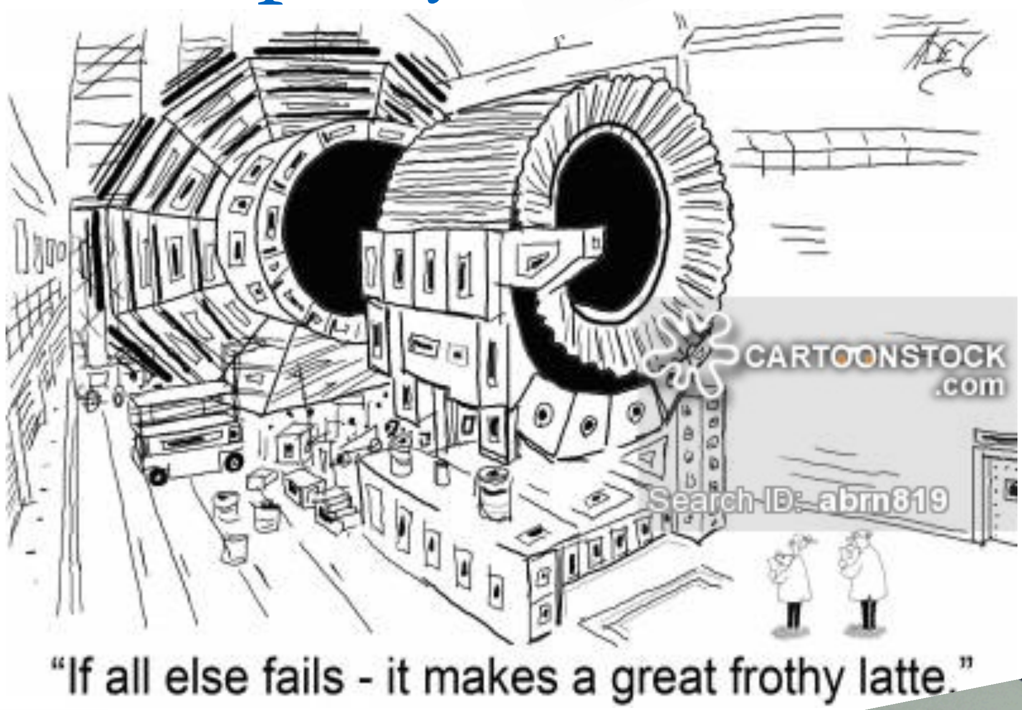
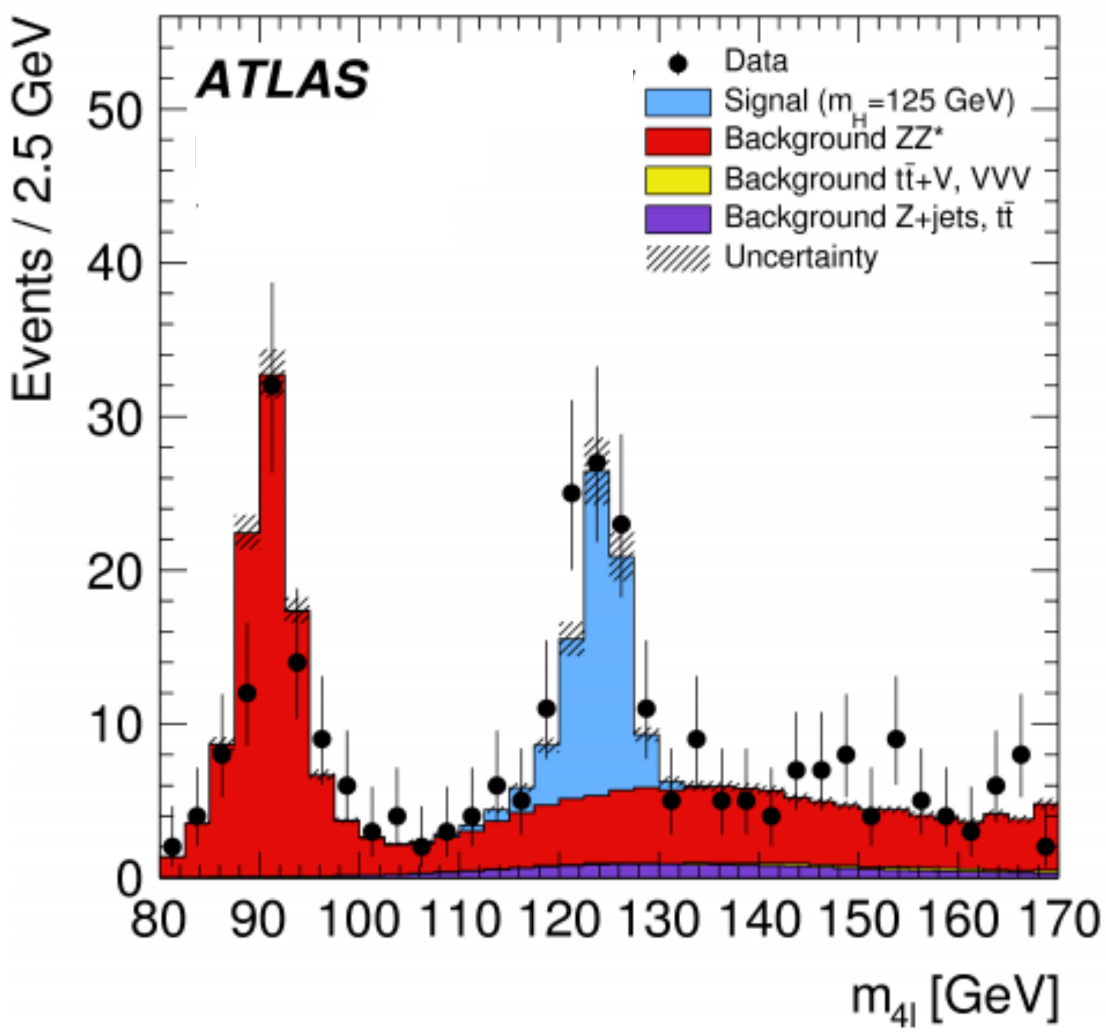
# The new ATLAS Fast Calorimeter Simulation

Hasib Ahmed

# We like to party!

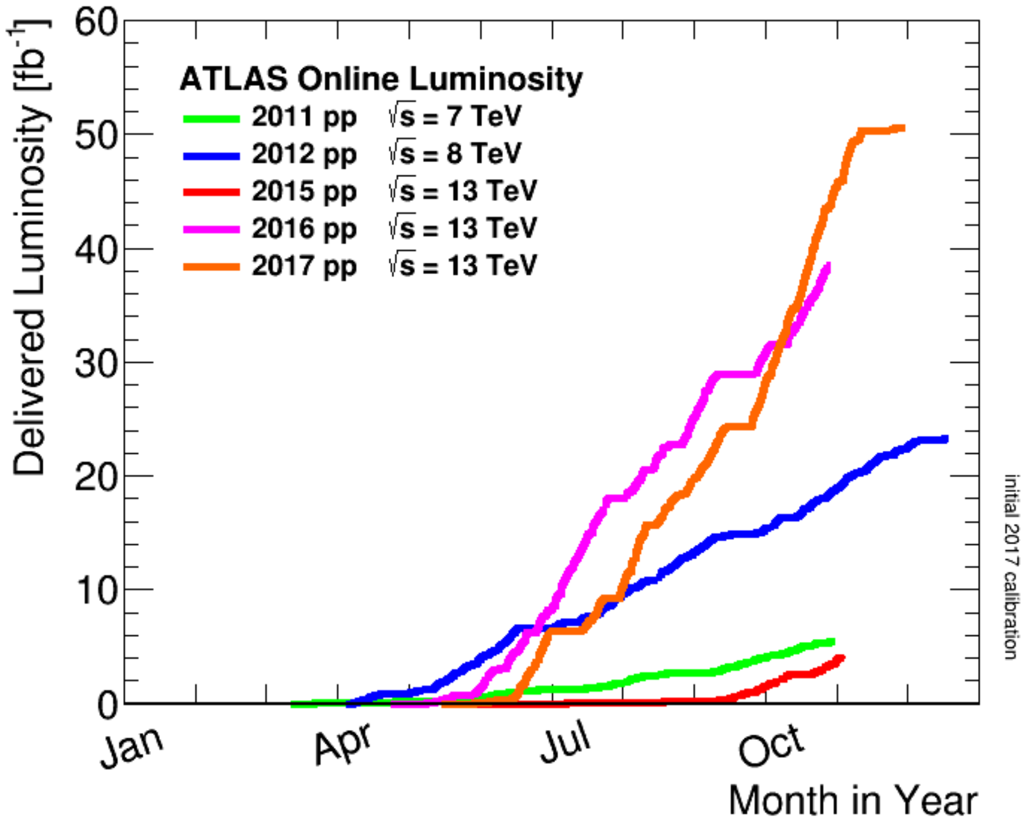



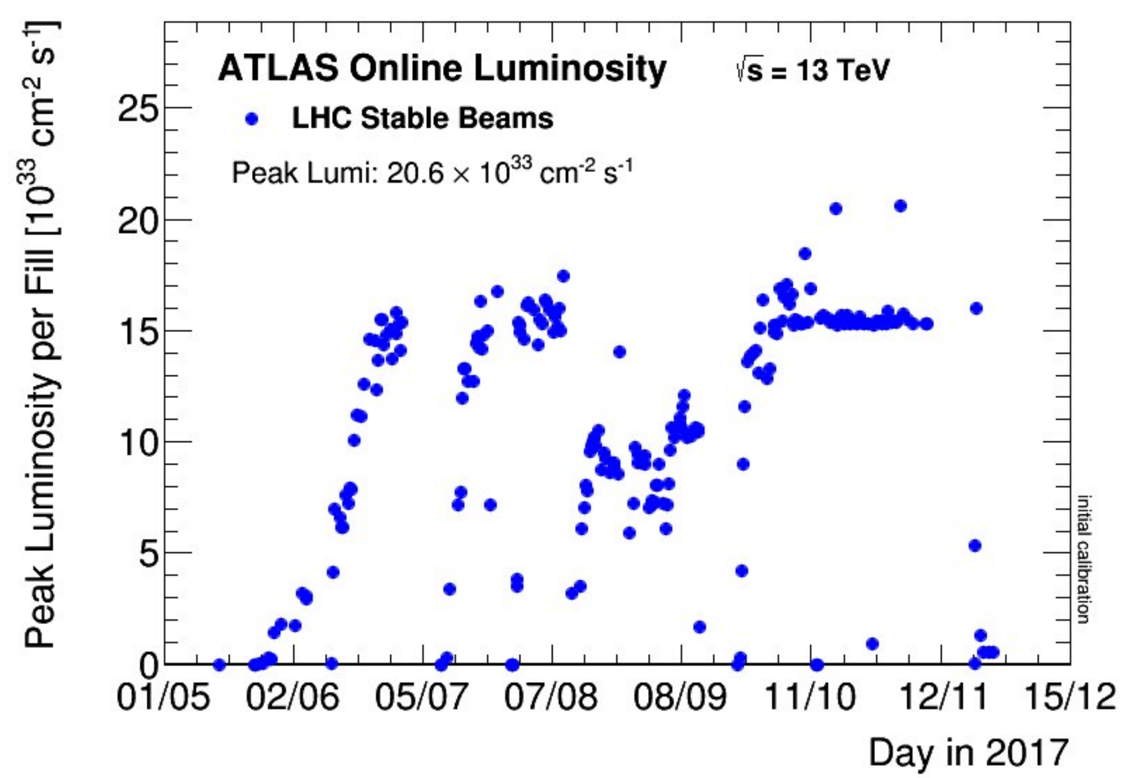

# Energy frontier: The Large Hadron Collider

Tremendous performance by the LHC: exploring the high energy frontier at 13TeV

Delivered ~93 $fb^{-1}$ of integrated luminosity at 13 TeV since 2015

Large number of interactions in each collision (pile-up) $<\mu>$ ~78 at peak luminosity

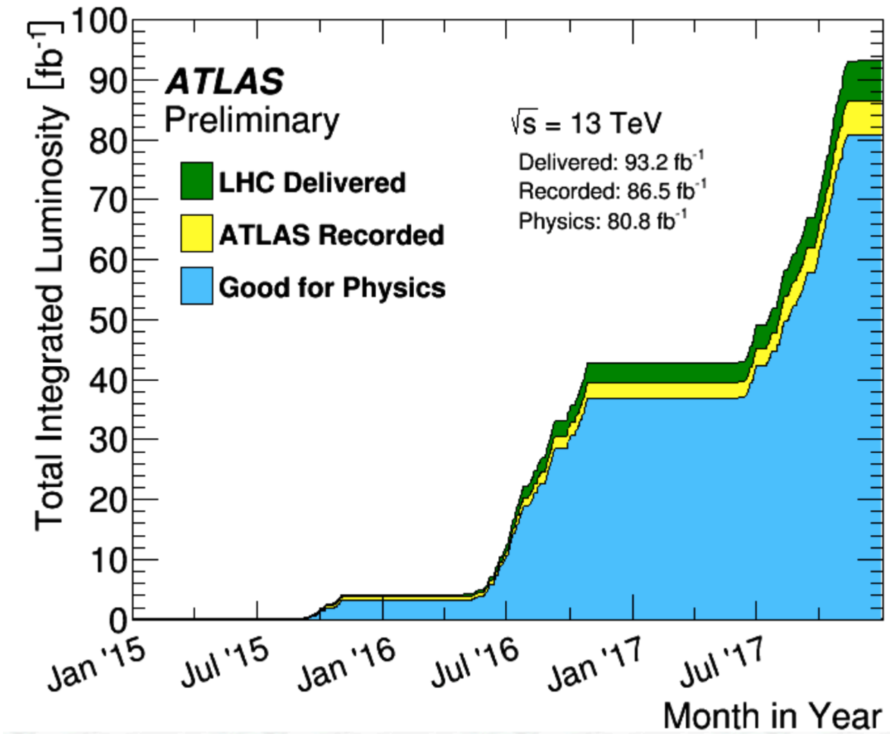



***At the designed energy LHC will deliver even higher integrated luminosity and larger pile-up interactions***

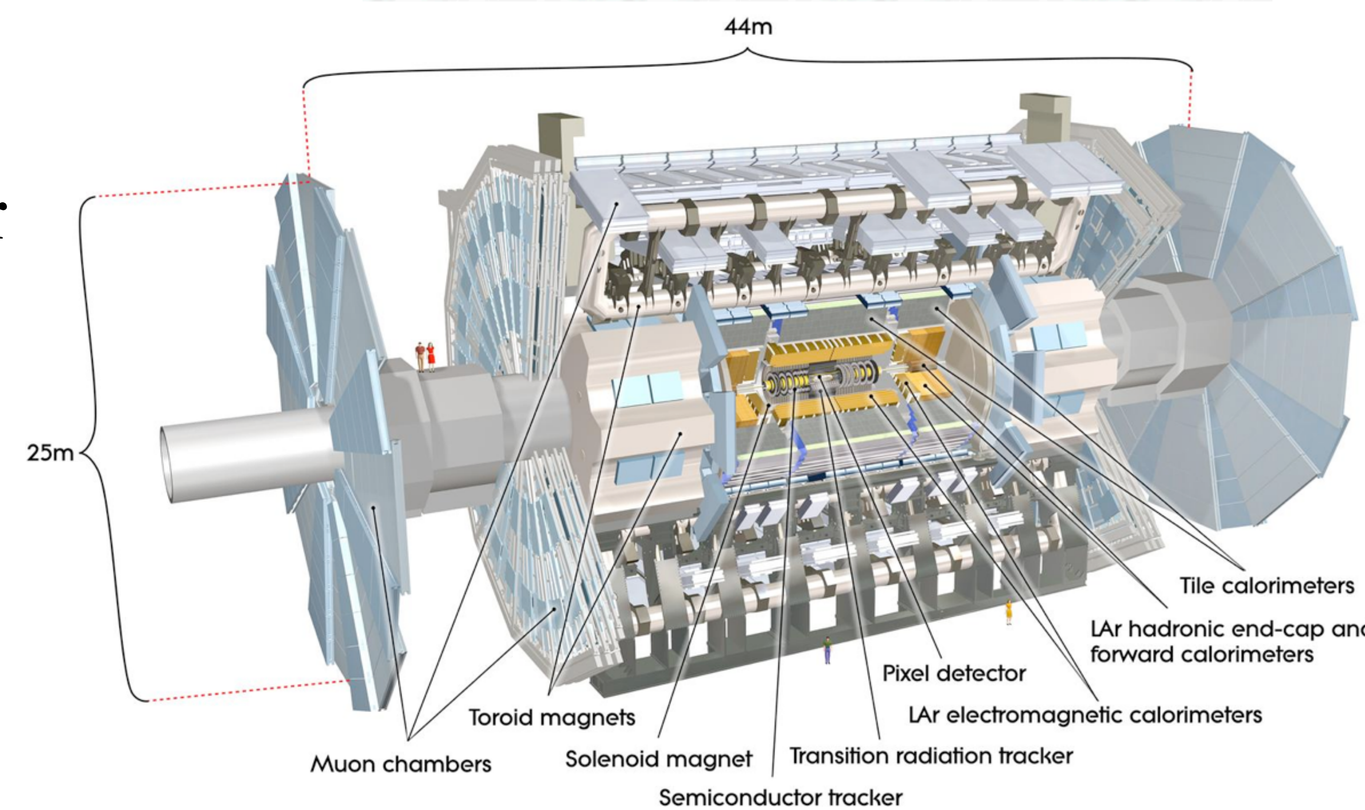
# Discovery machine: The ATLAS detector

Outstanding performance in recording delivered luminosity by the LHC




General purpose detector designed for discovery of physics beyond the Standard Model

Silicon and transition radiation for charge particle tracking, Calorimeter for energy measurement

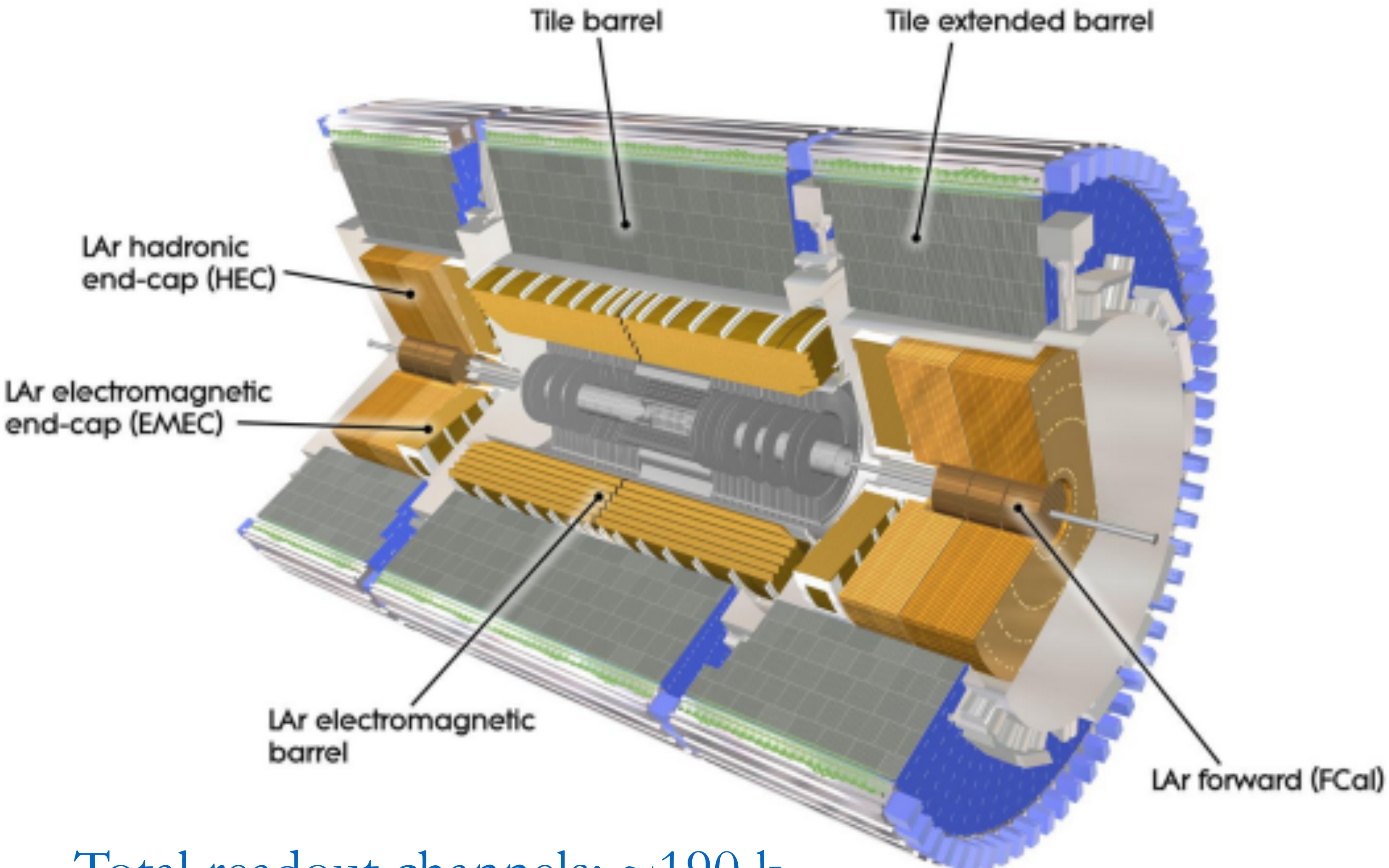


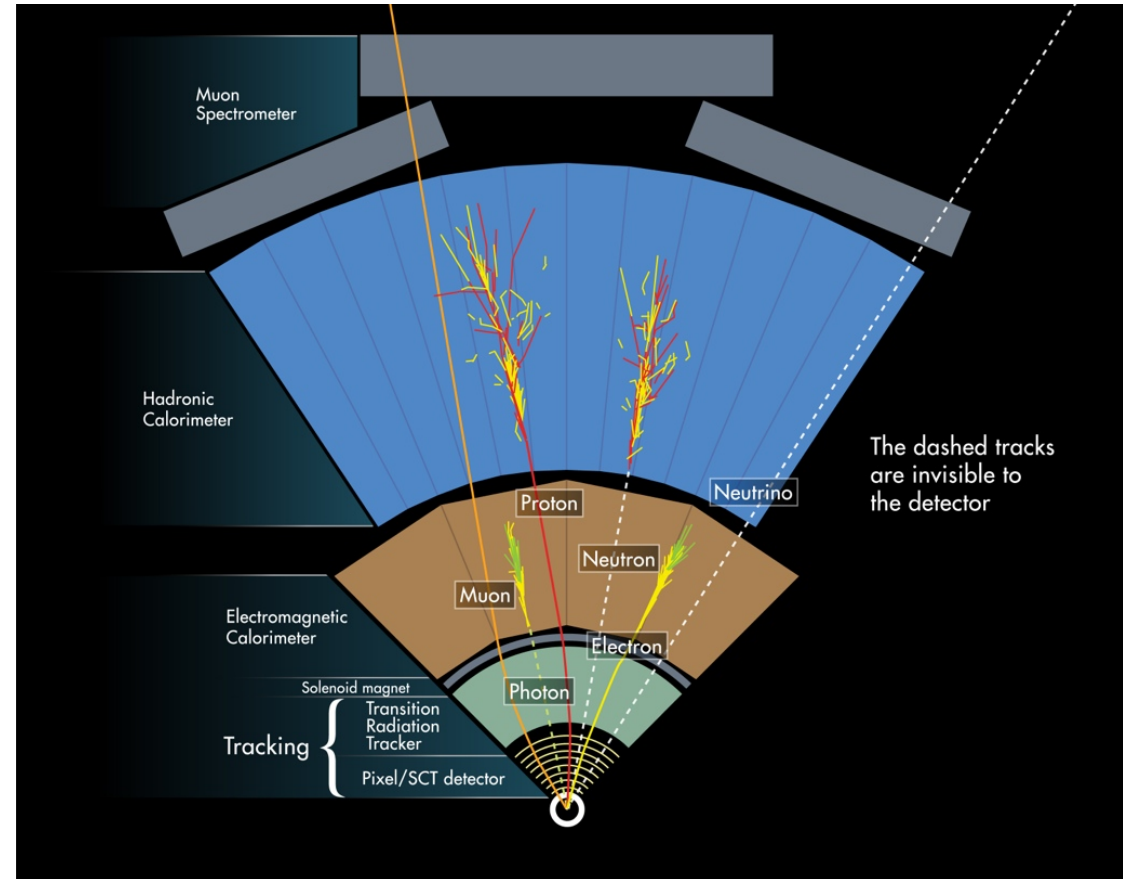

Successful physics program depends on large number of Monte Carlo simulated events

***Large collision data events require even larger number of simulated events for physics analysis***

# The ATLAS Calorimeter

Sampling calorimeter covering $|\eta| < 4.9$




Total readout channels: ~190 k

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

Electromagnetic (EM) Cal:

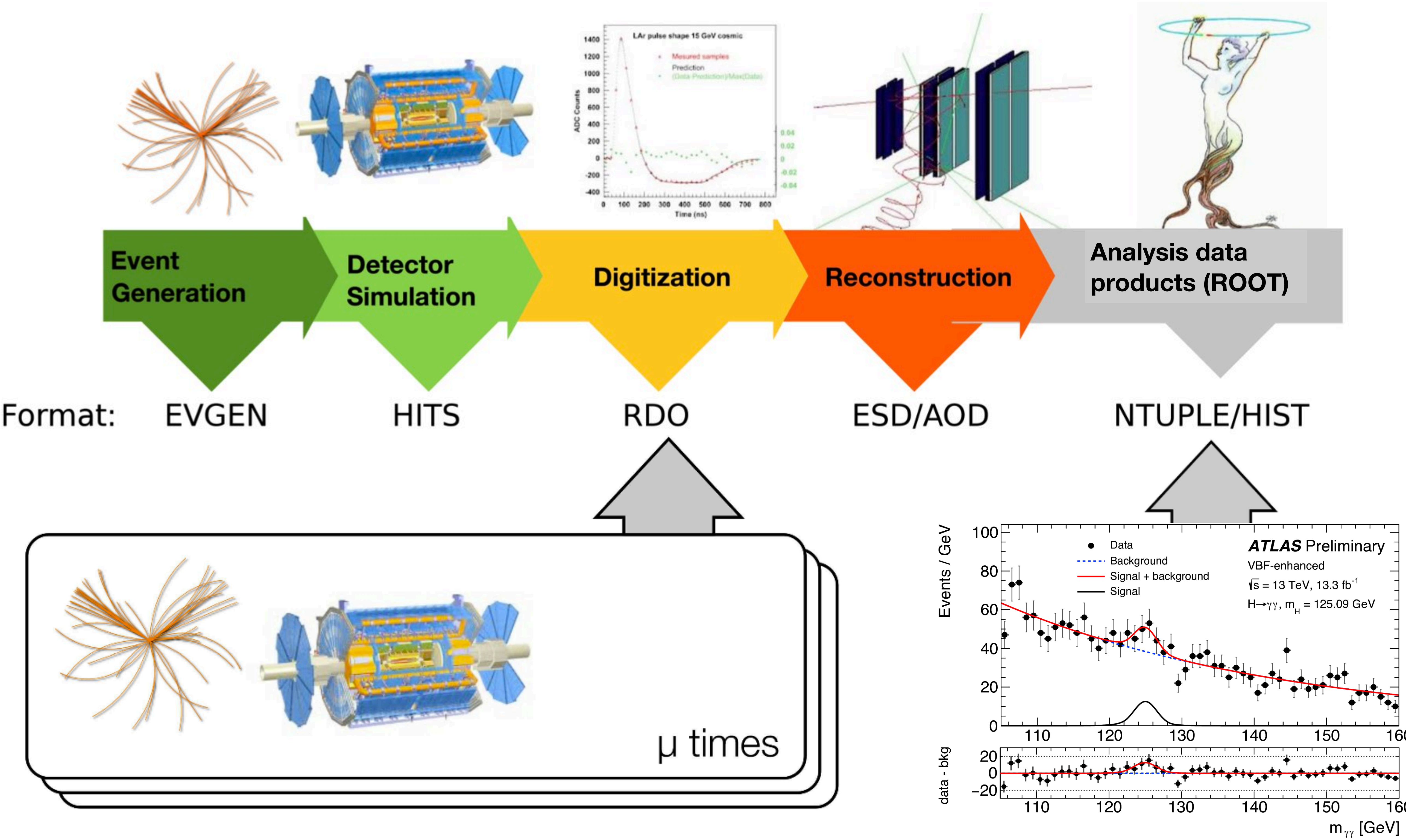
- Liquid Argon (active)
- Pb/Cu/Tungsten (absorber)

Hadronic/Tile Cal:

- Scintillating tiles (active)
- Steel (absorber)

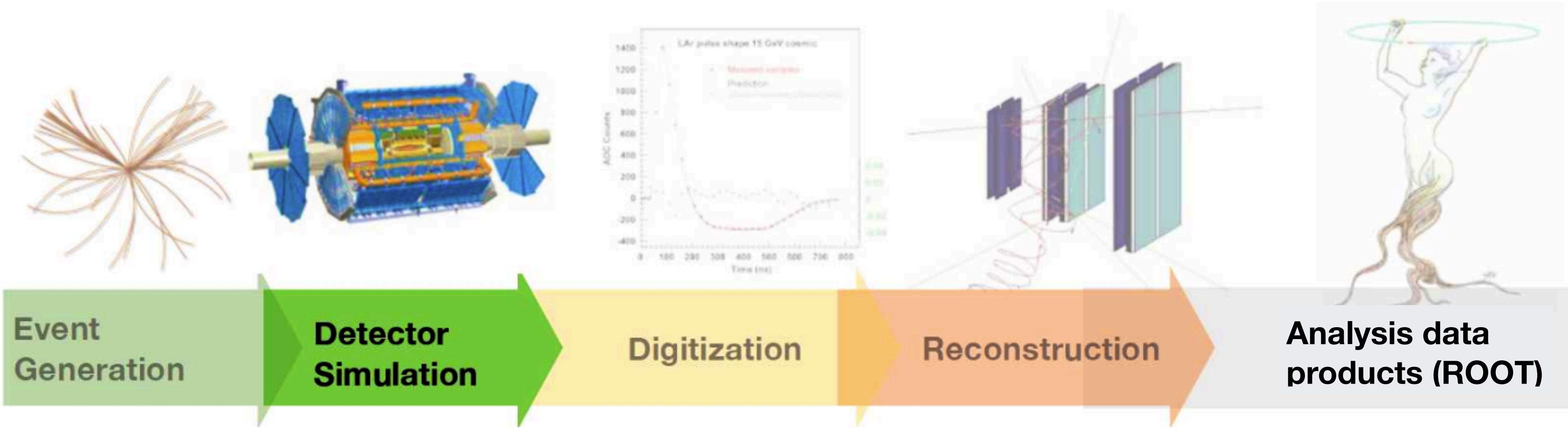
***Crucial for electrons, photons, jets and missing energy reconstruction***

# ATLAS Monte Carlo Production Chain



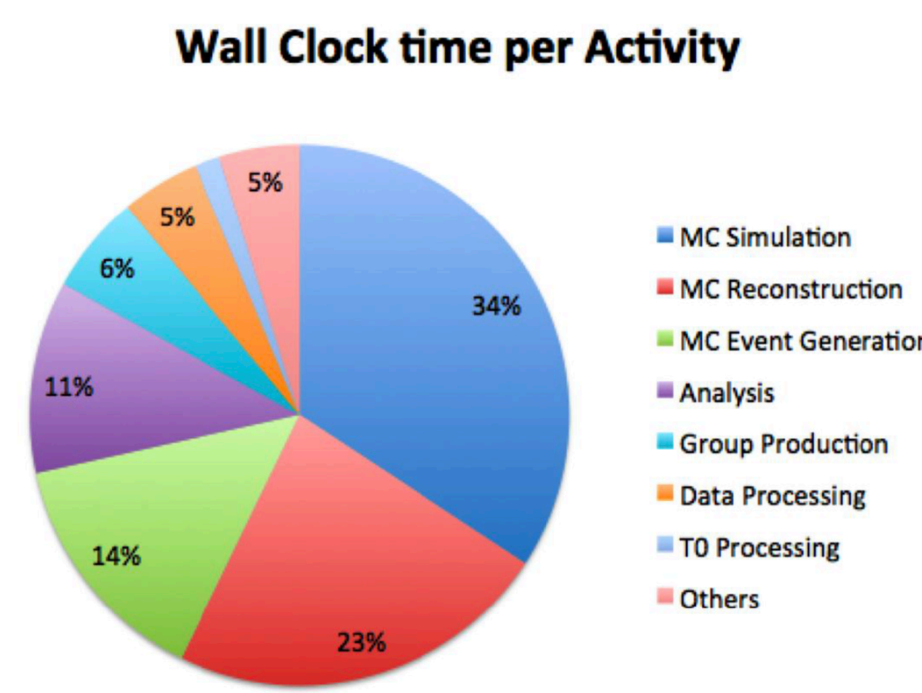

***Simulated events undergo the same conditions as reconstructed collision events***

# ATLAS Standard Simulation: Geant4



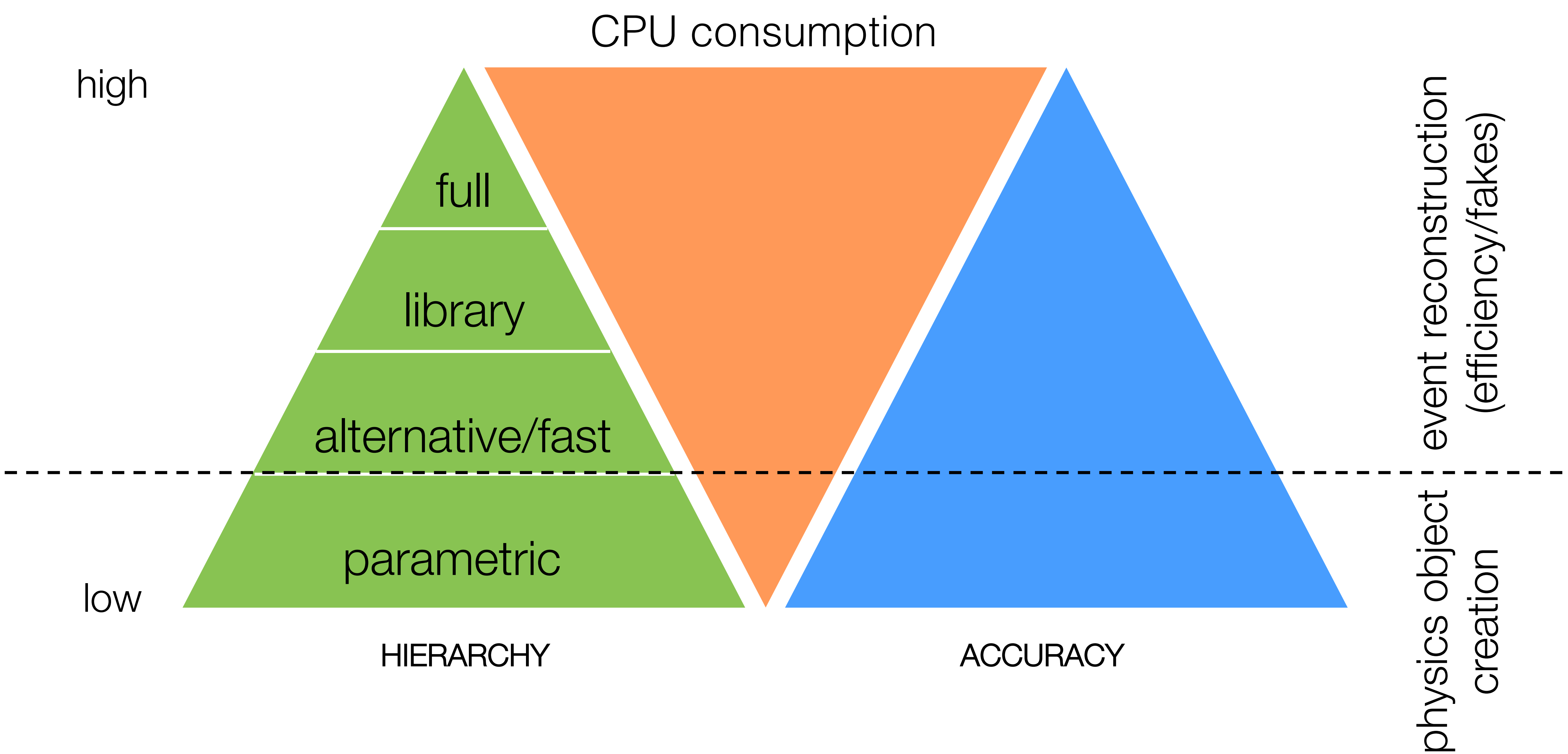

- Geant4 is the standard ATLAS simulator
- Full description of the detector and most precise
- Large CPU and disk space requirement
- Calorimeter simulation accounts for ~90% of total time
- Higher luminosity/pile-up requires larger MC production

Grid usage 2016:




***Fast simulation is essential with increasing luminosity!***

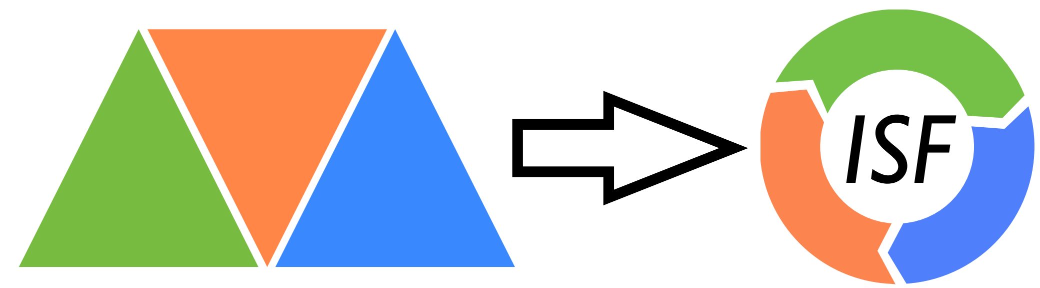
# The simulation hierarchy

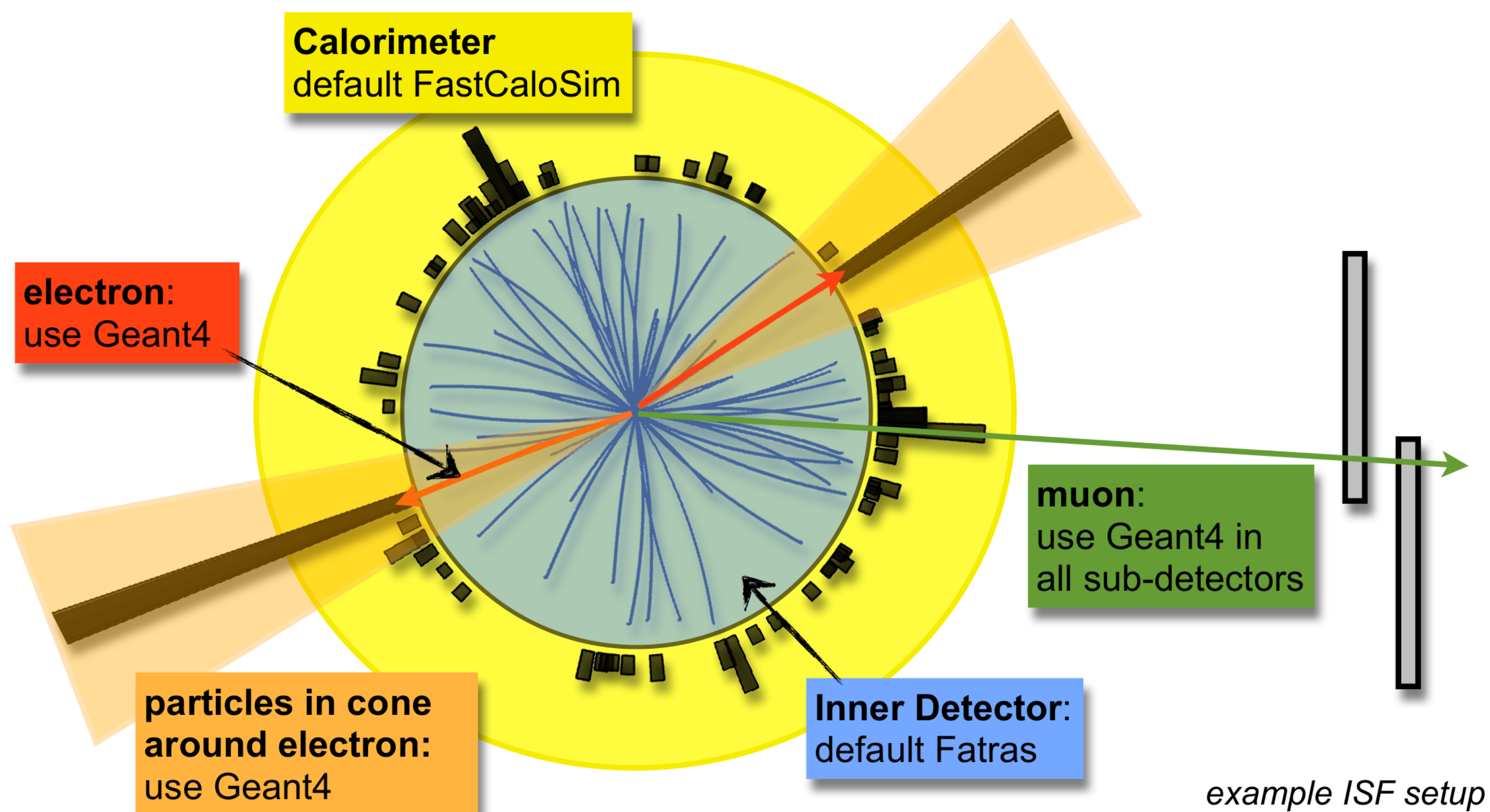



***A trade-off between accuracy and speed***

# ISF - Integrated Simulation Framework

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



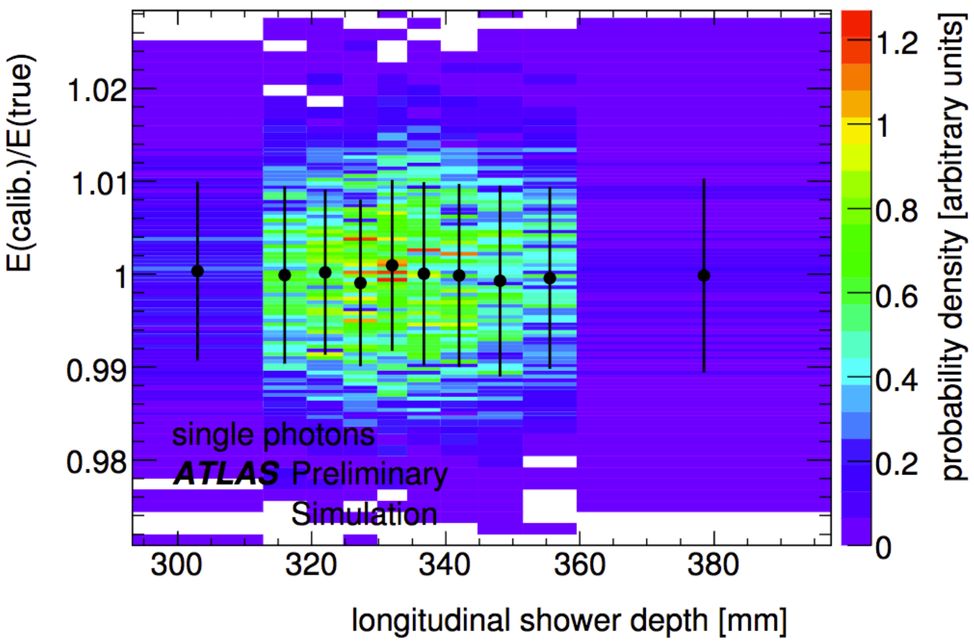




*example ISF setup*

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

# Current ATLAS Fast Calorimeter Simulation (FastCaloSim)

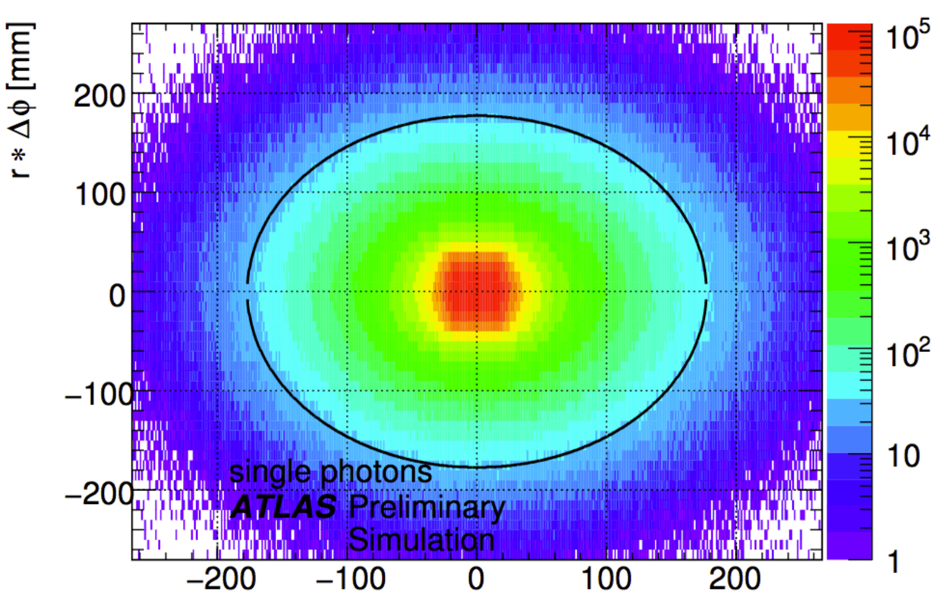
# Current FastCaloSim: Parametrization




Parametrized Calorimeter response in E-η grid

Geant4 simulated single particles:
e, γ (EM interaction ) and $\pi^{\pm}$(hadronic interaction)

Parametrization split into longitudinal and lateral shower development




Lateral energy distribution

Detailed parametrization of the energy as a function of longitudinal shower depth.

Average lateral shower parametrization obtained from a fit to the Geant4 lateral shower shape

Uses simplified geometry for hit to cell assignment

***Calorimeter response modeled with detailed longitudinal and average lateral shower parametrization***

# Current FastCaloSim: Performance

**ATLFASTII** = current FastCaloSim for calorimeters + Geant4 for inner detector and muon systems

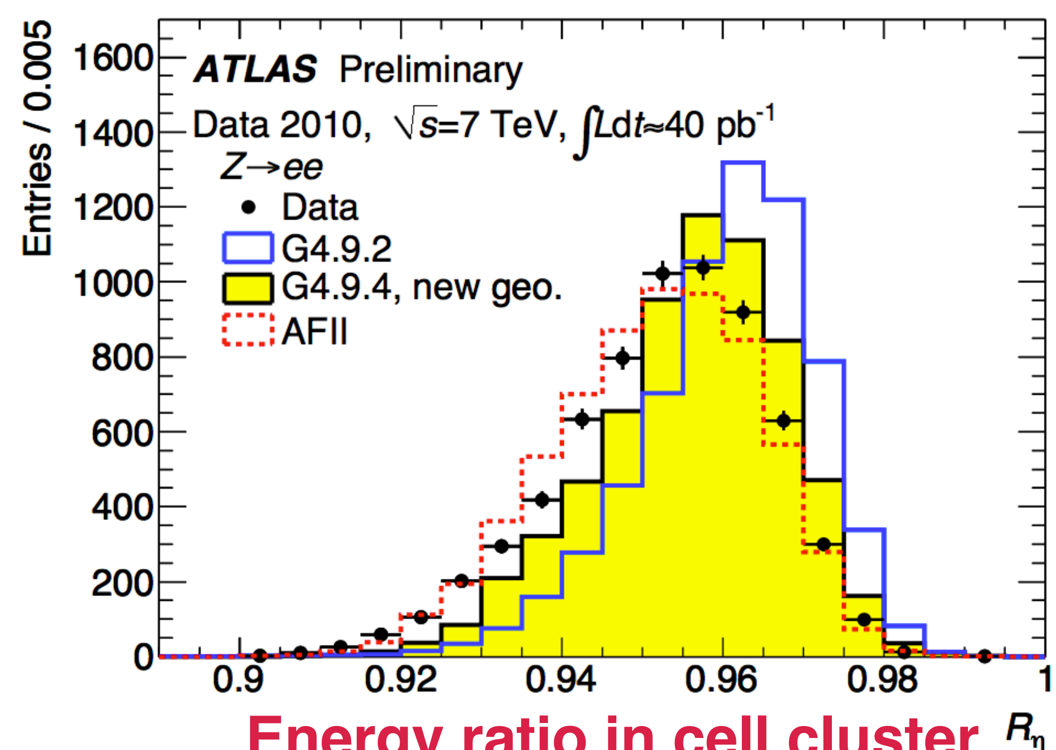
Widely used in ATLAS: e.g. SUSY signal samples for searches, ttbar samples systematic uncertainty studies

ATLFASTII ~10x faster than Geant4

Simulation time per event in seconds

| Sample | Full G4 Sim | Atlfast-II |
|---|---|---|
| Minimum Bias | 551. | 31.2 |
| $t\bar{t}$ | 1990 | 101. |
| Jets | 2640 | 93.6 |
| Photon and jets | 2850 | 71.4 |
| $W^{\pm} \rightarrow e^{\pm}\nu_e$ | 1150 | 57.0 |
| $W^{\pm} \rightarrow \mu^{\pm}\nu_{\mu}$ | 1030 | 55.1 |

* based on studies performed in 2010, current Geant4 has better performance




Energy ratio in cell cluster

***Good average of shower descriptions, however poor modeling of substructure inside the shower***

# New ATLAS Fast Calorimeter Simulation

# New FastCaloSim: FastCaloSimV2

## Goals:

Describe the physics better than ATLFASTII, esp. substructure of showers

Decrease the time required to simulate each event

Optimize I/O and memory consumption

## Developments:

Single particle ($\gamma$, e, $\pi^{\pm}$) samples on a fine E-$\eta$ grid produced with current ATLAS geometry in Geant4

New energy (longitudinal) and shower shape (lateral) parametrization

Reduce the amount of information to a compact form

Add lateral shower fluctuations

Assignment of hits to cells overcoming simplified geometry drawback

Use exact Forward Calorimeter geometry

***FastCaloSimV2 must improve over* ATLFASTII *in describing physics processes***

# FastCaloSimV2 Workflow

Geant4 simulated single particles

↓

Longitudinal (energy), lateral (shape) parametrization

↓

Simulation

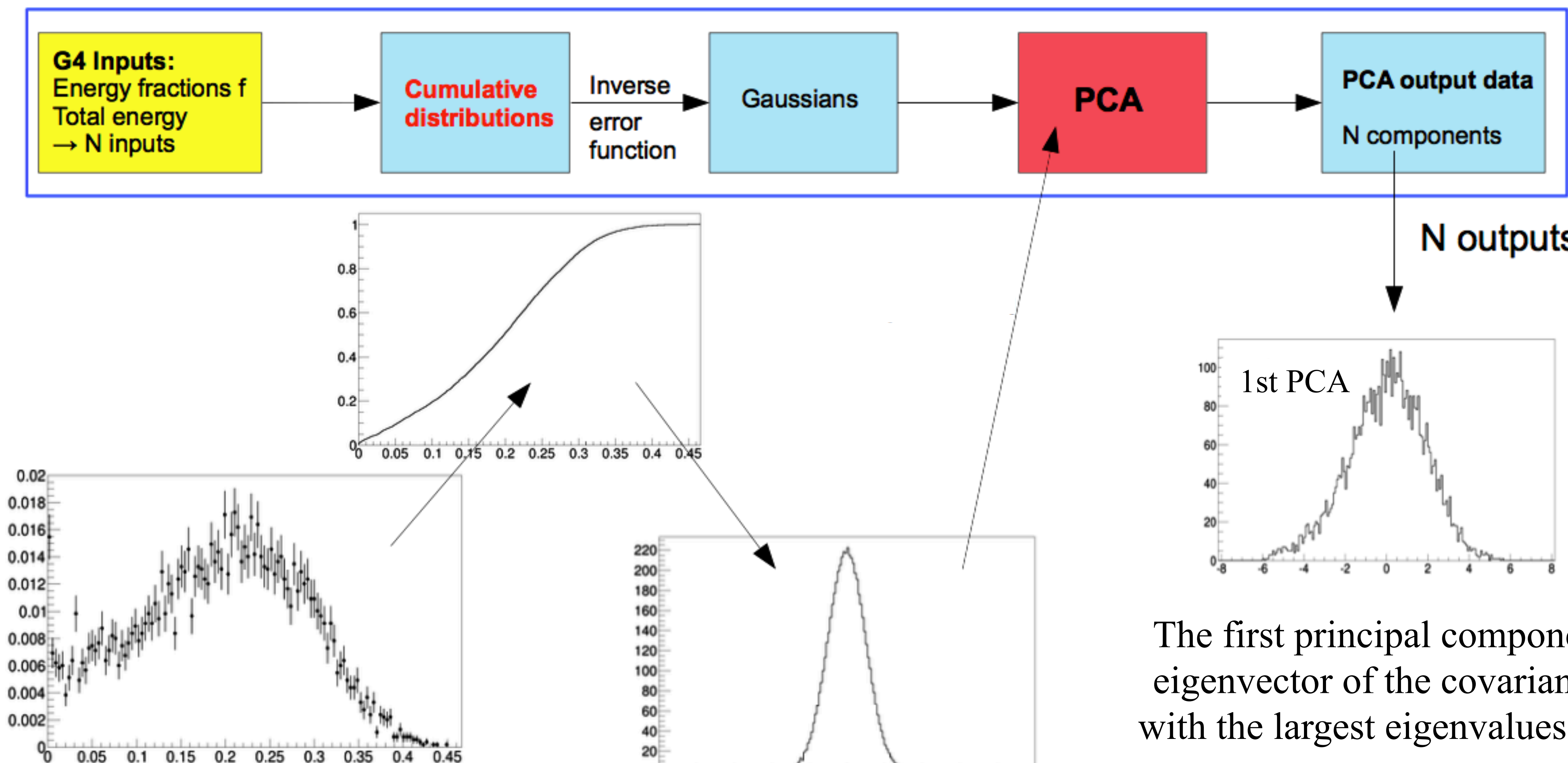
↓

Validation/Performance measurement

# Longitudinal Shower Parametrization

Goal: Model the total energy and energy fraction in each layer
Difficulty: Energy deposits each layer are correlated
Solution: Transform the correlated energy deposits into linearly uncorrelated ones through Principal Component Analysis (PCA)

**1st PCA chain:**



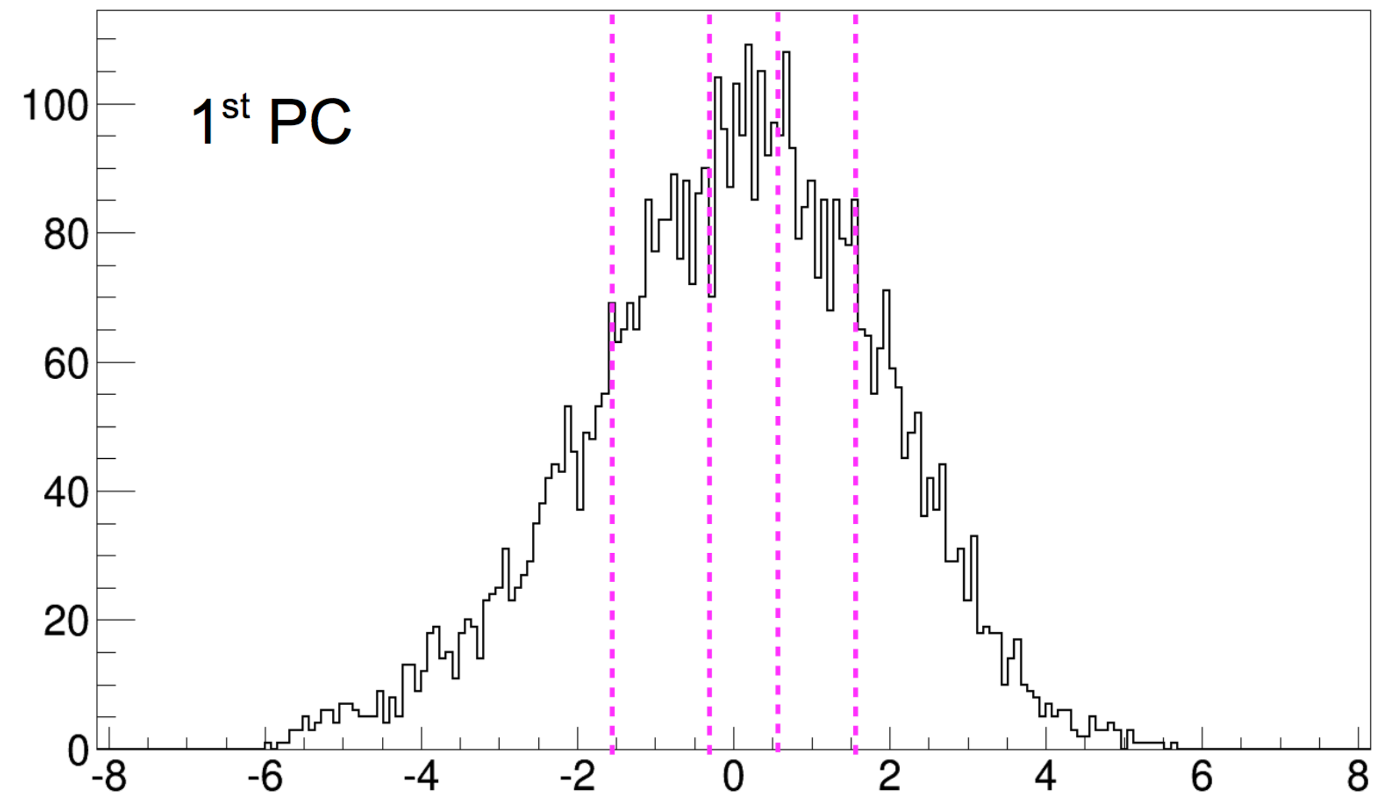

Energy fraction EM Barrel 1

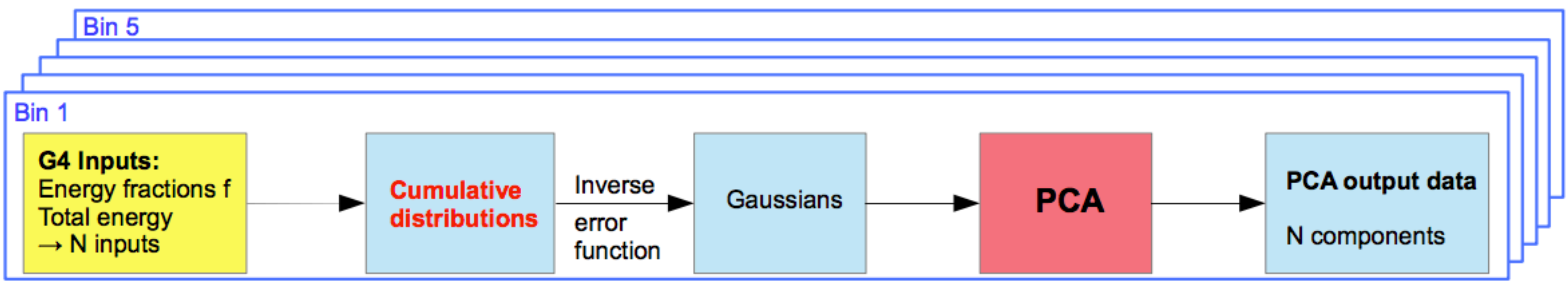
The first principal component is that eigenvector of the covariance matrix with the largest eigenvalues (variance)

# Longitudinal Shower Parametrization

Use the 1st PCA to divide the input data into quantiles of same size

Use a 2nd PCA to further decouple the correlation for each type of shower (i.e. in each PCA bin)



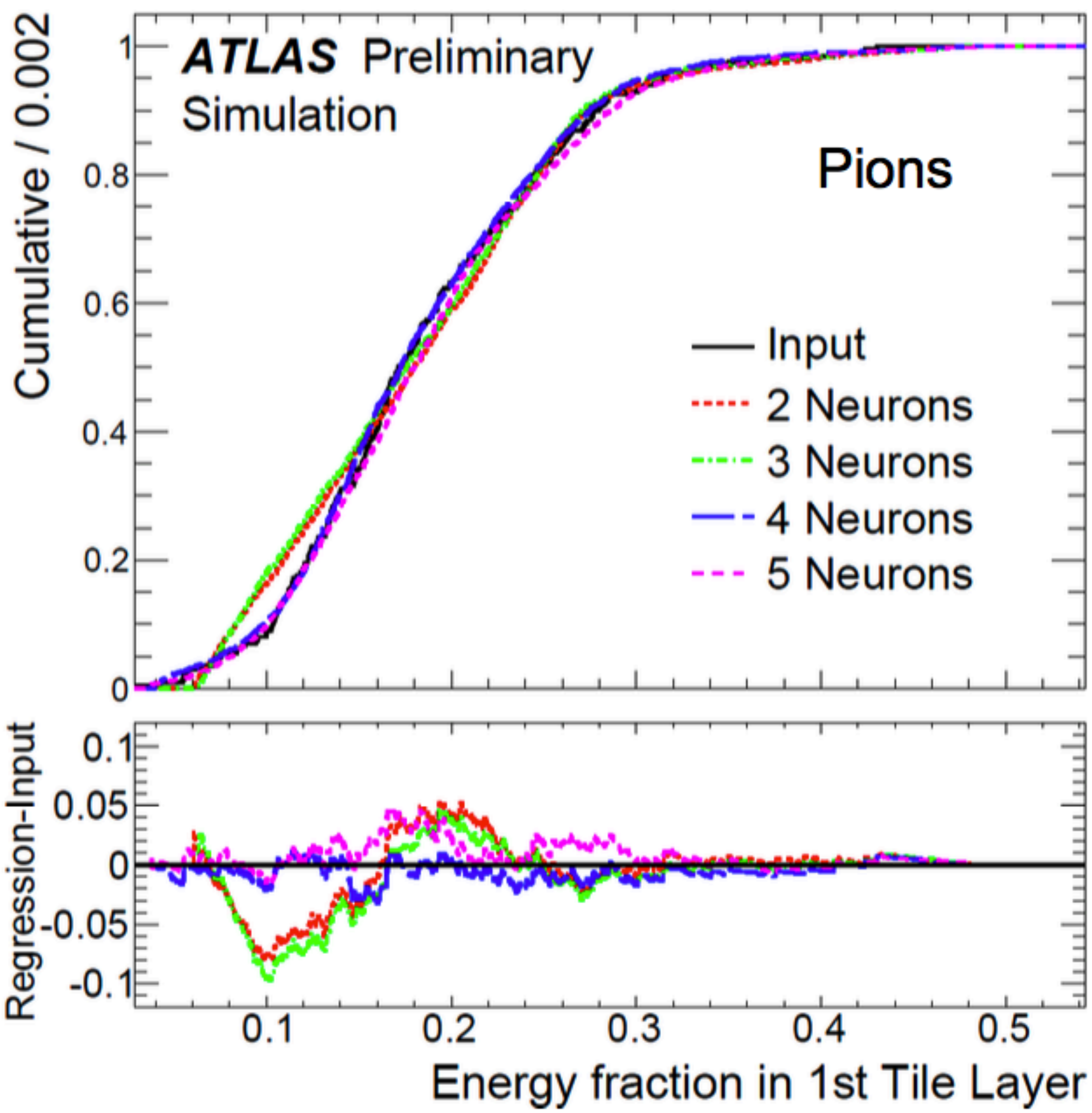




N linearly uncorrelated gaussian distributions

***Cumulative distributions, PCA matrices, mean and RMS of the Gaussians are stored for parametrization***

# Longitudinal Shower Parametrization

Goal: Memory optimization
Difficulty: Large memory consumption by the cumulative distributions
Solution: Multivariate regression to approximate the functional form




Multi Layer Perceptron (MLP) used for regression

Number of weights needs to be saved scales with the number of neurons

No of weights = $1 + n + (2n)$
$n$ = no. of neurons

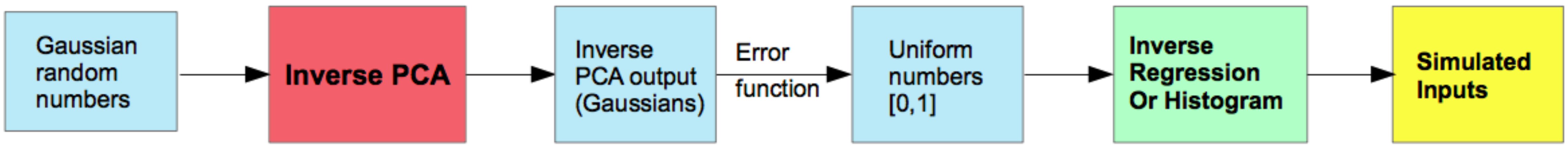
Iterative procedure to achieve good agreement and optimized number of neurons

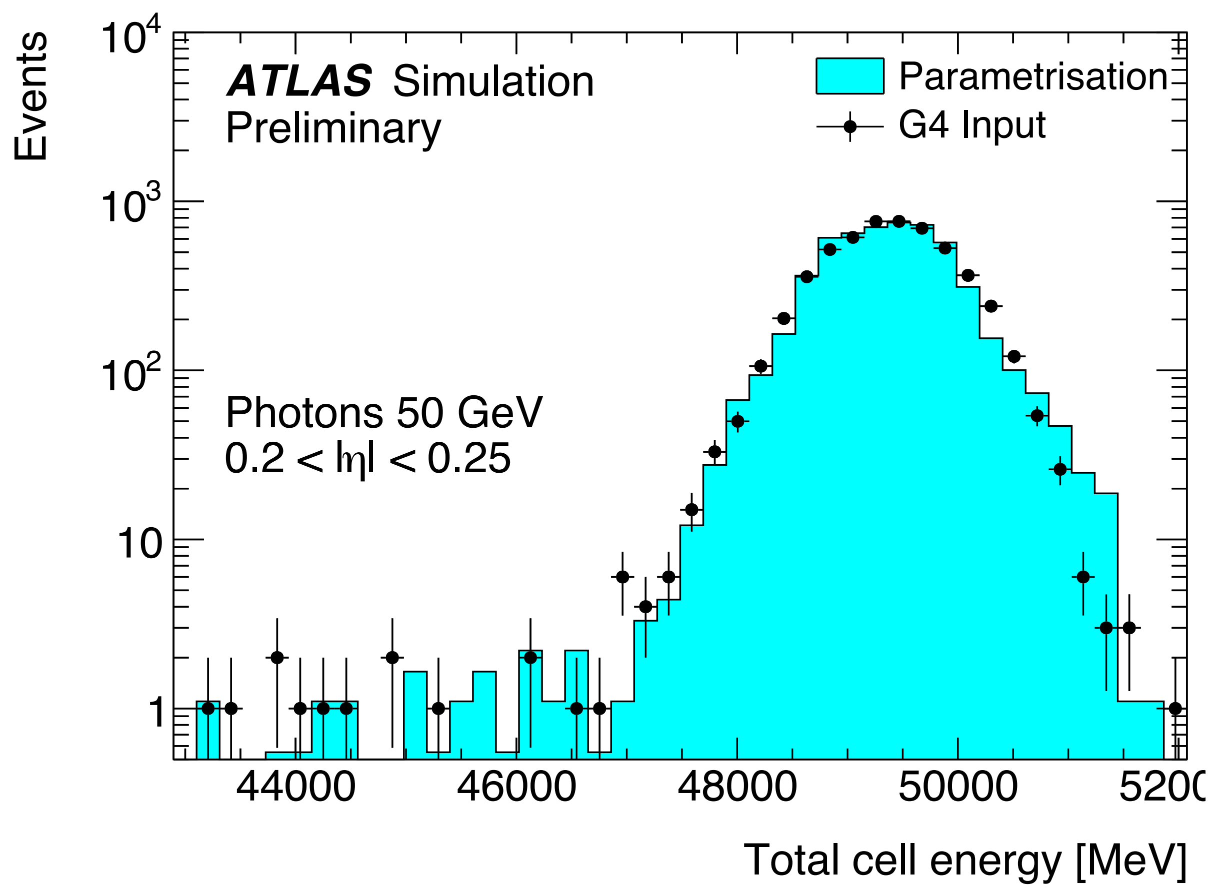
***Only weights from multilayer perceptron are stored!***

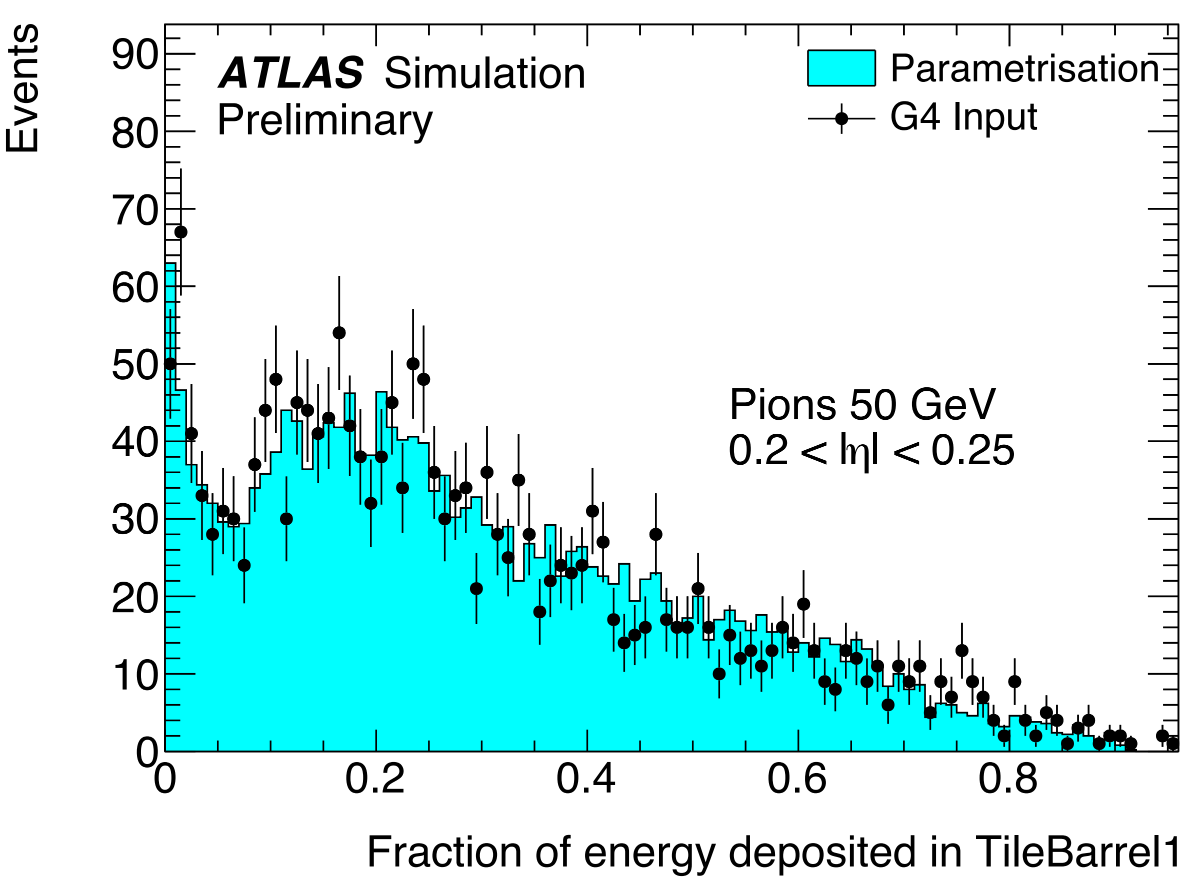
# Longitudinal Shower Simulation

Randomly determine the PCA bin: each bin has the same probability by construction

Perform inverse PCA analysis in each PCA bin



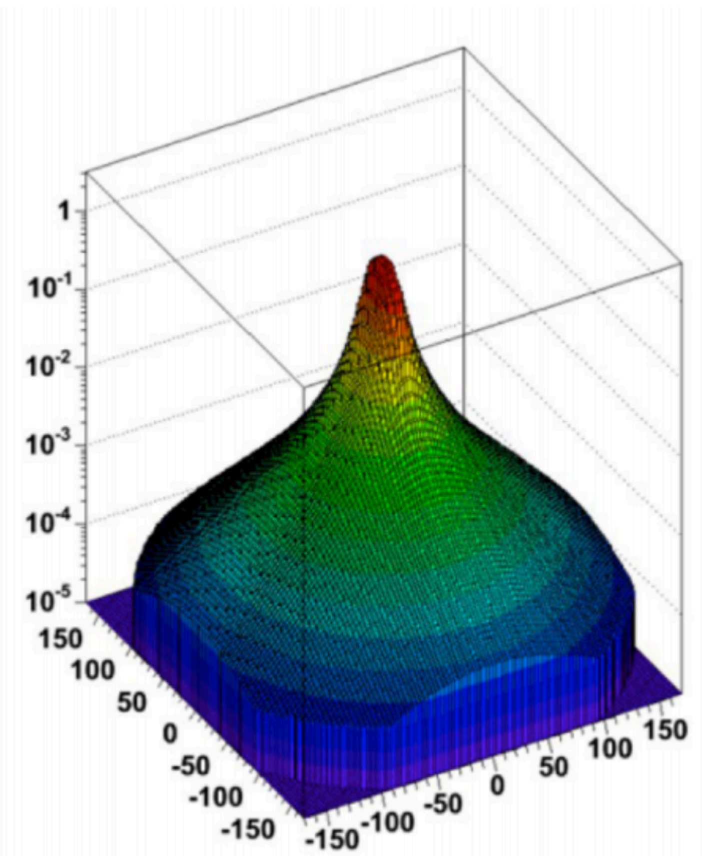

***Good agreement between Geant4 and energy parametrization!***

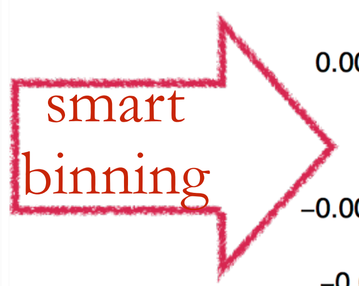
# Lateral Shower Parametrization

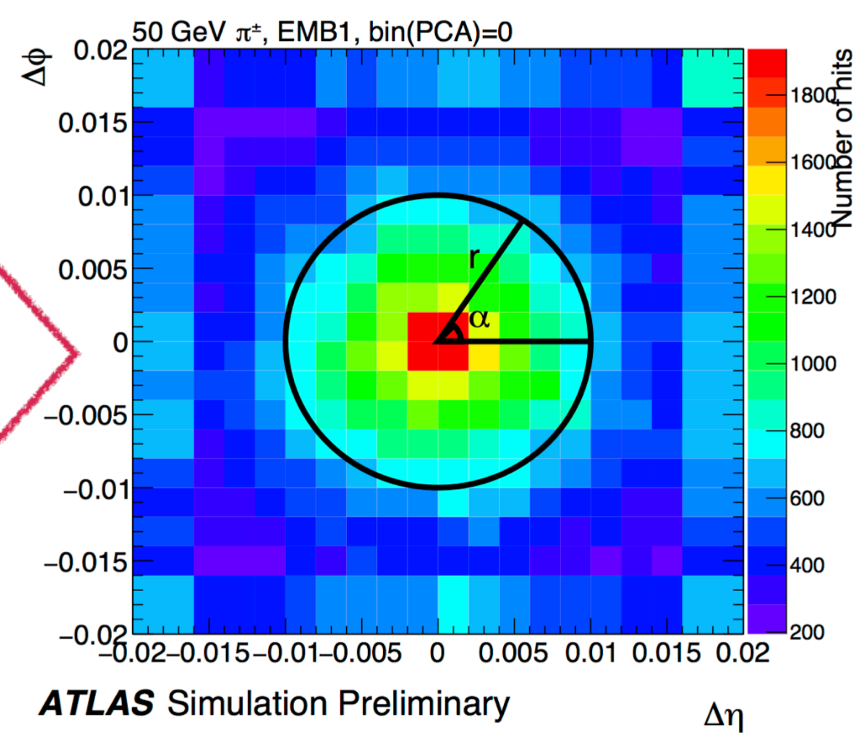
Derive shower shape parametrization for each shower type i.e. each PCA bin

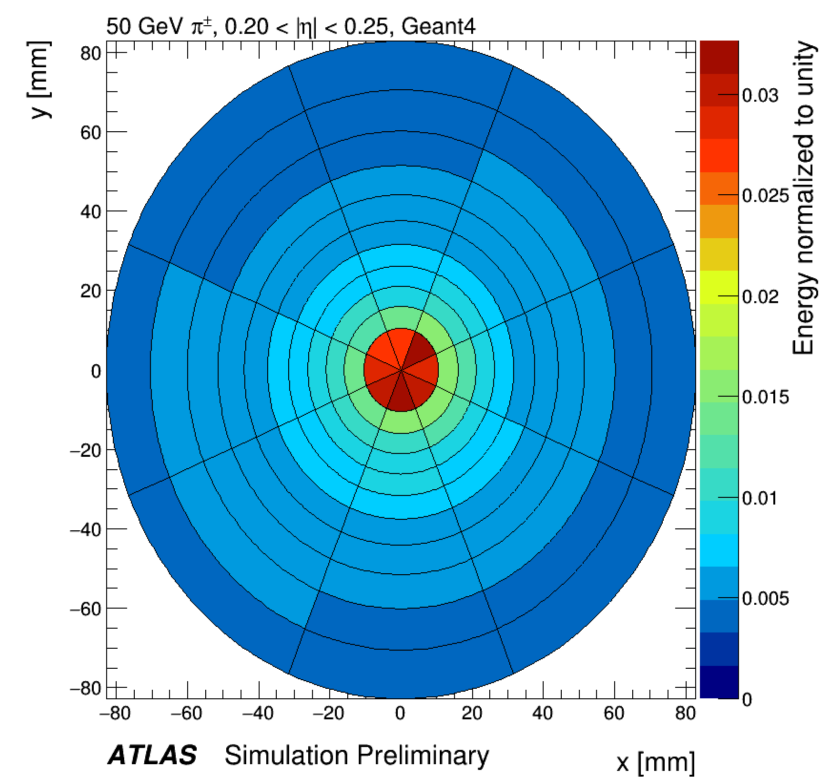
Utilize the symmetric shower topology around the center to refine the shower geometry: radial distance (r) and angle (α)

$$r[\text{mm}] = \sqrt{(\delta\eta[\text{mm}])^2 + (\delta\phi[\text{mm}])^2} \qquad \alpha = \arctan(\delta\phi[\text{mm}], \delta\eta[\text{mm}])$$






Binning defined iteratively in (α,r) using mm units to match the calorimeter quantities

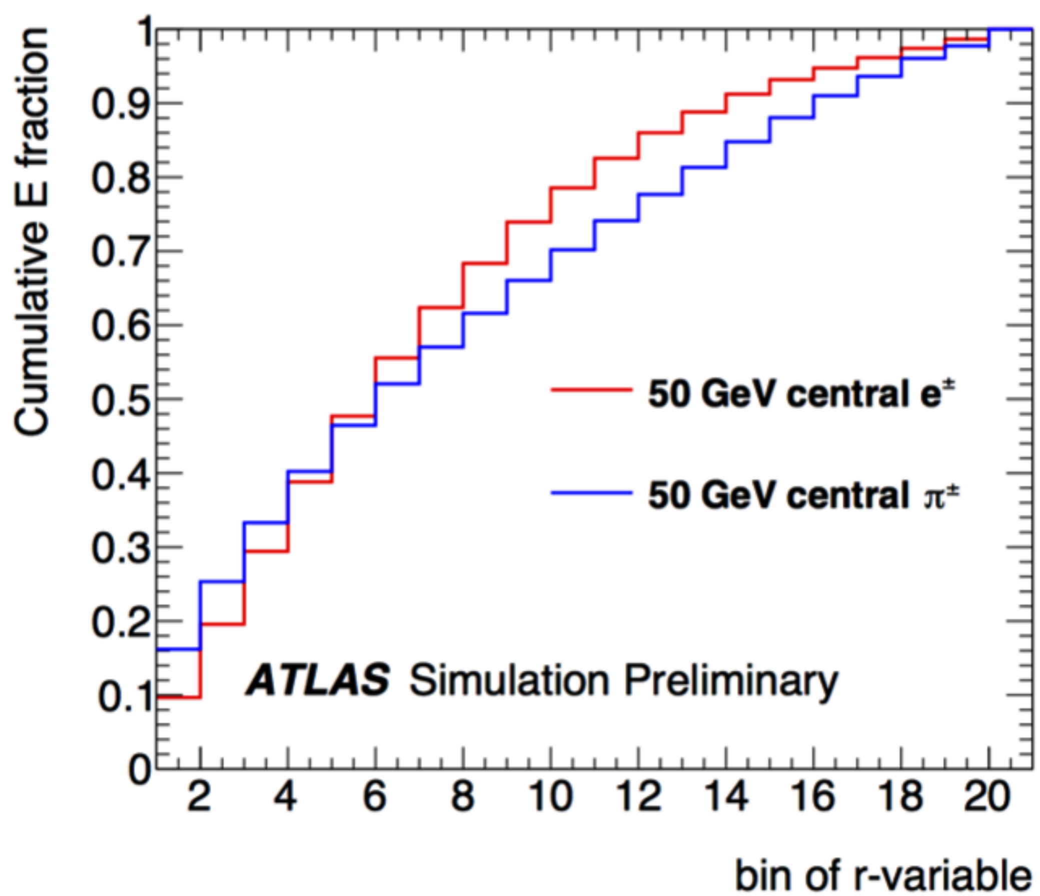
***Save the parametrized shower shape in a 2D histogram***

# Lateral Shower Regression

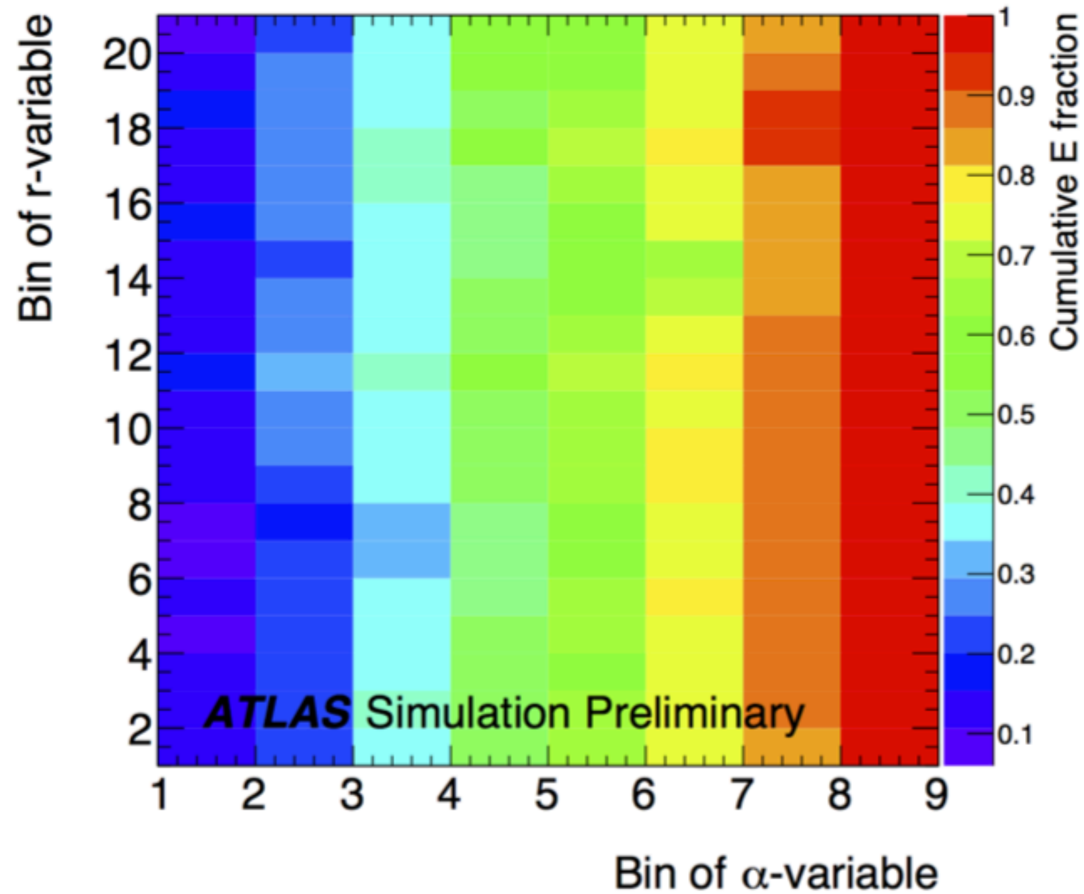
Goal: Optimize memory and I/O consumption

Solution: Get hit co-ordinates (α,r) from probability densities, without saving or reproducing the entire energy distribution

**P(r) calculated using the cumulative hit energy in bins or r, averaged for all α:**




**P(α) calculated using the cumulative hit energy in bins or α, for each r bin:**




Train the neural network with P(α) , P(r) as input variables and α, r as target variables

***Shape regression is under development: Coming soon!***

# Lateral Shower Simulation

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

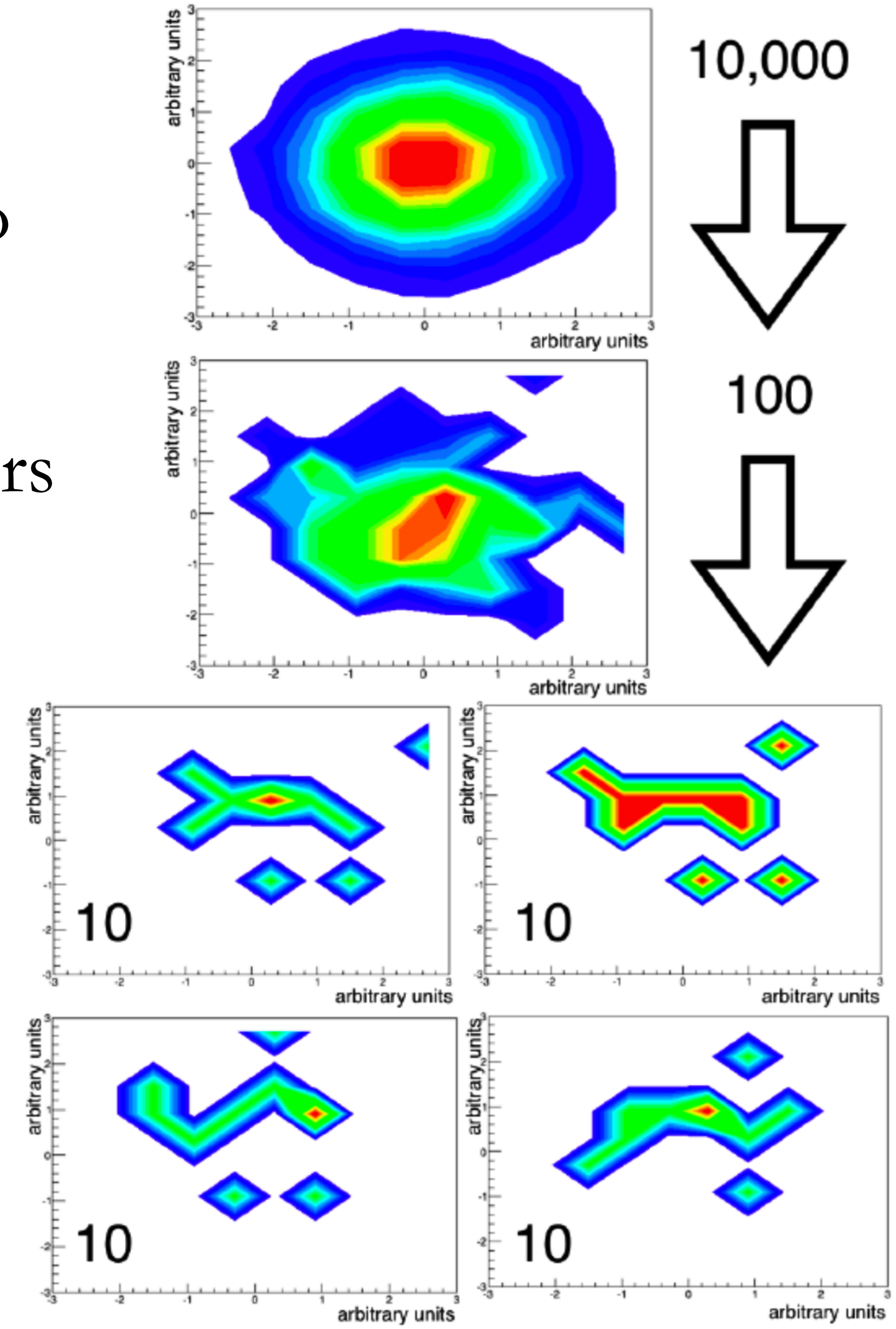
***Sufficient to describe fluctuations is electromagnetic showers***

# Lateral Shower Fluctuation

Hadronic shower characterized by multi particle production and particle emission

Complex and irregular shower formation compared to electromagnetic processing

Results into different number of reconstructed clusters i.e. large fluctuations




Mimic hadronic shower complexity/fluctuation:

Sample fewer hits for each shower

Split one pion shower into two pion showers (probability of splitting derived from Geant4)

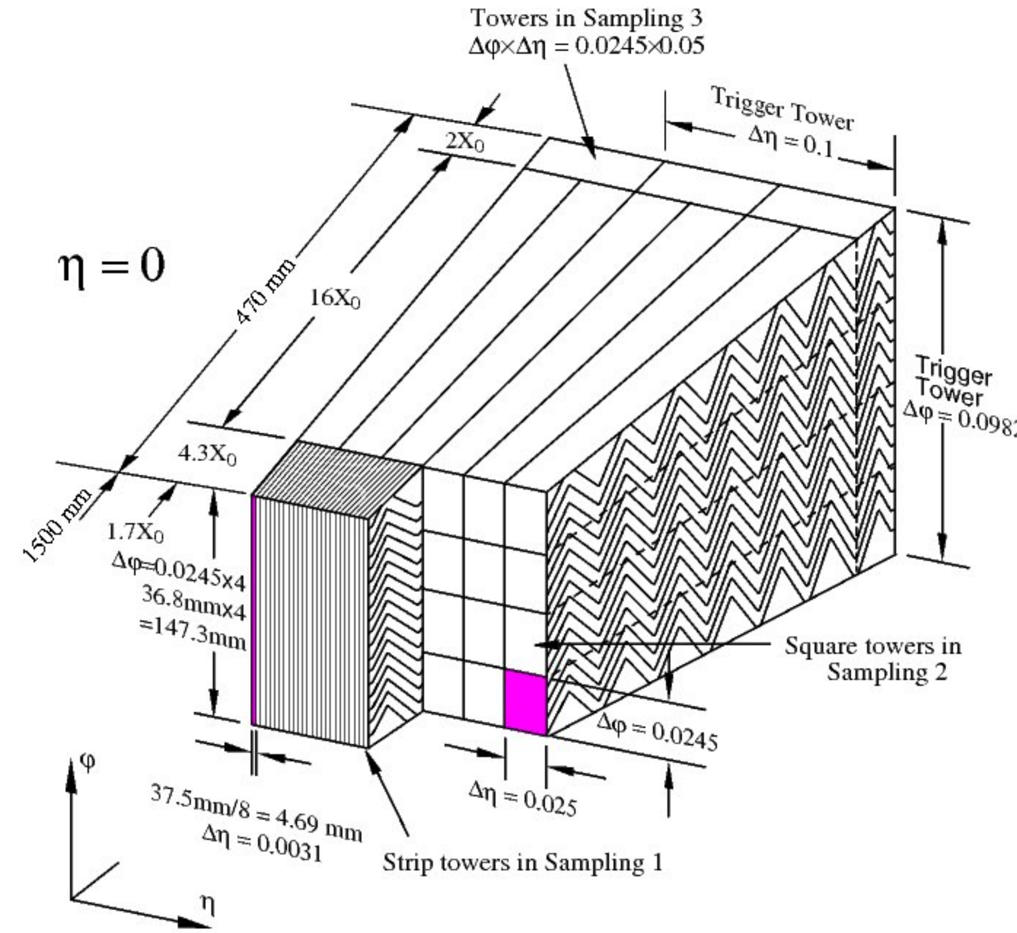
***Introduce fluctuation with fewer hits and pion splitting***

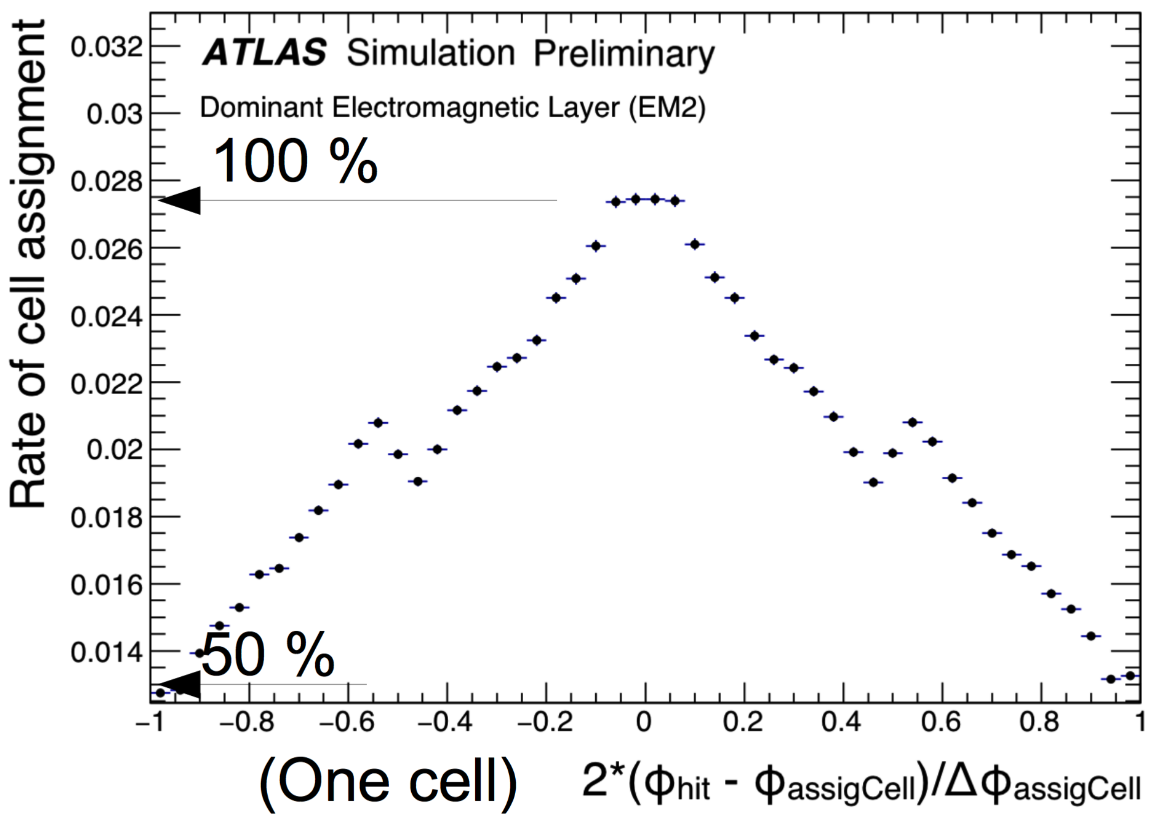
# Hit to Cell Assignment: Correct for simplified Geometry

Simulated hits assigned to cells assuming a simplified cuboid geometry

In reality, the calorimeter has a accordion geometry

Results in incorrect hit to cell assignment



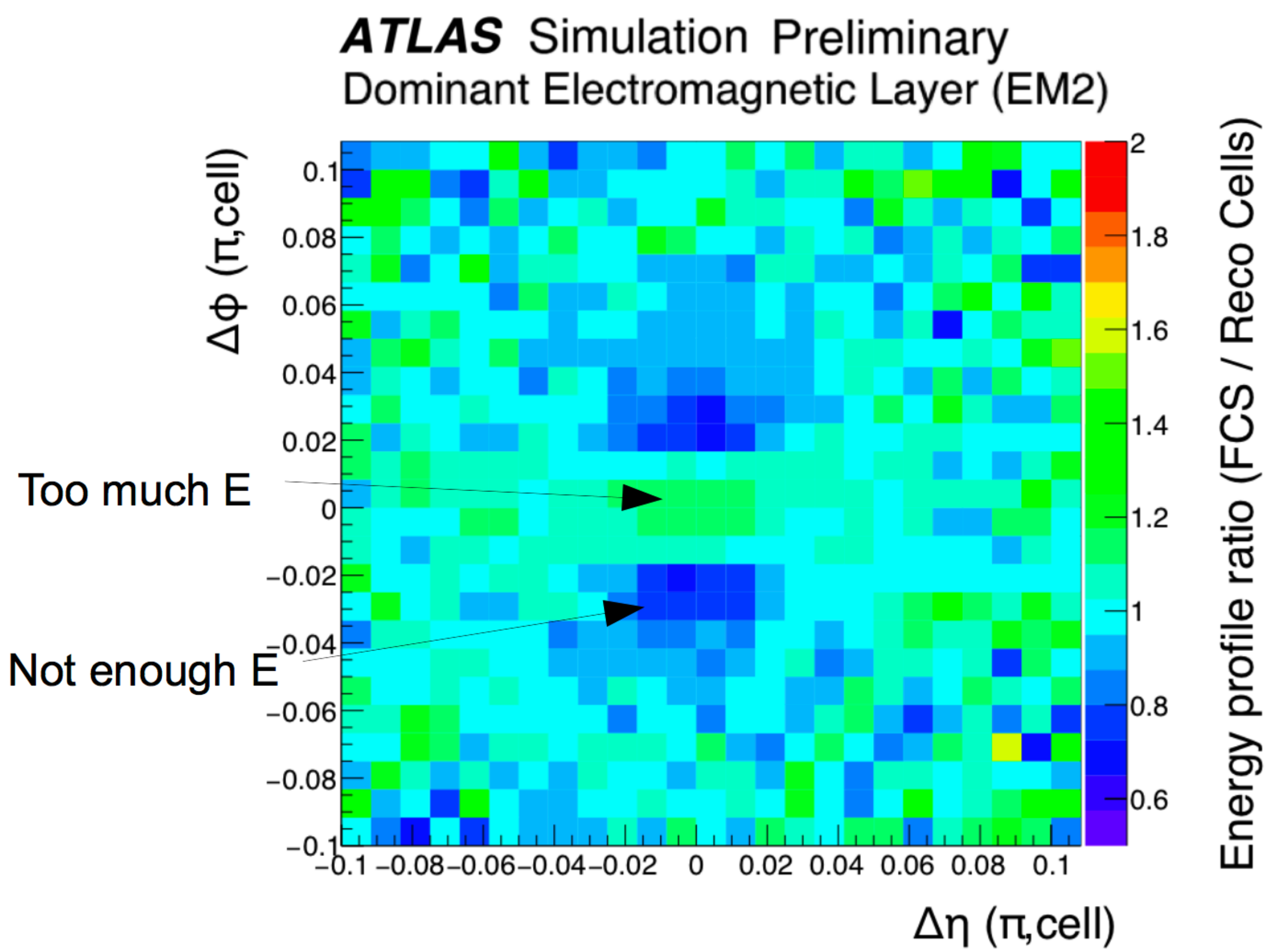




Define a function to describe the probability that a hit belongs to a neighboring cell

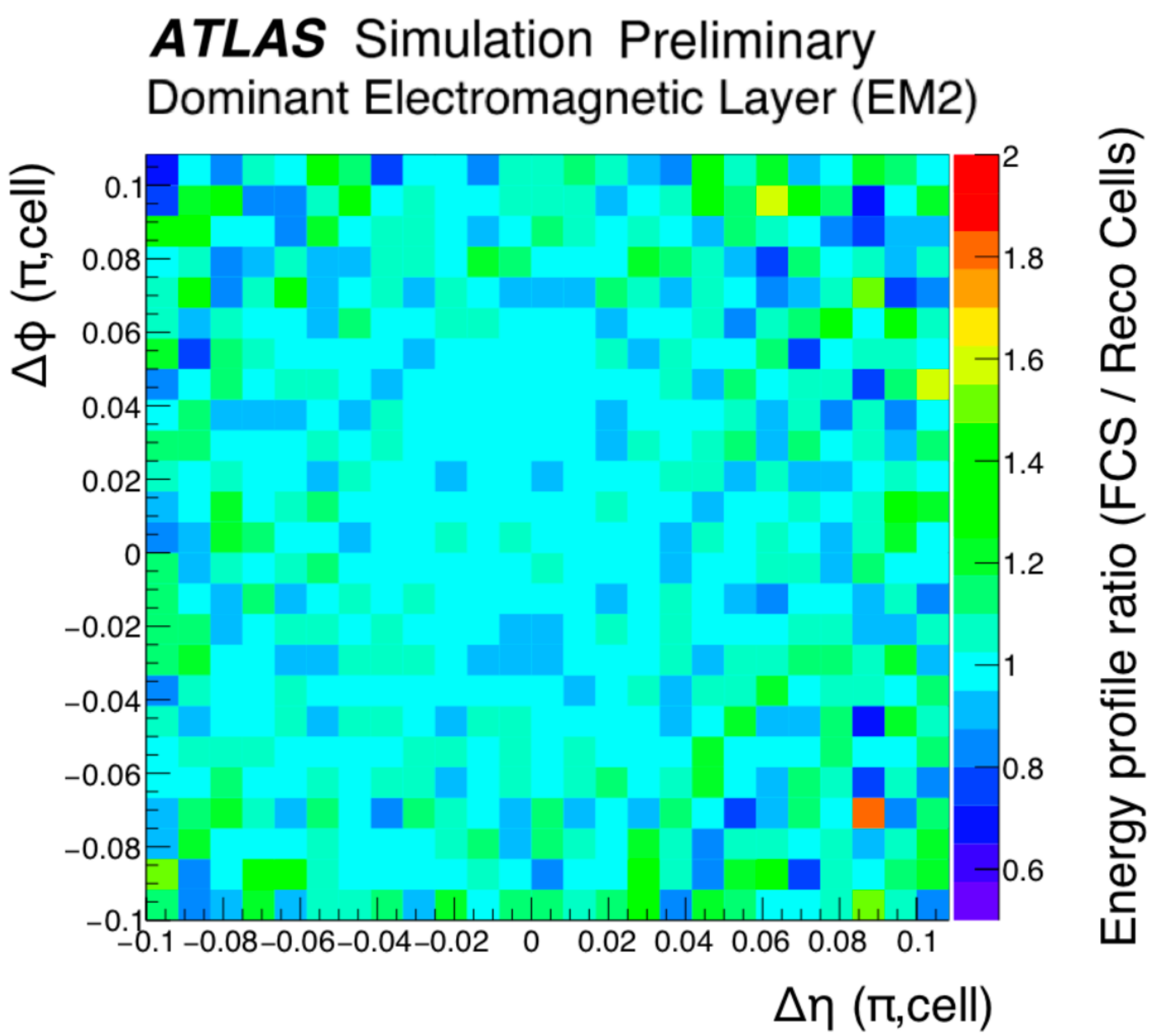
***Correct for simplified geometry with a hit displacement function***

# Hit to Cell Assignment: Correct for simplified Geometry

w/o hit displacement




w/ hit displacement




Reco cell: Geant4 cell
FCS cell: assigned cell

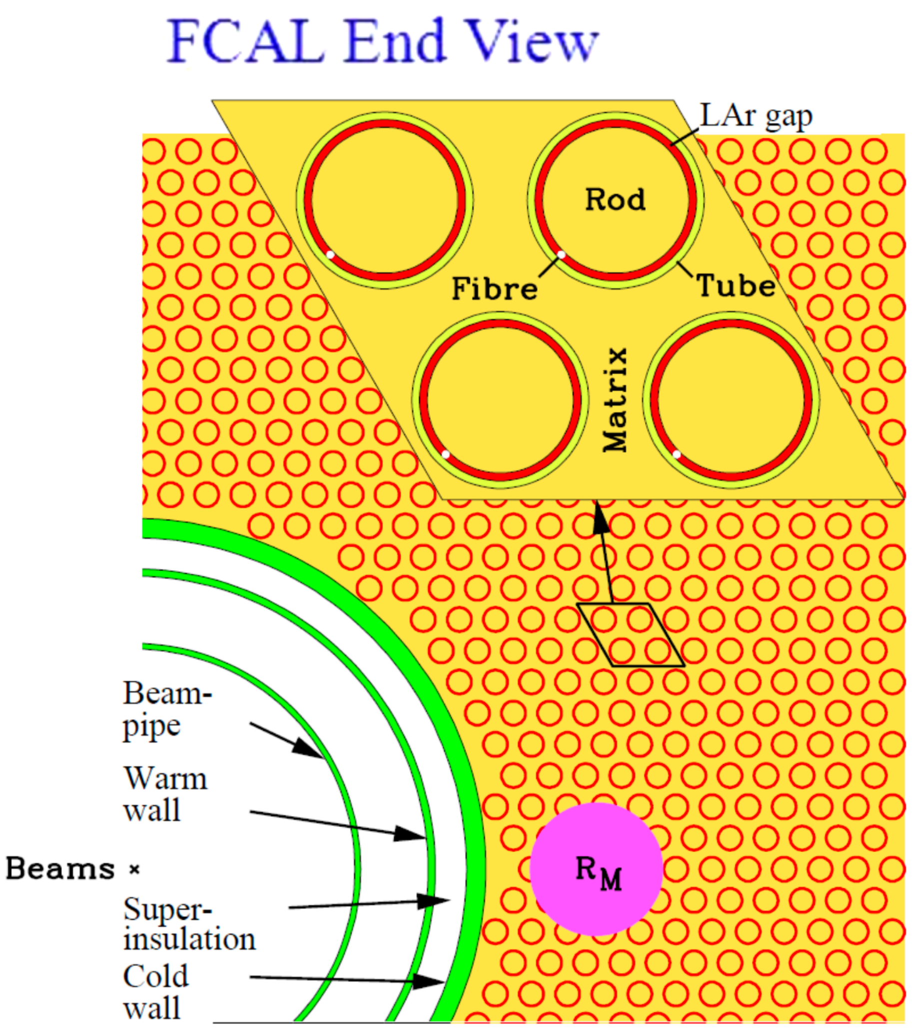
***Good agreement with Geant4 cell assignment!***

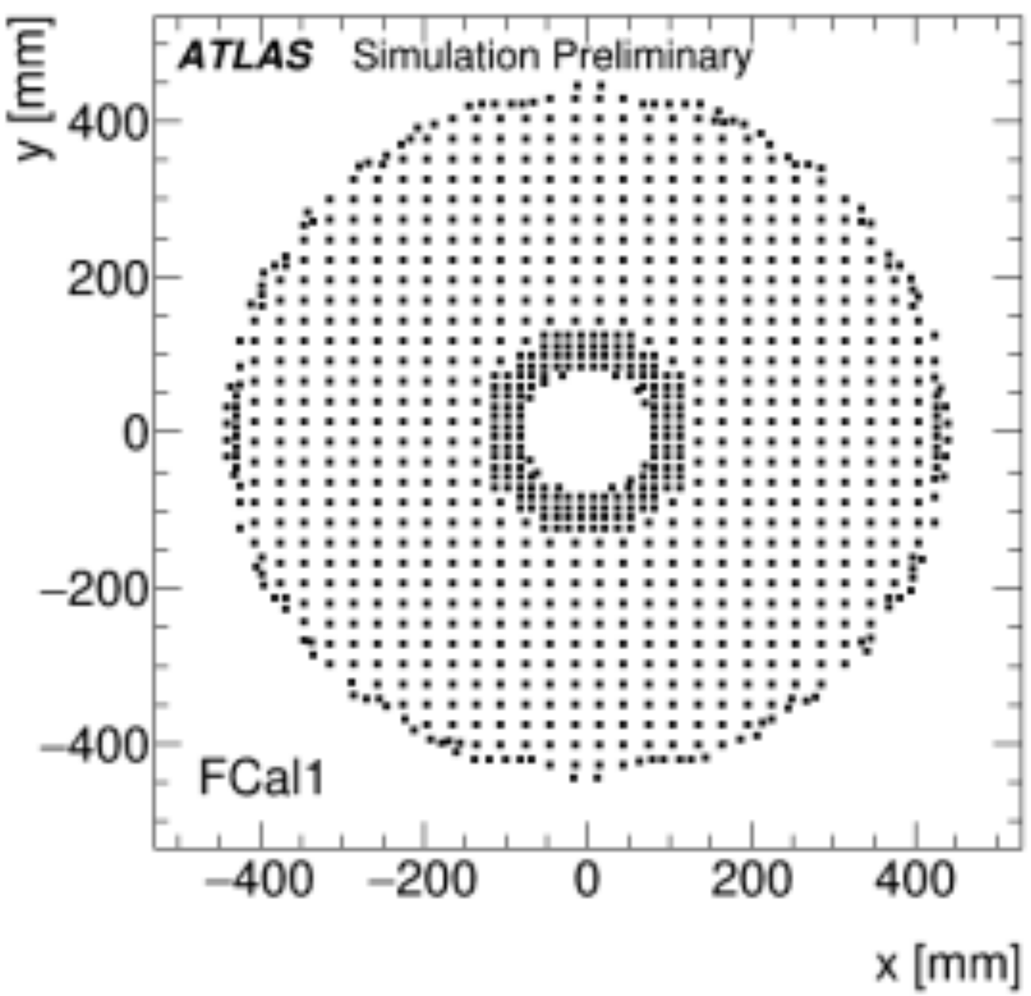
# Forward Calorimeter Geometry

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



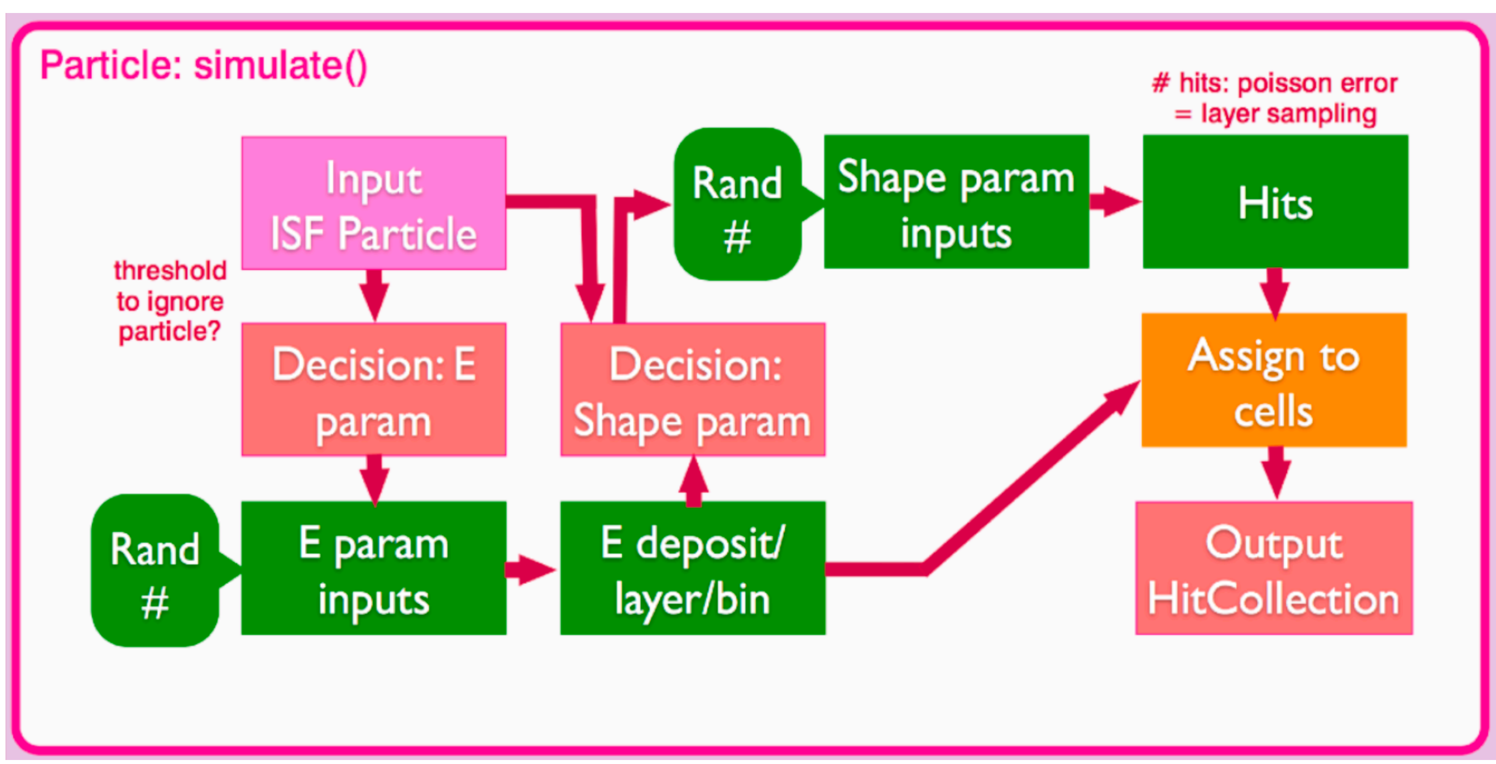

***Dedicated parametrization for* FCals**

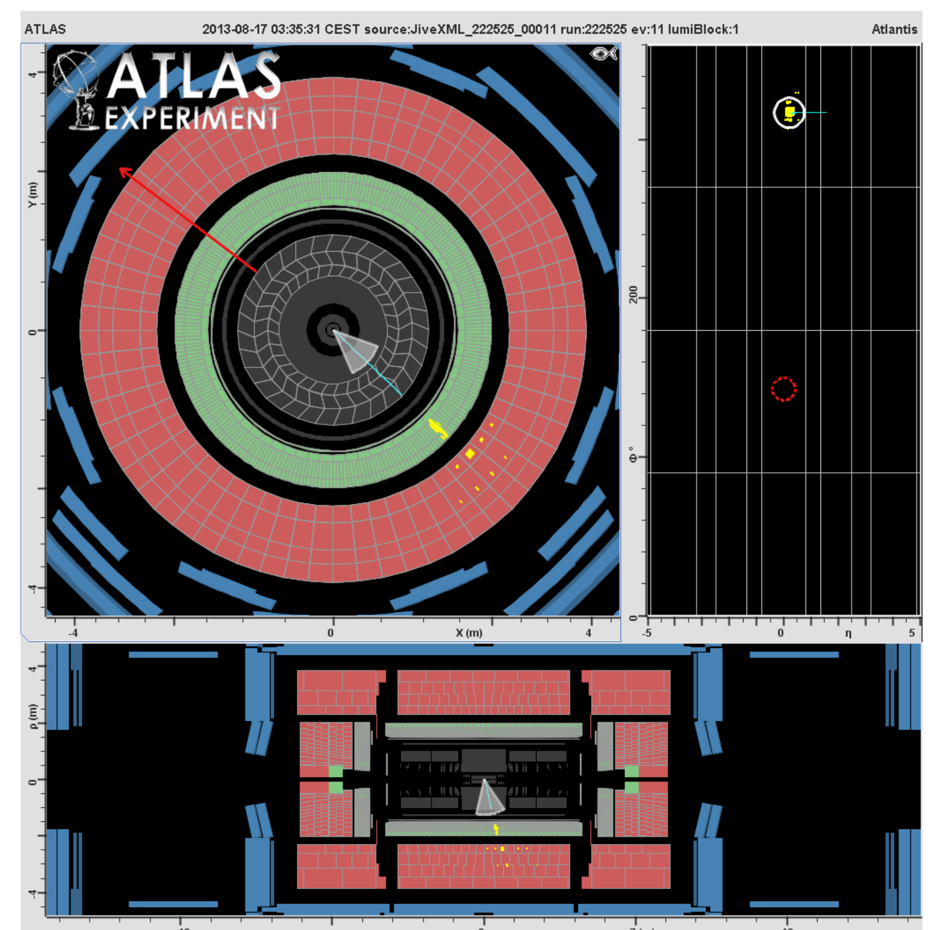
# Prototype: Putting it all together!

FastCaloSimV2 is part of the Integrated Software Framework (ISF)

Only the ATLAS Calorimeter is simulated with fast simulation, inner detector and muon systems are simulated with Geant4

Integrated in the ATLAS software development and production releases




***Utilizes the entire Monte Carlo production chain to produce reconstructed single particle events for validation***

# Validation: How well does it perform?

Current validation performed with single particle events

Study shower shape variables for electromagnetic showers and cluster variables for hadronic showers

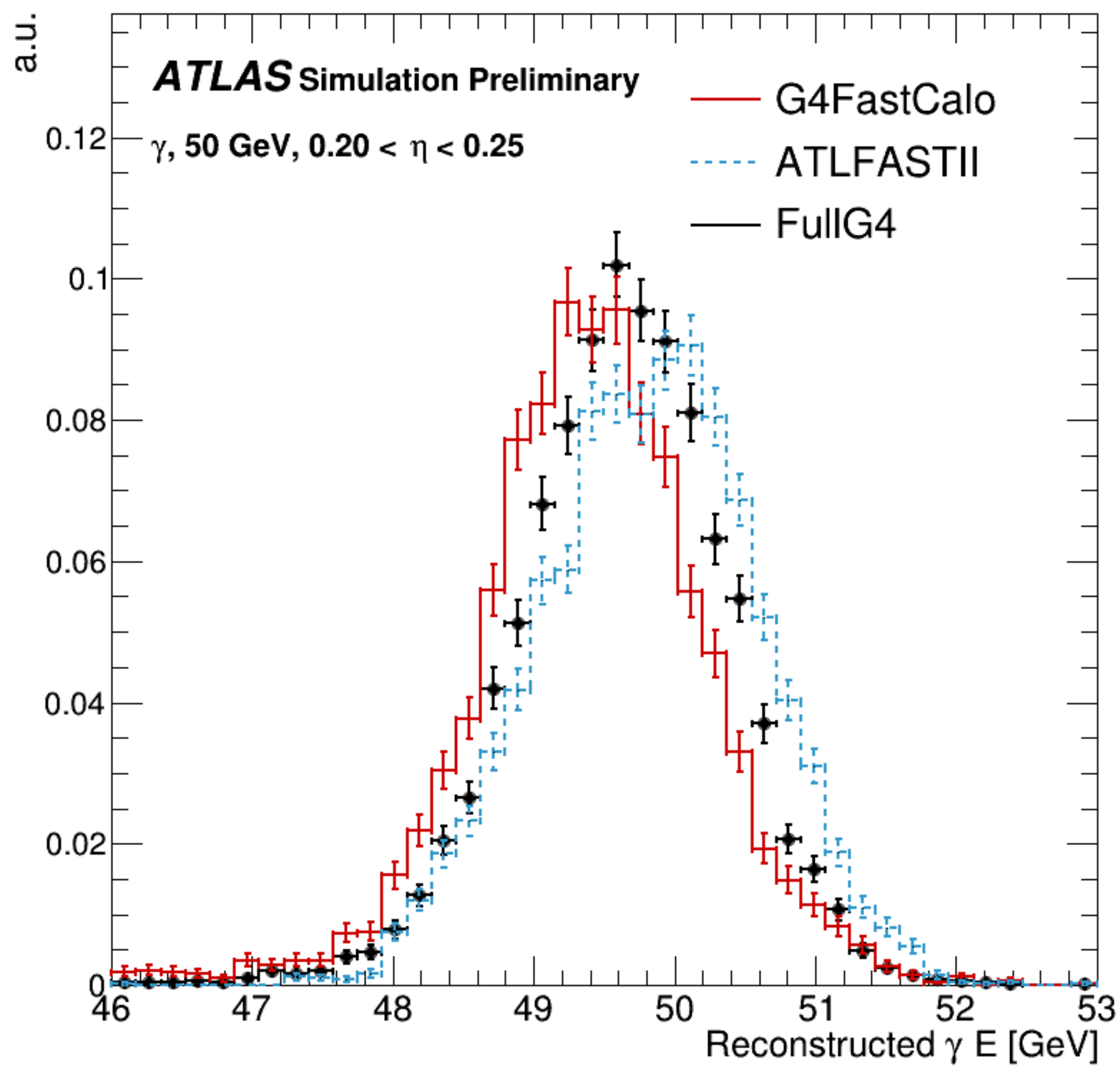
Compare the distributions with the ATLFASTII and Full Geant4 simulated distributions

G4FastCalo = FastCaloSimV2 for calorimeter + Geant4 for ID/muon
ATLFASTII = FastCaloSim for calorimeter + Geant4 for ID/muon
Full Geant4 = Geant4 for calorimeter, ID and muon

**ATLFASTII** ***events are tuned to data!***
**G4FastCalo** ***events are not yet tuned***

# Performance in electromagnetic shower

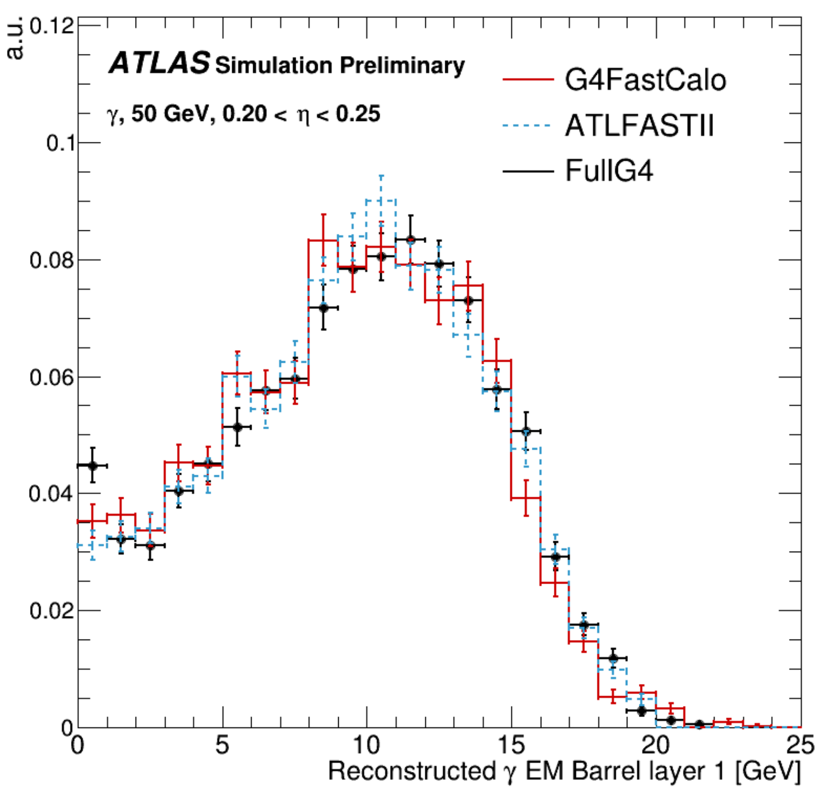
# Reconstructed photon energy

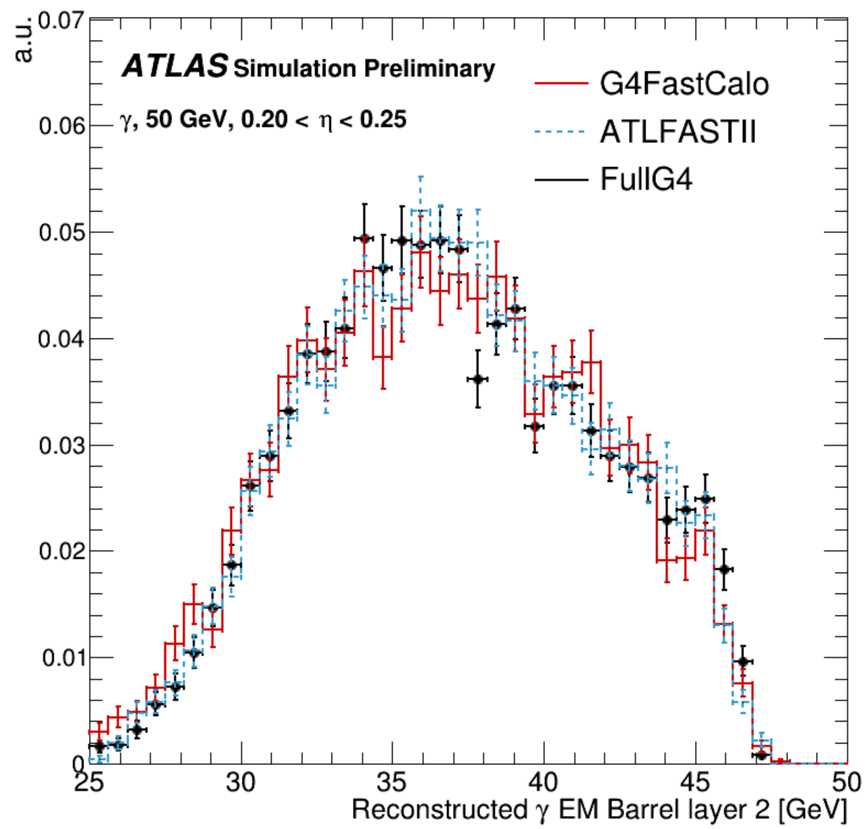


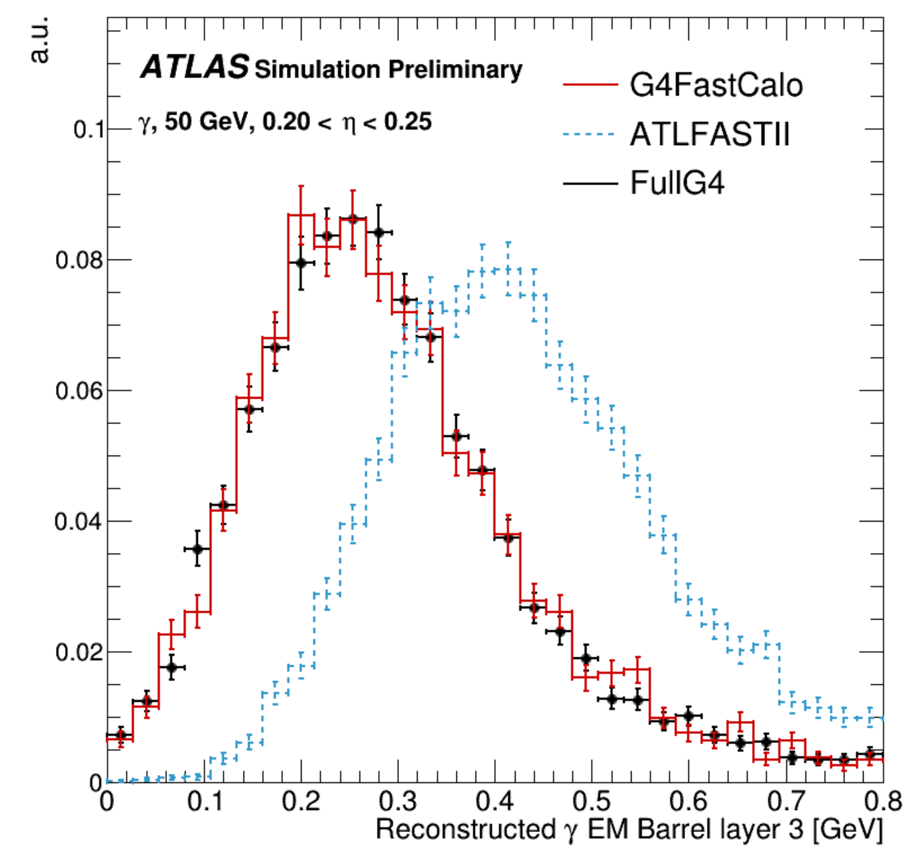

***Mean value closer to Geant4 compared to AF2***

# Reconstructed photon energy in each layer

Energy fraction in each electro magnetic (EM) barrel layer



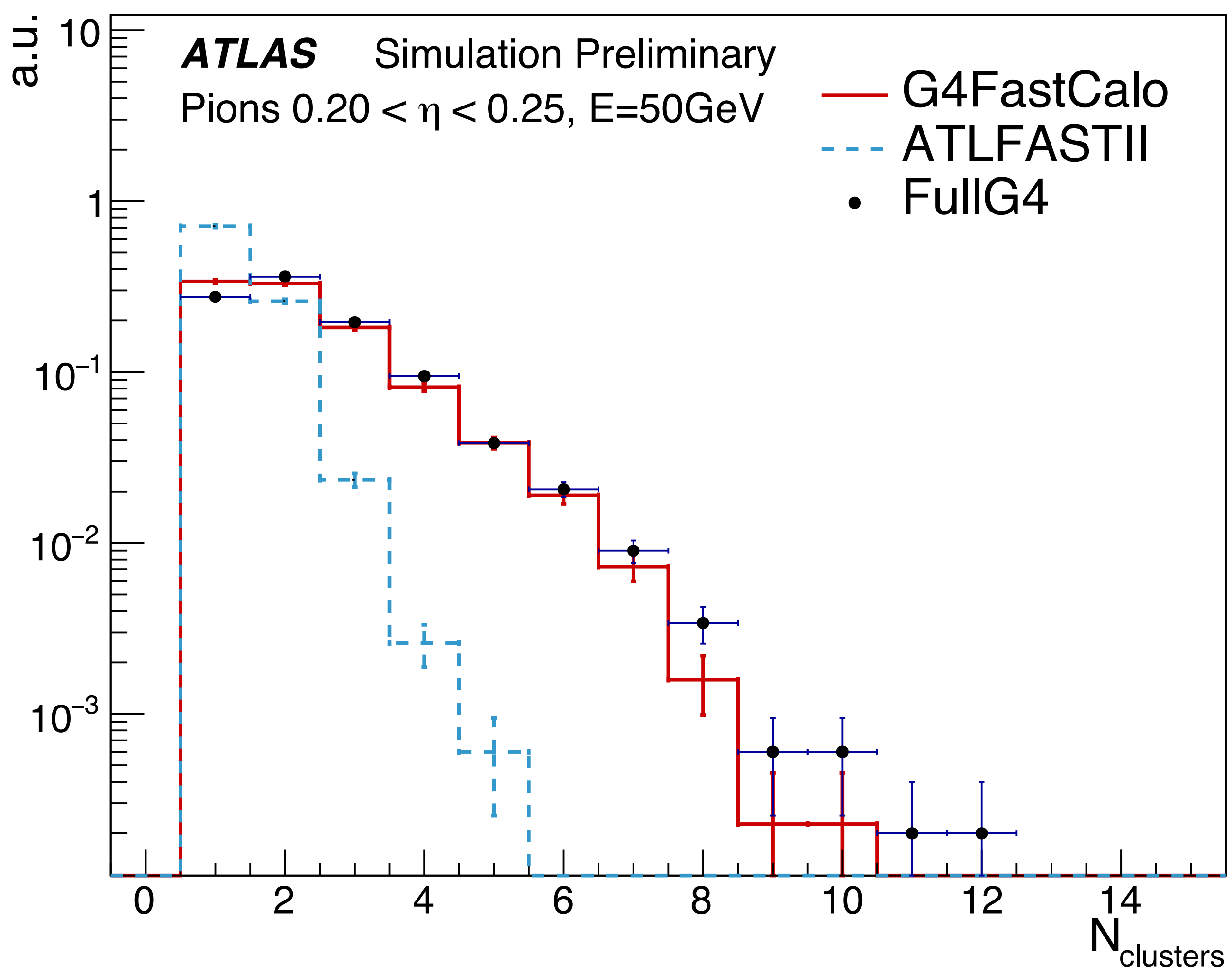



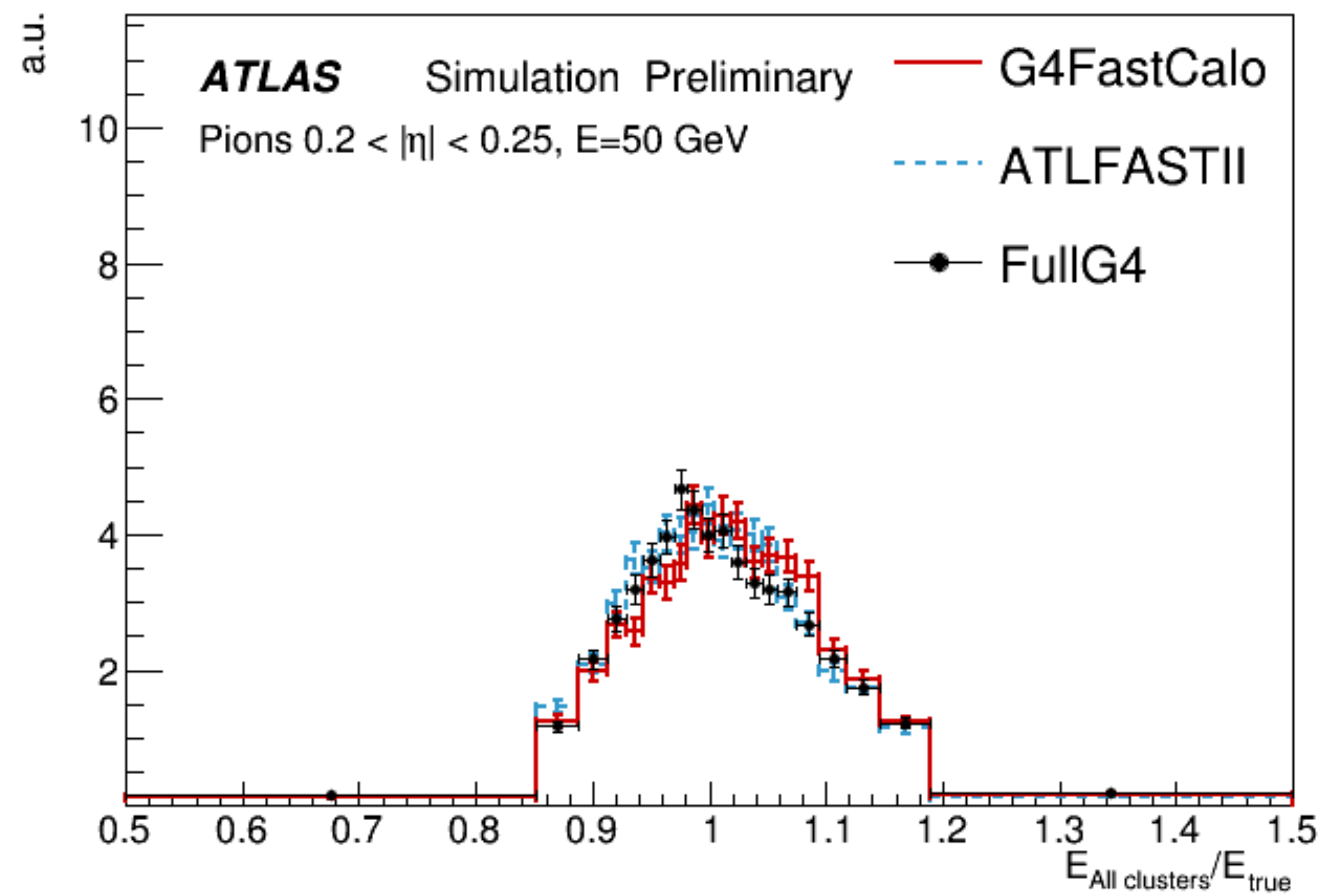




***Significant improvement in 3rd EM barrel layer compared to AF2***

# Performance in hadronic shower

# Number of Calorimeter Clusters & Energy



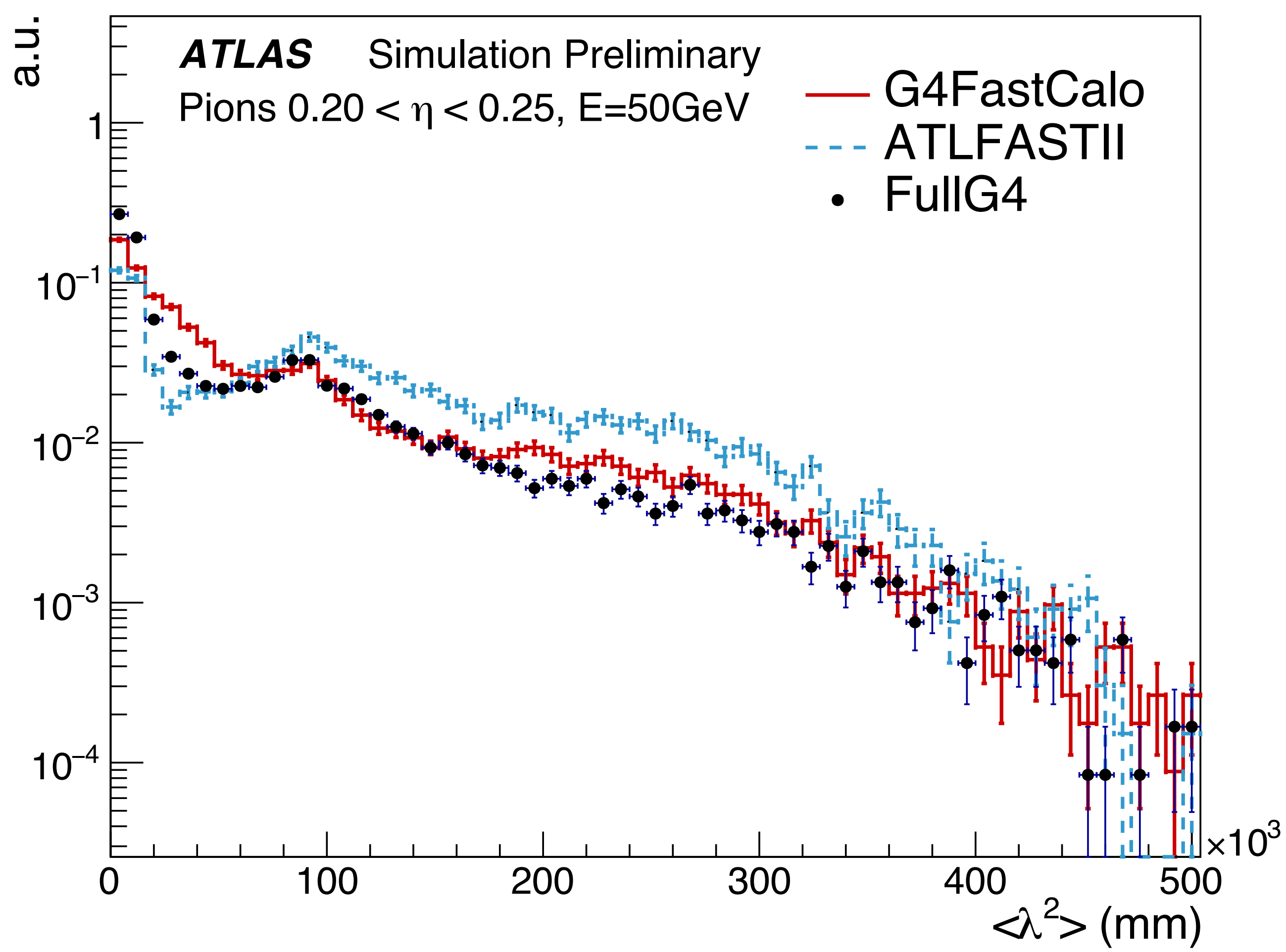




***Significant improvement in the no of clusters compared to AF2***

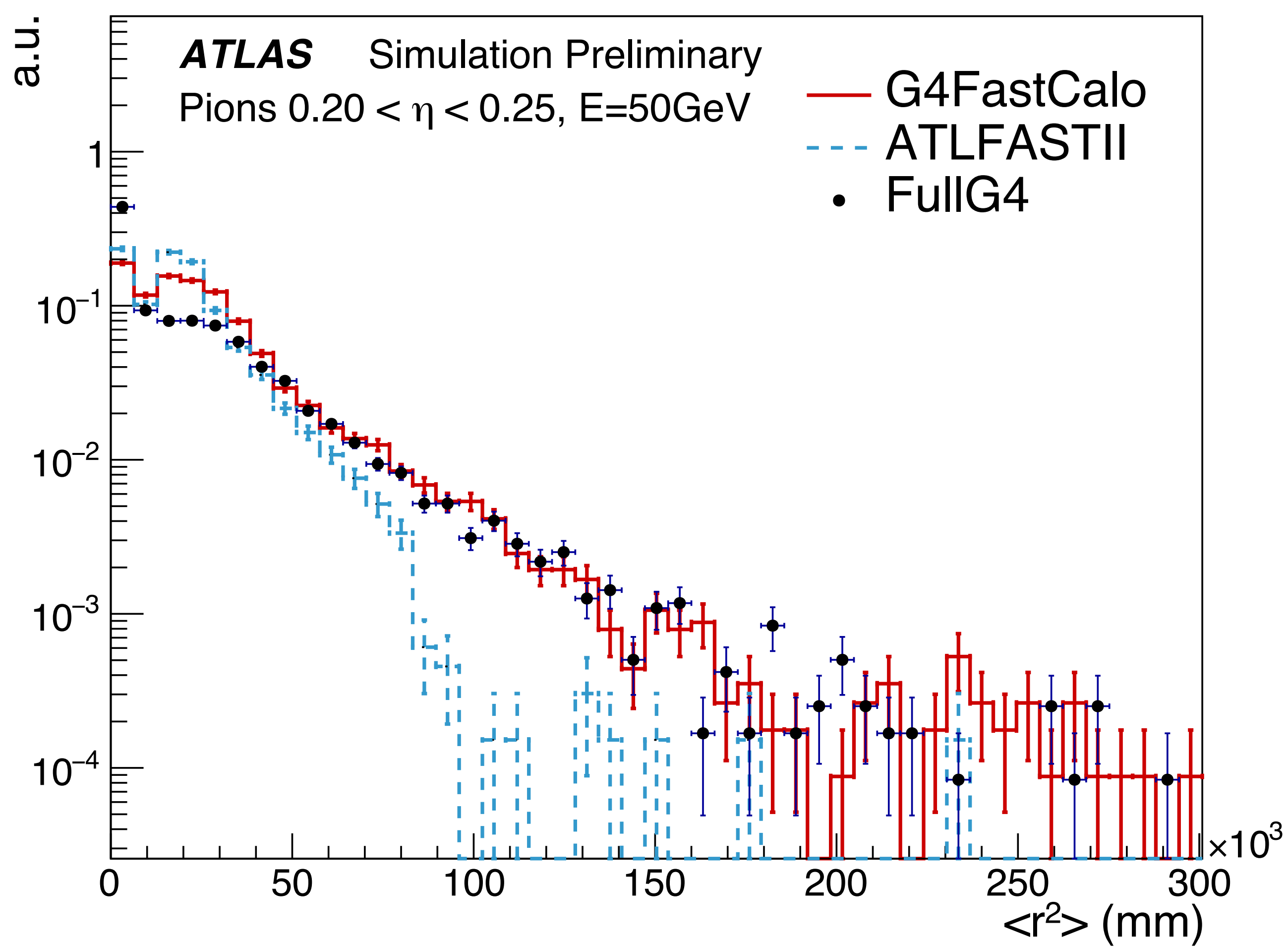
# Cluster moments

Describe inner structure of a cluster

$$< x^n >= \frac{\sum_{E_i>0} E_i x_i^n}{\sum_{E_i>0} E_i}$$




λ distance from the shower center along shower axis



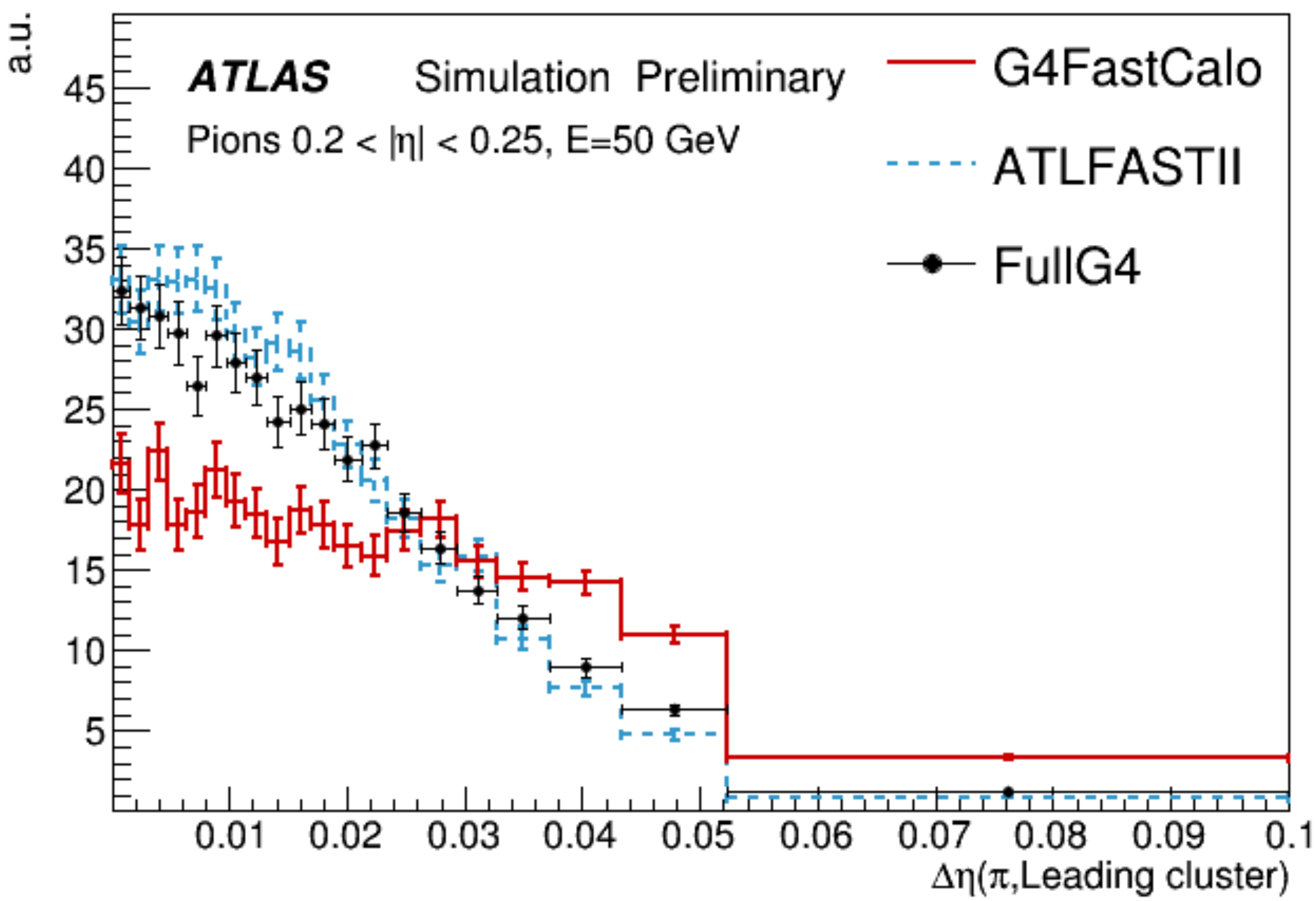

r perpendicular distance of the cell from the shower axis

***Agreement with Geant4 can be improved at low values***

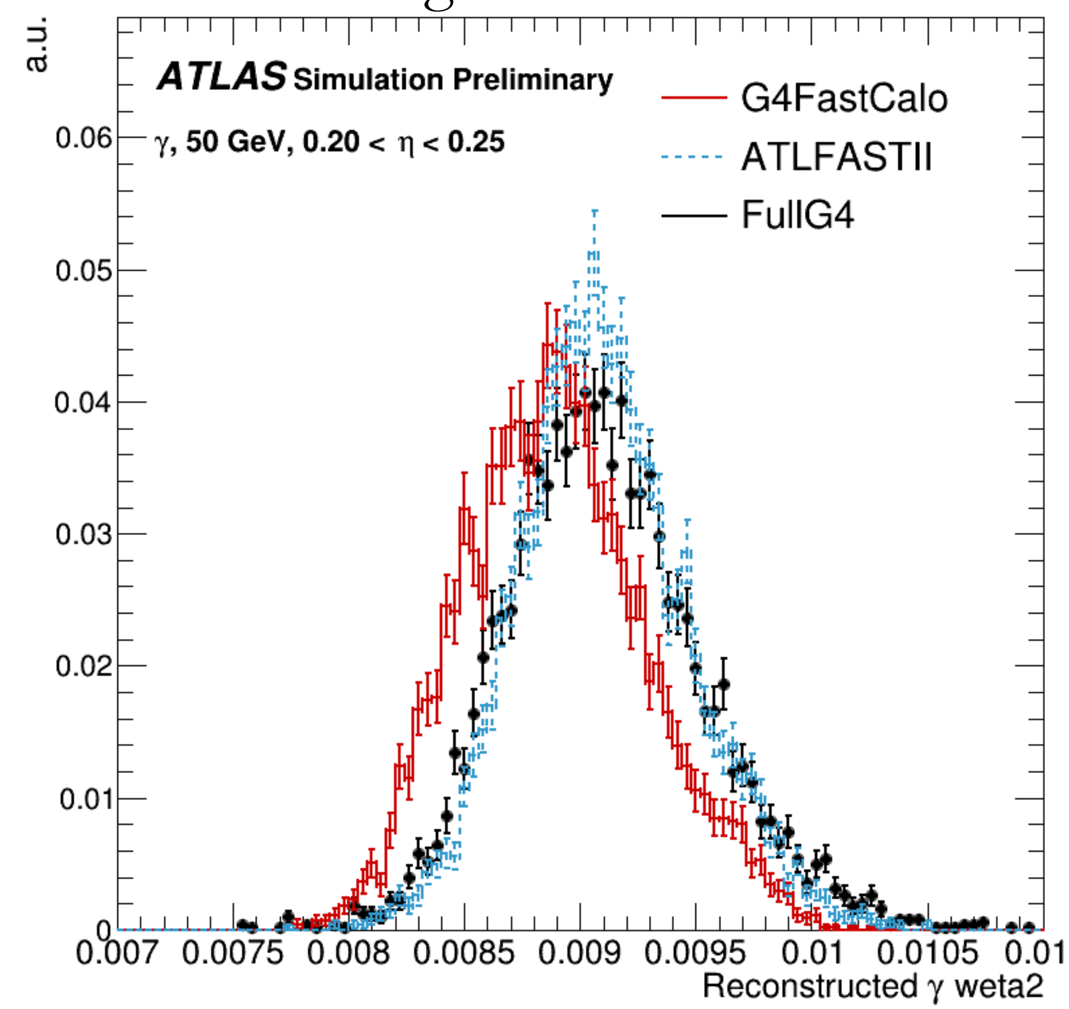
# Still room for improvement …

Some cluster/shape distributions show disagreement

$\Delta\eta$ between true pion and cluster with highest energy deposit




Lateral shower width calculated using 3x5 cells window




***Tuning to data can improve these distributions***

# Summary and Outlook

- Fast simulation essential for ATLAS physics program at 14 TeV and at HL-LHC
- Current fast calorimeter simulation does not describe collision data adequately to be used in precision measurements
- New fast calorimeter simulation has been developed and the first prototype is integrated in the ATLAS software release
- FastCaloSimV2 shows good agreement with Geant4 and in some cases outperforms the current FastCaloSim out of the box
- Current version only tested for a certain energy and rapidity region. Parametrization and validation needs to be performed for other $(E,\eta)$ points
- With complete $(E,\eta)$ parametrization, physics processes would be simulated for validation and tuning to collision data

***Fast Simulation is destined for greatness: believe it!***



***Coming in your town: December 23rd!***