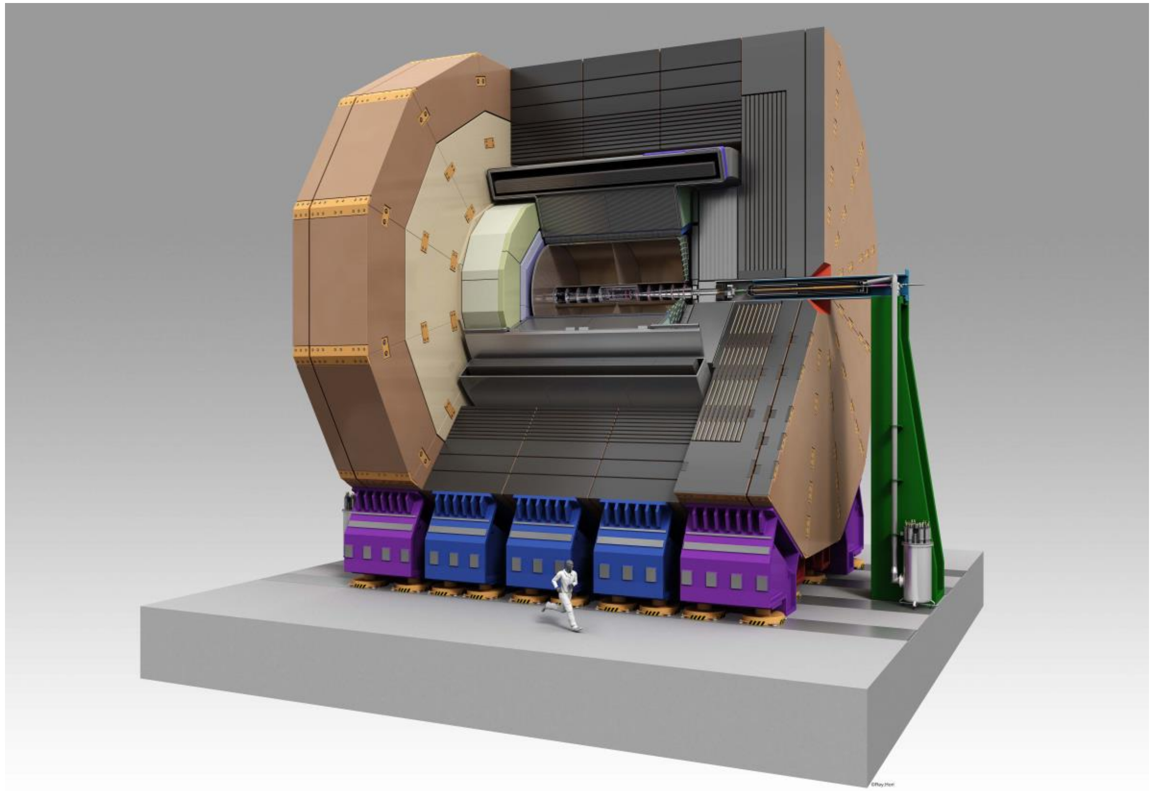
Fast Simulation of High Granularity Calorimeters with Deep Generative Models

Peter McKeown
Deutsches Elektronen-Synchrotron
16.03.2021

Outline

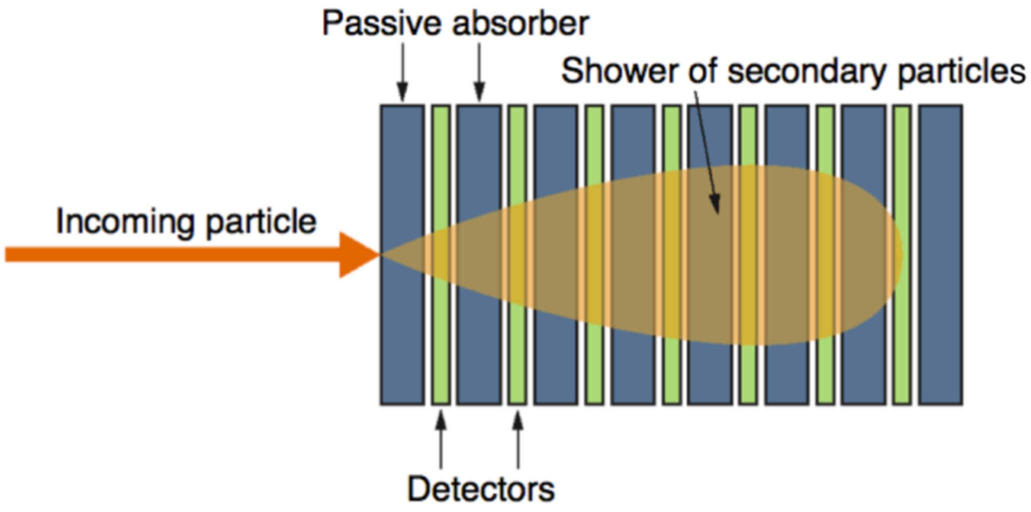
The ILD Concept

- International Large Detector (ILD) concept for the International Linear Collider (ILC)
    - Higgs Factory
- Optimized for Particle Flow
    - Reconstruct each individual particle in subdetector
    - Obtain optimal detector resolution
- Key features:
    - Precise tracking and vertexing
    - Excellent hermeticity
    - High granularity calorimeters



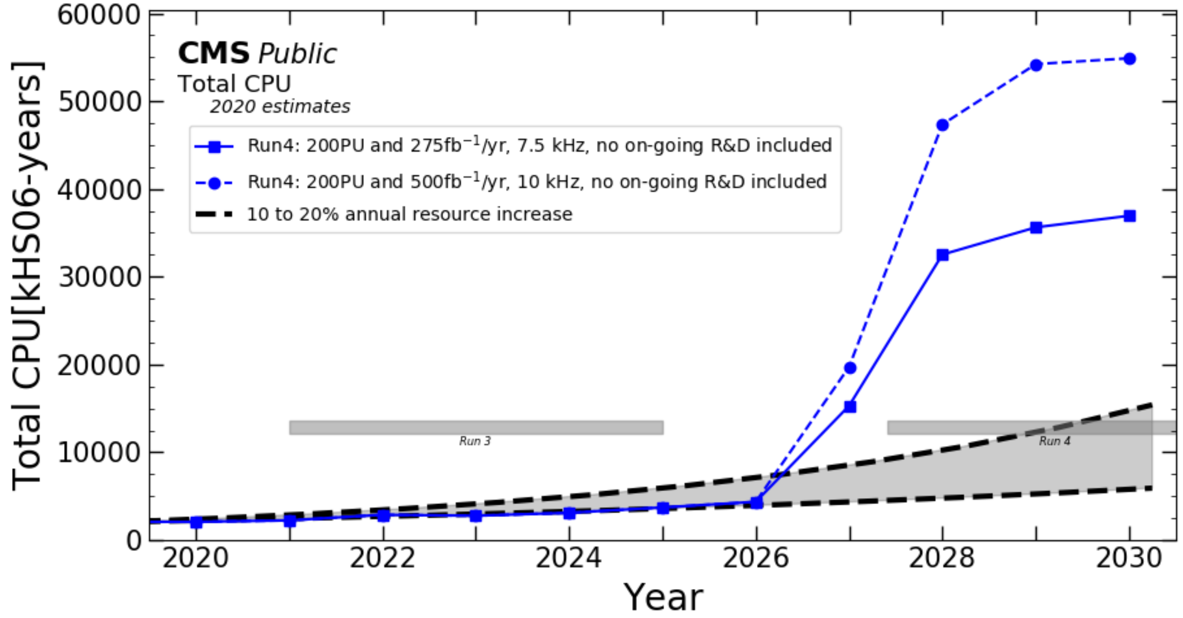
The ILD Electromagnetic Calorimeter

- Destructively measure particle’s energy
  - Produce shower of secondary particles until totally absorbed
- Sampling calorimeter- measure fraction of energy
- ILD proposal: Si-W ECAL
  - 30 layers of alternating active and passive material
  - 20 layers: 2.1 mm thick
  - 10 layers 4.2 mm thick
  - 5x5 mm^2 silicon cells
  - ~ 80 million channels for barrel
- Geometry is not perfectly regular… ⟶ **Project onto regular grid**

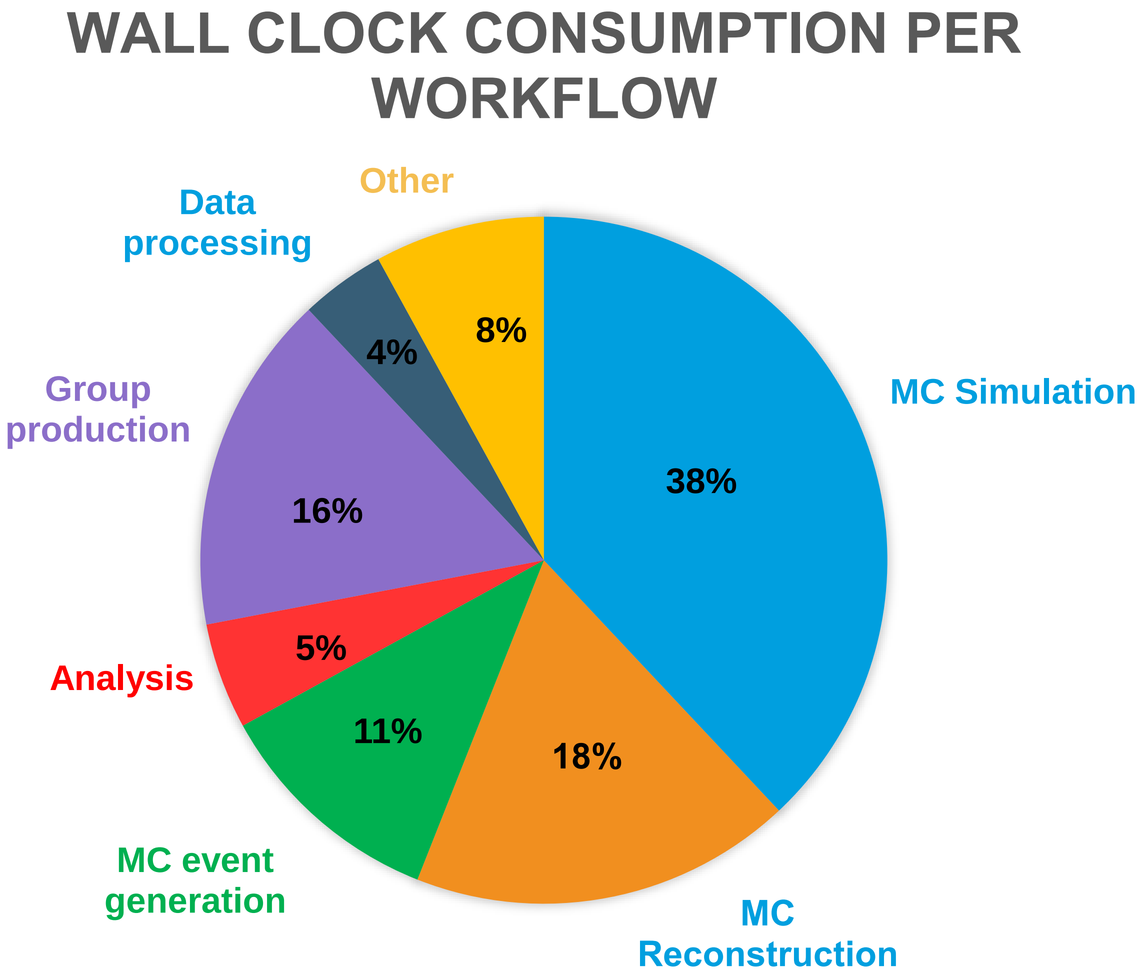


The Strain on HEP Computing Resources

- Ever increasing demand for computing resources
  - MC generation largest fraction
  - Calorimeters most intensive part of detector simulation



https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults

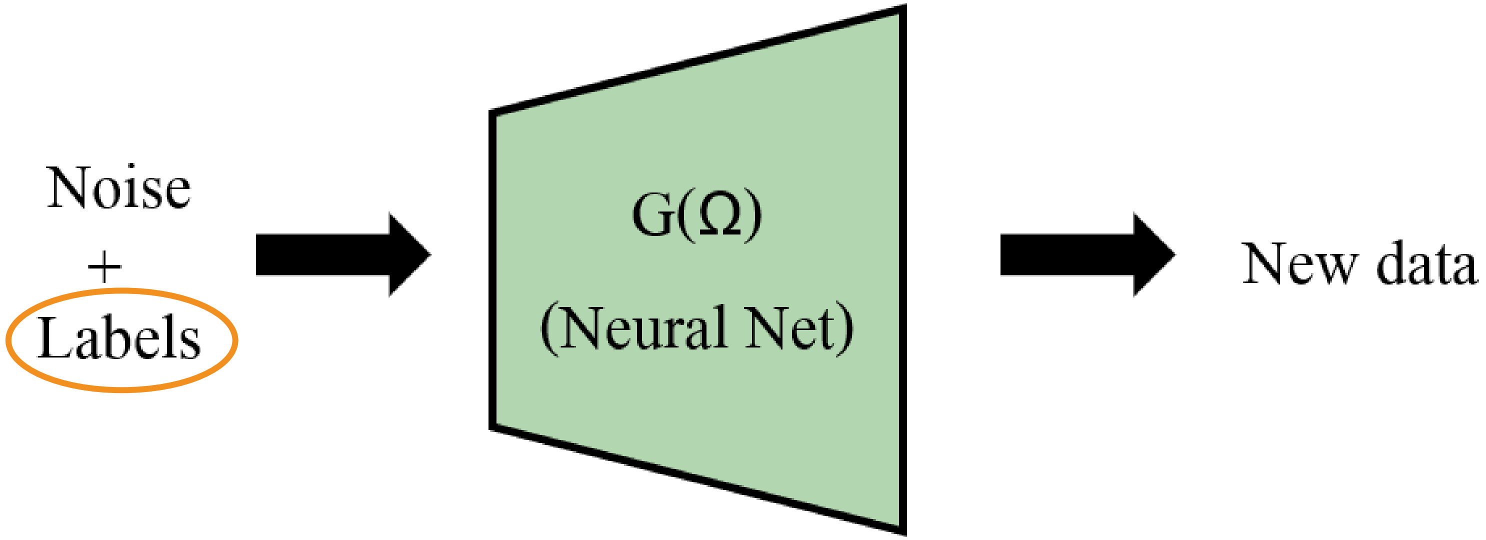


D. Costanzo, J. Catmore, ATLAS Computing update, LHCC meeting , 2019

Generative Models

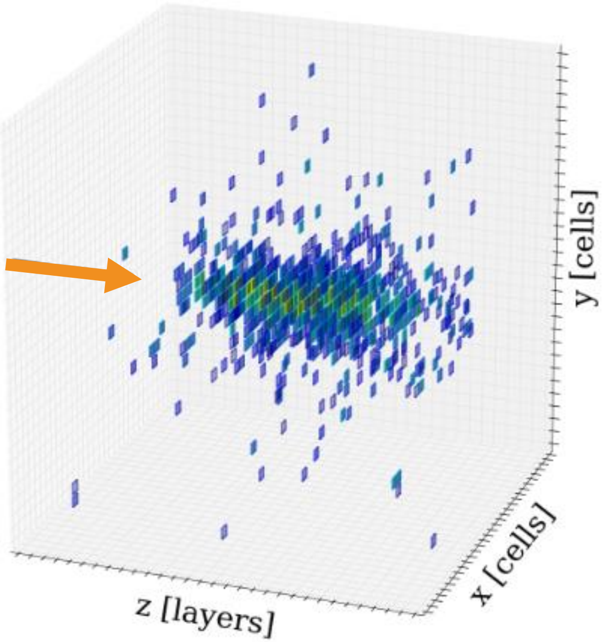
- Approach to fast simulation- amplify statistics
- Promising solution: **generative models**
  - Generate new samples following the distribution of original data
  - Map random noise to data
  - Highly parallelizable
  - Conditioning

See preceding talk by Sascha outlining proof of principle:
T 38.8- GANplifying Event Samples
arXiv:2008.06545



Getting High: Simulating the ILD ECAL with Deep Learning

- Previous work in our group: ILD ECAL simulation with 3 generative models
  - 30x30x30 regular grid of calorimeter cells
  - 950k photon showers
  - Continuous range of energies: 10-100 GeV
  - Fixed incident point
  - **Perpendicular to the calorimeter face**

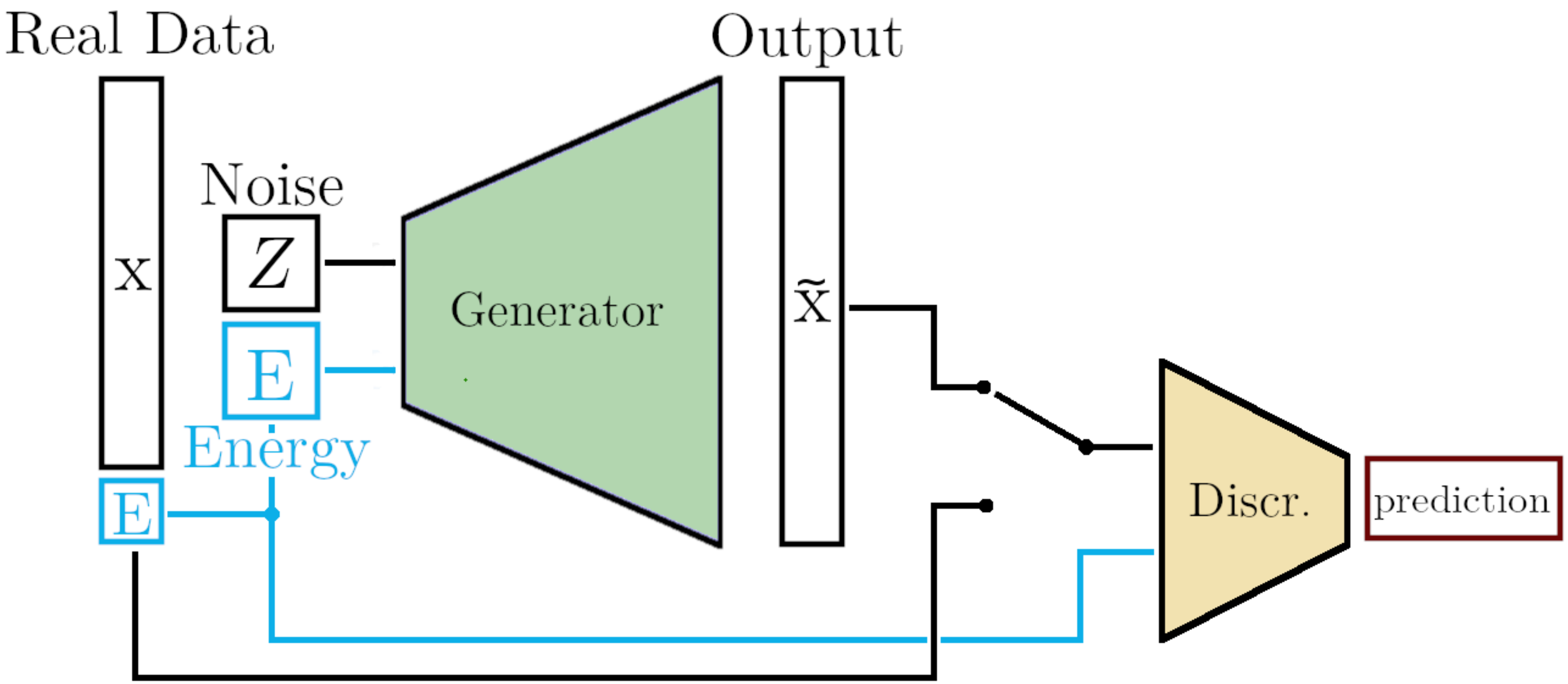


Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed

Erik Buhmann · Sascha Diefenbacher · Engin Eren · Frank Gaede · Gregor Kasieczka · Anatolii Korol · Katja Krüger

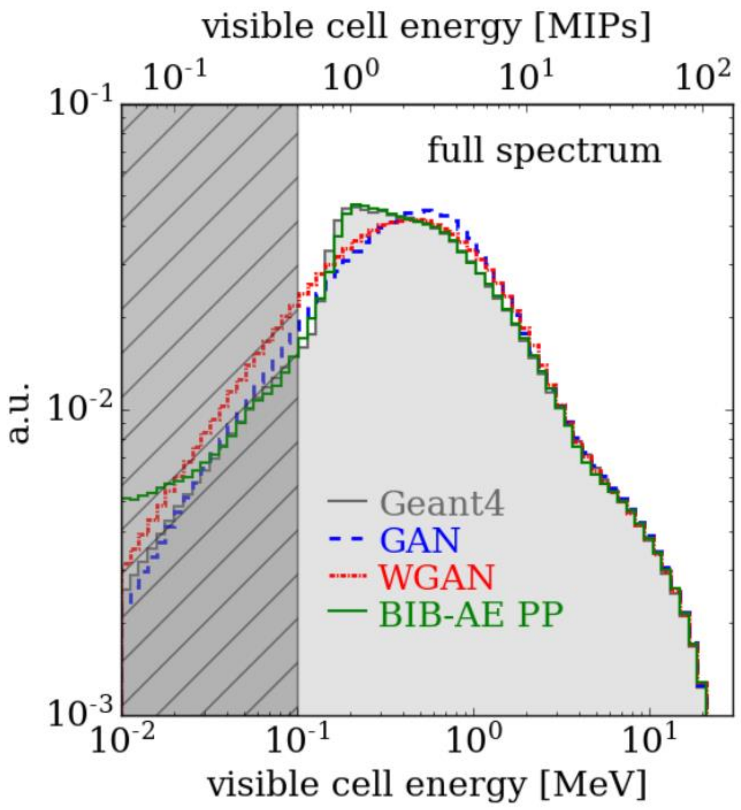
arXiv:2005.05334

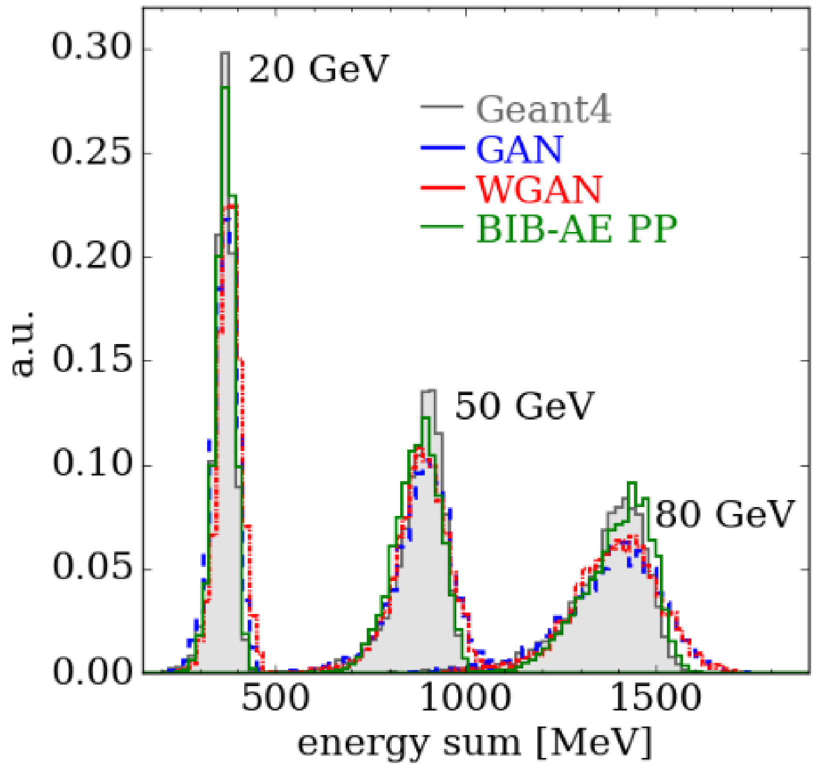
Getting High: Generative Adversarial Network



* For details on the other architectures see: T21.5 from Erik (Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network, arXiv:2102.12491) and T57.1 from Sascha

Getting High: High Fidelity Simulation



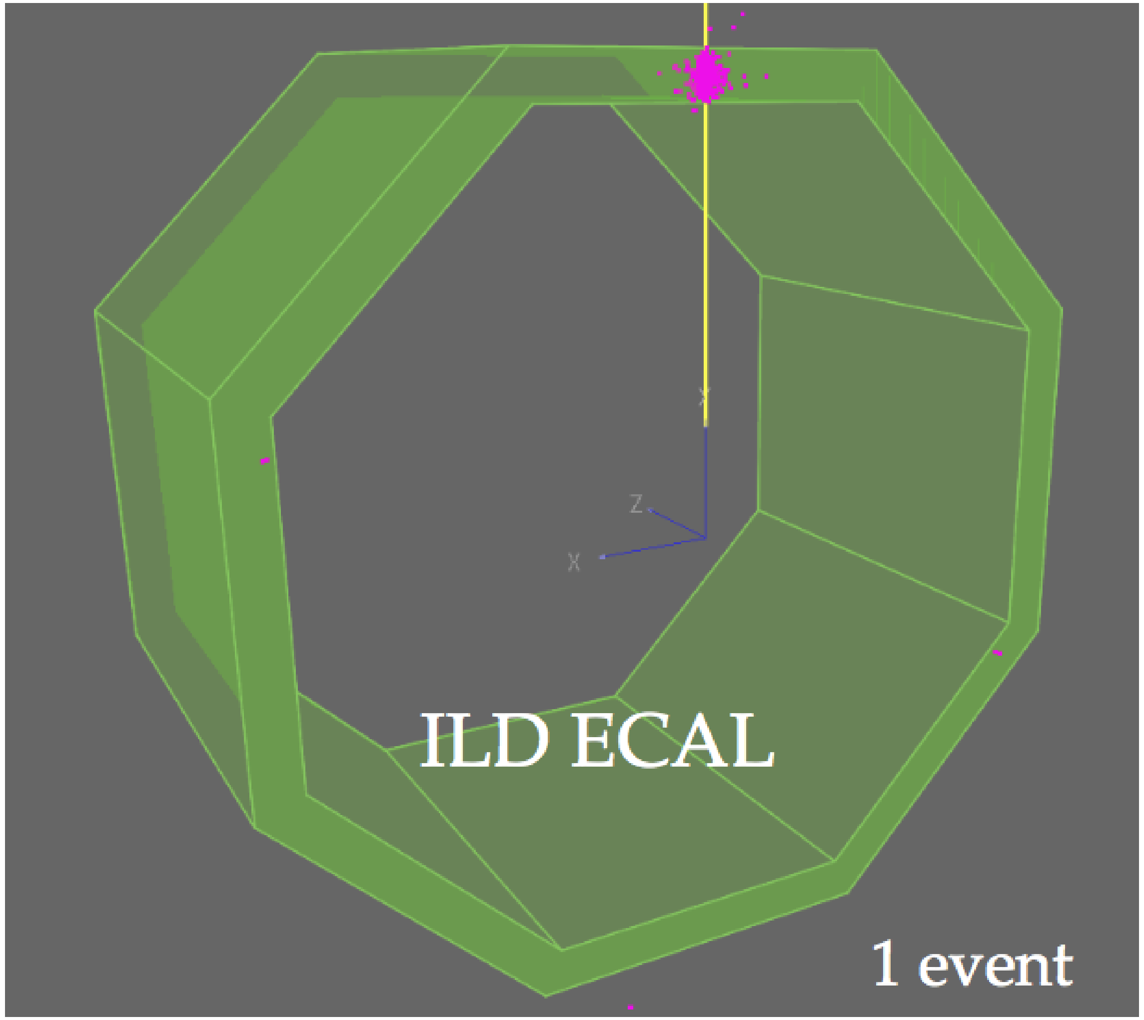


- BIB-AE successfully reproduces MIP peak at 0.2 MeV (post-processing)

- Mean and width well reproduced

Conditioning requirements for a general simulation

- Conditioning for a general calorimeter simulation:
  - Energy
  - Particle type
  - Incidence point
  - Two angles
    - Polar angle: θ
    - Azimuthal angle: ϕ



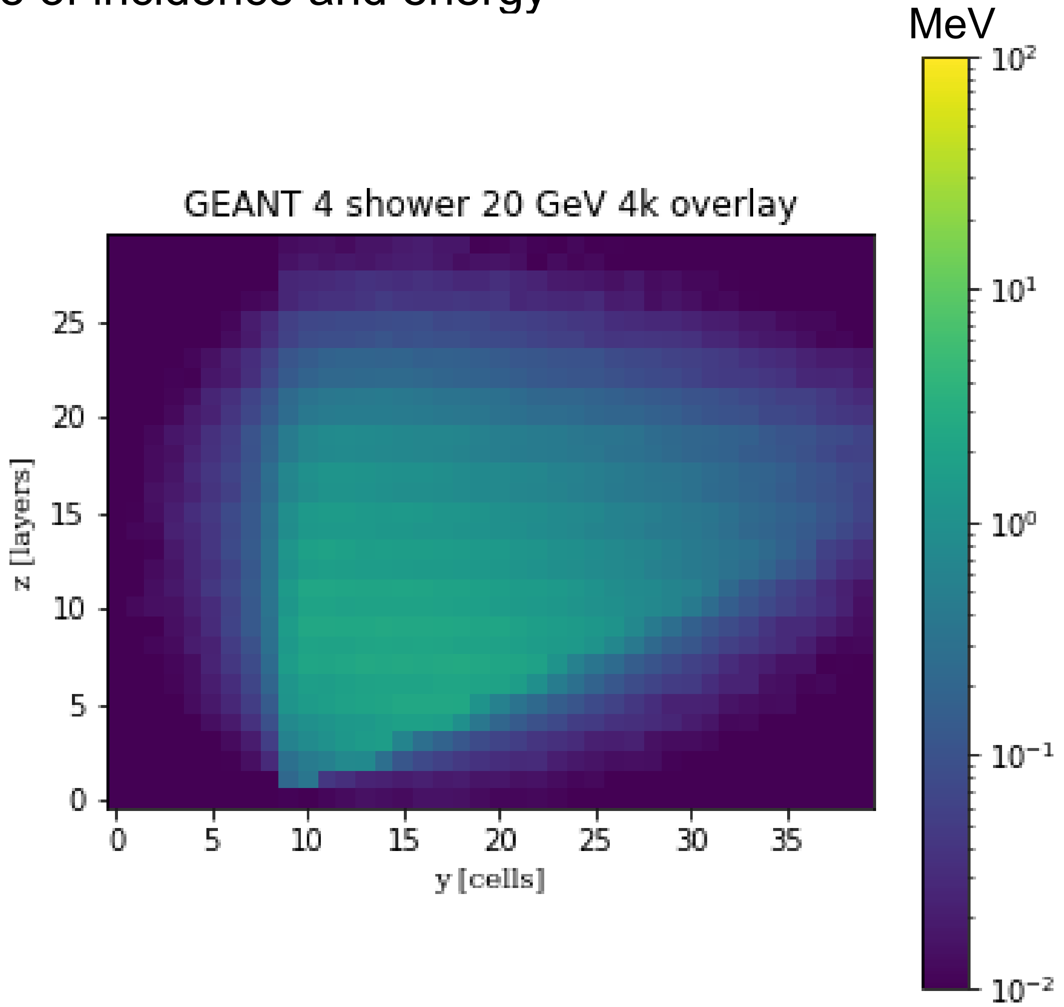
Angular conditioning- Training data

- In Progress: condition generative networks on particle's angle of incidence and energy

- Start simple:
  - Fixed energy- 20 GeV
  - Only vary polar angle in one direction- from 90°-30°
  - Fixed particle type- photons

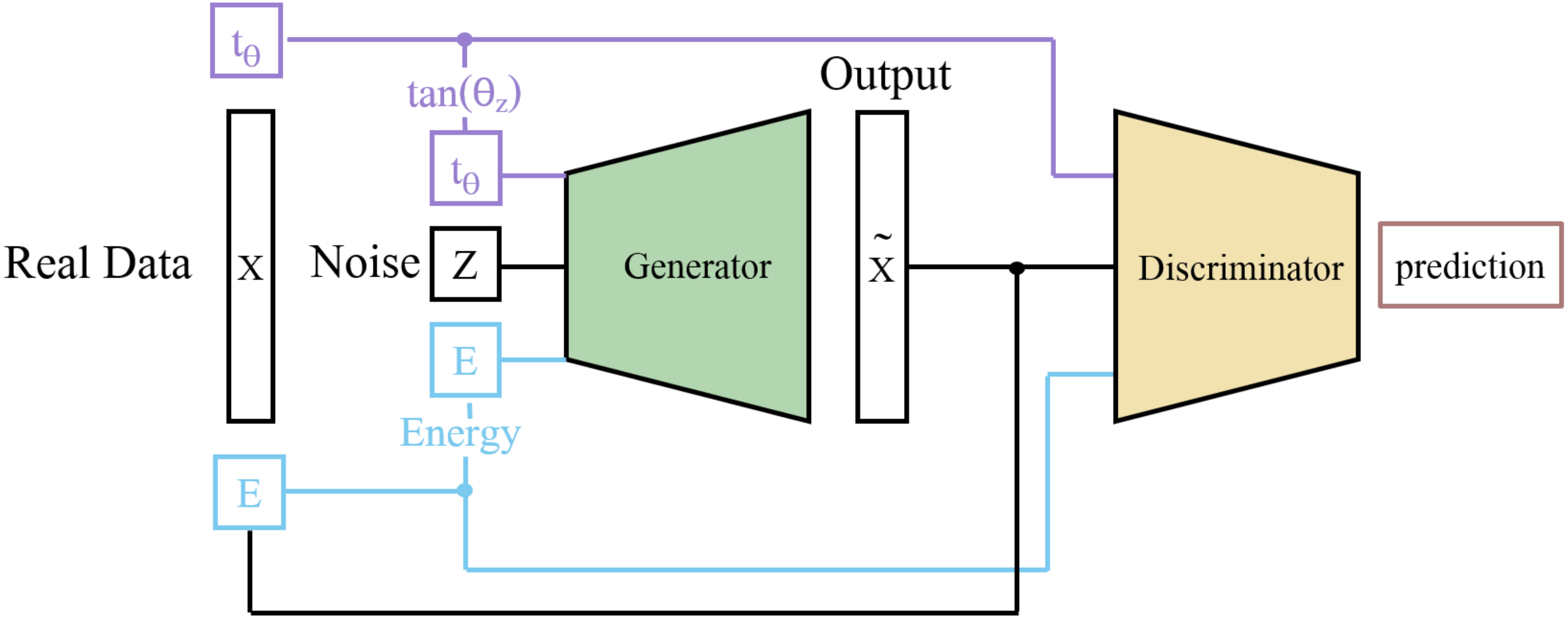
- Problem: How to make sure the full shower is contained?
  - Extend the grid in y: shape (30,30,40) (z,x,y)
  - Shift gun position

- Currently have 132k showers for training



Angular conditioning- GAN network architecture

- Begin with the simplest architecture
- Give $\tan(\theta_z)$ and E as separate network parameters
- Multiply both by noise before passing to network
- Generator learning rate 10^2 larger than discriminator



Angular conditioning- Preliminary results

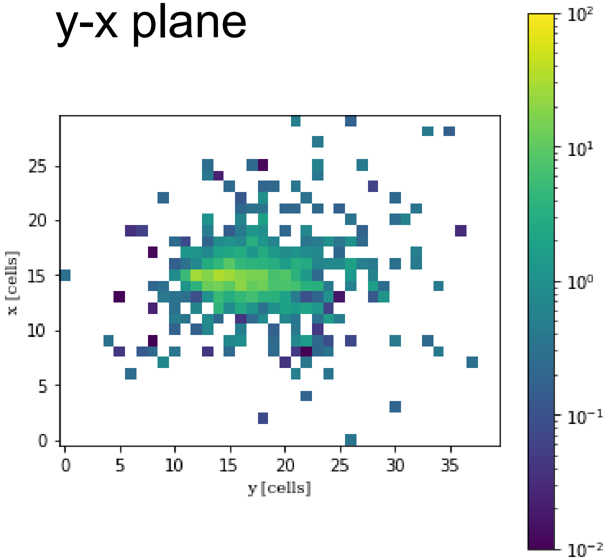
y-x plane

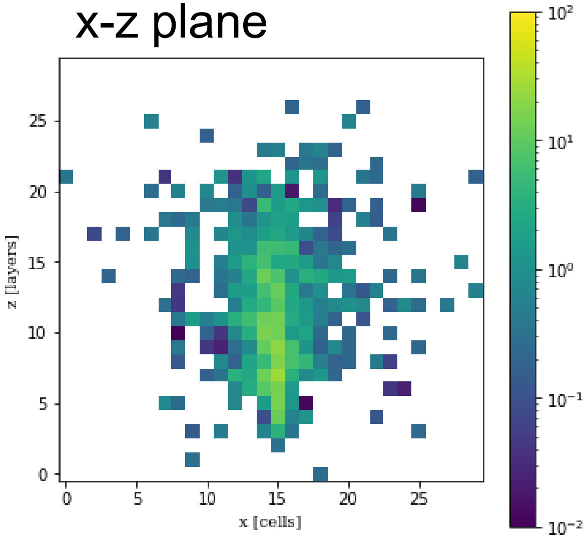
MeV

x-z plane

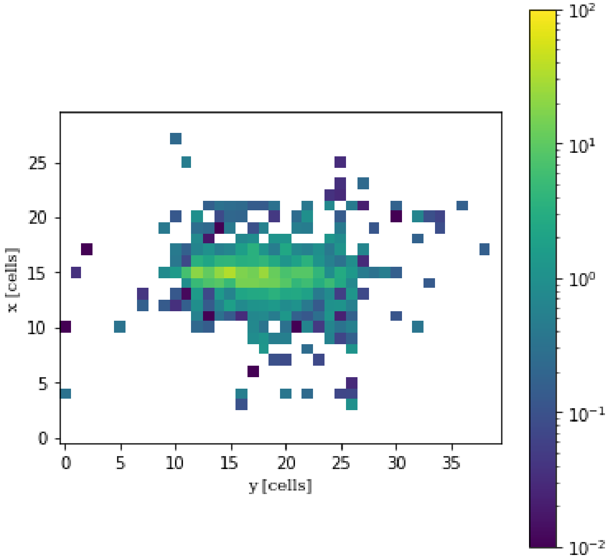
MeV

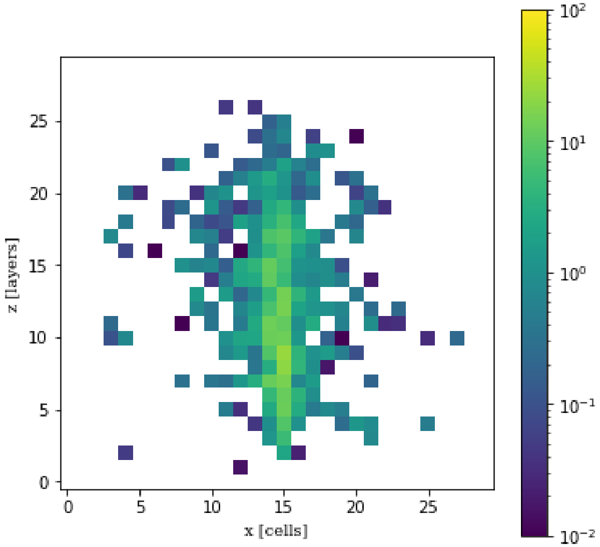
GEANT 4





GAN

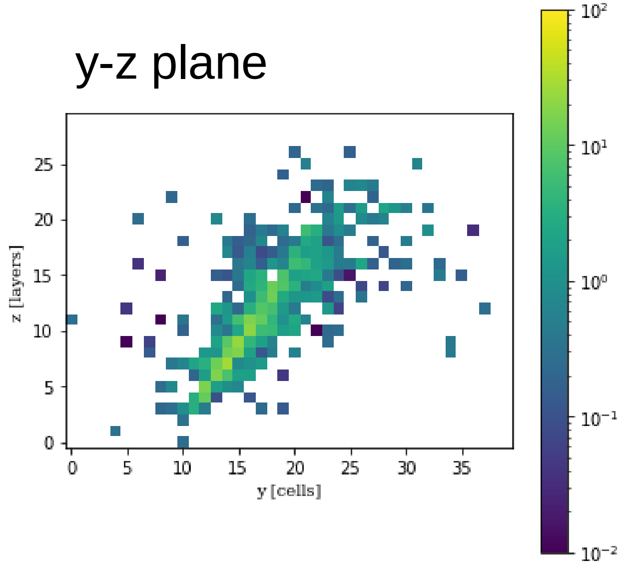


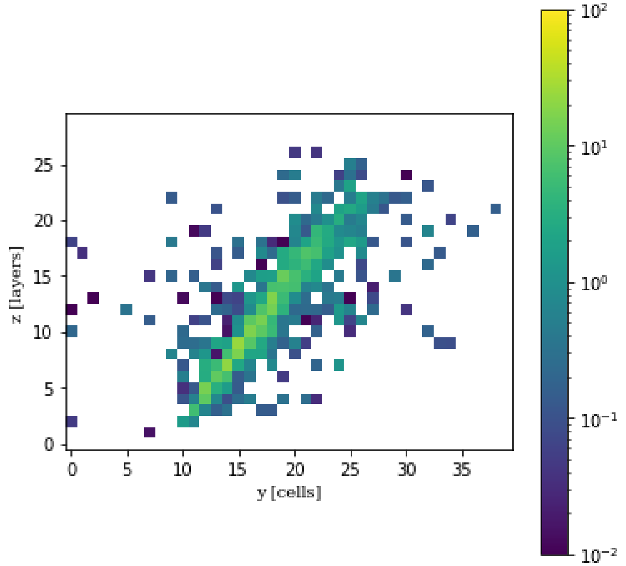


60 degree showers

MeV

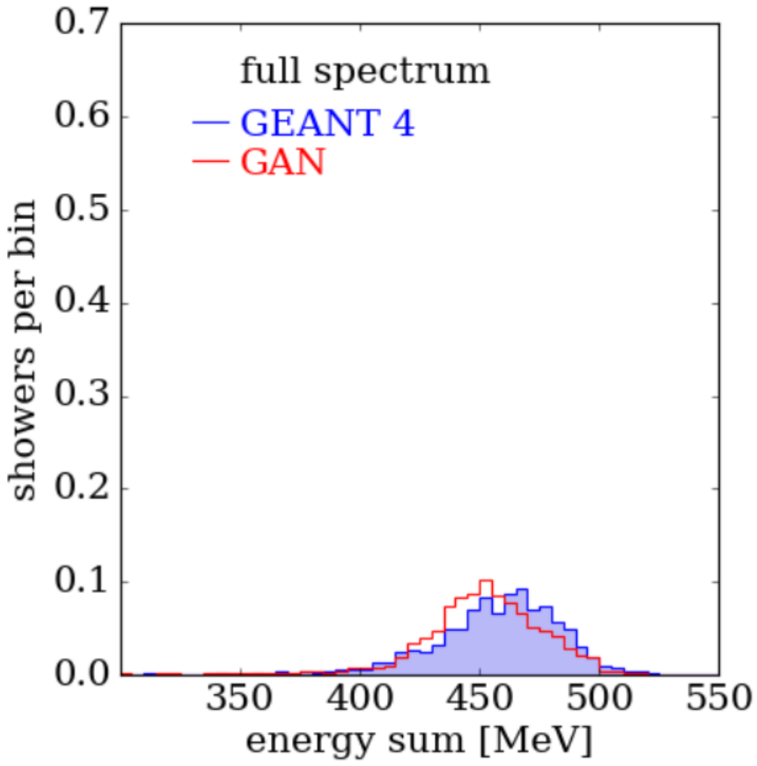
y-z plane

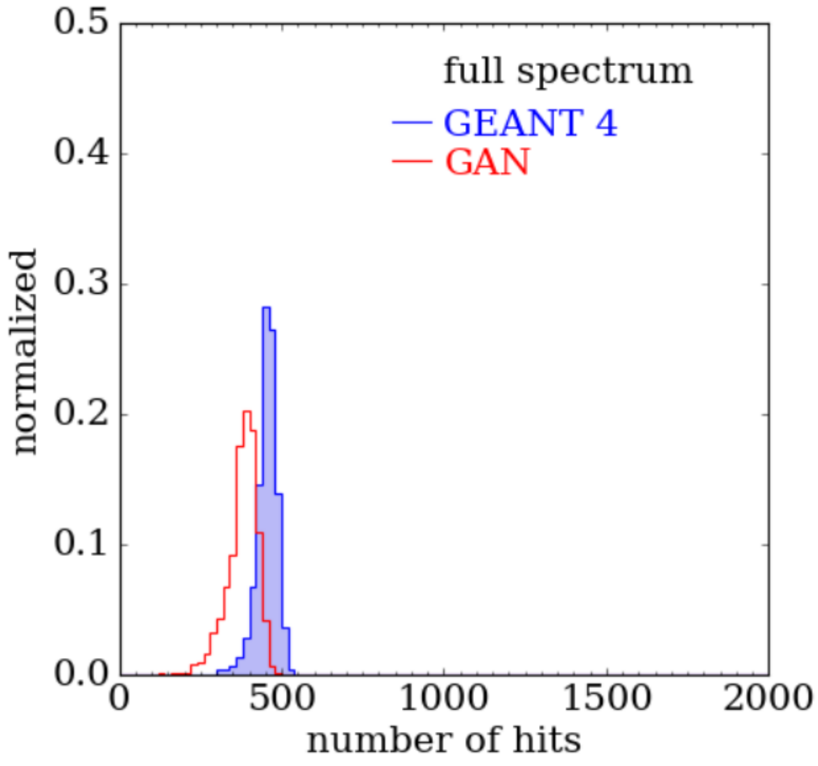


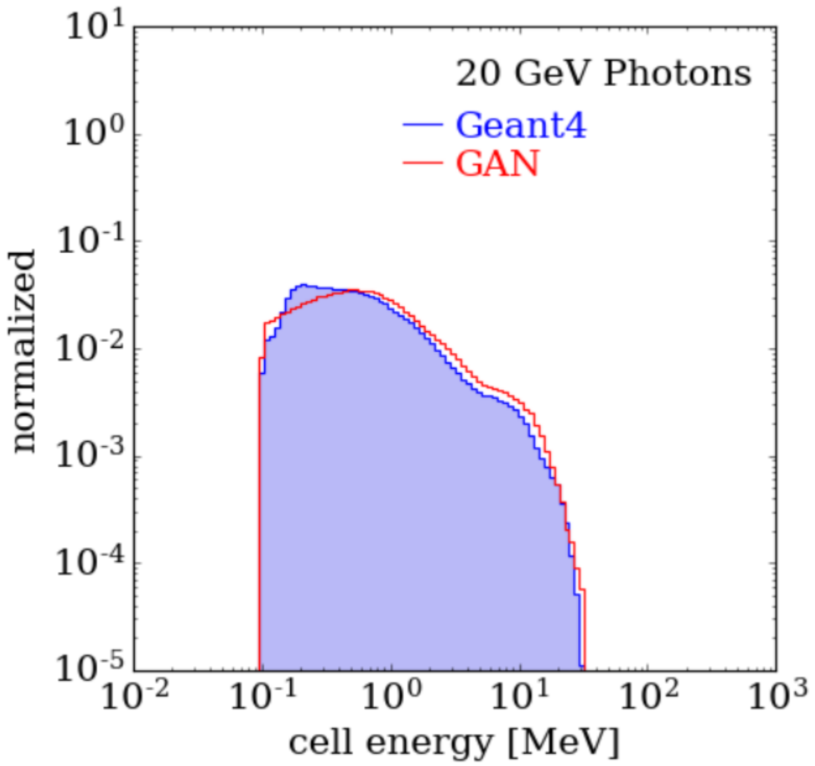


Angular conditioning- Ongoing progress

- Compare generated and GEANT4 distributions for a fixed angle of 60 degrees

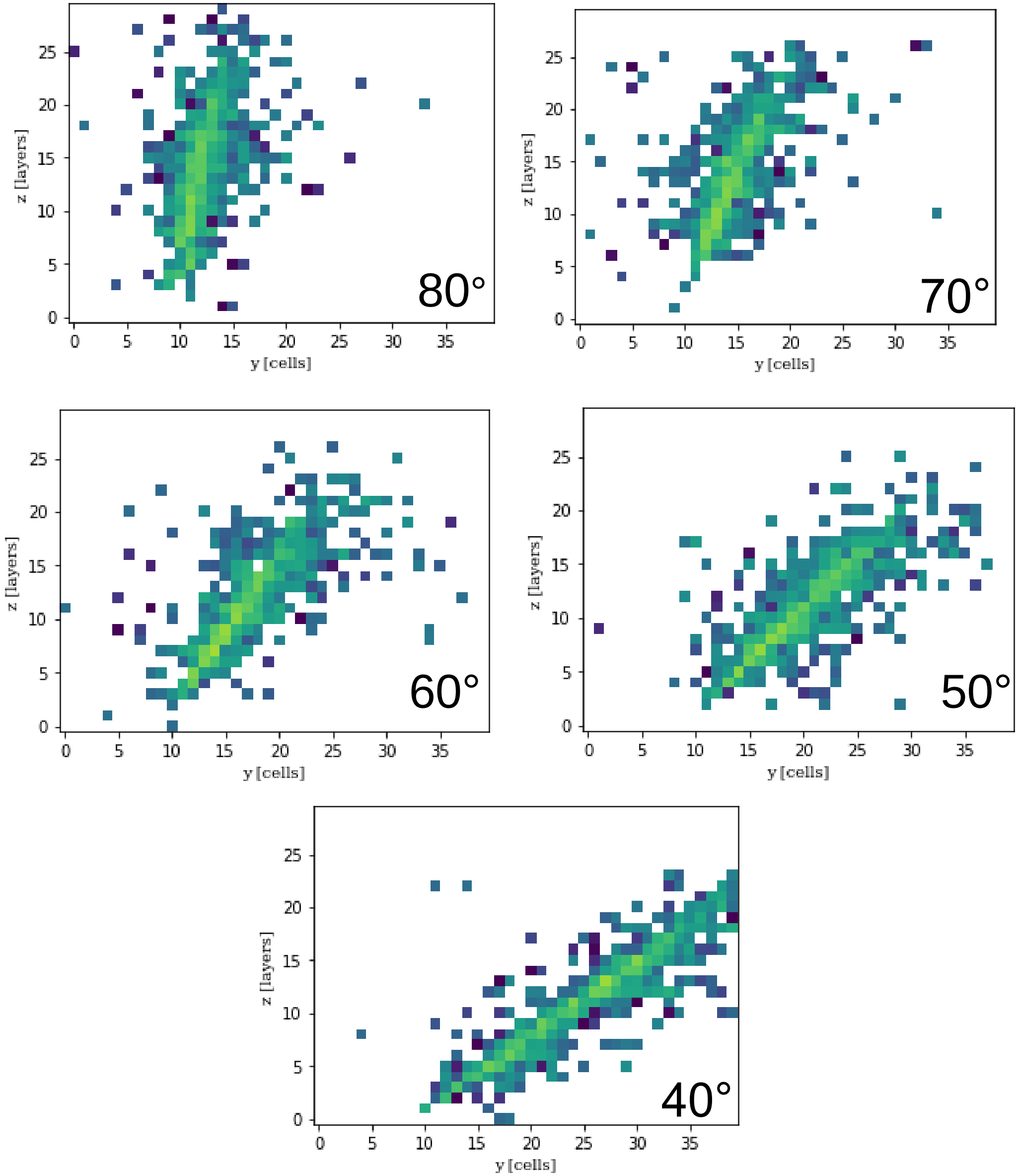




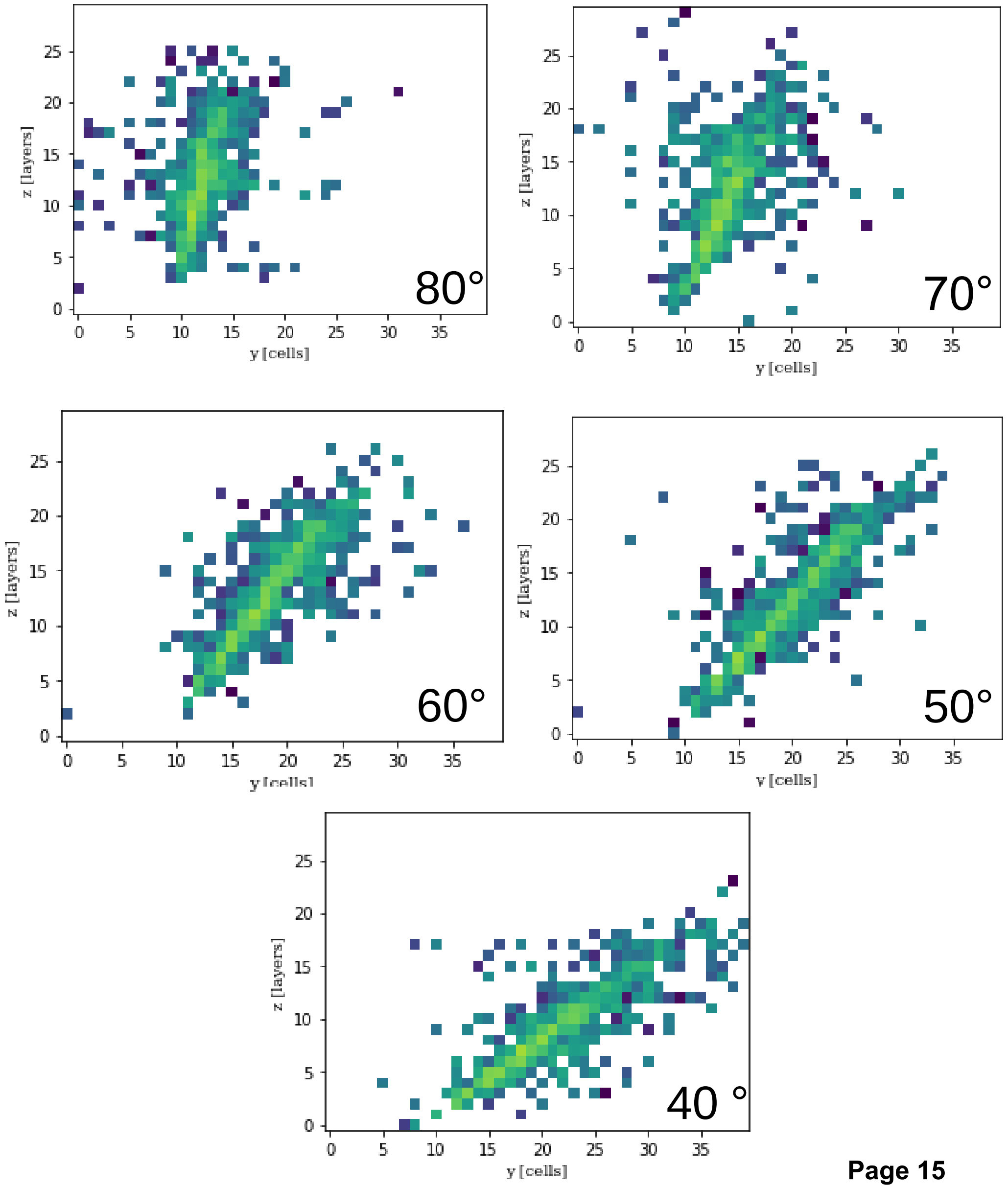


Angular conditioning- Preliminary results

GEANT4



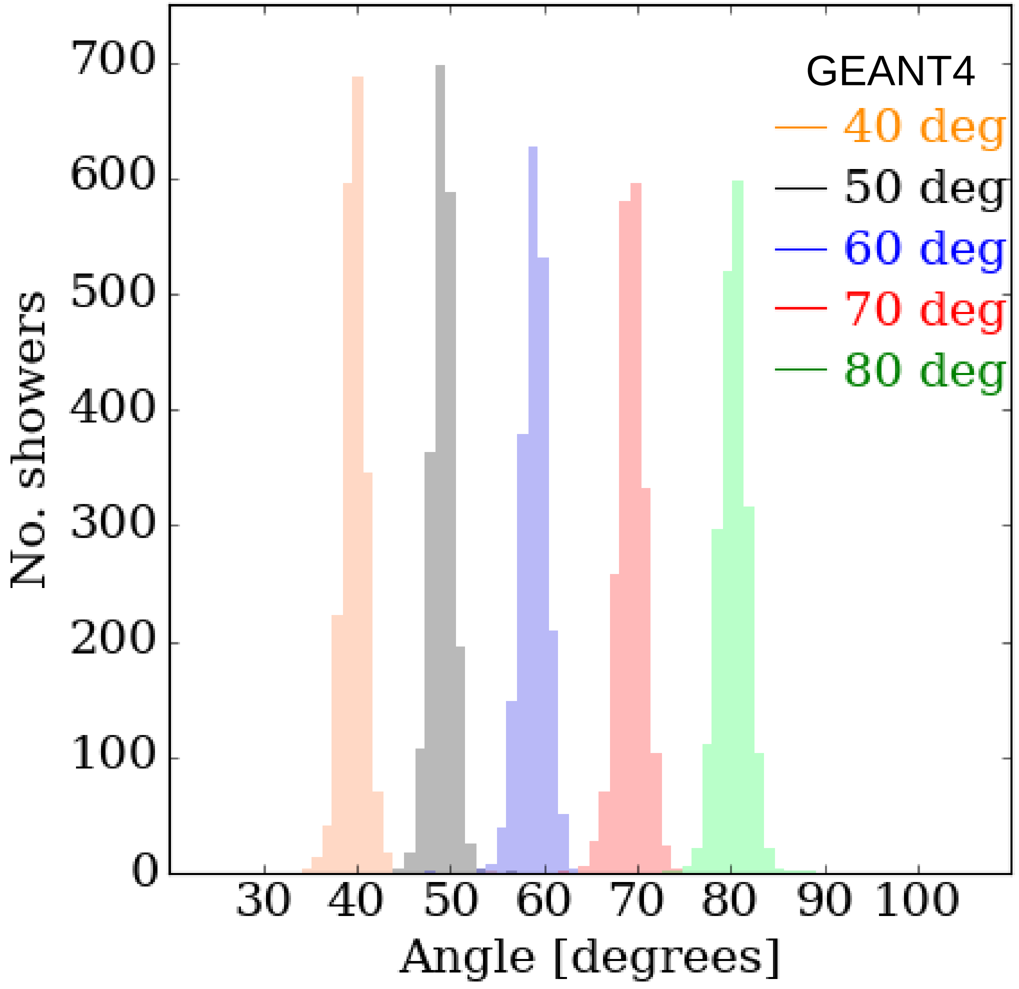
GAN



Angular conditioning- Angular benchmark

- Find principal axis of showers- benchmark for model
- Diagonalize analogue inertia tensor
- Principal axis is eigenvector with the smallest eigenvalue

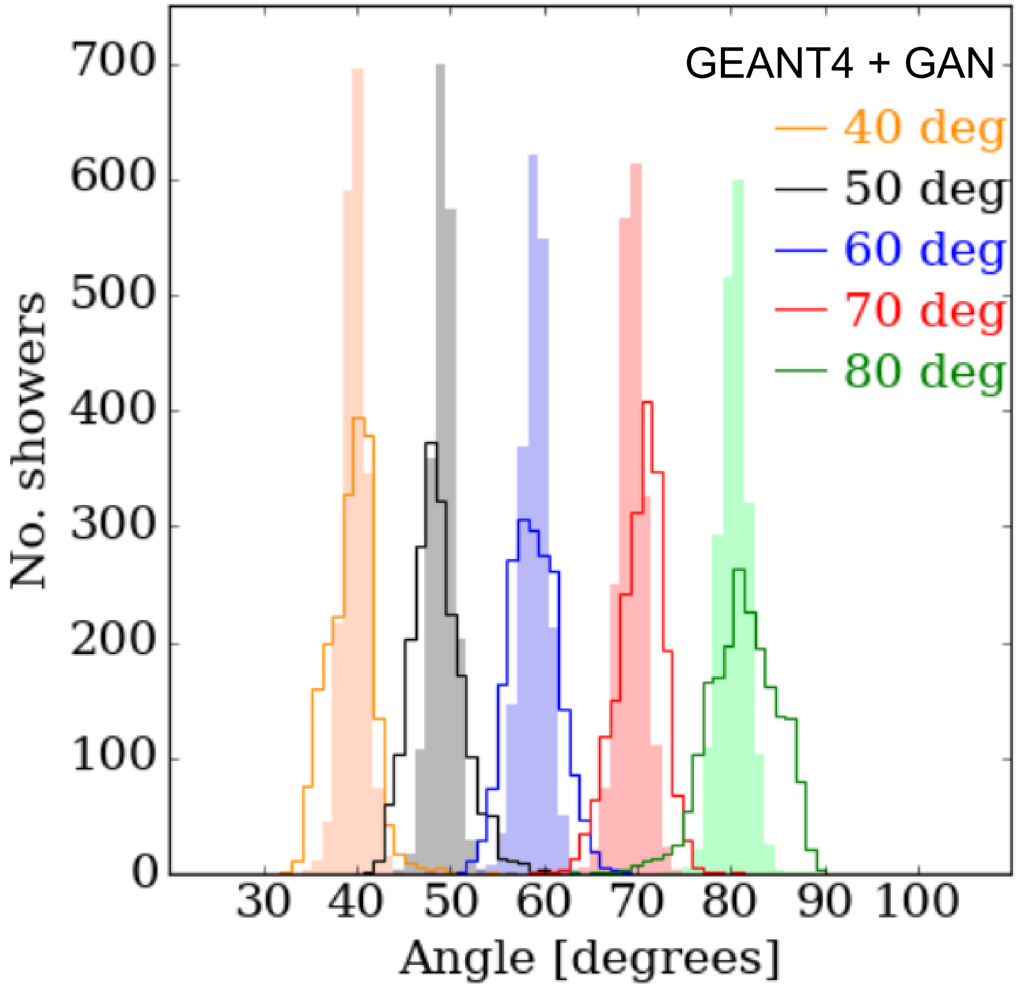
- In order to reconstruct the true angle:
    - Layers in ILD ECAL are not regularly spaced in z
    - Take this into account in positions of hits used in the inertia tensor



Angular conditioning- Angular benchmark

- Find principal axis of showers- benchmark for model
- Diagonalize analogue inertia tensor
- Principal axis is eigenvector with the smallest eigenvalue

- In order to reconstruct the true angle:
  - Layers in ILD ECAL are not regularly spaced in z
  - Take this into account in positions of hits used in the inertia tensor



Conclusion

- Demonstrated angular and energy conditioning in a GAN architecture

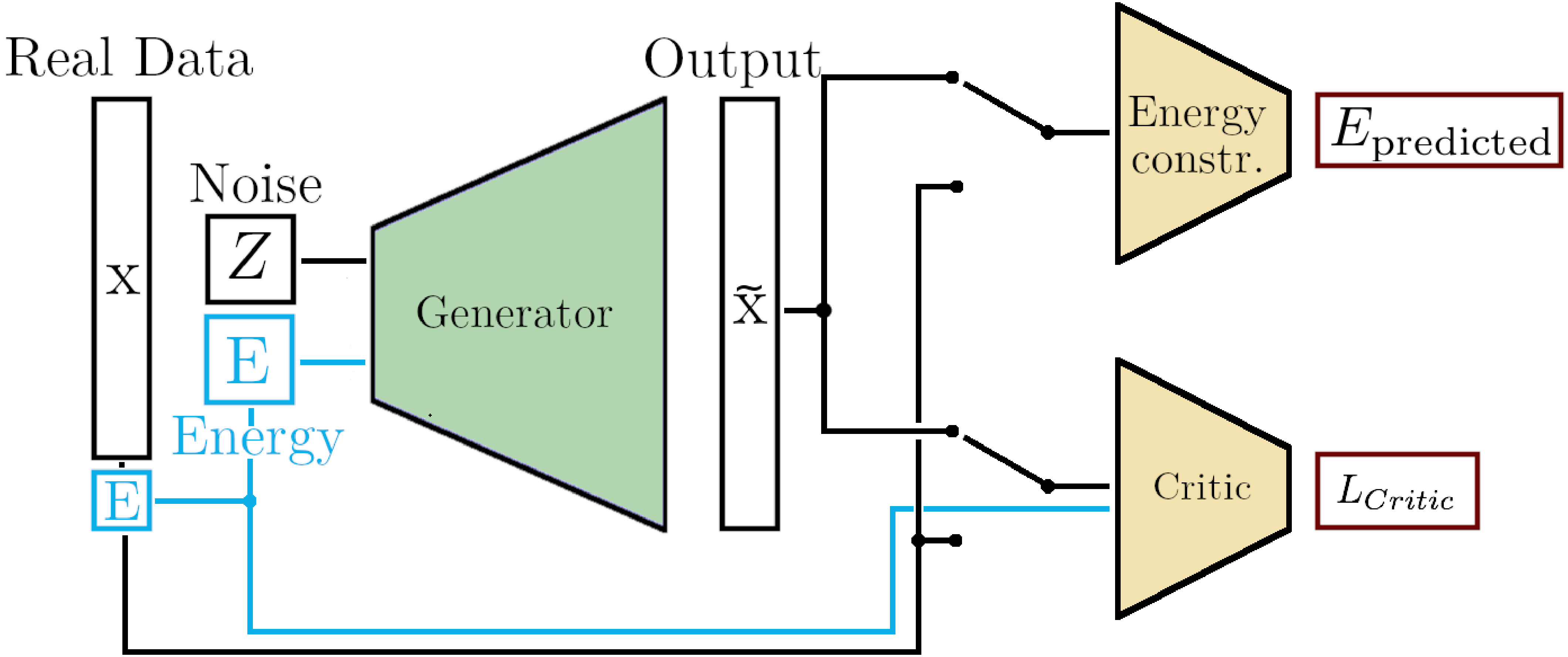
Next Steps

- Train on more data
- Quantitative metric for epoch selection
- Study trade off in physics distributions when angular conditioning is introduced

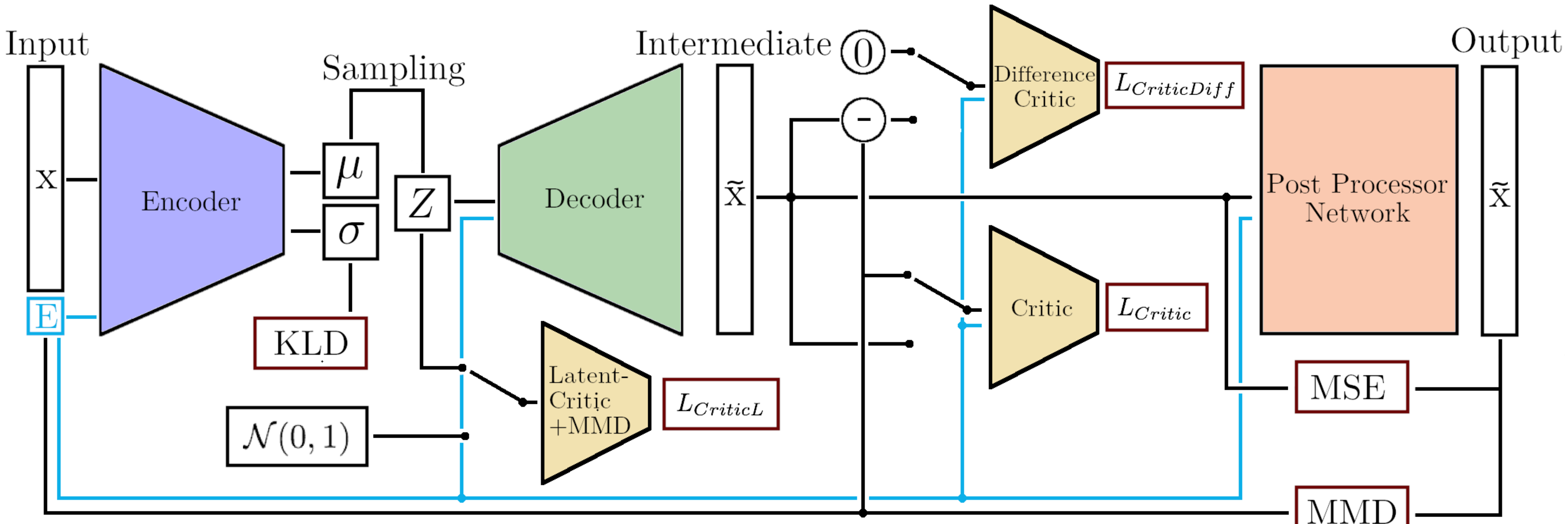
- More sophisticated architectures e.g. BIB-AE
- Vary energy and study effect on performance

Backup

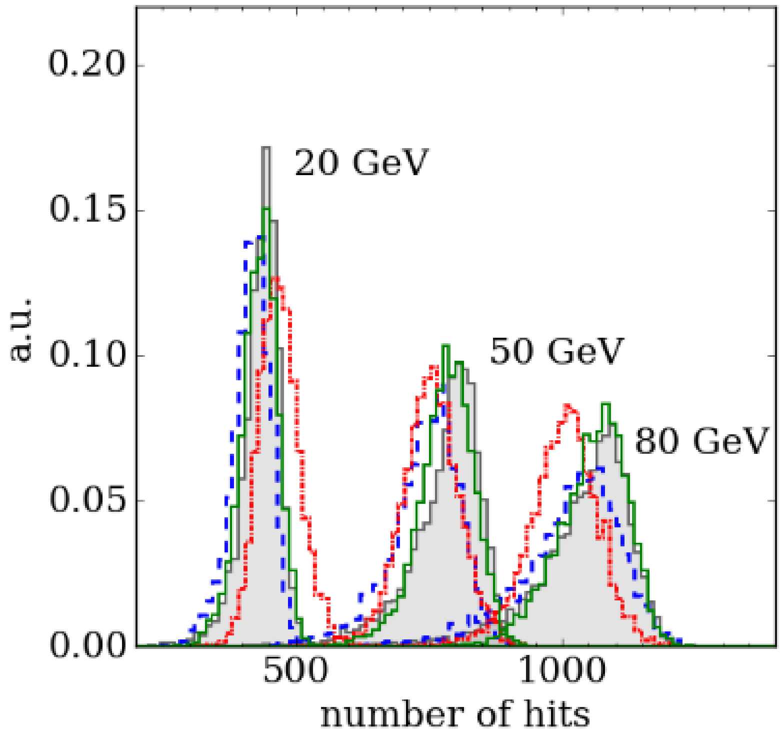
Getting high: Generative models- Wasserstein GAN

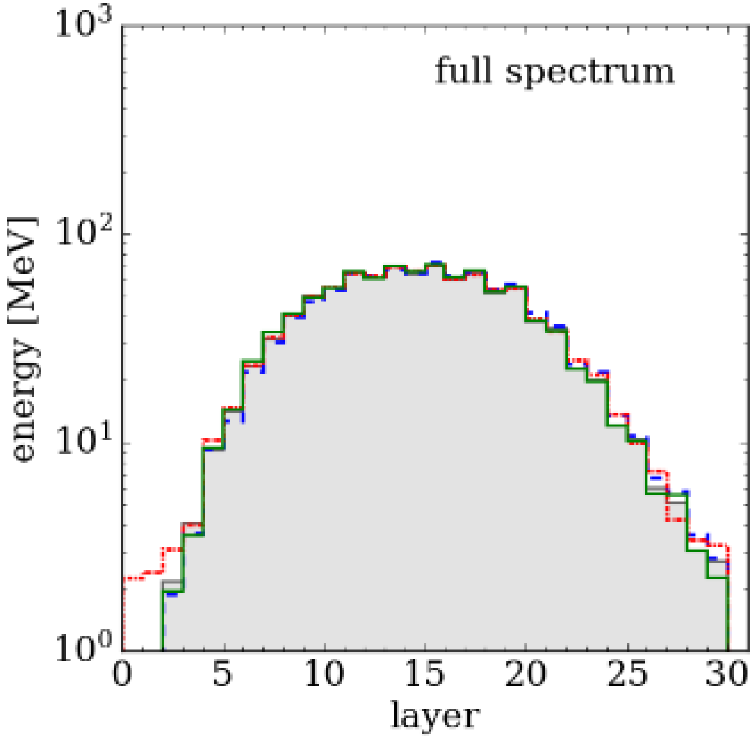


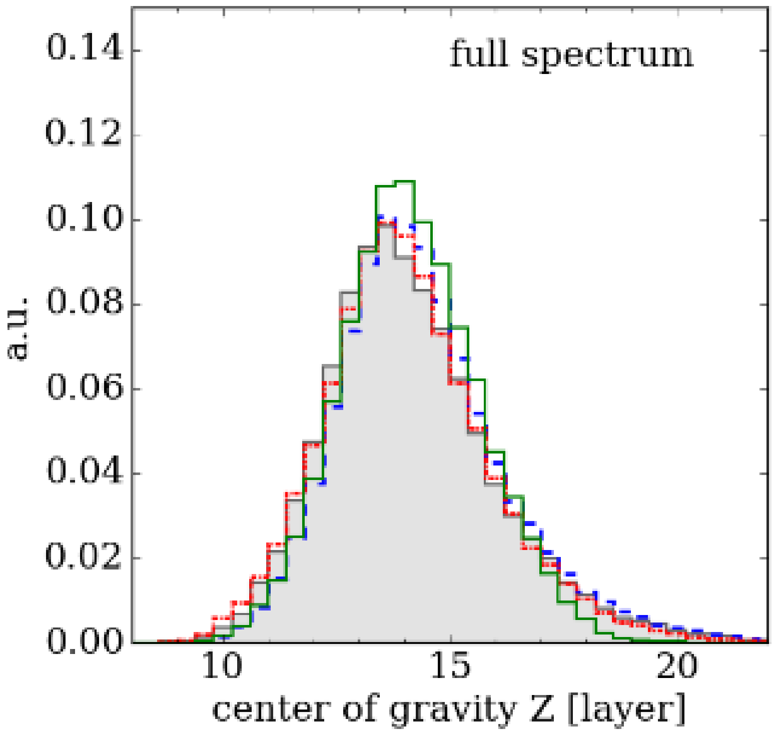
Getting high: Generative models- BIB-AE



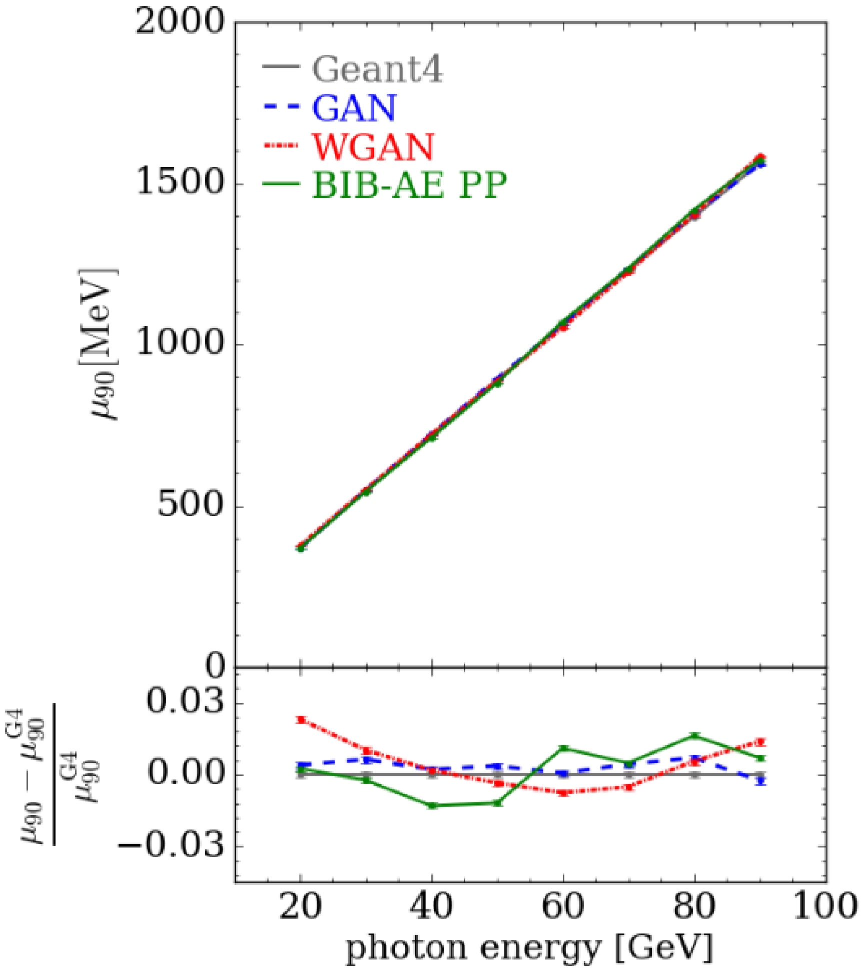
Getting high- other physics distributions

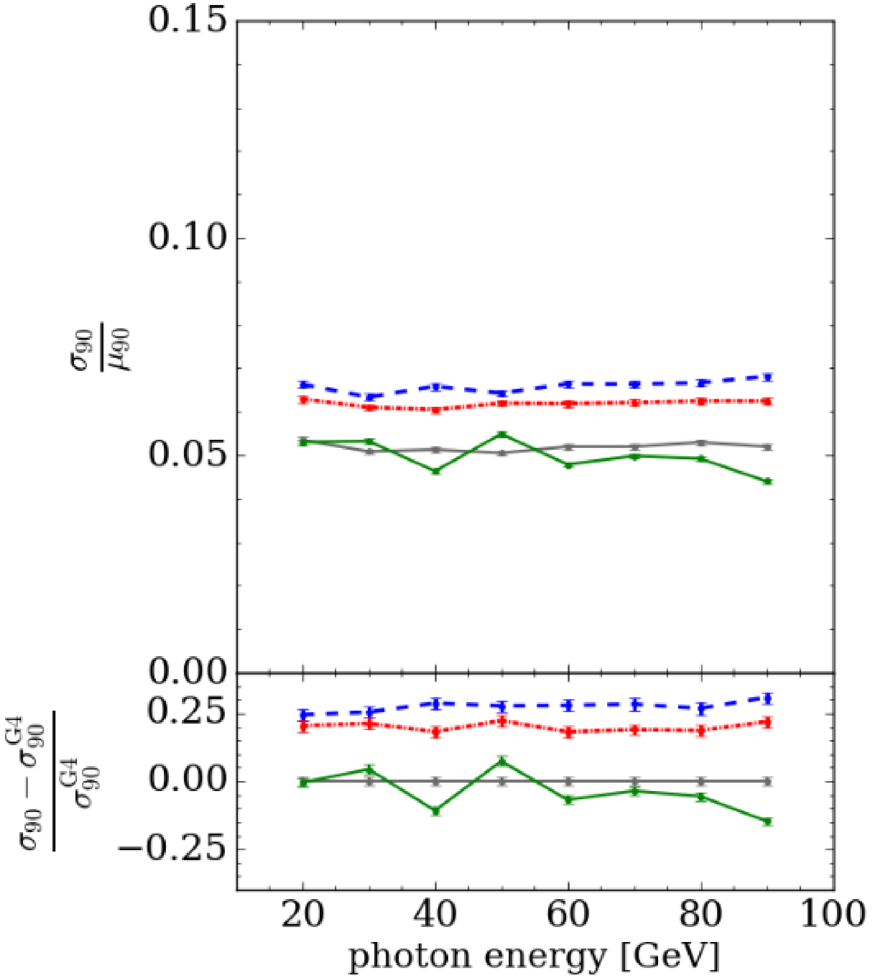






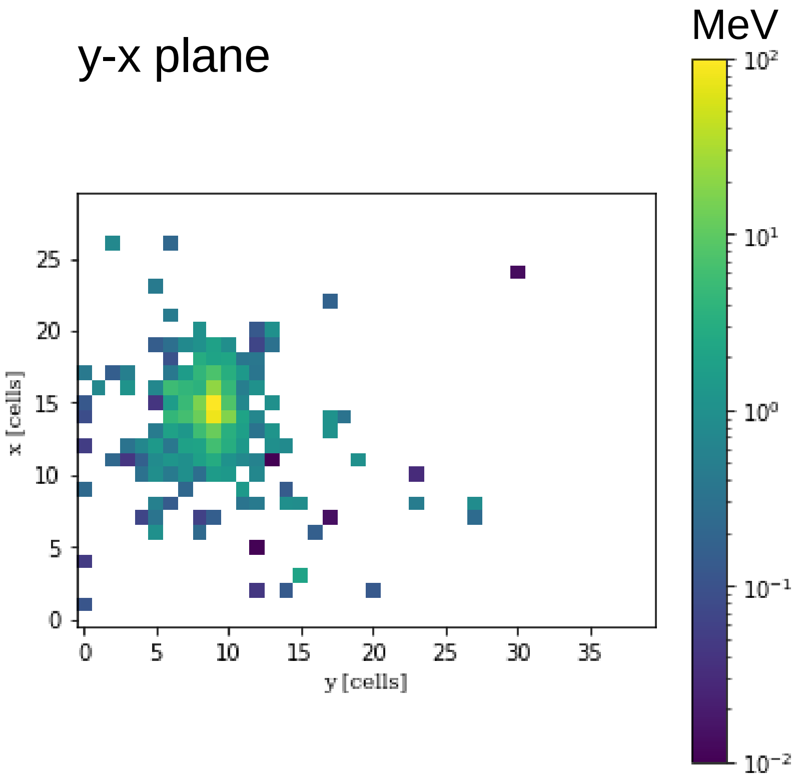
Getting high- other physics distributions



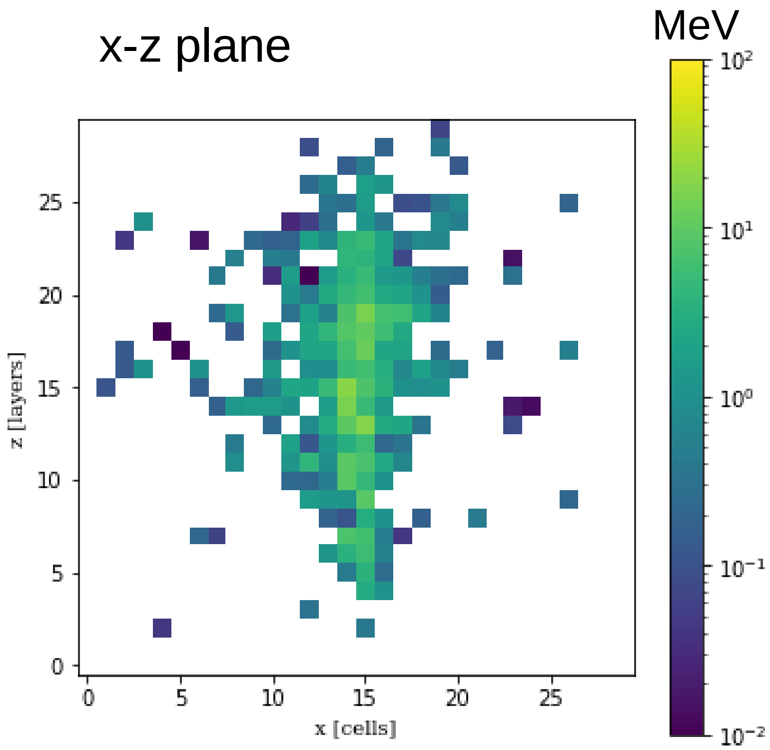


Angular conditioning- Preliminary results

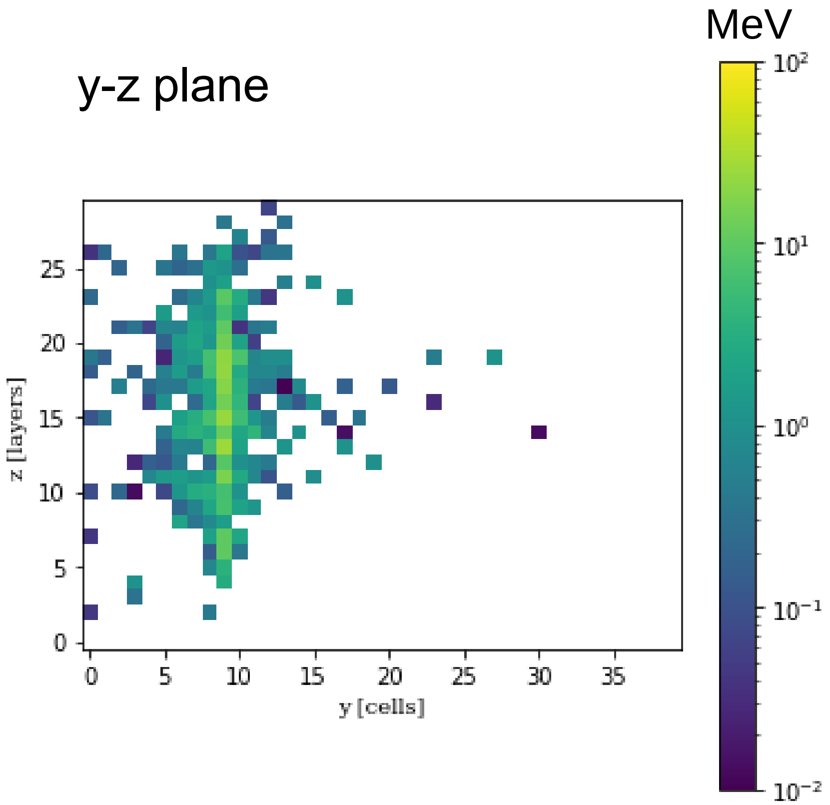
y-x plane



x-z plane

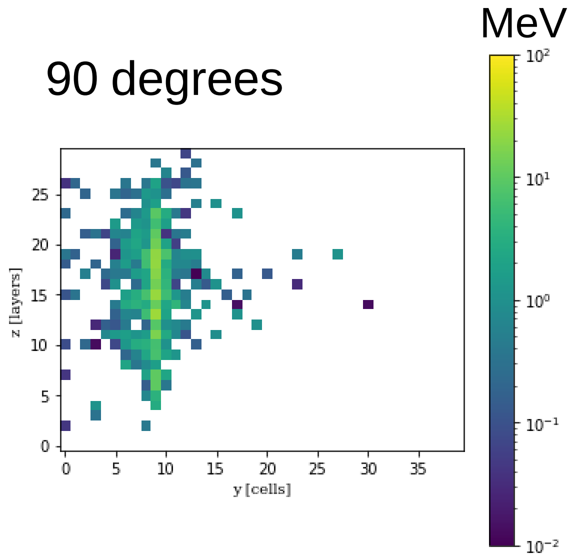


y-z plane

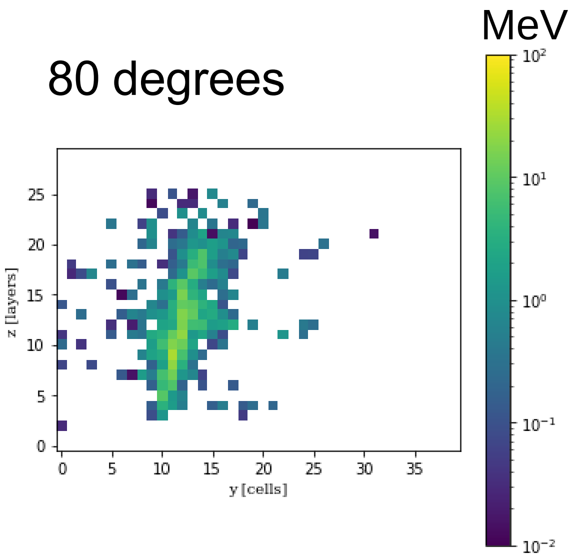


Angular conditioning- Preliminary results

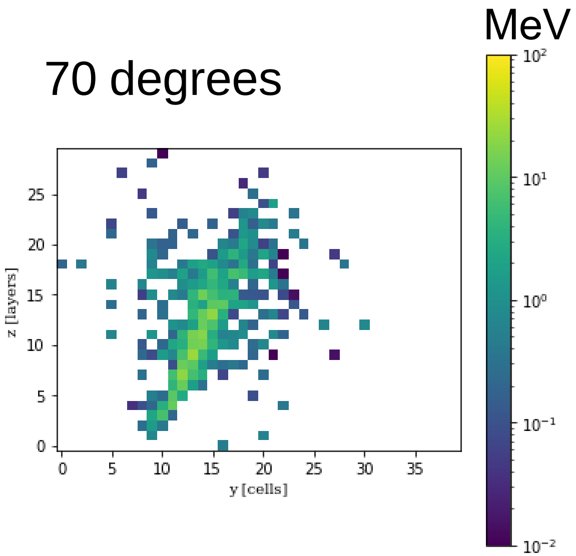
90 degrees



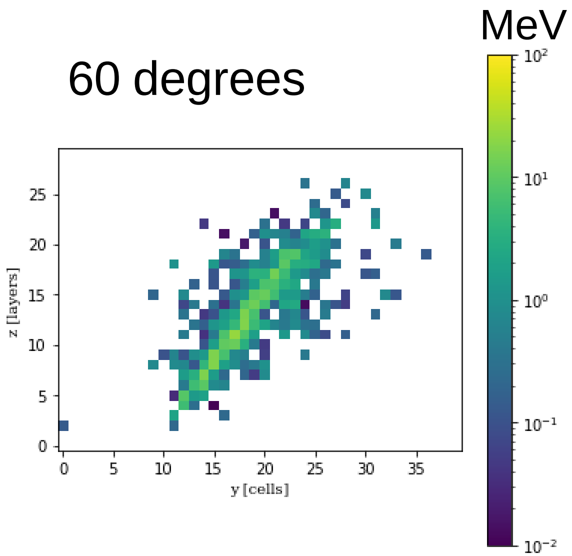
80 degrees



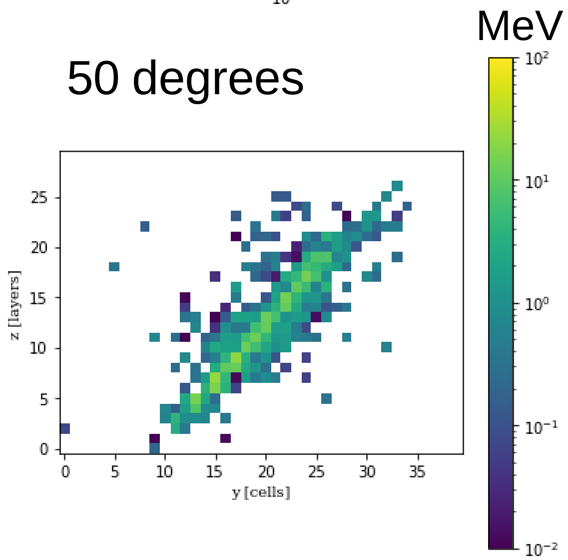
70 degrees



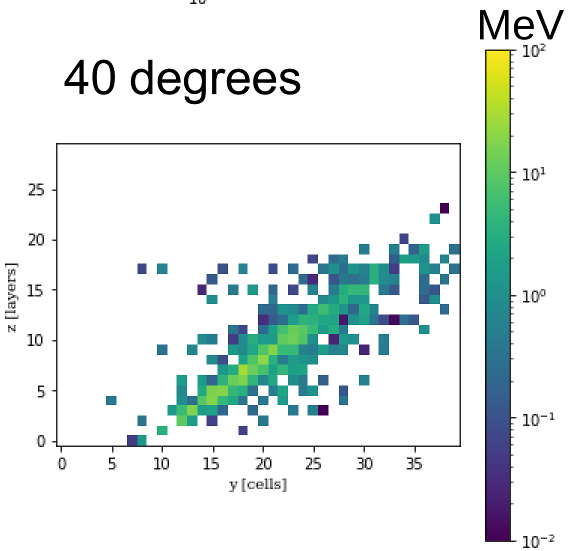
60 degrees



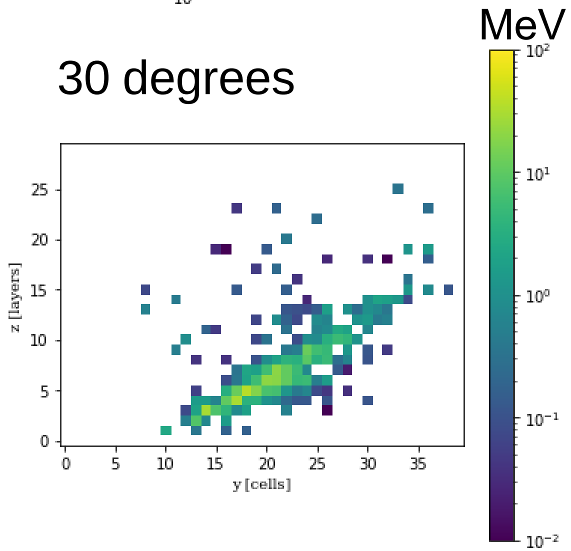
50 degrees



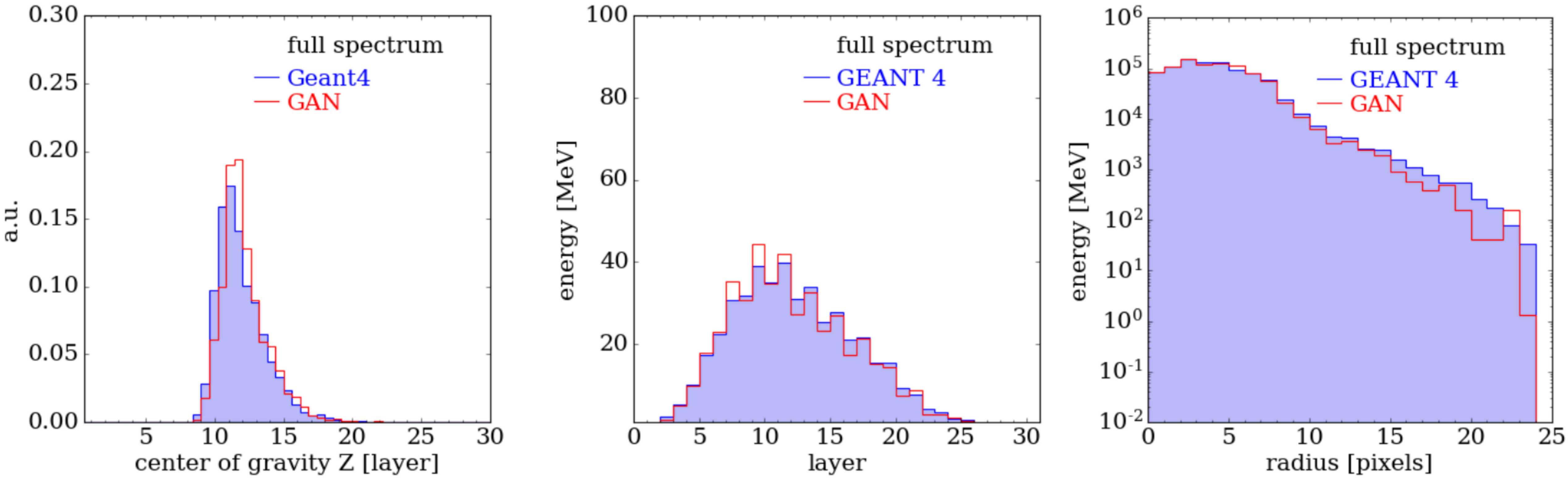
40 degrees



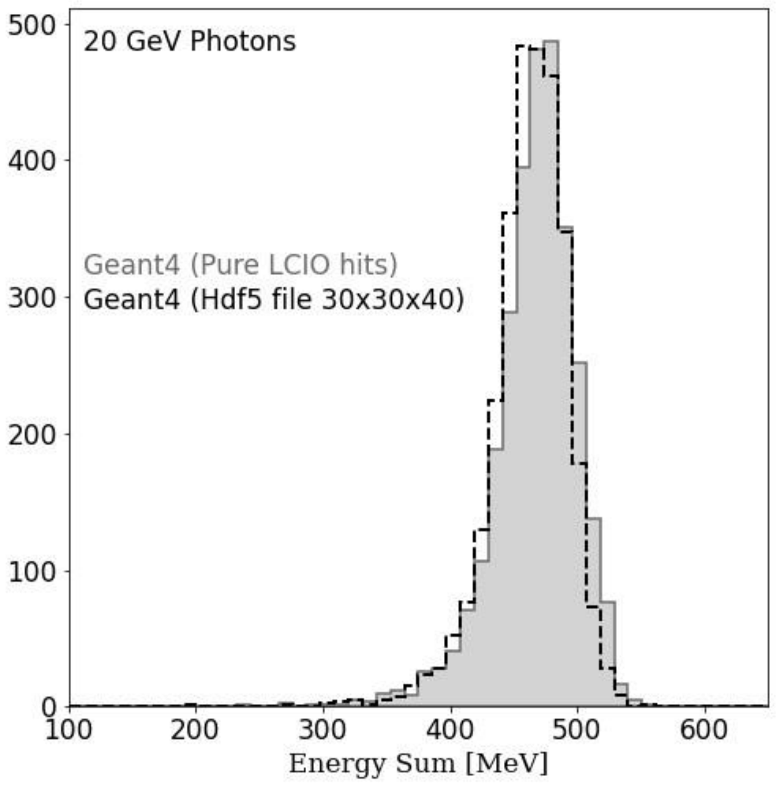
30 degrees

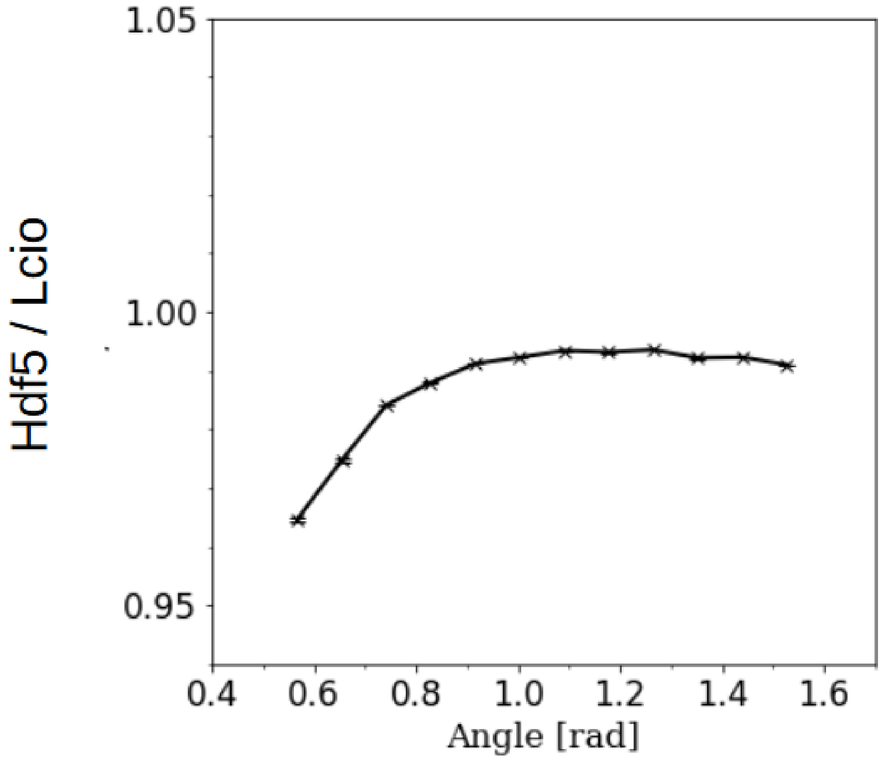


Angular conditioning- other physics distributions (preliminary)



Angular conditioning- Checking full shower is contained





Angular conditioning- GAN network architecture: more details

- Generator:
  - Noise vector of length 100, uniformly between -1 and 1
  - 4 transposed 3D convolutional layers
  - ReLU activation functions
  - Learning rate: 1×10^{-3}
- Discriminator:
  - Five 3D convolutional layers- batch normalization except final layer
  - Two fully connected layers for readout
  - LeakyReLU; sigmoind in the final layer
  - Learning rate: 1×10^{-5}
- Trained for 50 epochs
- Use Adam optimizer with a learning rate of 2×10^{-5}