Fast Simulation in ATLAS: From Classical to Generative Models

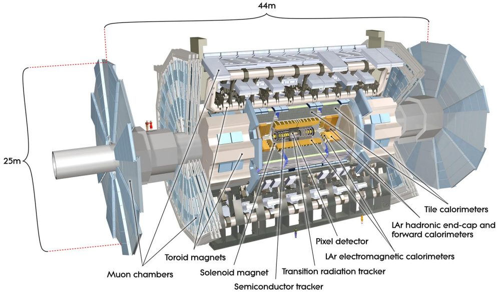
CHEP 2019, Adelaide

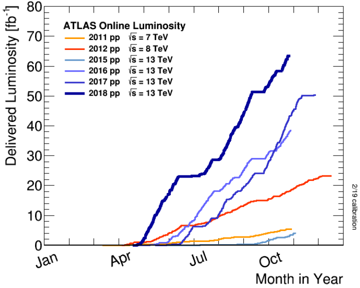
Johnny Raine (Université de Genève)

On behalf of the ATLAS collaboration

4th November, 2019

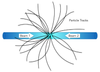
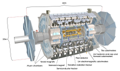
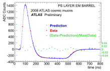
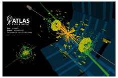
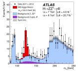
ATLAS Detector



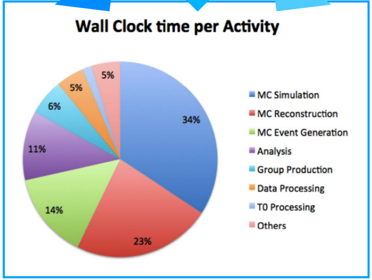


- General purpose detector, targets wide range of physics
  - Comprises different subsystems focussing on different physics objects
  - Complex calorimetry for energy measurements
- Operating at energy and intensity frontier
- For predictions from simulation need accurate modelling of detector
  - With increase in data collection will need to simulate more events!

ATLAS MC Production Chain

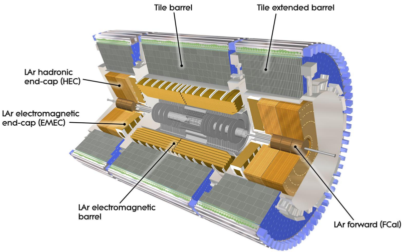
Event Gen → Detector Sim → Digitisation → Reconstruct → Analysis



2016 values

- Simulation is largest use of the grid
- Of which calorimeter sim is 85%!
- Geant4 used for full sim of detector
  - Current default
  - Highly accurate
  - CPU intensive
- Cannot rely on full sim for all samples in Run 3
  - Need to use fast simulation
  - Trade off in accuracy
  - Target improvements here

ATLAS Calorimeter

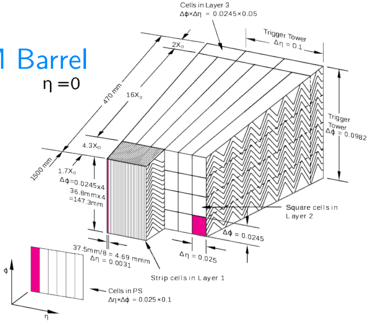


- Sampling calorimeter covering $|\eta| < 4.9$
- Two technologies employed:
  - LAr + Pb/Cu/W
  - Scintillating tiles + steel
- $\sim$190k readout channels

- Crucial for reconstruction of jets, γ, $e^{\pm}$ and E_{T}^{miss}
- None-trivial geometry
- Accordion shape, varying cell sizes
- Takes a long time to simulate full interactions in Geant4 $\mathcal{O}(\text{mins})$

EM Barrel

η =0



Fast Simulation

Due to the intensive CPU requirement of G4 calorimeter simulation

focus effort on faster methods of simulating calorimeter

Single particle G4 events → Extract Representation → Fast Simulation

Classical Approach

Parametrise calo response in $E - \eta$

Split in long/lat shower development

Machine Learning

Learn functional approximation of G4

Use Neural Networks

- GAN
- VAE

Fast Simulation
Classical Approach

Overview

- Use single particle G4 events for γ, $e^{\pm}$ and $\pi^{\pm}$ starting at calo surface
- $\sim$5k parametrisations in total
  - 17 energy bins (log spacing) from 60 MeV ($\pi^{\pm}$ 256 MeV) – 4.2 TeV
  - 100 bins in $|\eta|$ from 0 – 5.0
- Split into energy (longitudinal) and shower shape (lateral) parametrisations
- Use simplified geometry to assign hits to cells

ATL-SOFT-PUB-2018-002

Fast Simulation
Classical Approach

Longitudinal Parametrisation

N Inputs: E_{layer}/E_{tot}, E_{tot}

↓ Transform into decorrelated inputs

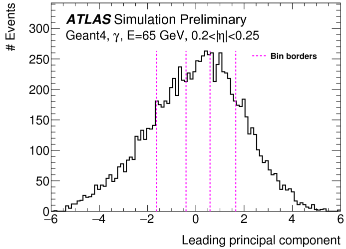
Principal Component Analysis

↓ Divide leading comp. into 5 equal bins

2nd PCA in each bin

↓ Transform into decorrelated inputs

5 x N linearly uncorrelated $\mathcal{G}$



Store histograms of PCA2 cumulative distributions.

Additionally store: PCA matrices, Means+RMS of all Gaussians

Interpolate between energies with a spline function

Fast Simulation
Classical Approach

Lateral Parametrisation

Energy/layer in PCA bins

↓ Convert 1mmx1mm G4 hit grid to polar coord.

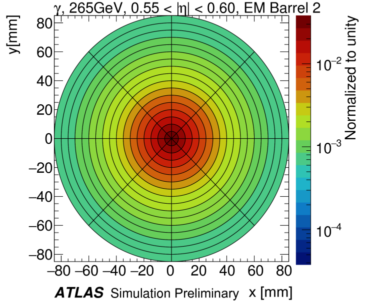
Parametrise energy in r-α

↓ Store parametrisation, use as PDF

Sample N "hits" from PDF

↓ $E_{hit} = E_{layer}/N$

Match hits to cells



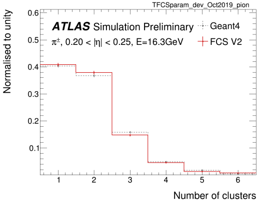
Simplified geometry used for cell lookup

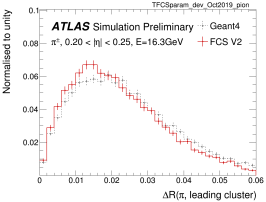
Number of hits calculated from expected energy resolution

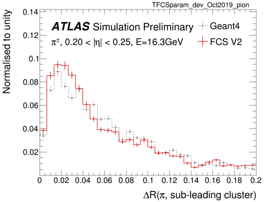
Reduce N_{hits} for increased fluctuations in hadronic showers

Fast Simulation
Classical Approach - Updated Performance

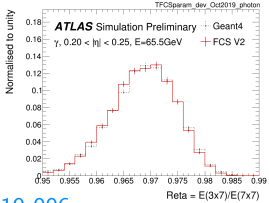
New fluctuations derived for pions, improves modelling of clusters
(Geant4 samples used for evaluating stochastic and constant fluctation term)

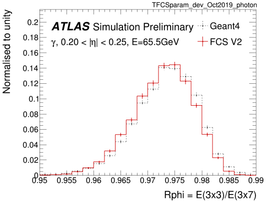






Improved treatment of cross-talk between cells $\rightarrow$ better fractional energy deposits in η, ϕ for photons





ATL-PLOT-SIM-2019-006

Fast Simulation
Generative Models

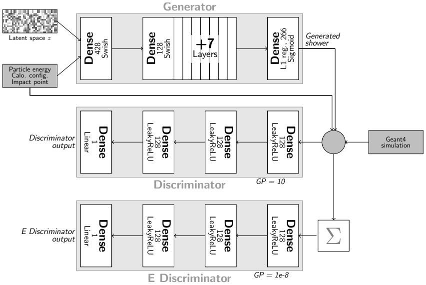
Overview

- Can we train a network to approximate showering in Geant4?
- At present, focus only photons in a single $|\eta|$ slice, $0.2 < |\eta| \leq 0.25$
- Train the network conditional on photon energy to reproduce Geant4 shower
- Two different approaches:
  - Generative Adversarial Networks (GANs)
  - Variational Auto-encoders (VAEs)
- Train on energy deposits in cells from G4 events, generate new showers
  (In comparison to detailed hits used by FCS)

ATL-SOFT-PUB-2018-001

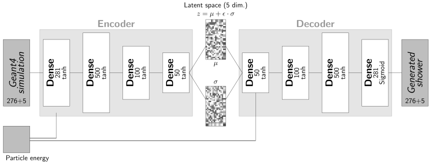
Fast Simulation
Generative Models

GAN



- Train a generator for new showers
- Play off against critic discriminating generated from Geant4

VAE

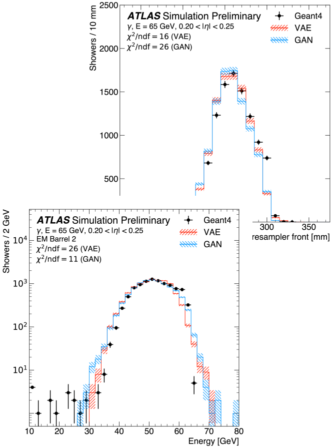


- Encode representation of Geant4 showers into latent space
- Use decoder to generate new showers

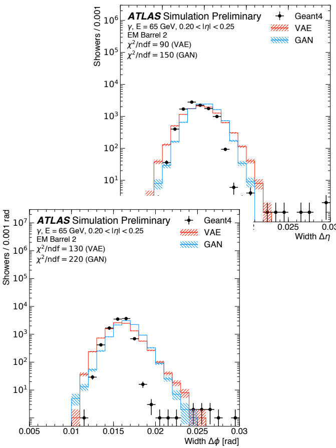
Showing new improved architectures

Generative Models
Previous Performance

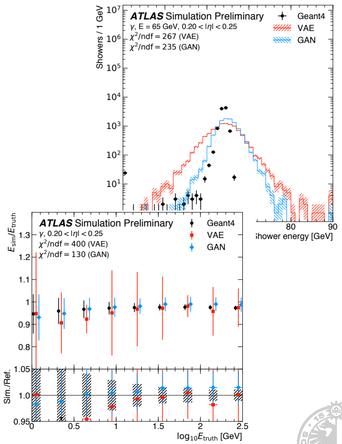
Longitudinal



Lateral



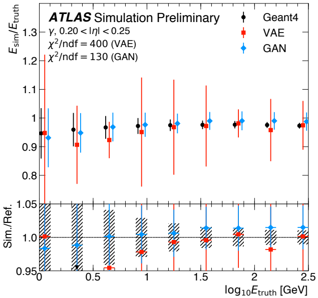
Total Energy



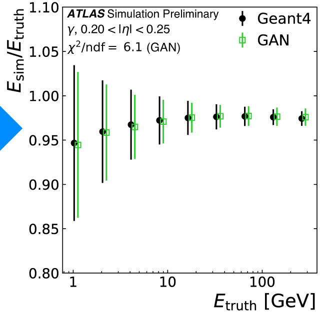
Previous performance of GAN and VAE for photons compared to Geant4

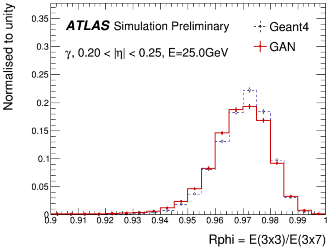
Problematic

Generative Models
Updated Performance



GAN



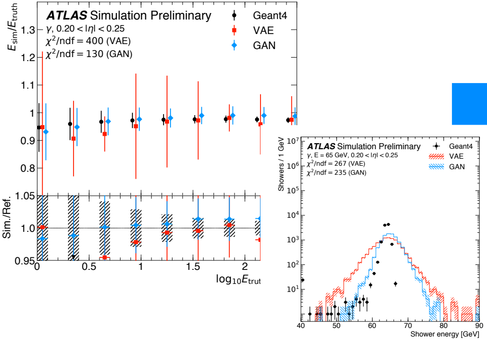


- Addition of second critic for total energy
- Condition on particle position in target cell
- Optimisation of generator architecture

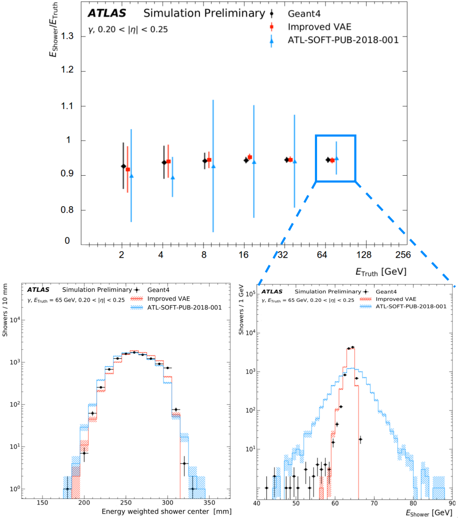
- Improved mean and width of total energy
- Good performance for interpolated energies

ATL-PLOT-SIM-2019-004

Generative Models
Update Performance



VAE



- Train on relative cell energies $^{E_{cell}}/_{E_{layer}}$, $^{E_{layer}}/_{E_{tot}}$ and $^{E_{tot}}/_{E_{truth}}$
- Optimisation of loss function and architecture
- Large improvement in total energy

ATL-PLOT-SIM-2019-007

Conclusion

- As more data is recorded, need more simulated events to compare against
- Will only increase need for faster simulation of the detector
- Current fast simulation doesn't adequately describe data for precision measurements
- Improvements in the classical approach will play a key role in future
  - FastCaloSim v2 is very mature and will soon enter production
- Generative models with machine learning show great promise for use
  - Studies ongoing in larger η range and with pions

Advert

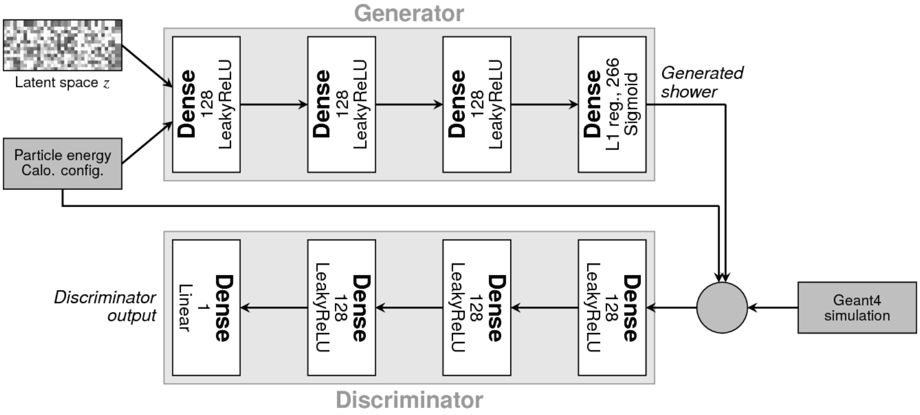
For more information and discussion come and see our two posters!

Backup

Links

Deep Generative models for fast shower simulation in ATLAS	ATL-SOFT-PUB-2018-001
The new Fast Calorimeter Simulation in ATLAS	ATL-SOFT-PUB-2018-002
Energy resolution with a GAN for Fast Shower Simulation	ATL-PLOT-SIM-2019-004
Updated plots for Fast Calorimeter Simulation	ATL-PLOT-SIM-2019-006
Updated Variational Autoencoder for Fast Shower Simulation	ATL-PLOT-SIM-2019-007

Previous GAN Architecture



Previous VAE Architecture