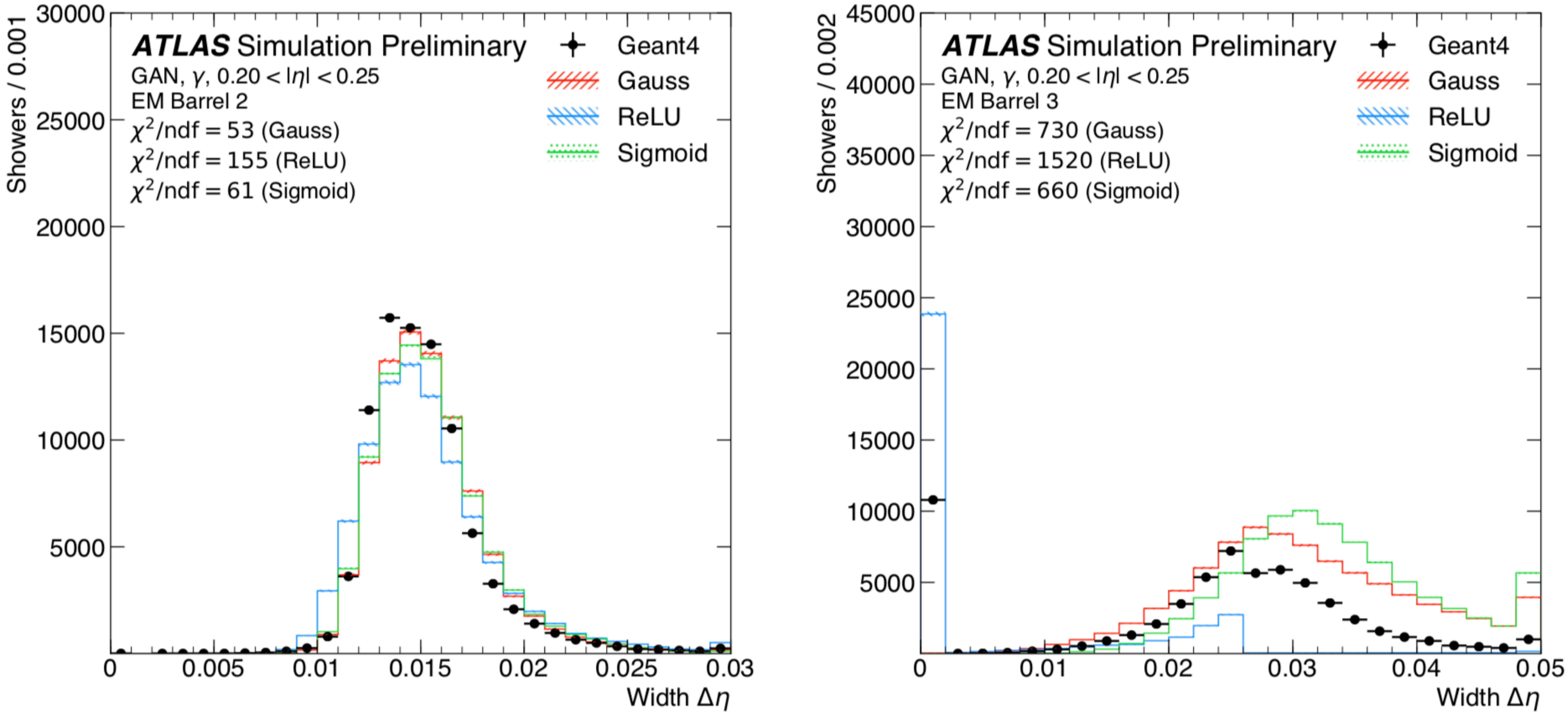
Dedicated parametrization for* FCals *are foreseen

Activation functions(1)



(a) Presampler (b) Front layer

Activation functions(2)



(c) Middle layer

(d) Back layer

Hyperparameters: VAE

Hyperparameter	Values
Latent space dim.	[1, . . . , **10**, . . . , 100]
Reco. weight	(0, . . . , **1**, . . . , 3]
KL weight	(0, . . . , $\mathbf{10^{-4}}$, . . . , 1]
E_{tot} weight	[0, . . . , $\mathbf{10^{-2}}$, . . . , 1]
E_i weights	[0, . . . , $\mathbf{8 \times 10^{-2}}$, . . . , 1]
	[0, . . . , $\mathbf{6 \times 10^{-1}}$, . . . , 1]
	[0, . . . , $\mathbf{2 \times 10^{-1}}$, . . . , 1]
	[0, . . . , $\mathbf{10^{-1}}$, . . . , 1]
Hidden layers (encoder)	1, 2, 3, **4**, 5
Hidden layers (decoder)	1, 2, 3, **4**, 5
Units per layer	[180, . . . , **200**, . . . , 266]
	[120, . . . , **150**, . . . , 180]
	[80, . . . , **100**, . . . , 120]
	[10, . . . , **50**, . . . , 80]
Activation func.	**ELU** [22], ReLU [22], SELU [30] , LeakyReLU [31], PReLU [32]
Kernel init.	zeros, ones, random normal, random uniform, truncated normal, **variance scaling**, `glorot_normal` [33]
Bias init.	zeros, **ones**, random normal, random uniform, truncated normal, variance scaling, `glorot_normal` [33]
Optimizer	**RMSprop** [28], Adam [34], Adagrad [35], Adadelta [36], Nadam [37, 38]
Learning rate	[10^{-2}, . . . ,$\mathbf{10^{-4}}$, . . . , 10^{-6}]
Mini-batch size	50, **100**, 150 , 1000

Table 1: Summary the results of the grid search performed to optimize the hyperparameters of the VAE for simulating calorimeter showers for photons. The optimal parameter is typeset in bold font.

Hyperparameters: GAN

Hyperparameter	Values
Hidden layers	1, **3**, 5, 10
Units per layer	64, **128**, 512, 1024
Activation func.	SELU [30] + Sigmoid, **LeakyReLU** [31] **+ {Sigmoid**, Gauss, ReLU [22], Sigmoid + ReLU, clipped ReLU, softmax, softmax + ReLU}
Activity `L1_REG_WEIGHT` (Gen.)	0, $\mathbf{10^{-5}}$, 10^{-2}
Kernel init.	**`glorot_uniform`** [33], `lecun_normal` [47]
Gradient penalty	one-sided, **two-sided**
Gradient penalty weight	0, **10**, 20
Training ratio	20, 10, **5**, 3, 1
Learning rate	$\mathbf{5 \times 10^{-5}}$, 5×10^{-6}, 1×10^{-6} (training ratio 5) 5×10^{-5}, 5×10^{-6}, 1×10^{-5}, 1×10^{-7} (training ratio 3) 1×10^{-6} (training ratio 1)
Mini-batch size	**64**, 1024
Preprocessing (all norm. to E_γ)	$\log_{10} E_{\text{cell}}$, $\log_{10}(E_{\text{cell}} \times 10^{10})$, $\boldsymbol{E_{\text{cell}}}$
Conditioning	$\{E_\gamma$, $\boldsymbol{\log_{10} E_\gamma}\}$ **+ multi-hot encoding of cell alignments**

Table 2: Summary the results of the grid search performed to optimize the hyperparameters of the GAN for simulating calorimeter showers for photons. The optimal parameter is typeset in bold font. In addition to the architectures summarized in the table, generators and discriminators with differing number of hidden layers and units per layer were tested.