南京大學
NANJING UNIVERSITY

# Machine learning in ATLAS simulation

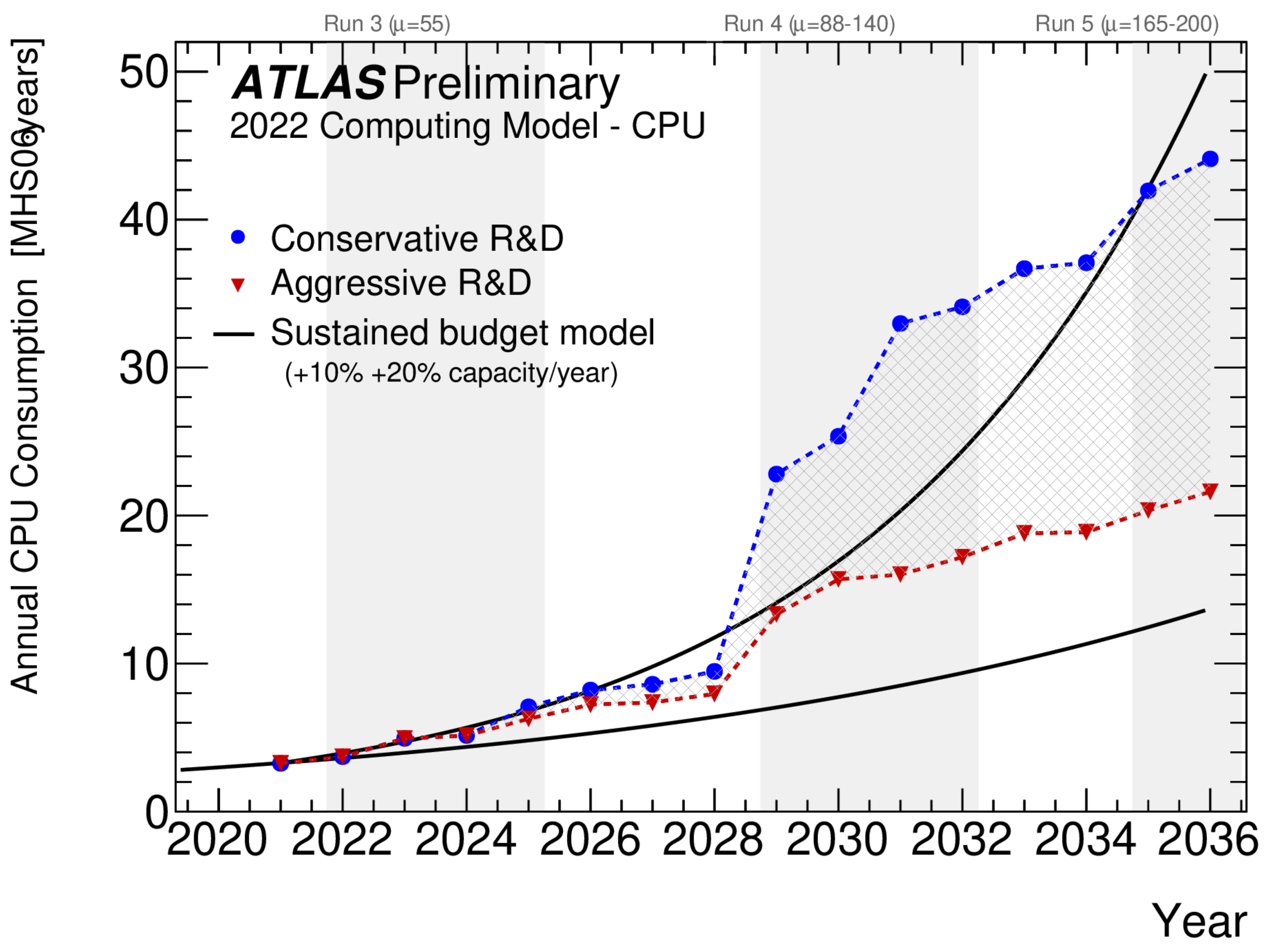
张 瑞 (南京大学)

2026年超级陶粲装置研讨会

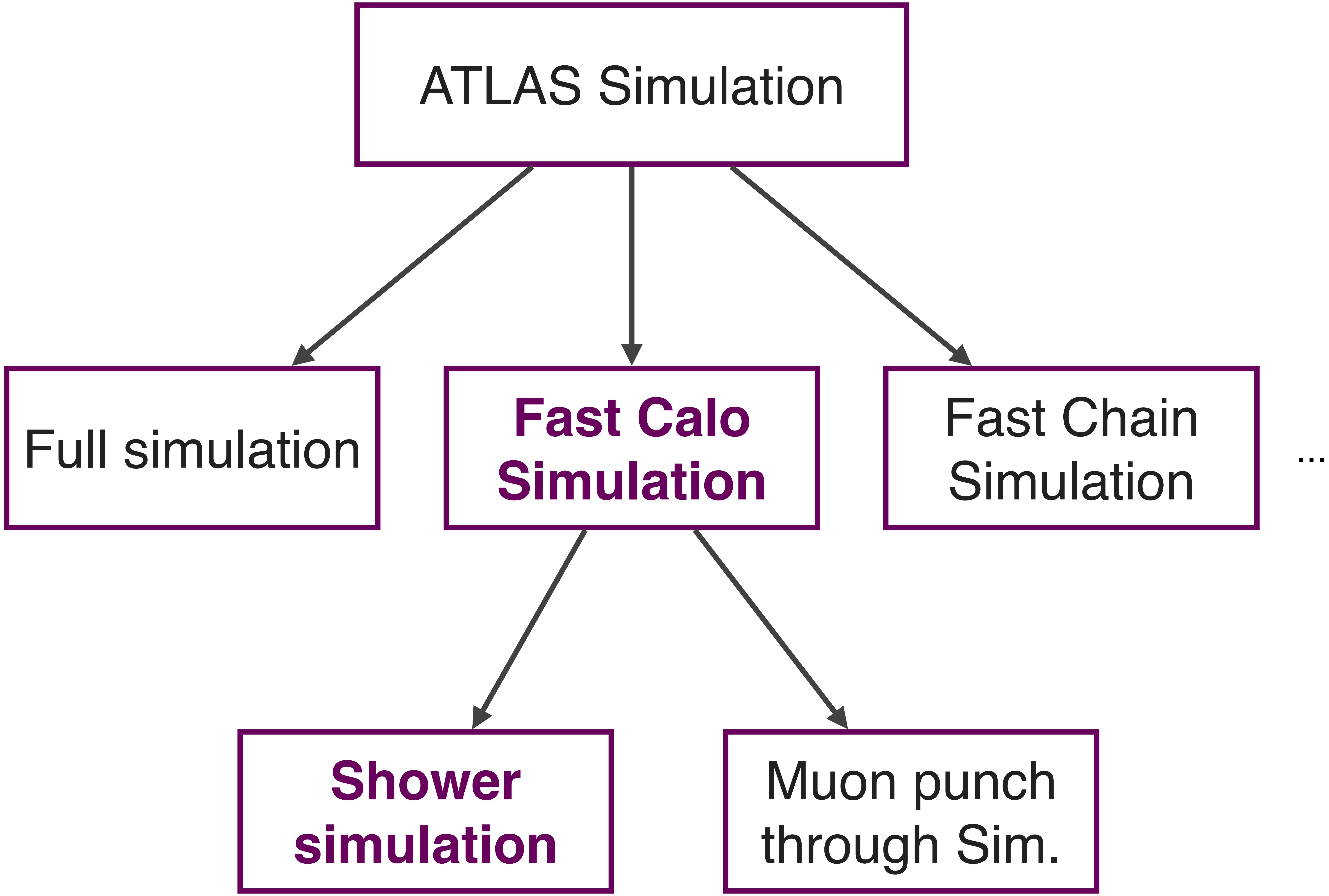
2026.6.30—2026.7.5，西安

# Introduction

- Machine learning (ML) has wide applications in high energy physics
- One of the applications is in detector simulation
- Geant 4 is the most precise tool for detector simulation, however, it requires considerable CPU resources
- ML can assist in various places in fast simulation tools
- Present today:
  - Machine learning to acceleration simulation at ATLAS



# Contents


ATLAS Simulation
Full simulation
Fast Calo Simulation
Fast Chain Simulation
...
Shower simulation
Muon punch through Sim.

# ATLAS Fast Calorimeter Simulation (AF)

| AltFast | → 5 years → | AF2 | → 10 years → | AF3 Run2 version | → 2.5 years → | AF3 Run3 version |
|---|---|---|---|---|---|---|
| (Before LHC data taking) | | (Run 1, Run 2) Much improved wrt AF | | (Run 2) Much improved wrt AF2 | | (Run 3) Improved wrt AF3 Run2 |

**AltFast**

Simplistic, e.g. Gaussian smearing

**AF2**

**FastCaloSim**
- Based on parametrisation
- 2D splines for lateral shape

------------

- ■ Limited shower fluctuations
  - ■ Not usable for boosted topologies
- ■ No FCAL parametrisation
  - ■ Not usable for VBF
- ■ No Muon Punch-Through

**AF3 Run2 version**

**FastCaloSimV2**
- PCA for longitudinal energy deposits (better correlation)
- Hit-based lateral shape with more fluctuations to resemble energy resolution

**FastCaloGAN**
- First use of GANs in a large-scale HEP experiment)
- Used for mid-energy pions

**Muon Punch-Through**
- Simple (based on a look-up)

------------

- ■ $0 < |\eta| < 5$ fully covered
- ■ Derived from G4 10.1
- ■ Released in Spring 2021

**AF3 Run3 version**

**FastCaloSimV2**
- Same as Run2 version

**FastCaloGANV2**
- Improved GANs
- Introduce proton GANs

**Muon Punch-Through**
- Improved based on DNN

------------

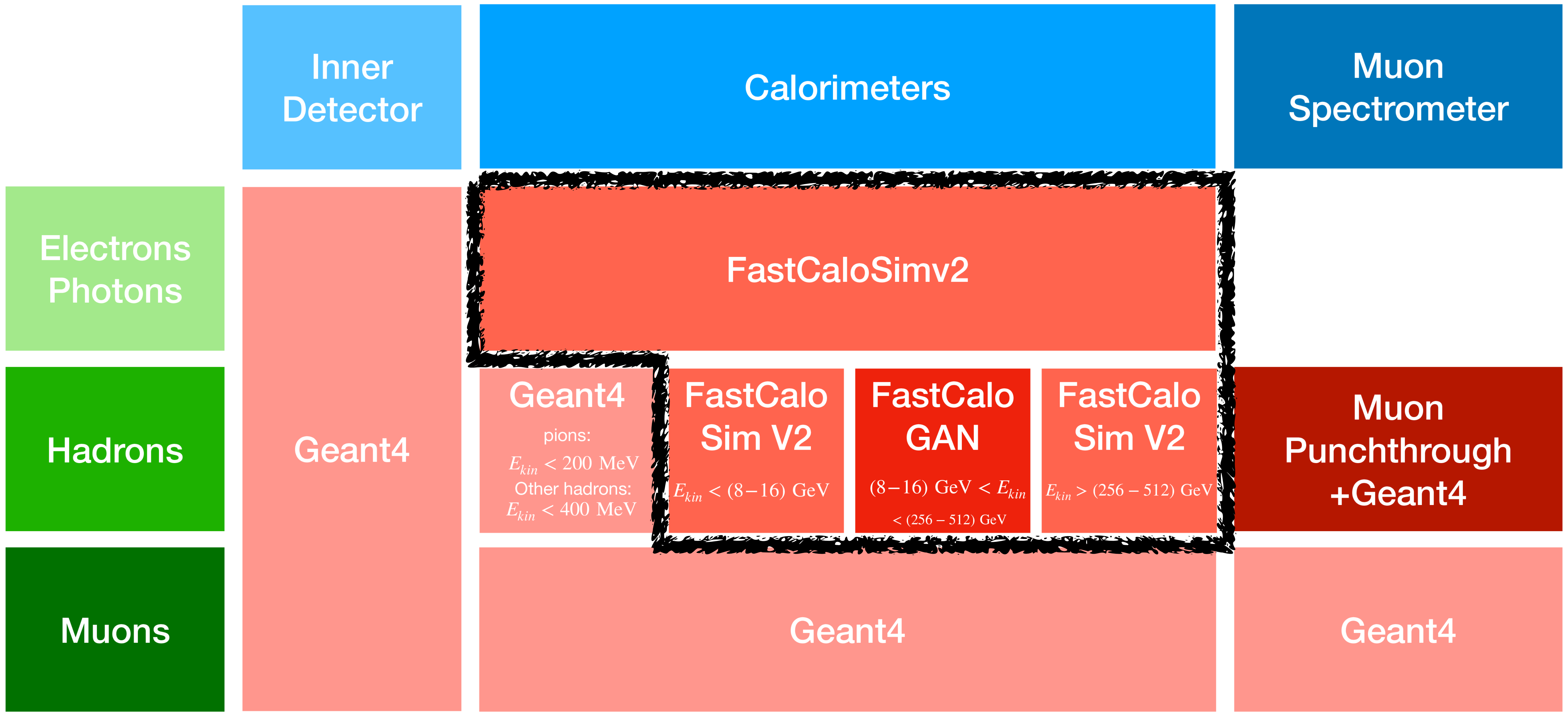
- ■ $0 < |\eta| < 5$ fully covered
- ■ Derived from G4 10.6
- ■ Released in Fall 2023

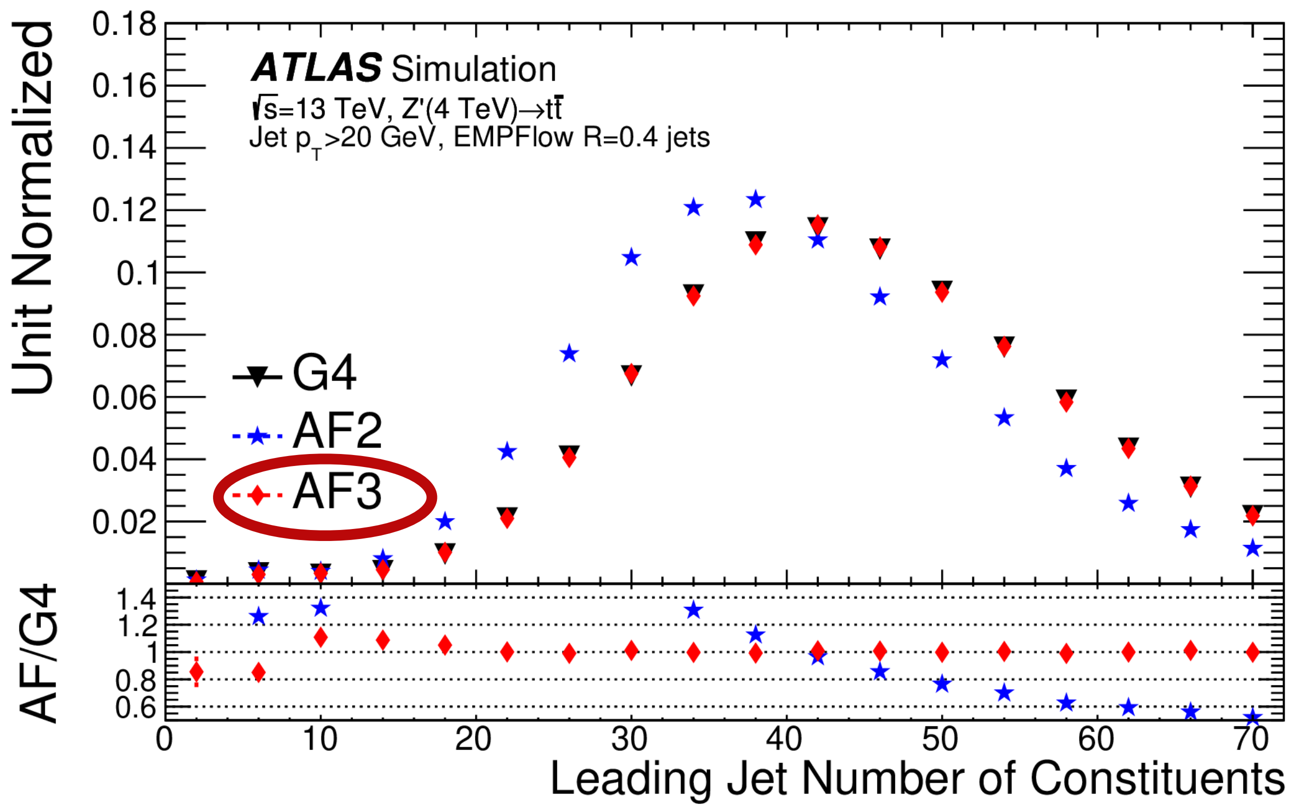
# AF3 Run2 version

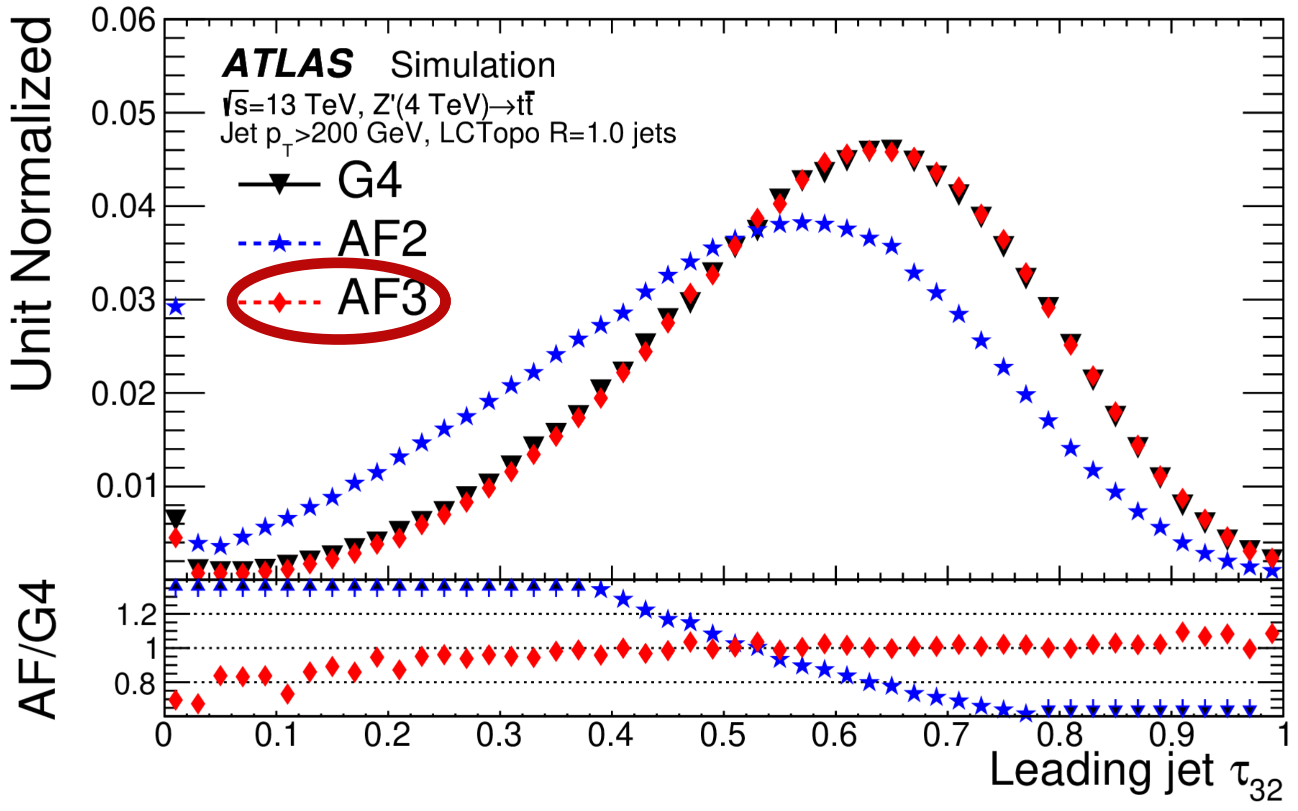
- Much improved over AF2
  - Has forward calo parametrisation, can be used for **VBF**
  - Better correlation modelling, can be used for **jet substructure and boosted** topology
  - First attempt to use ML in production

Note: Inner Detector, low energy hadrons, muons are simulated with Geant4

Comput Softw Big Sci 6, 7 (2022)



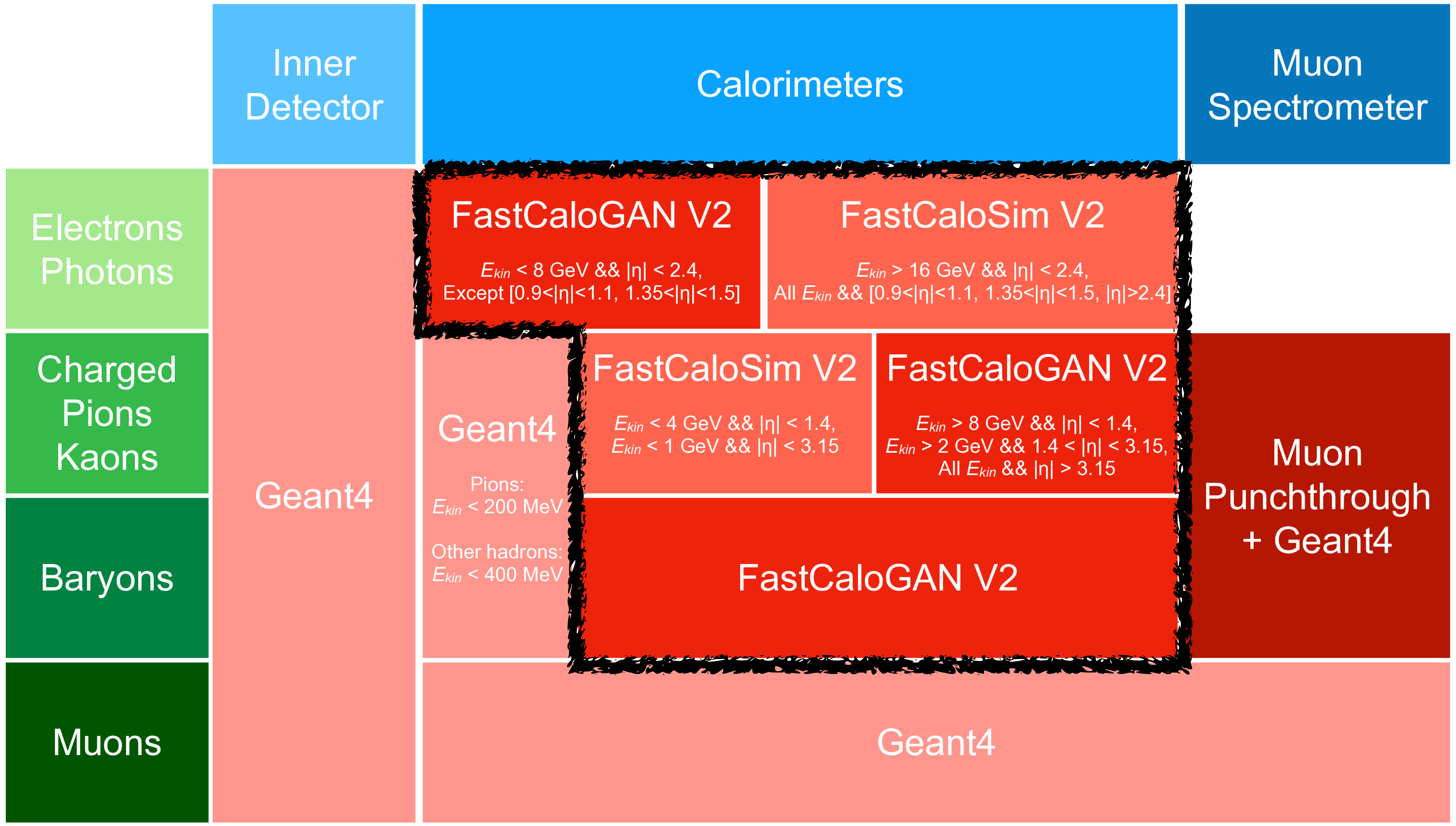



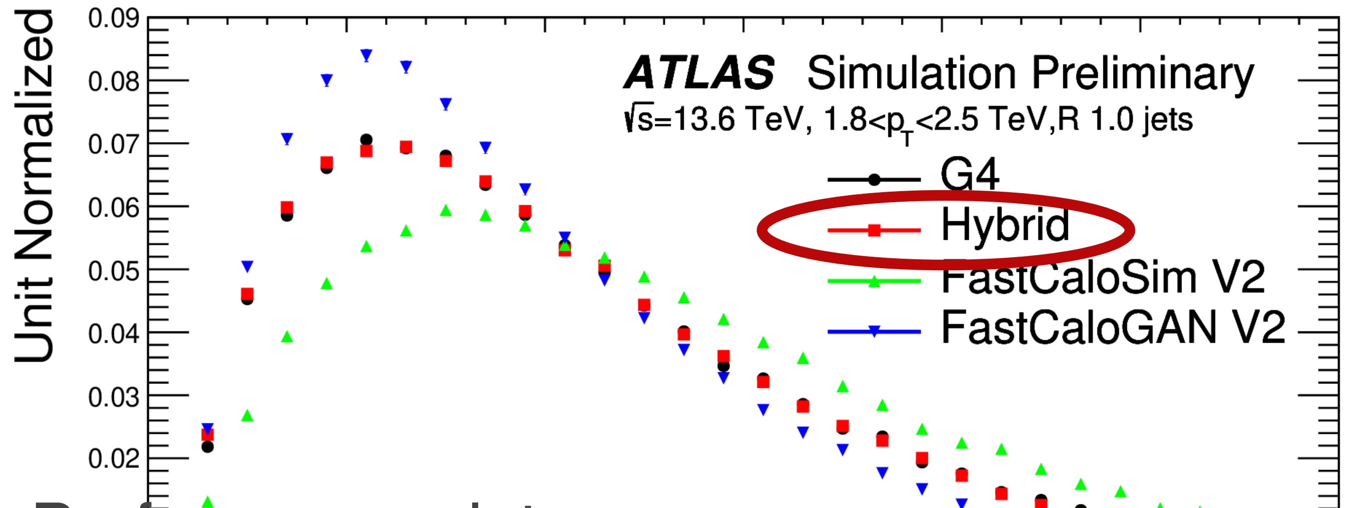
# AF3 Run3 version

- Improved over AF3 Run2 version
  - **FastCaloGAN V2** is better in increased phase space
  - Better **Muon Punch-Through** modelling

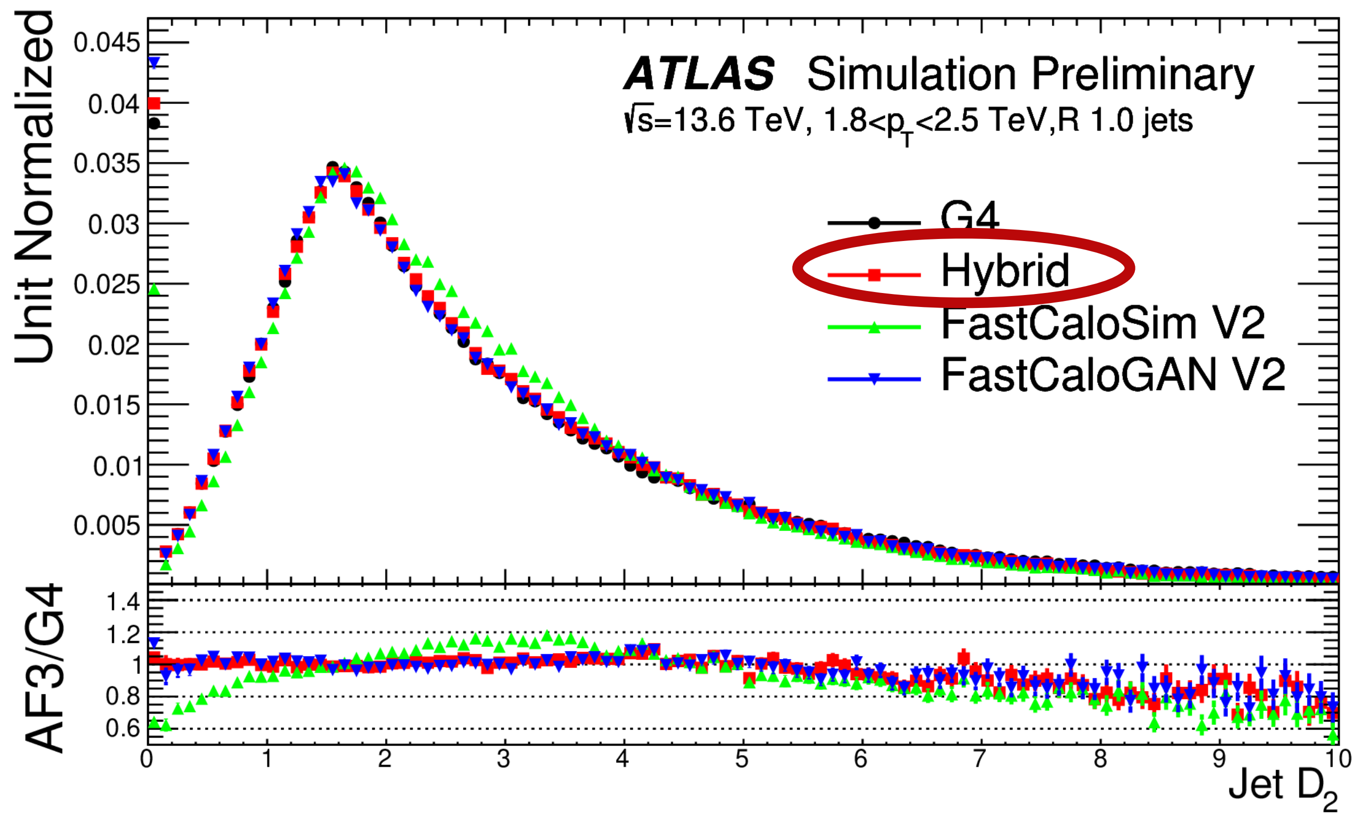
Note: Inner Detector, low energy hadrons, muons are simulated with Geant4

Run3 S&C paper



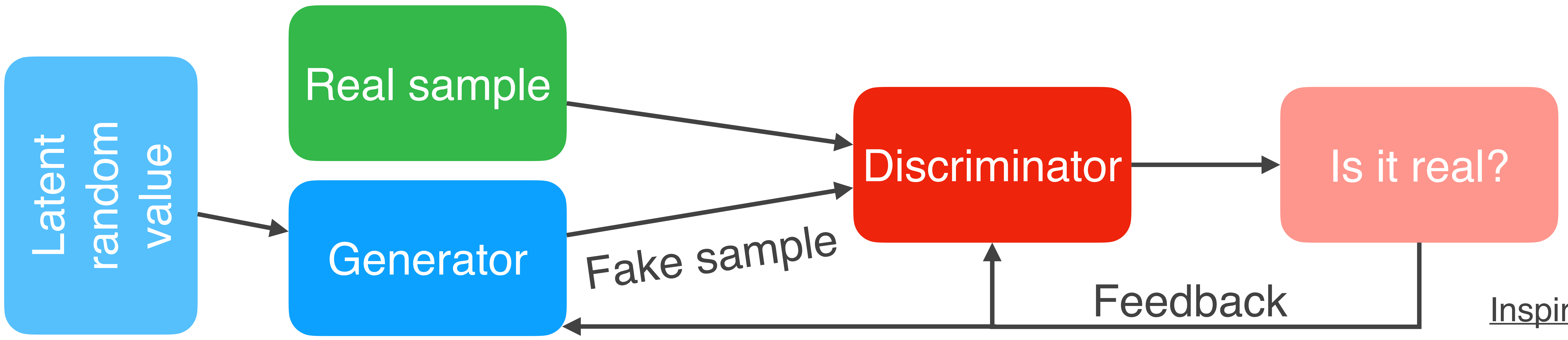




Performance plots



# What are GANs?

- Generative Adversarial Network (GAN) is a generative approach to generate new data that look like training data.
- Realised by adversarially training between a generator and discriminator. Once trained, generator is used for simulation

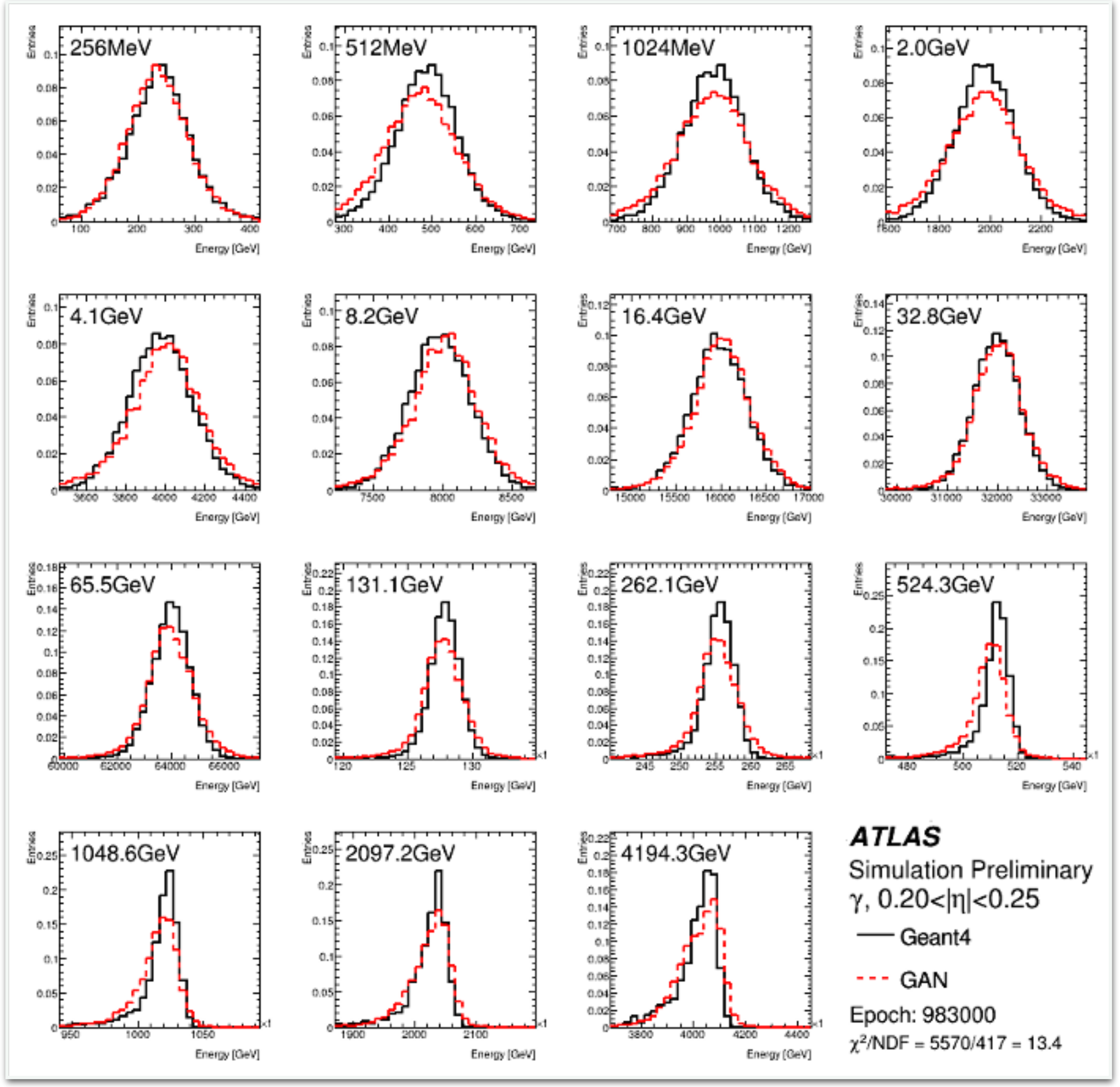



Inspiration

- FastCaloGAN is trained on Geant4 calorimeter shower data
- Due to very different behaviour, FastCaloGAN is designed and trained differently for different particle and |η|

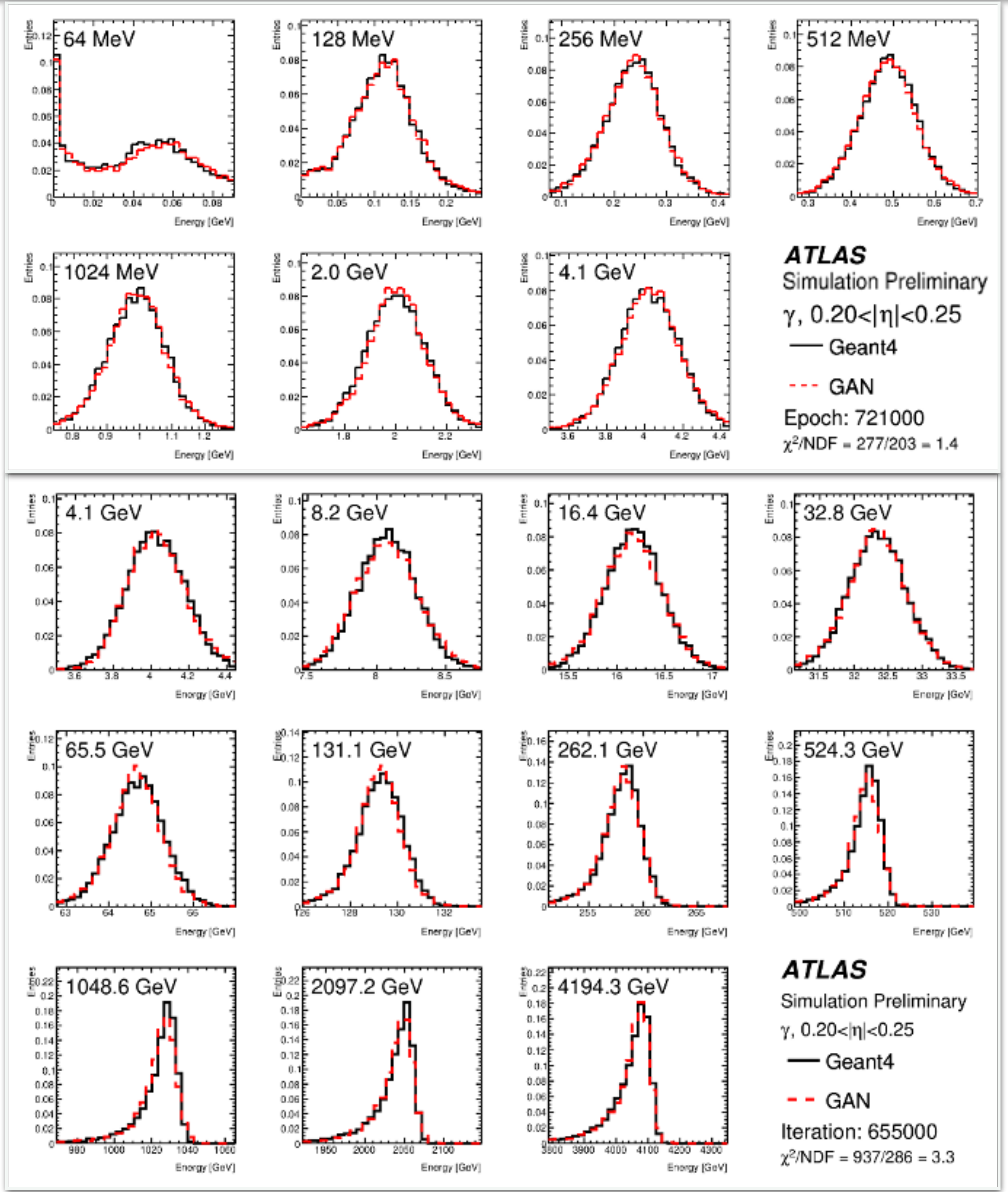
# Photon (γ) FastCaloGAN

V1




$\chi^2$/NDF = 12.3

V2

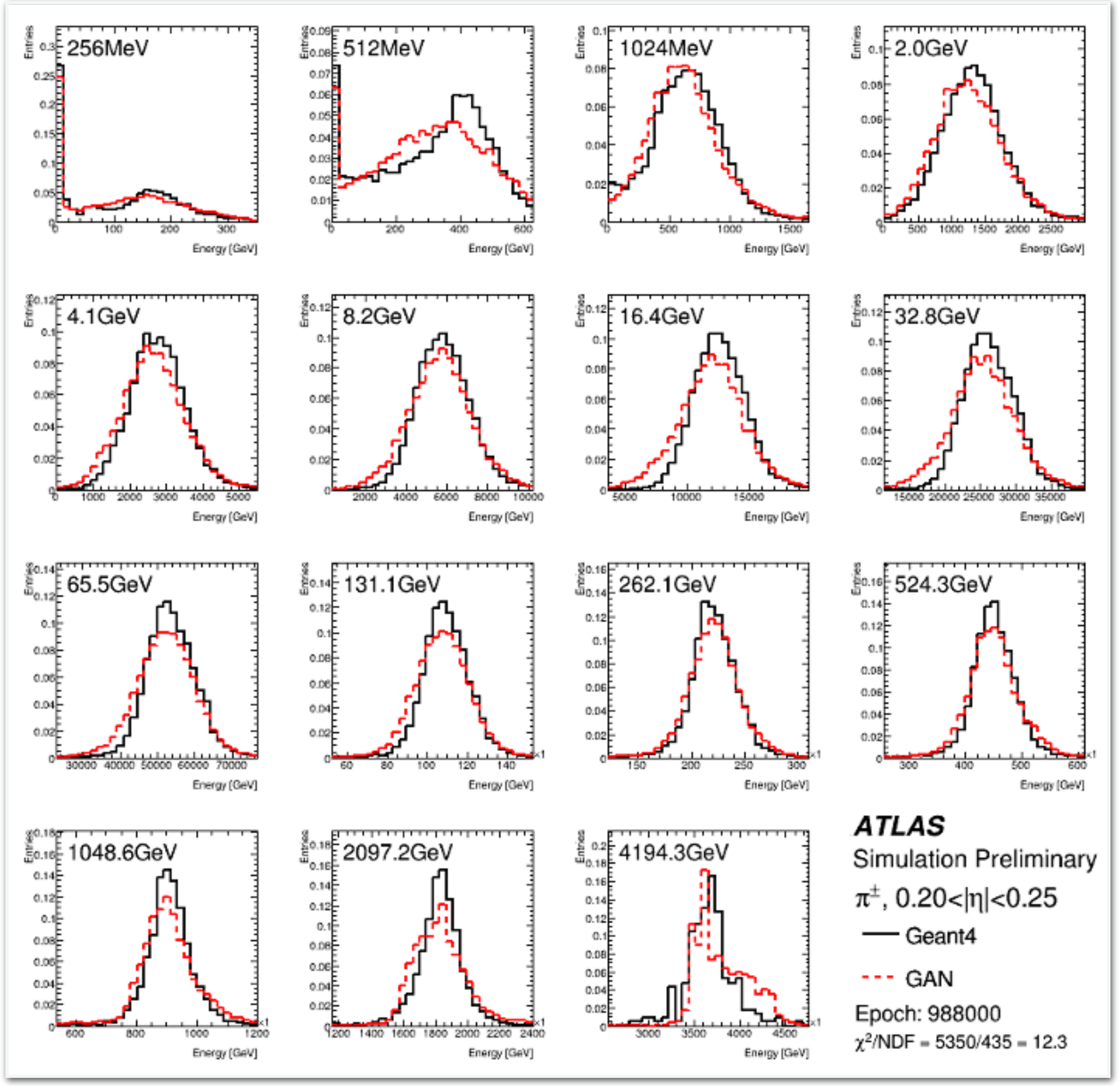



$\chi^2$/NDF = 2.5

Gaussian shape in low energy makes training easier.

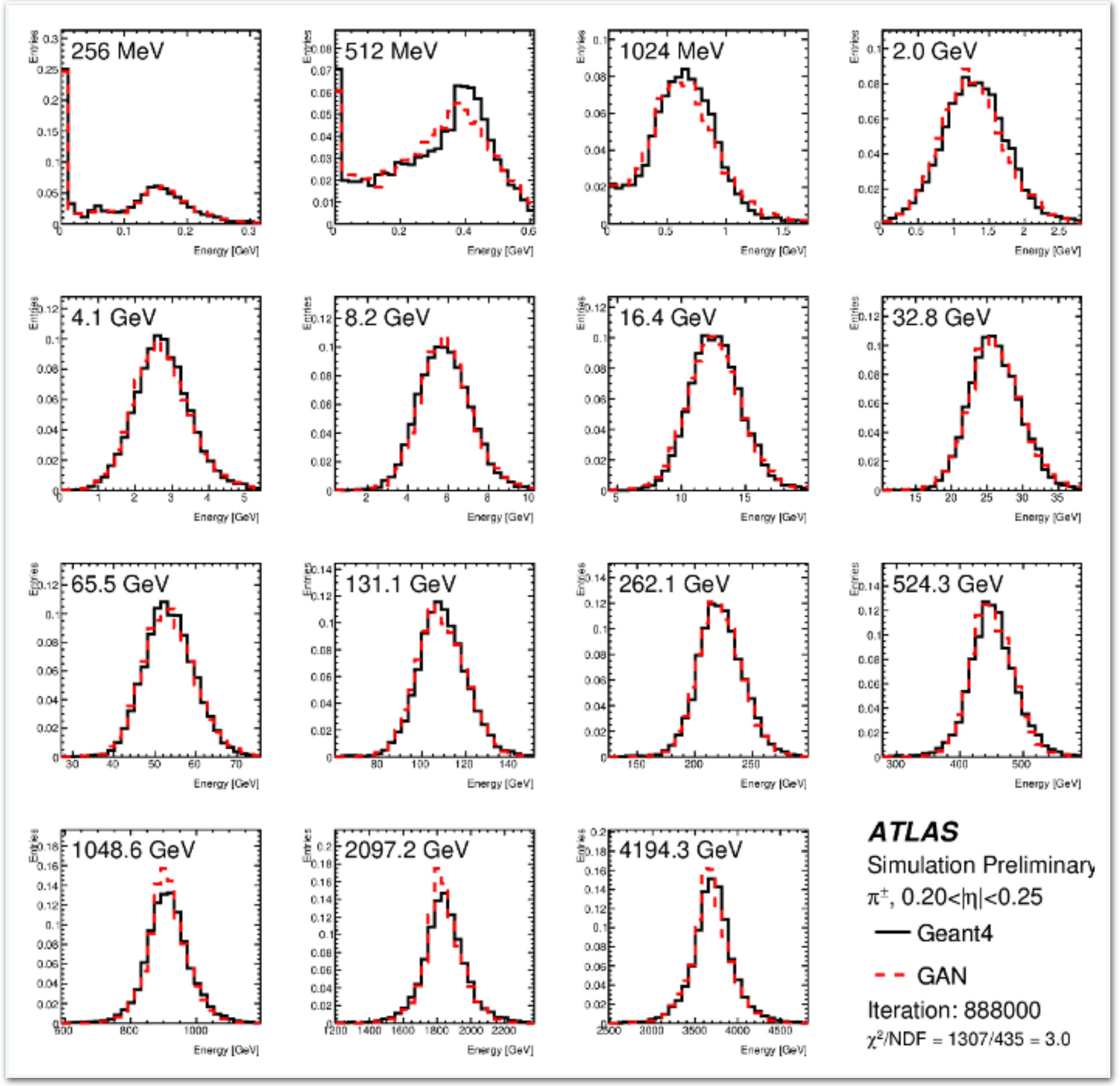
# Pion (π±) FastCaloGAN

V1




$\chi^2$/NDF = 12.3

V2




$\chi^2$/NDF = 3.0

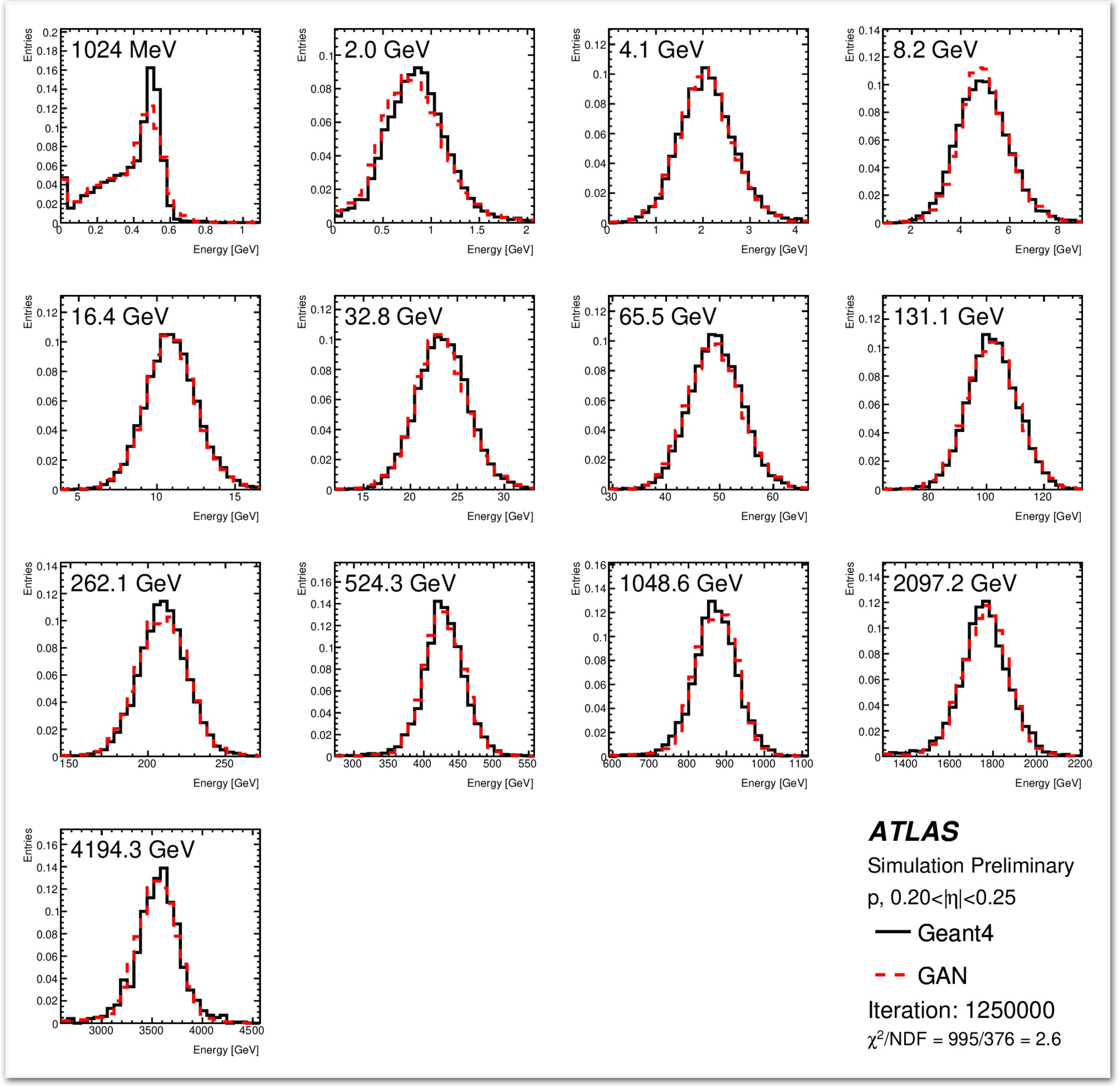
Significant improvement for charged pions

# Proton (p) FastCaloGAN

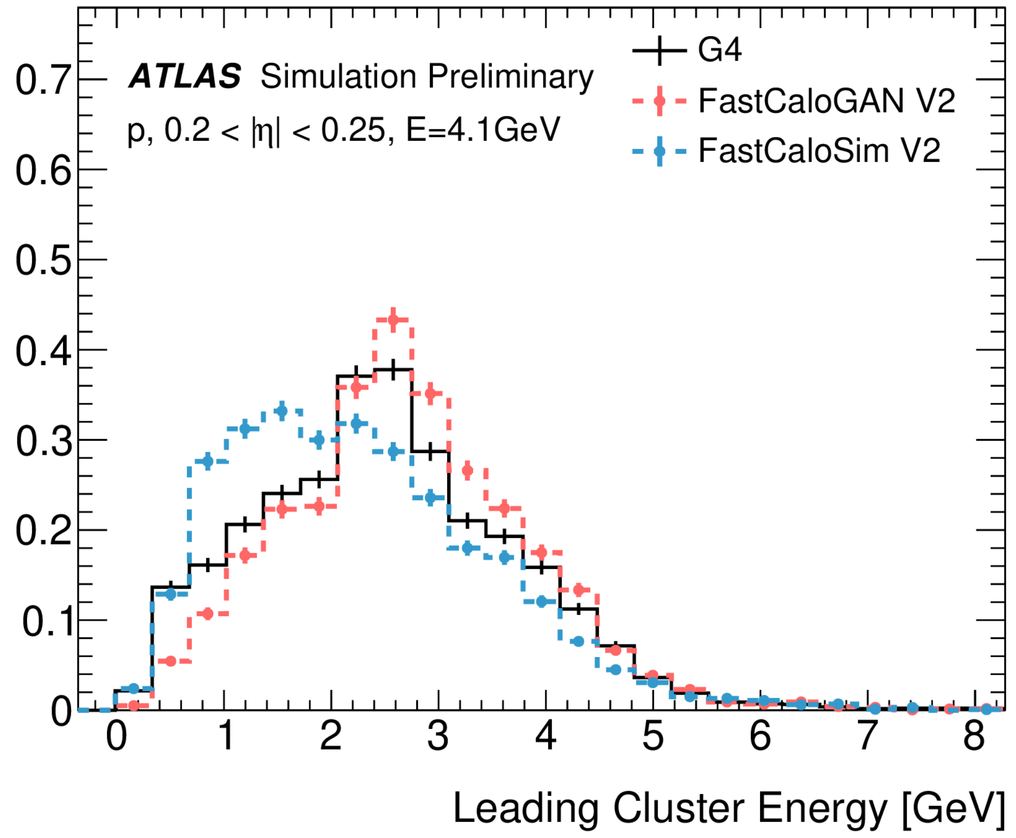
V1

$\emptyset$

V2



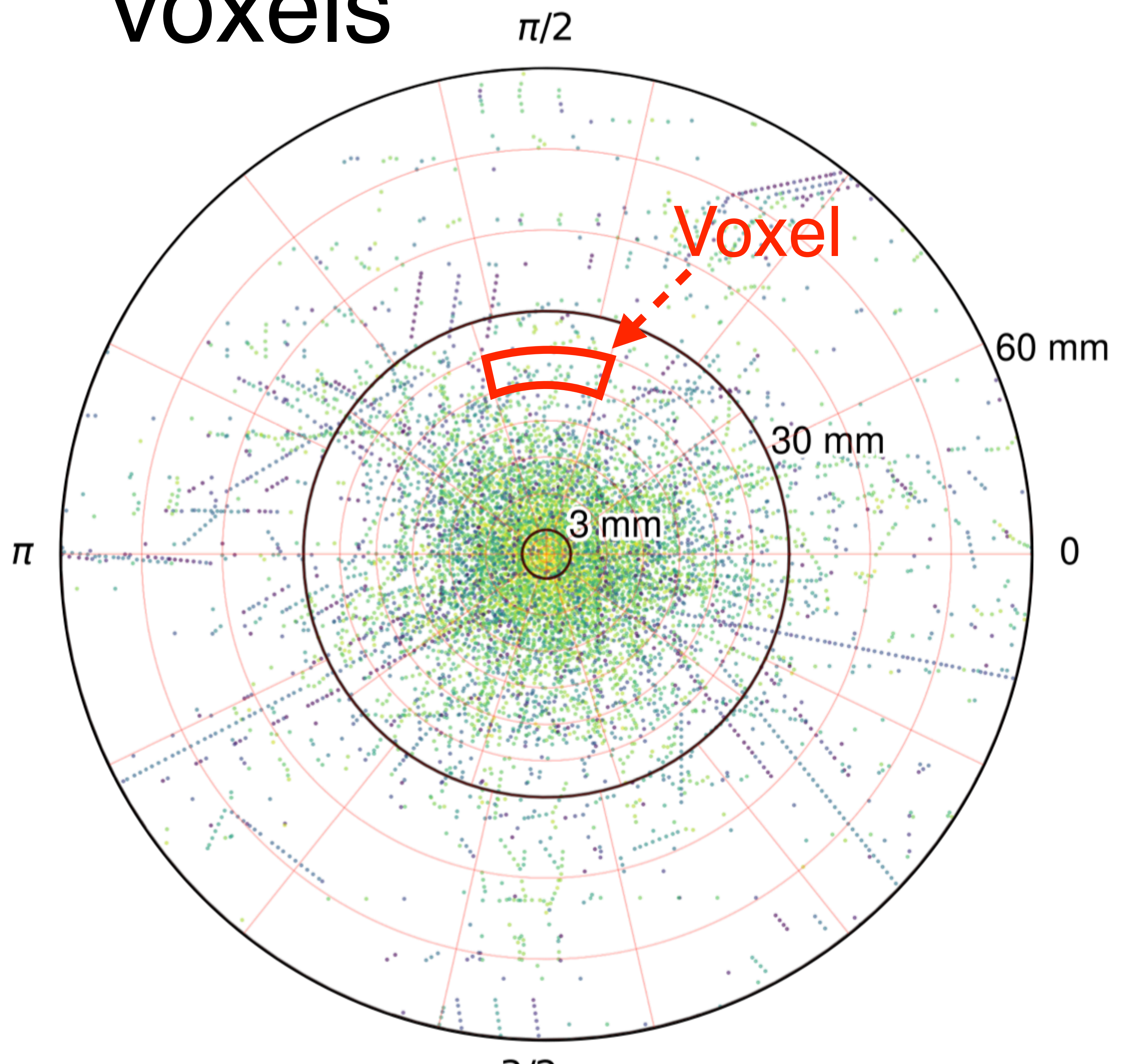

Dedicated GANs for protons: not possible to implement in FastCaloSim approach due to memory limitation



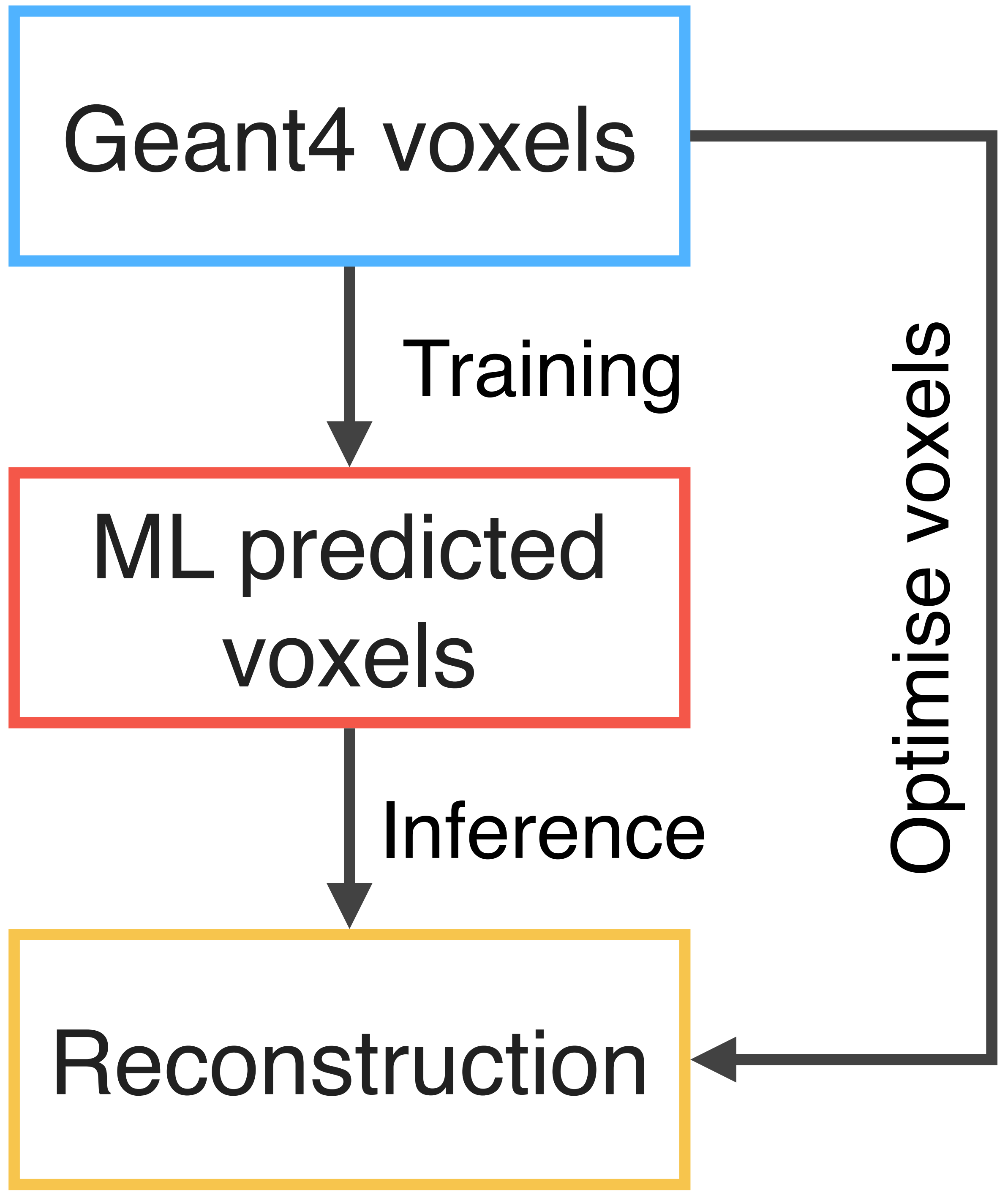
# Preparing for AF4 and Run4

## Improve ML input -- voxels

- Balancing voxel granularity
- Optimised by simulating directly on Geant4 voxels

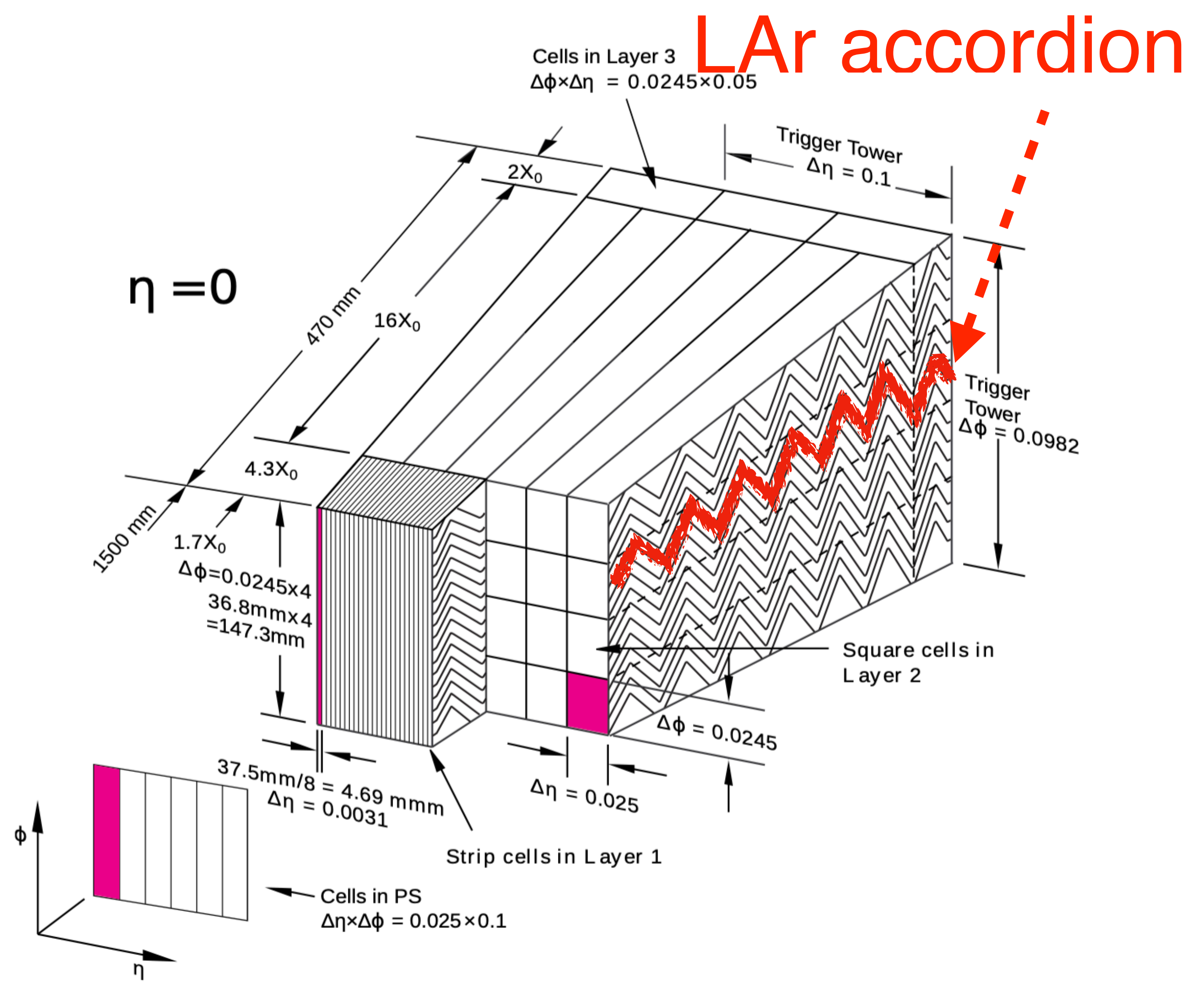



***ATLAS*** Simulation Preliminary
Single shower, Photons,
$E_{\text{kin}} = 65GeV$, $0.20 < |\eta_{\text{center}}| < 0.25$




## Remove modulation

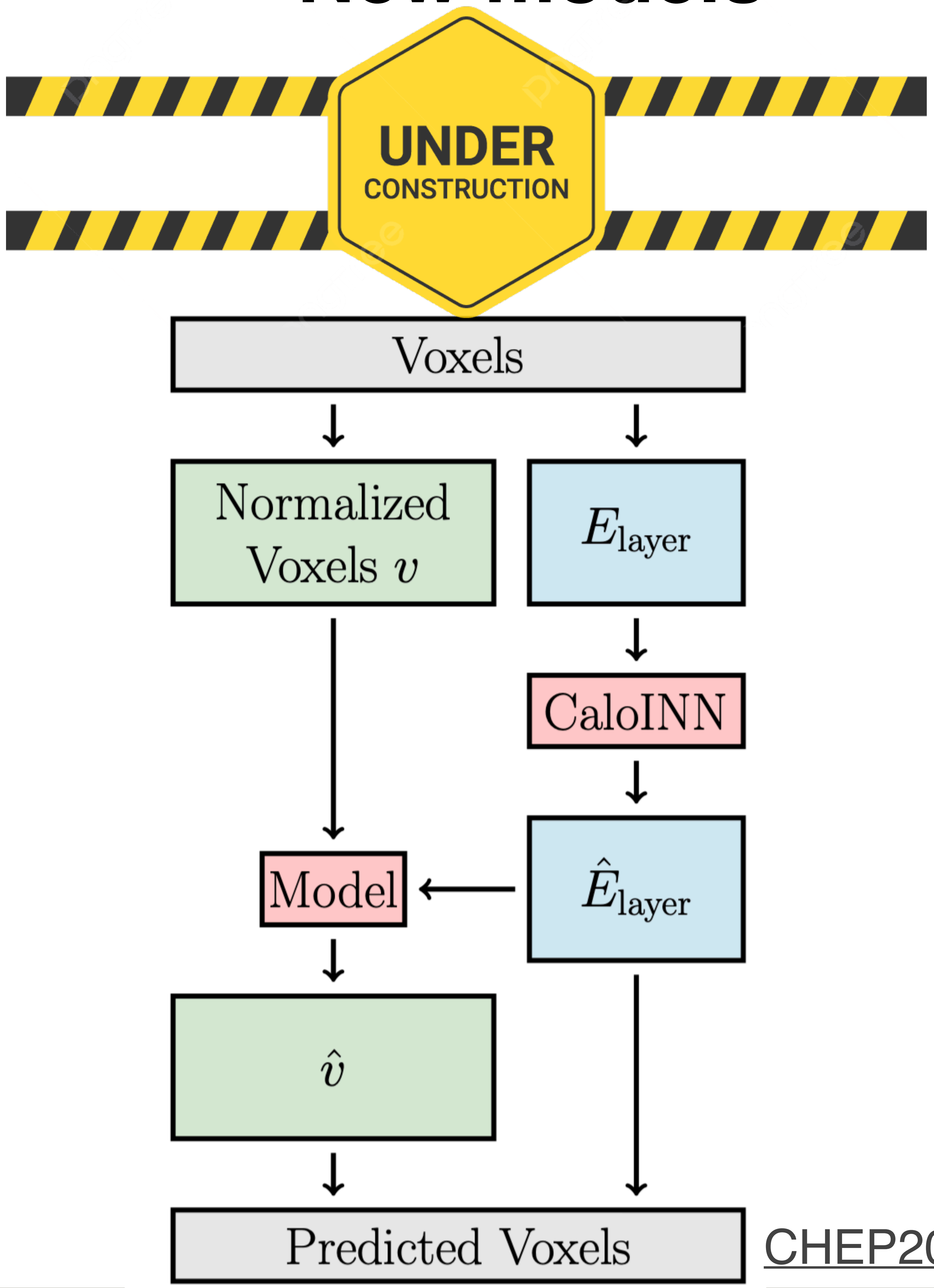
- Liquid Argon accordion shape breaks calo's uniformity
- Effect taken in Geant4 voxels, and removed from ML training




**Data is public as CERN Open Data**

# Preparing for AF4 and Run4

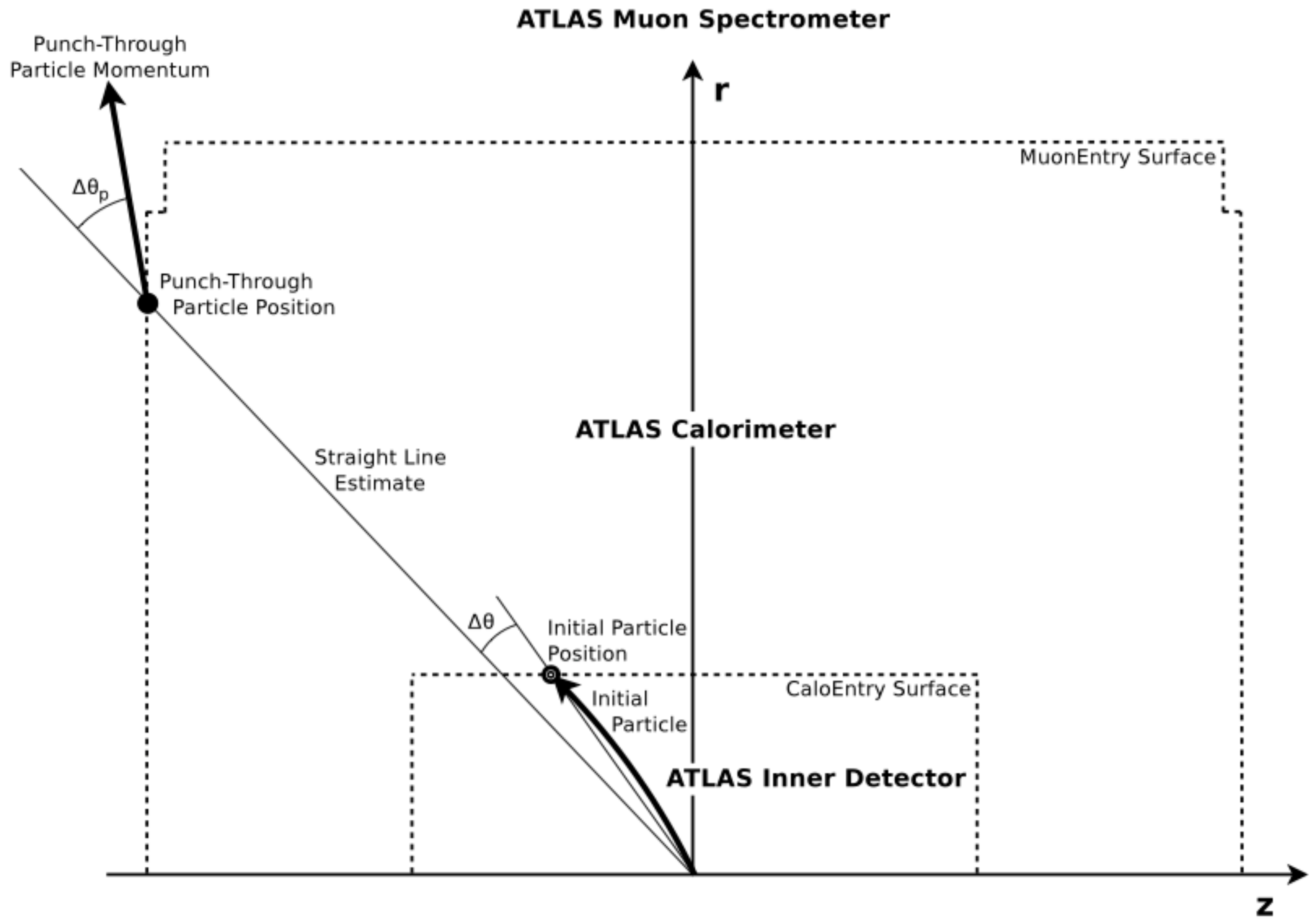
## New models




CHEP2026

- Move to flow and diffusion
- Separate models for layer energies and layer shape prediction

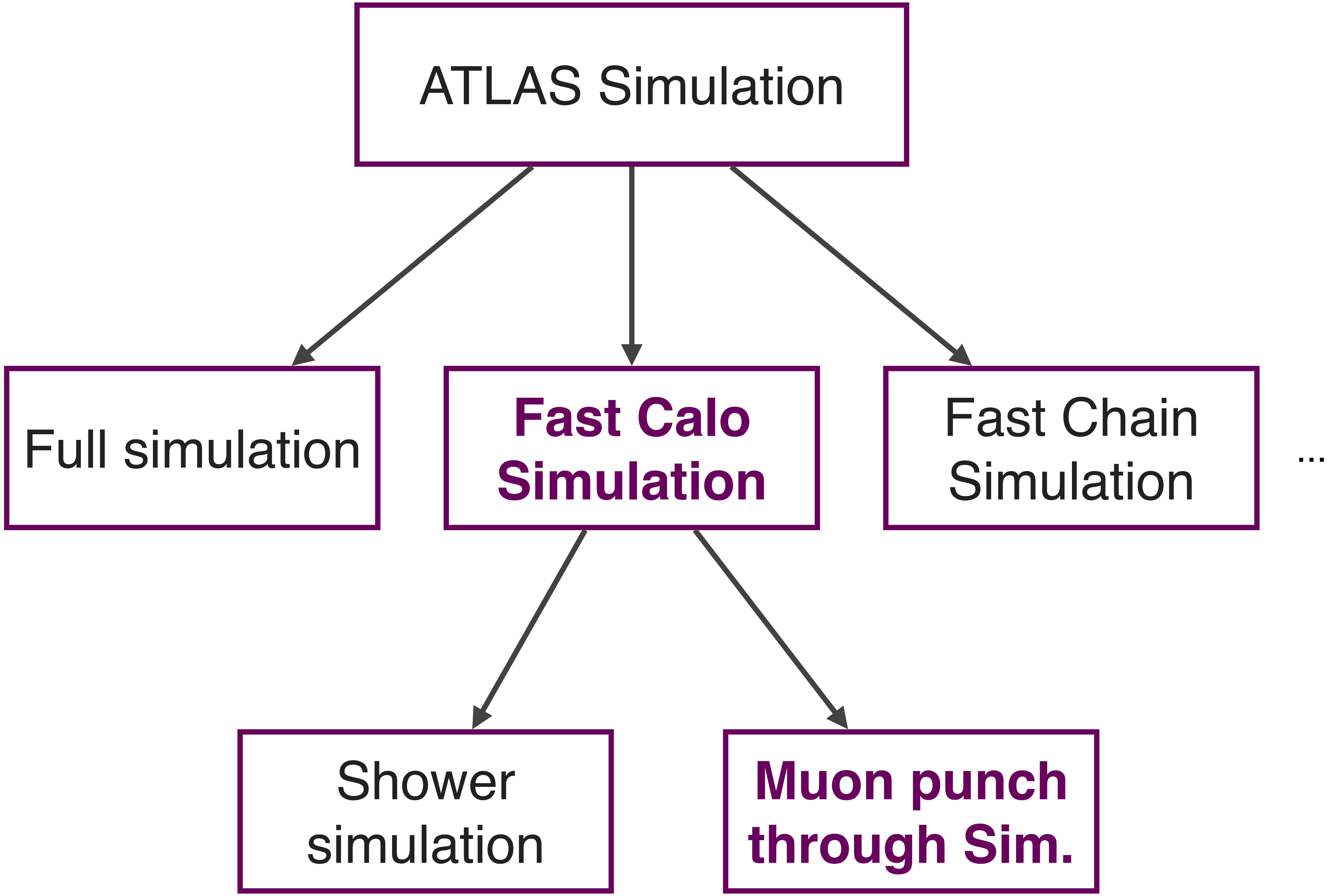
## New muon punch through




- Use better network than NN on slide 15 to predict true distribution

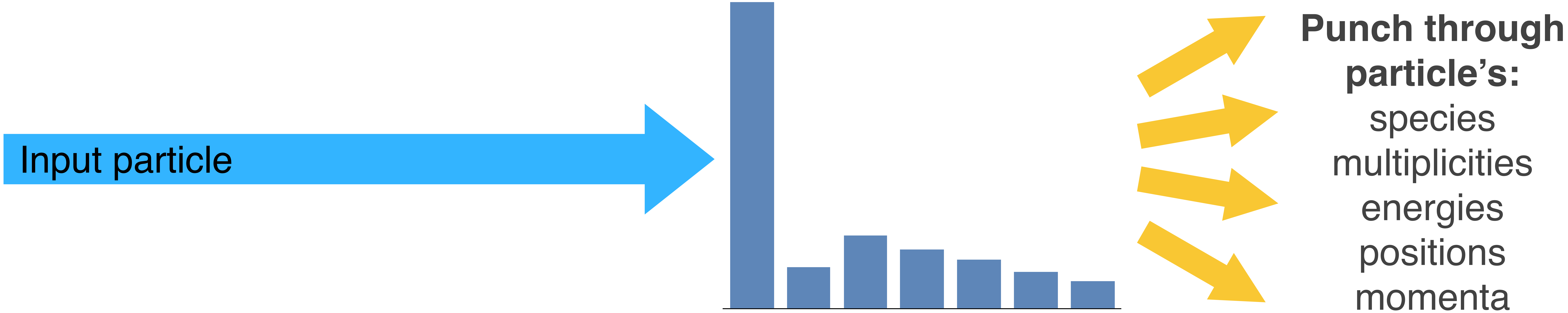
CHEP2026

# Contents


ATLAS Simulation
Full simulation
Fast Calo Simulation
Fast Chain Simulation
...
Shower simulation
Muon punch through Sim.

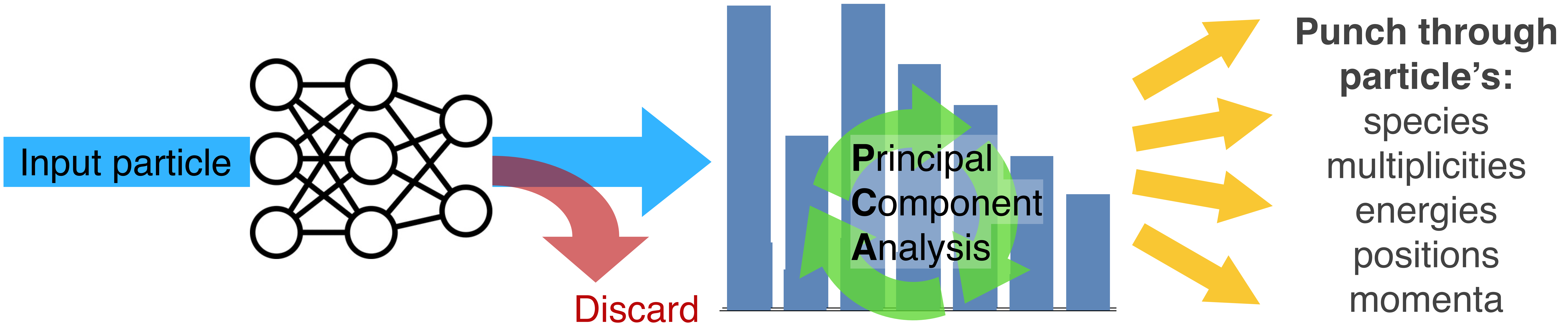
# Muon Punch-Through in AF3 Run2

- (Secondary) Particles escaping calorimeter can generate hits in muon detector: punch-through effect
- Fast simulation generates energy deposits but not create secondary particles.
- In AF3 Run2 done by look-up histograms according to initial particle's momentum and |η|
- Produce punch-through particles' (e, μ, γ, π, p) multiplicities, energies, positions and momenta
- Missing info on initial particle's shower: given momentum and |η|, if deposited energy is large, chance to punch through should be lower

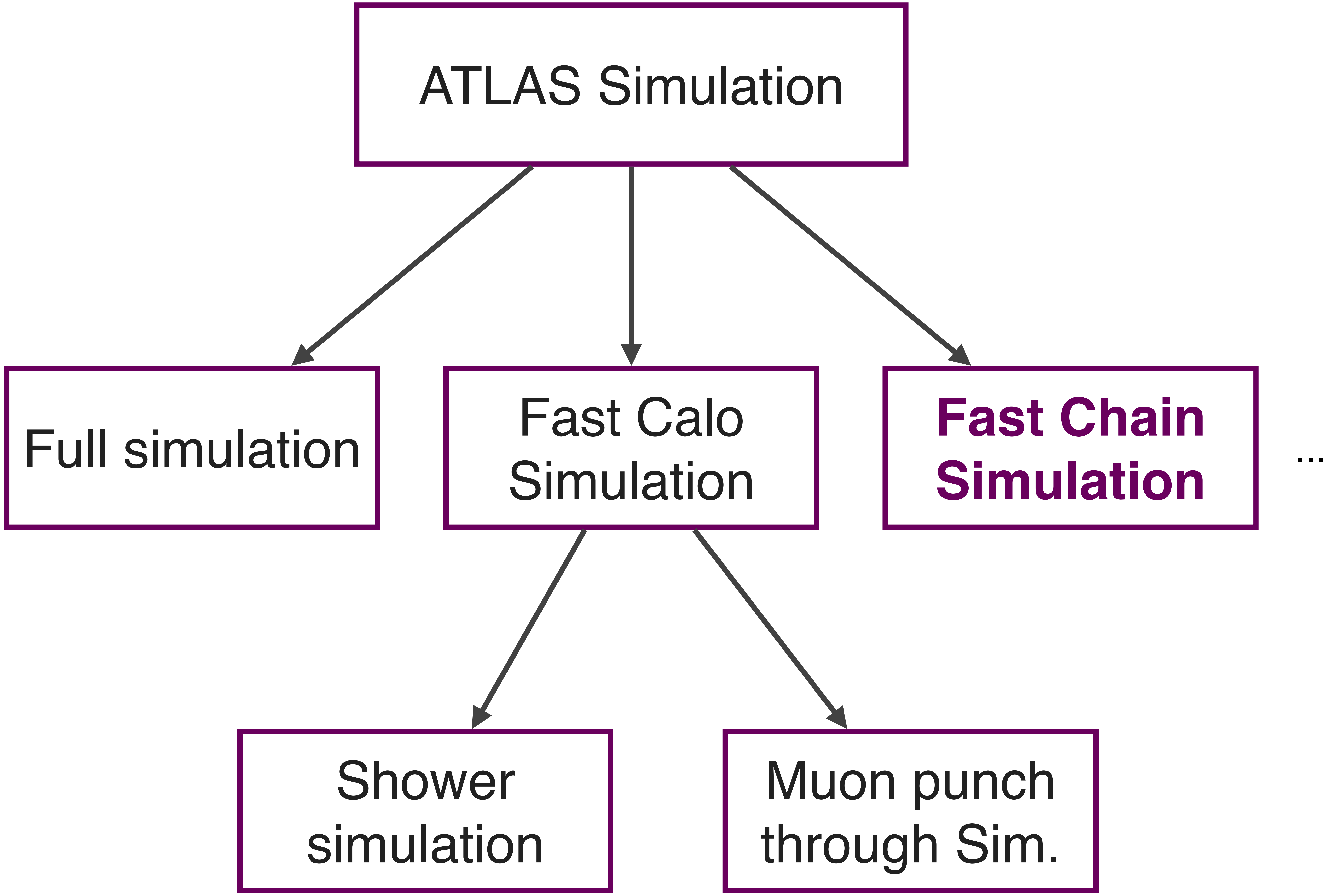


# Muon Punch-Through in AF3 Run3

- In AF3 (MC23), a ML model is used to predict the probability of a primary particle to have a reconstructed muon segment in Geant4 simulation.
- A three-layer neural network to discover such primary particle, **based on its shower info**
- Draw punch-through particle (e, μ, γ, π, p, K) multiplicities from look-up histograms
- Simulate 5-dim kinematics of the punch-through particle: energy, exist position (θ, ϕ) and momentum direction (θ, ϕ); use PCAs for better modelling
- Independent PCAs for different initial pion energy-|η|, different punch-through particles (p/ n v.s. others), inside/outside crack

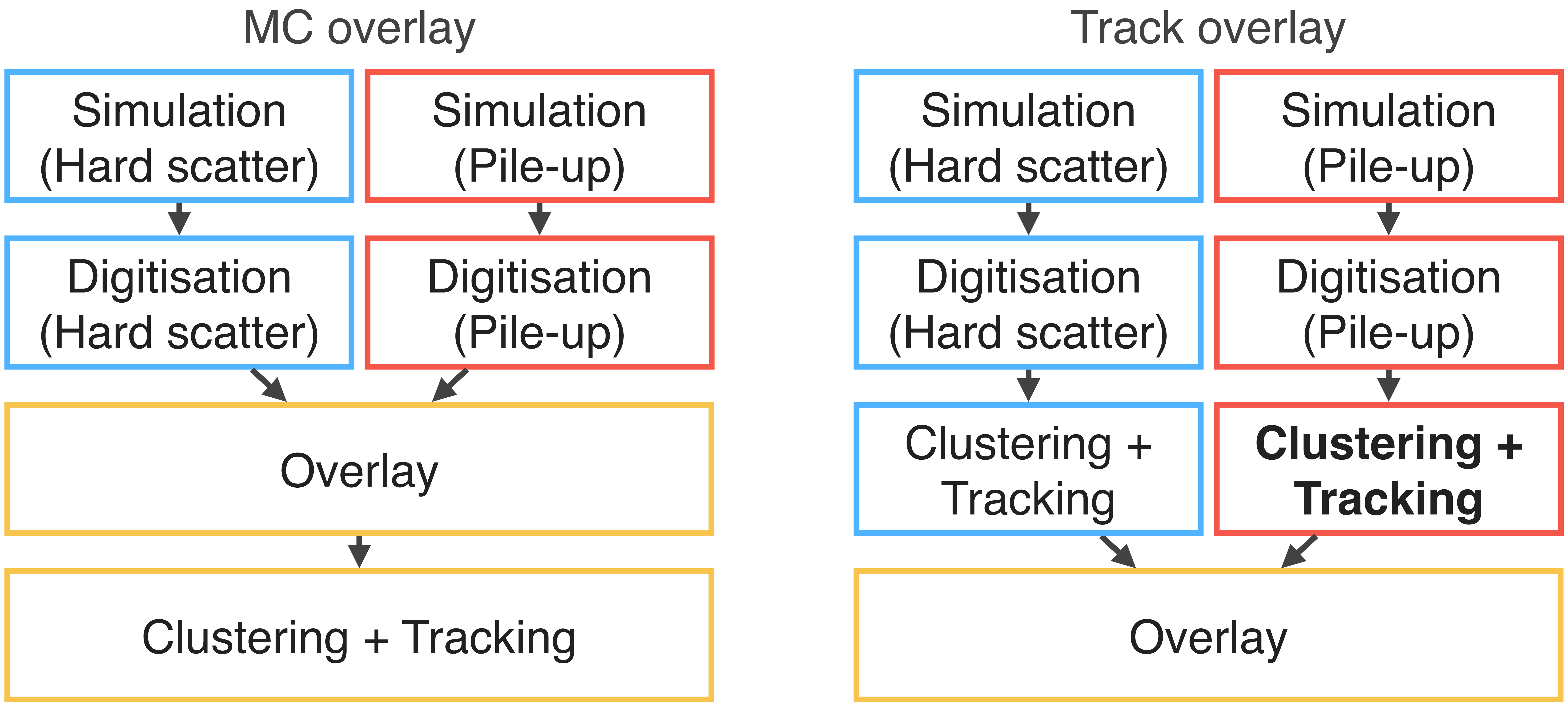


# Contents


ATLAS Simulation
Full simulation
Fast Calo Simulation
Fast Chain Simulation
...
Shower simulation
Muon punch through Sim.

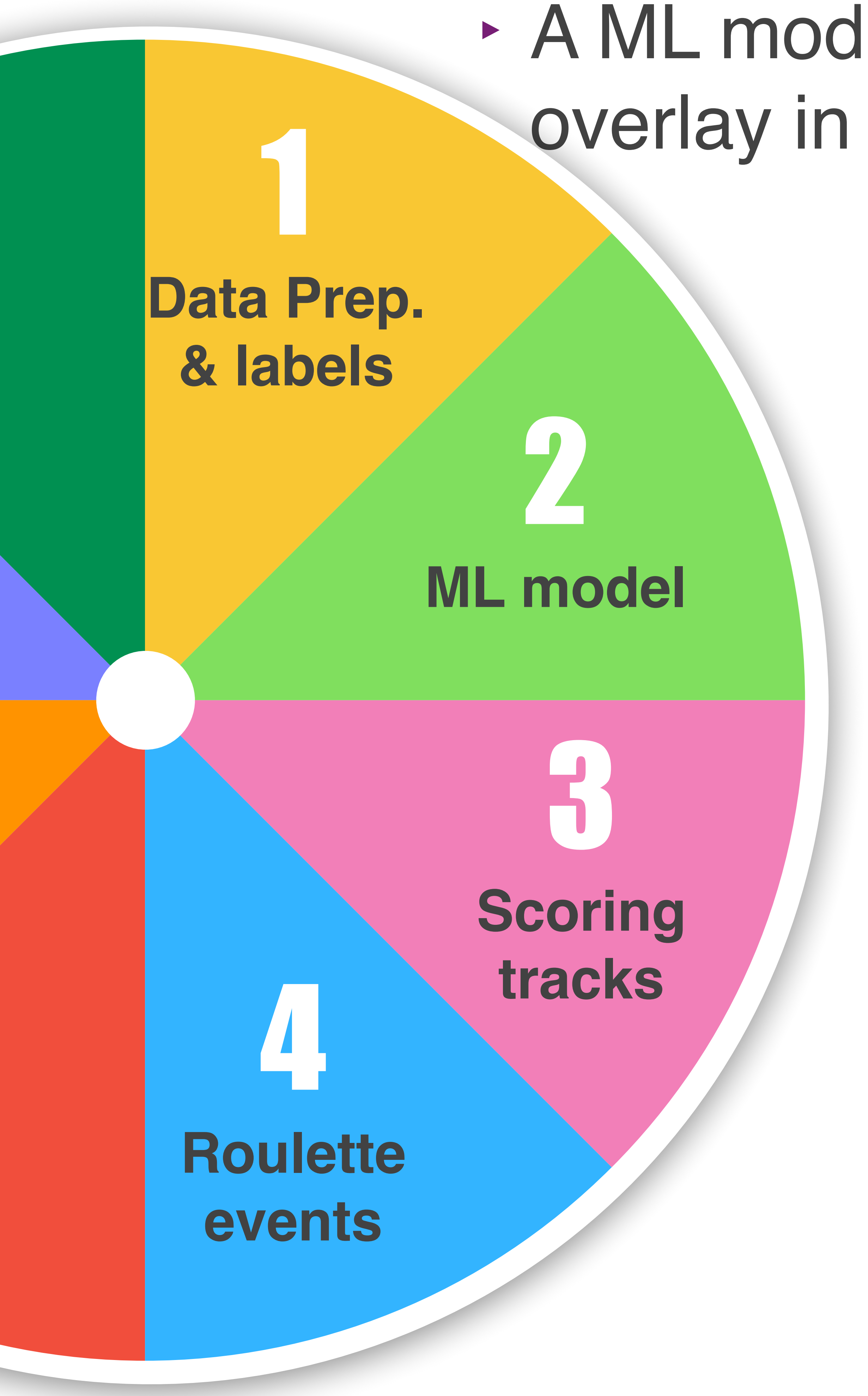
# Track overlay

- Track-overlay is a fast alternative to MC overlay
  - Overlay **pre-reconstructed pile-up tracks** to mitigate the computational overhead in reconstruction of pile-up tracks



# Roulette ML model in track overlay

- Track overlay performance is affected by event kinematics
  - A ML model is trained to label events in dense environment that need to be sent to MC-overlay in order to maintain MC-overlay accuracy




**1.** Input: truth particle info (px, py, pz, E, pT), density (R02, sumpTR02, …) and conditions (number of PU, TruthMultiplicity, event pT).
Label: “Bad track” if the track has a truth match in the track overlay but not in the MC overlay.

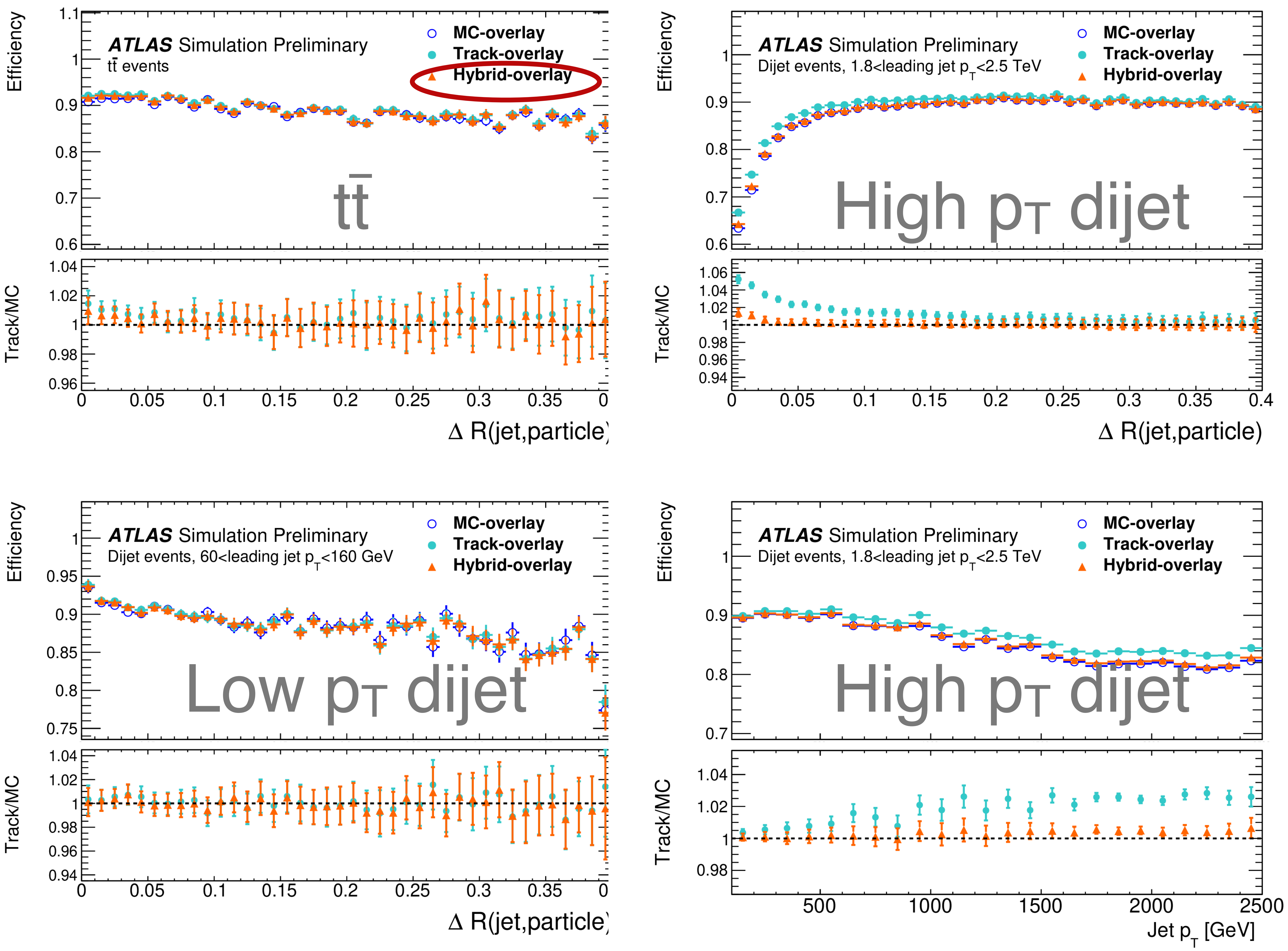
**2.** Train a supervised ML model using tracks in high pT jets (JZ7). Use this model to predict the “Badness” of the track.

**3.** A track receives a high score (> 0.742) will be considered as a “bad track”.

**4.** Calculate Roulette score for each event: N_Bad tracks / N_total tracks. Events with Roulette score = 0 will perform track overlay, while those with > 0 will perform MC overlay.

# Improvement from ML

- Reconstruction efficiency agreement with MC overlay improves when using Roulette ML (hybrid overlay)




Event fractions for track overlay

- Low $p_T$ dijet: 93.5%
- High $p_T$ dijet: 35.3%
- $t\bar{t}$: 90.15%

# Summary

- Simulation is a great field for ML applications
  - Suitable for both supervised (classification) and unsupervised (generative) applications
  - Rich information due to low-level inputs (hits, cells, …), more than just 4-momentum in a stable particle
- A few selected mature examples presented; more ideas in pipeline to test, implement, and verify
  - For example in recent #CaloChallenge a lot of cutting-edge ML models were proposed to simulate calorimeter showers
  - We are far from reaching the limit …