Generative Models for Fast Electromagnetic and Hadronic Shower Simulation

Erik Buhmann, Sascha Diefenbacher, Engin Eren, Frank Gaede, Daniel Hundhausen, Gregor Kasieczka, William Korcari, Anatolii Korol, Katja Krüger, **Peter McKeown**[1], Lennart Rustige

[1] Deutsches Elektronen-Synchrotron

23.11.2021

peter.mckeown@desy.de

FTX HELMHOLTZ RESEARCH FOR GRAND CHALLENGES ILD ilc UH CLUSTER OF EXCELLENCE QUANTUM UNIVERSE DESY.

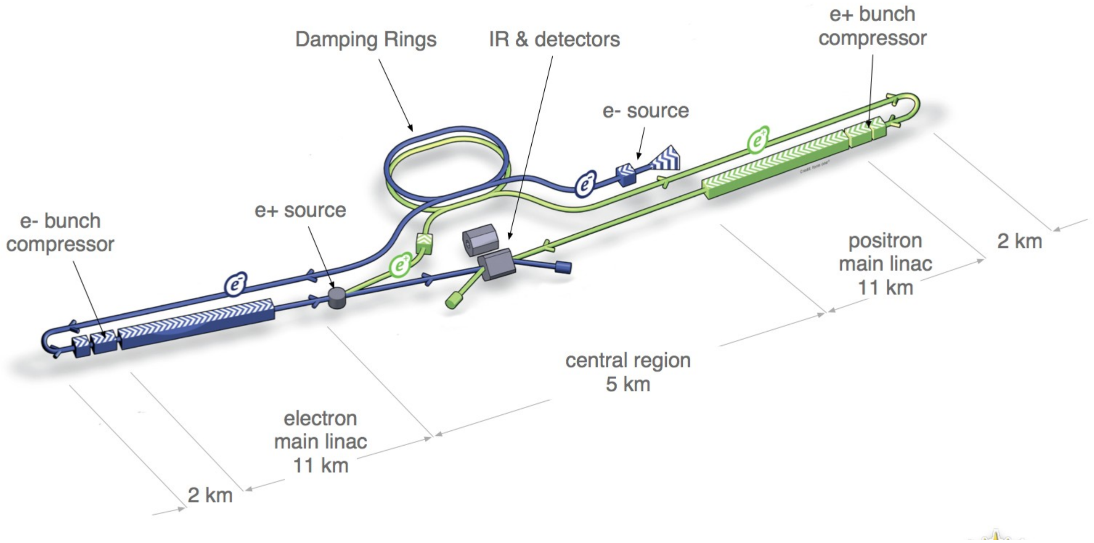
Outline

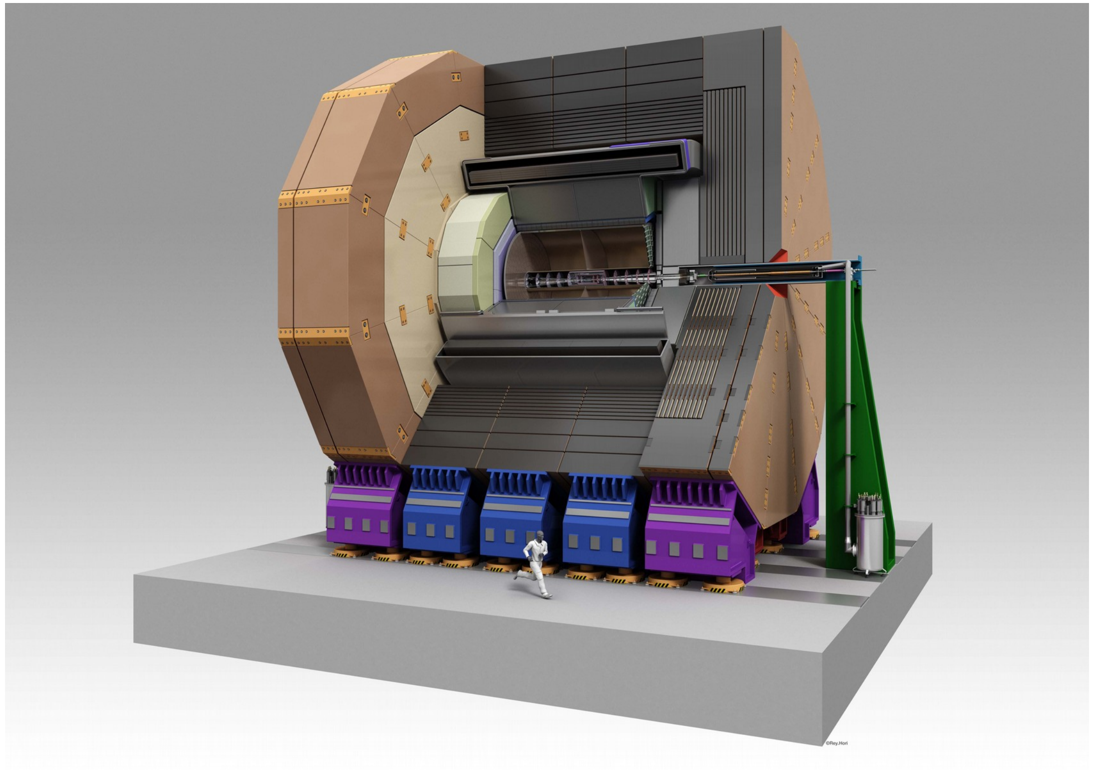
The ILD Concept

- International Large Detector (ILD) concept for the International Linear Collider (ILC)
  - Higgs Factory (initial 250 GeV stage)
  - High energy e^+e^- linear collider

- Optimized for Particle Flow
  - Reconstruct each individual particle in subdetector
  - Obtain optimal detector resolution

- High granularity calorimeters:
  - Sampling calorimeters
  - SiW Ecal: 30 layers, 5x5 mm^2, 2 sampling fractions
  - FeSci Hcal: 48 layers, 3x3 cm^2



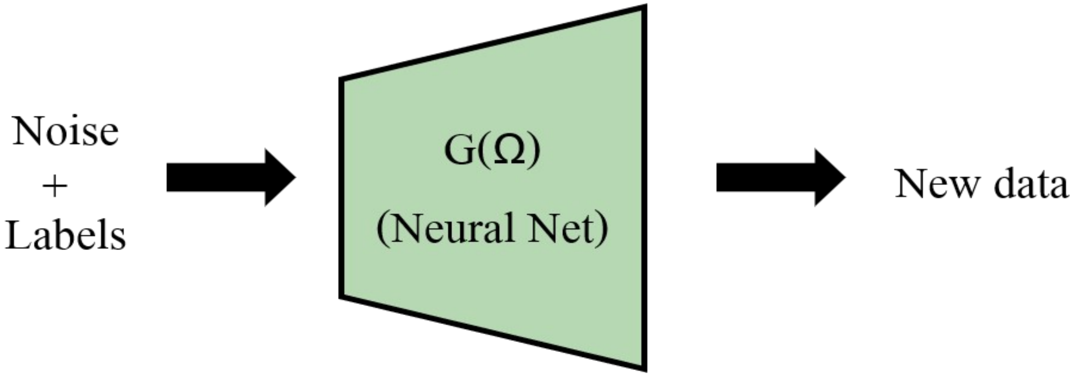


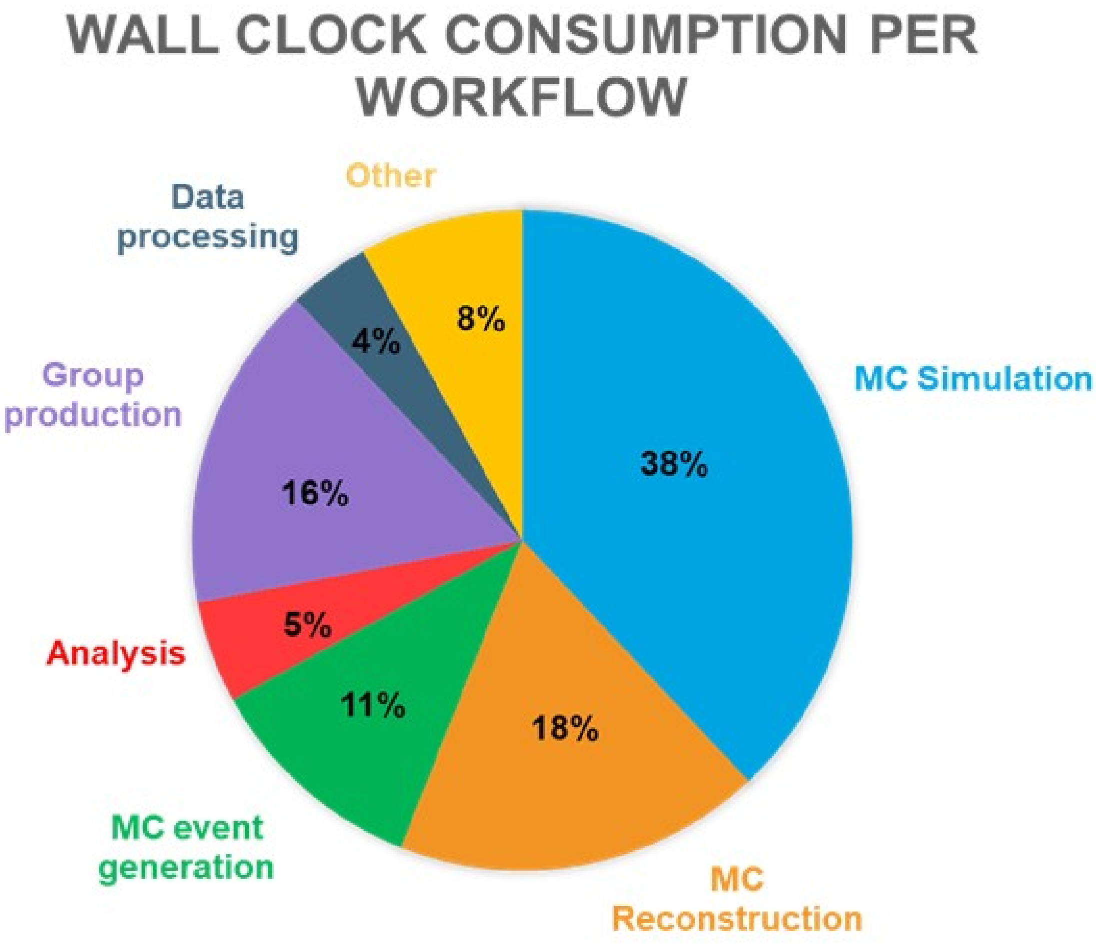
Reducing the Strain on HEP Computing Resources

- MC simulation is computationally expensive
  - Calorimeters most intensive part of detector simulation

- **Generative models** potentially offer orders of magnitude speed up

- Amplify statistics of original data set
  - Generate new samples following distribution of original data
  - Significant speed up



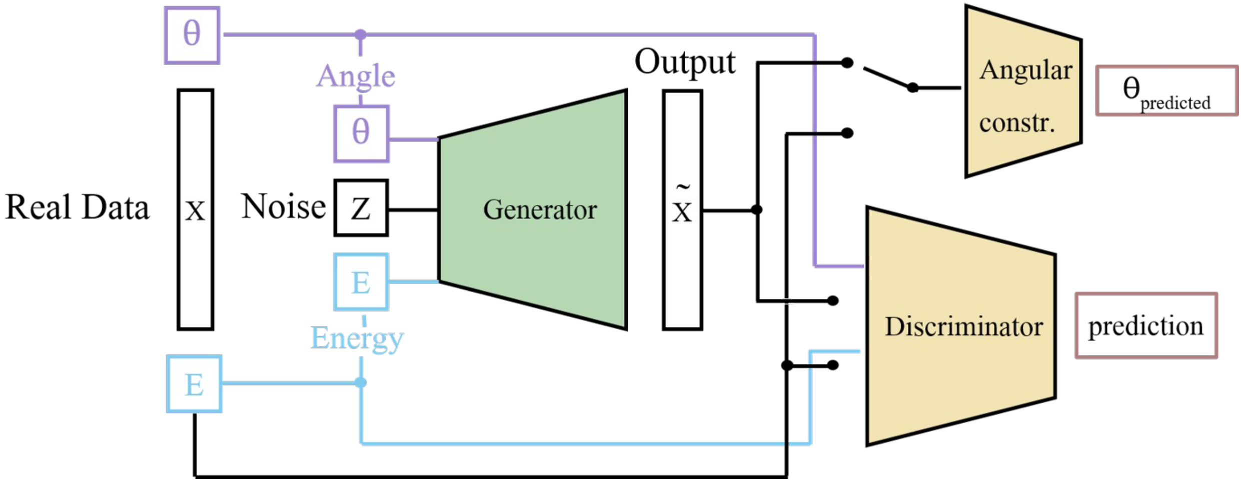


D. Costanzo, J. Catmore, ATLAS Computing update, LHCC meeting, 2019

Architectures: GAN and WGAN

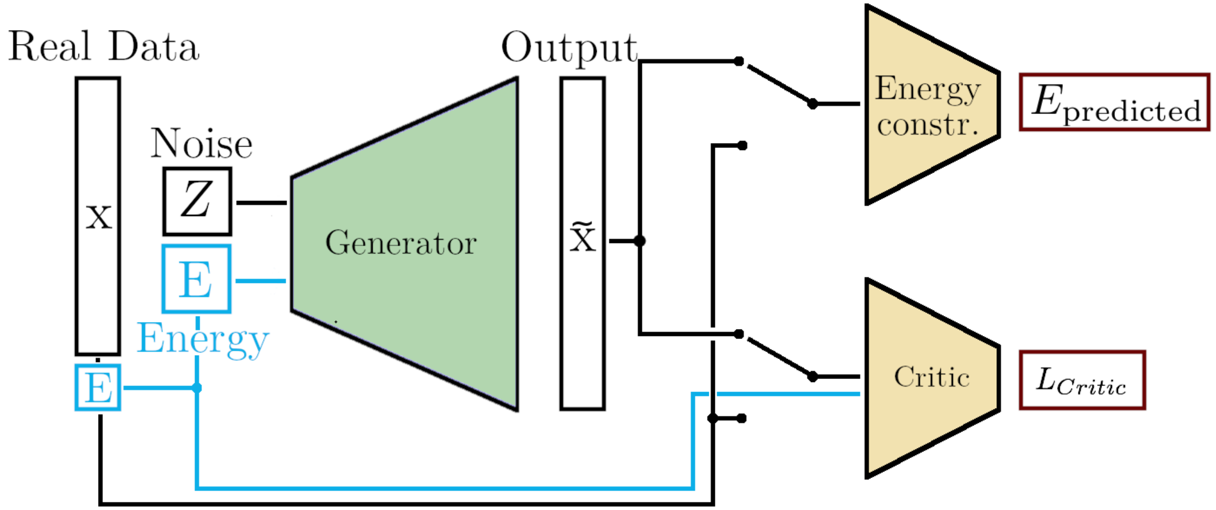
GAN- Angular photons

- Original Generative architecture applied for shower generation
- Discriminator and Generator play a min-max game



WGAN- Pions

- Alternative to classical GAN training
- Wasserstein-1 distance as loss with gradient penalty: **improve stability**



- **Addition of auxiliary constrainer networks for improved conditioning performance**

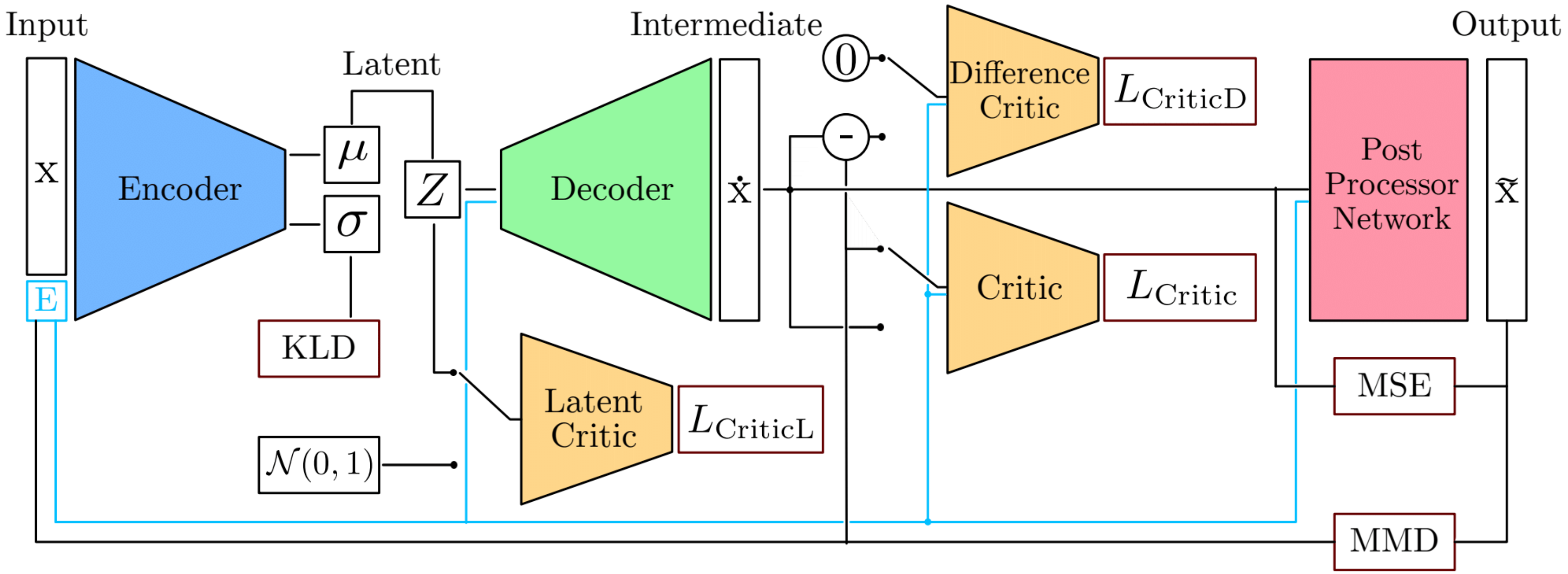
Architectures: BIB-AE

Bounded-Information Bottleneck Autoencoder (BIB-AE)- Pions

- Unifies features of both GANs and VAEs
- Post-Processor network: Improve per-pixel energies; second training
- Multi-dimensional KDE sampling: better modeling of latent space

Voloshynovskiy et. al: **Information bottleneck through variational glasses**, arXiv:1912.00830

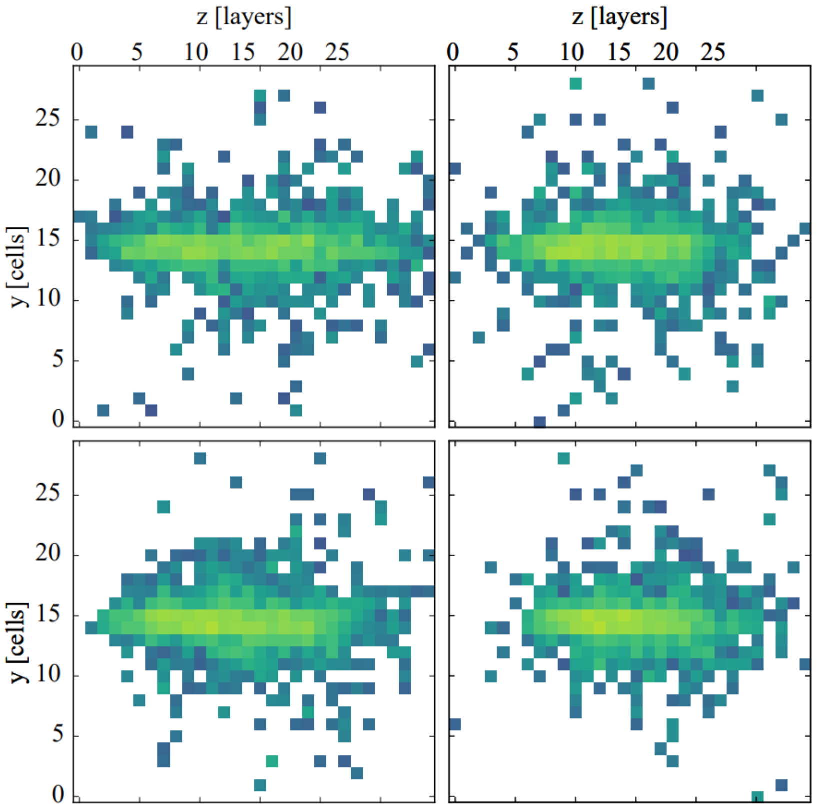
Buhmann et. al: **Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed**, CSBS 5, 13 (2021)

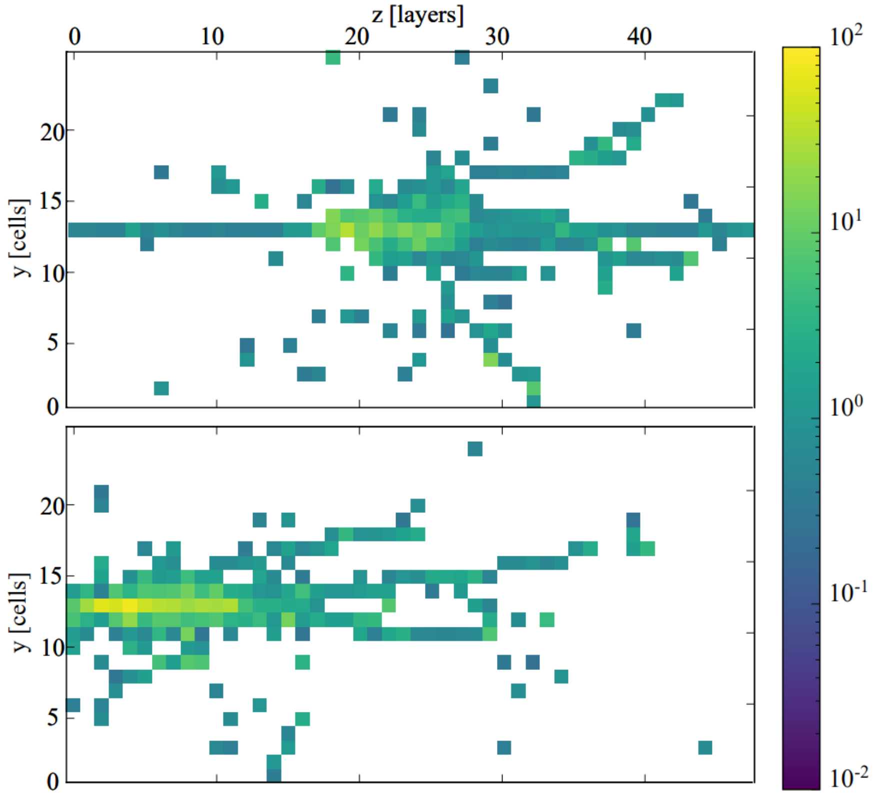


$$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}} + L_{\mathrm{CriticD}}$$

$$L_{\mathrm{Post}} = \mathrm{MMD} + \mathrm{MSE}$$

From Photons to Pions





Photon showers

- Predominantly governed by EM interactions
- Homogeneous structure → **Easy to generalise**

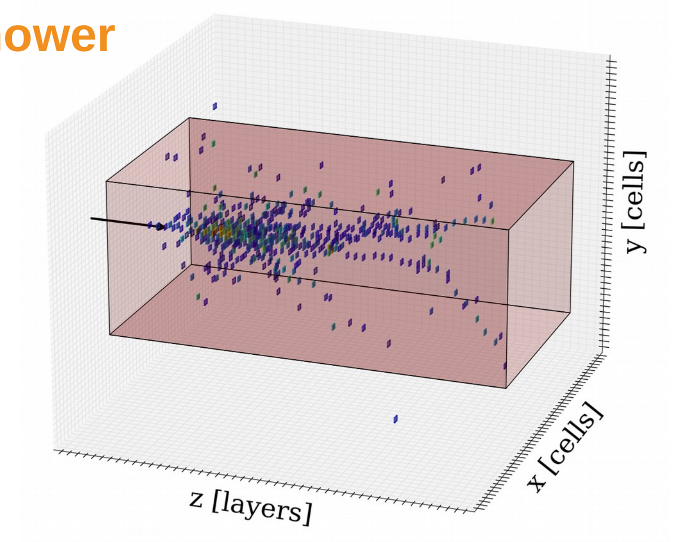
Pion showers

- Hadronic and EM interactions
- Complex structure
- Large event-to-event fluctuations

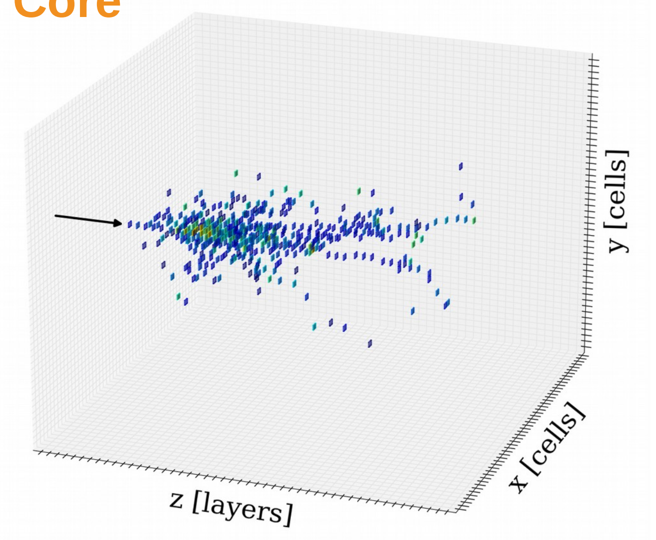
→ **Hard to learn**

Pion dataset

Full Shower



Shower Core

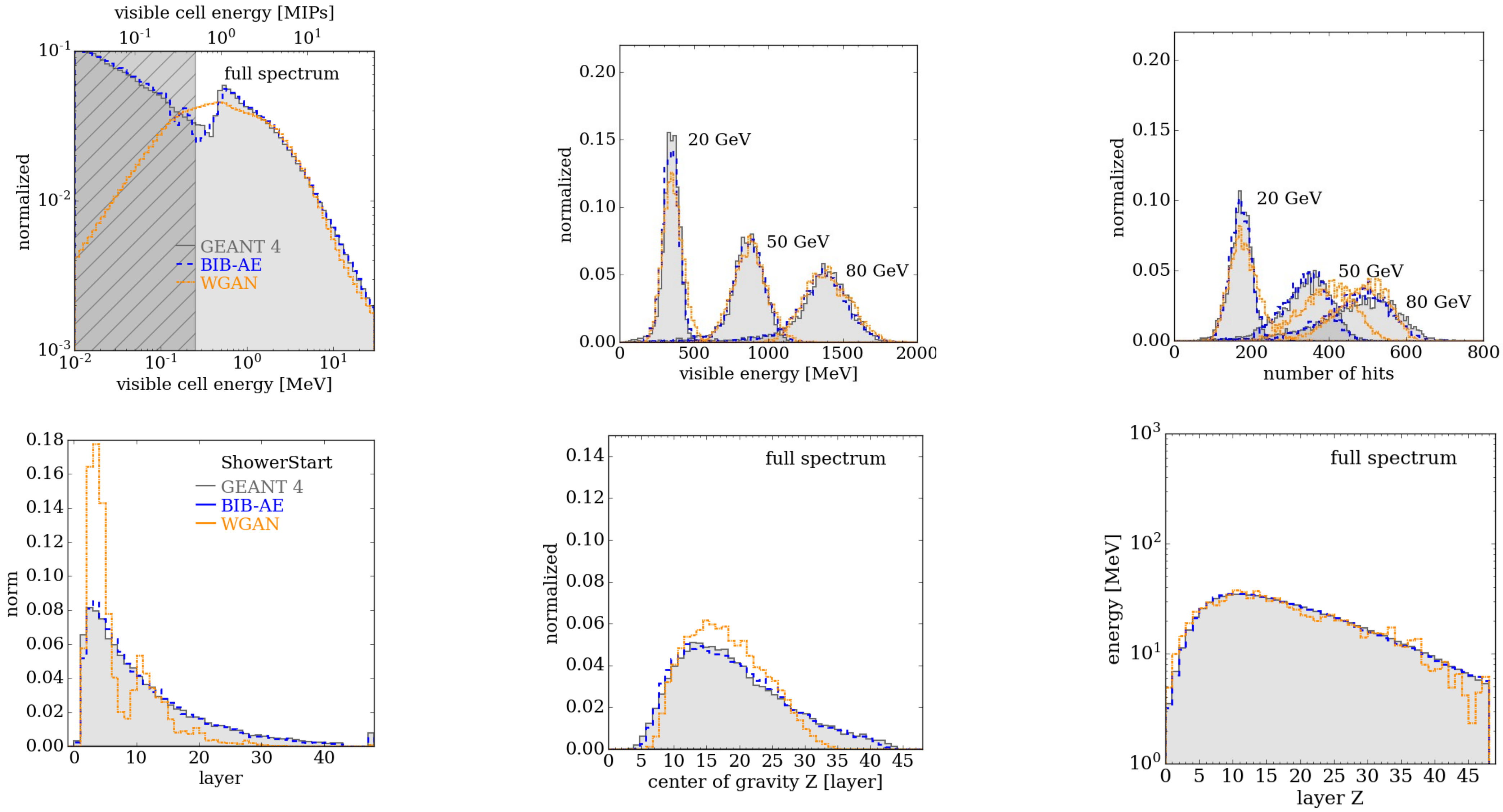


- AHCAL Option
- Remove ECal from geometry

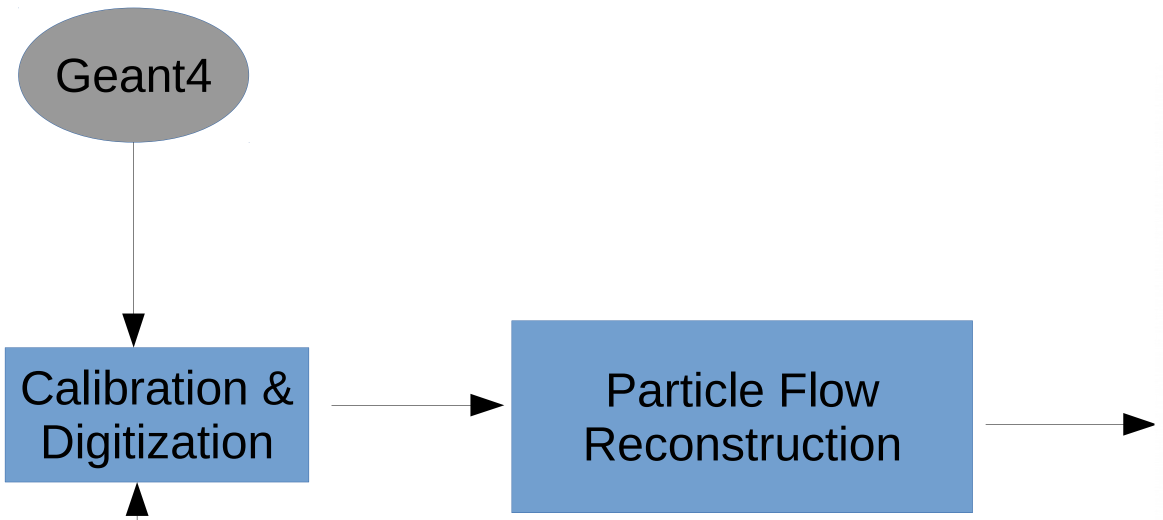
- Significant sparsity in data
  - Use shower core
  - Barely lose any hits

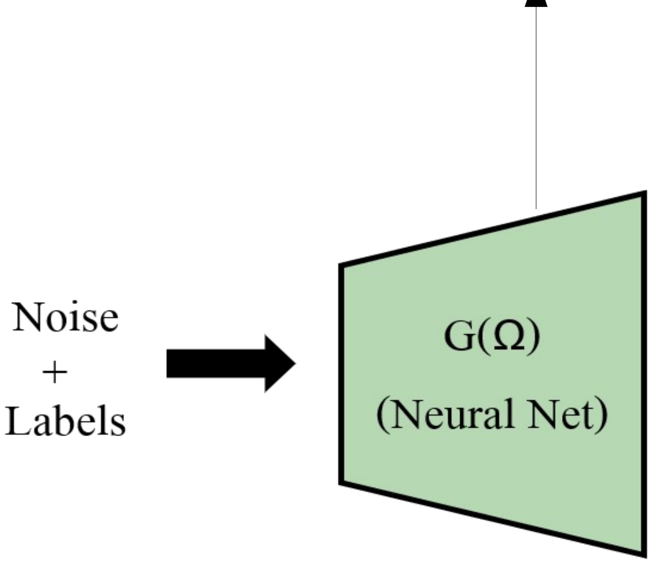
- 500k showers
- Fixed incident point and angle
- Irregular geometry projected into 25x25x48 regular grid
- Uniform energy: 10-100 GeV

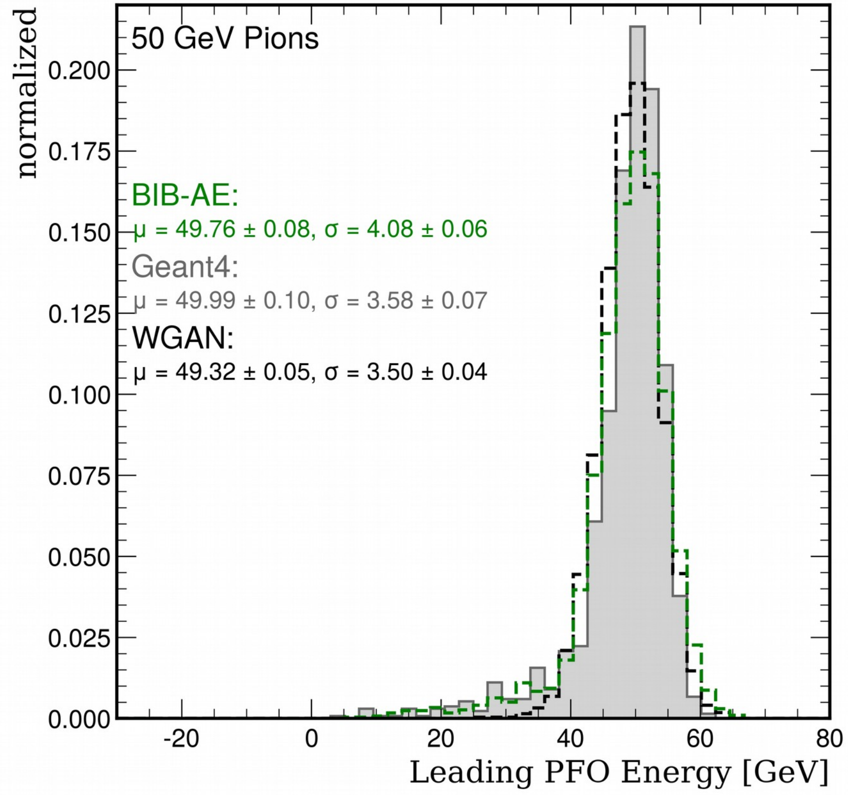
Pion Showers: Sim Level Results



Pion Reconstruction

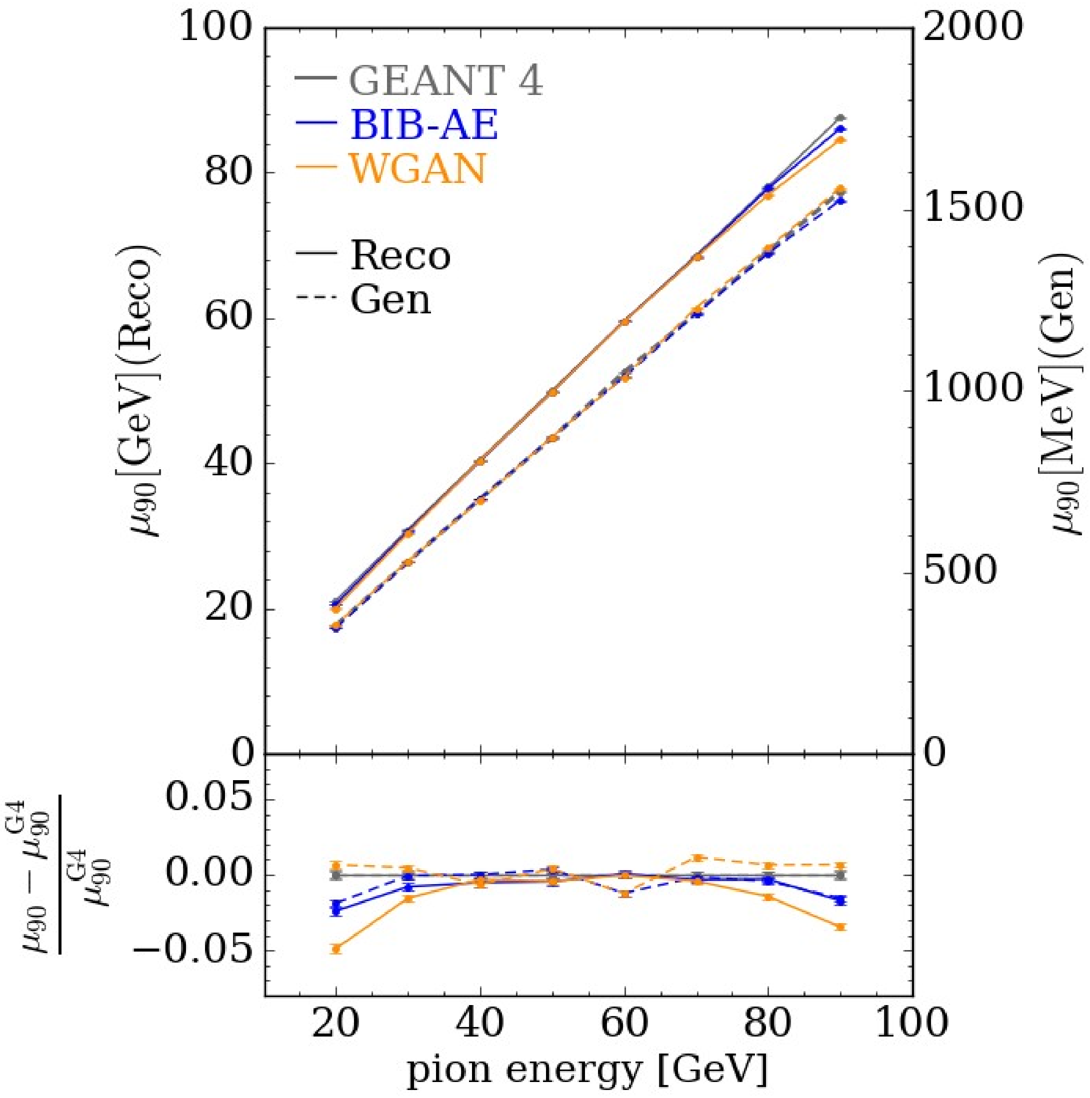


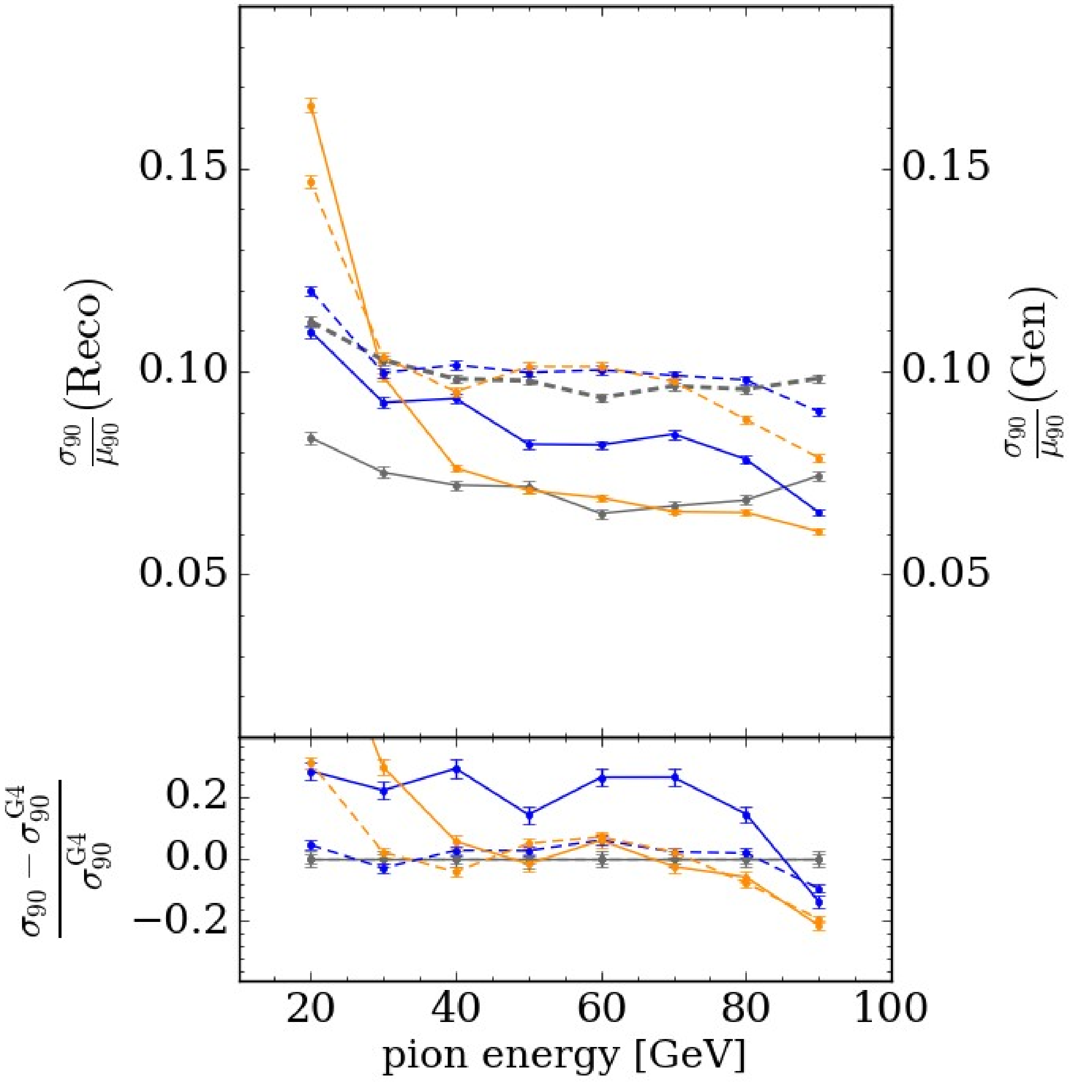




Final performance measure for networks will be physics performance after reconstruction!

Pion Showers: Linearity and resolution





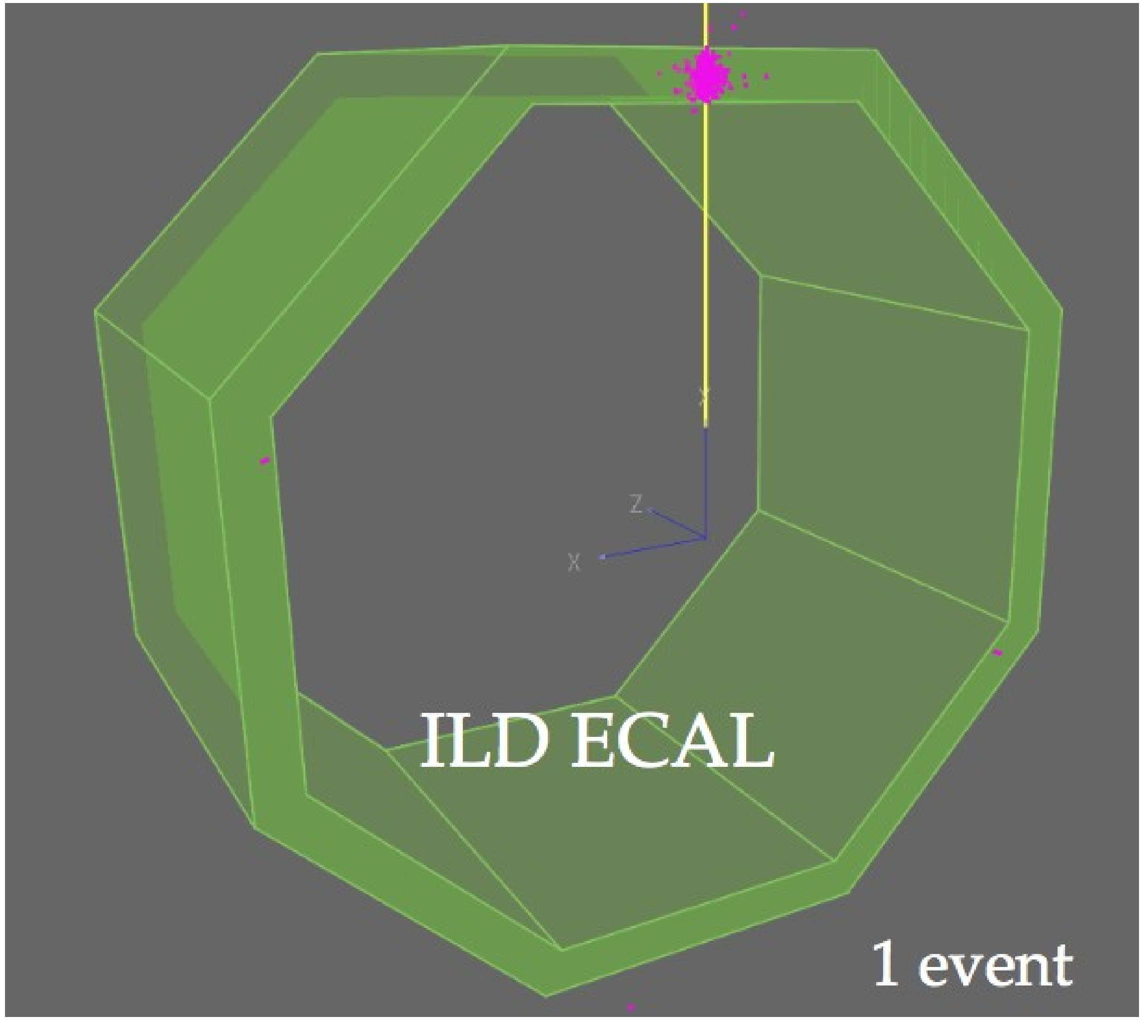
Pion Showers: Computing Time for Inference

Hardware	Simulator	Time / Shower [ms]	Speed-up
CPU	GEANT4	2684 ± 125	$\times 1$
	WGAN	47.923 ± 0.089	$\times 56$
	BIB-AE	350.824 ± 0.574	$\times 8$
GPU	WGAN	0.264 ± 0.002	$\times 10167$
	BIB-AE	2.051 ± 0.005	$\times 1309$

Speed-up of as much as four orders of magnitude on single core of Intel® Xeon® CPU E5-2640 v4 and NVIDIA® A100 for batch size 10000

Conditioning requirements for a general simulation

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



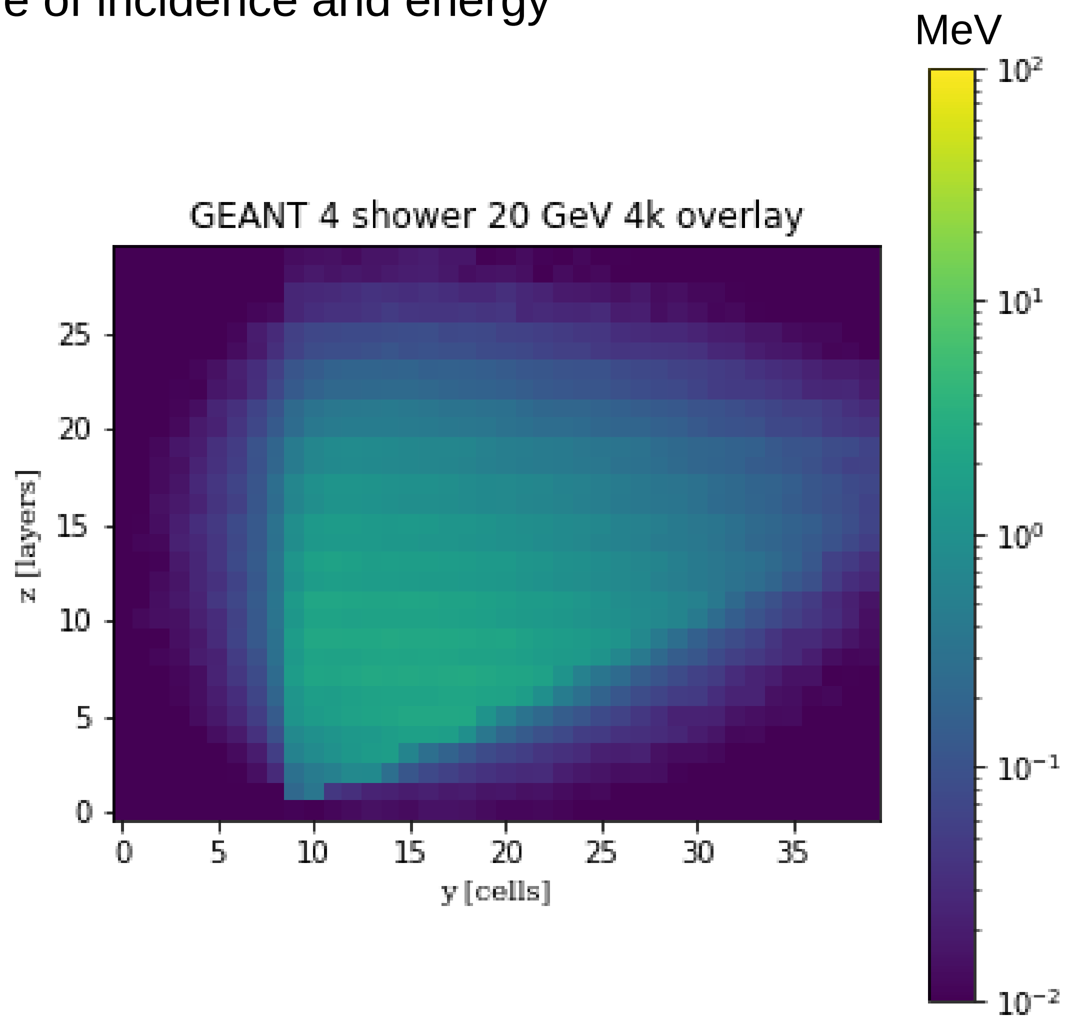
Angular conditioning- Training data

- 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

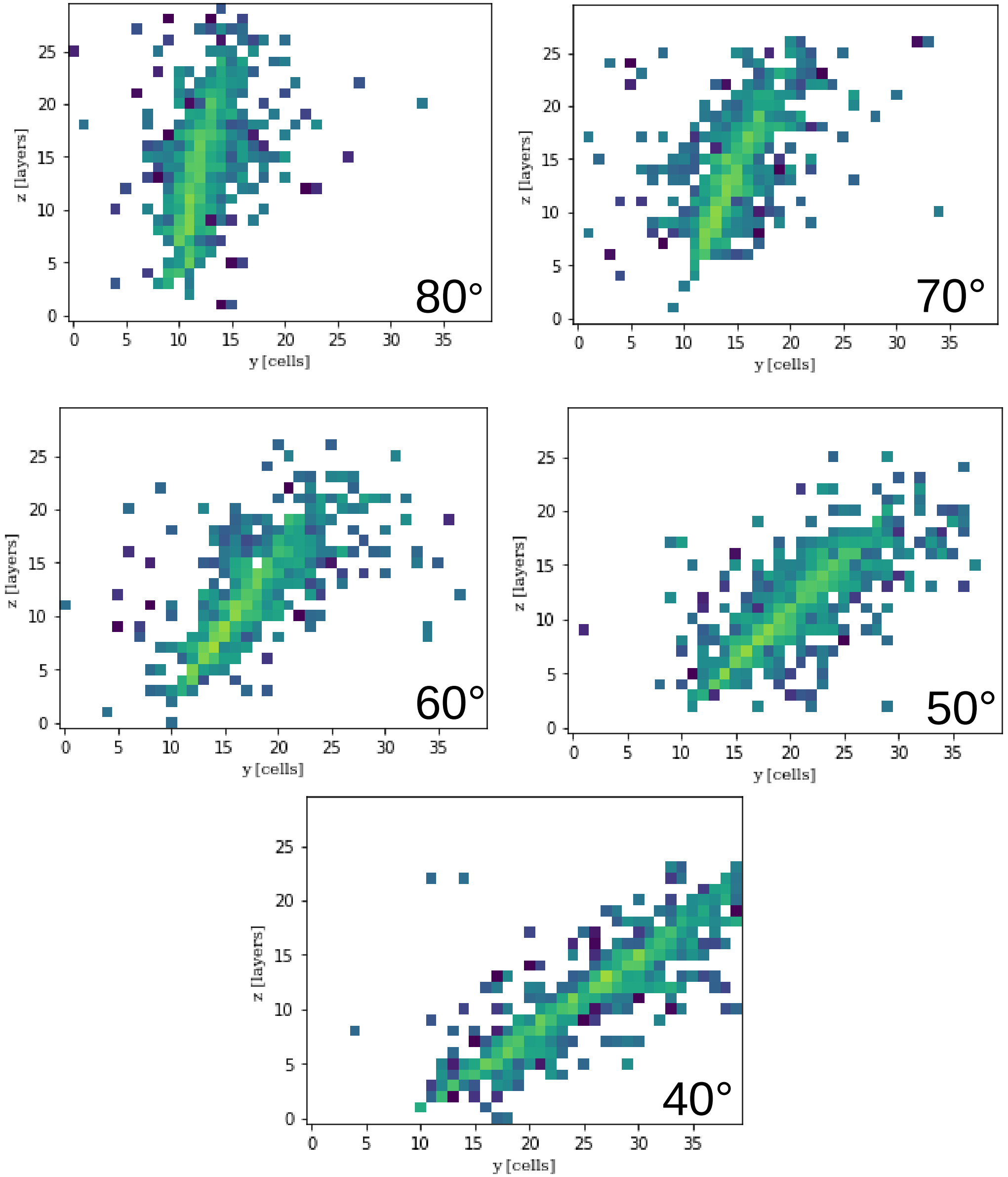
- Problem: How to make sure the full shower is contained?
  - Extend the selected grid in y: shape (30,30,40) (z,x,y)
  - Shift gun position

- Using 132k showers for training

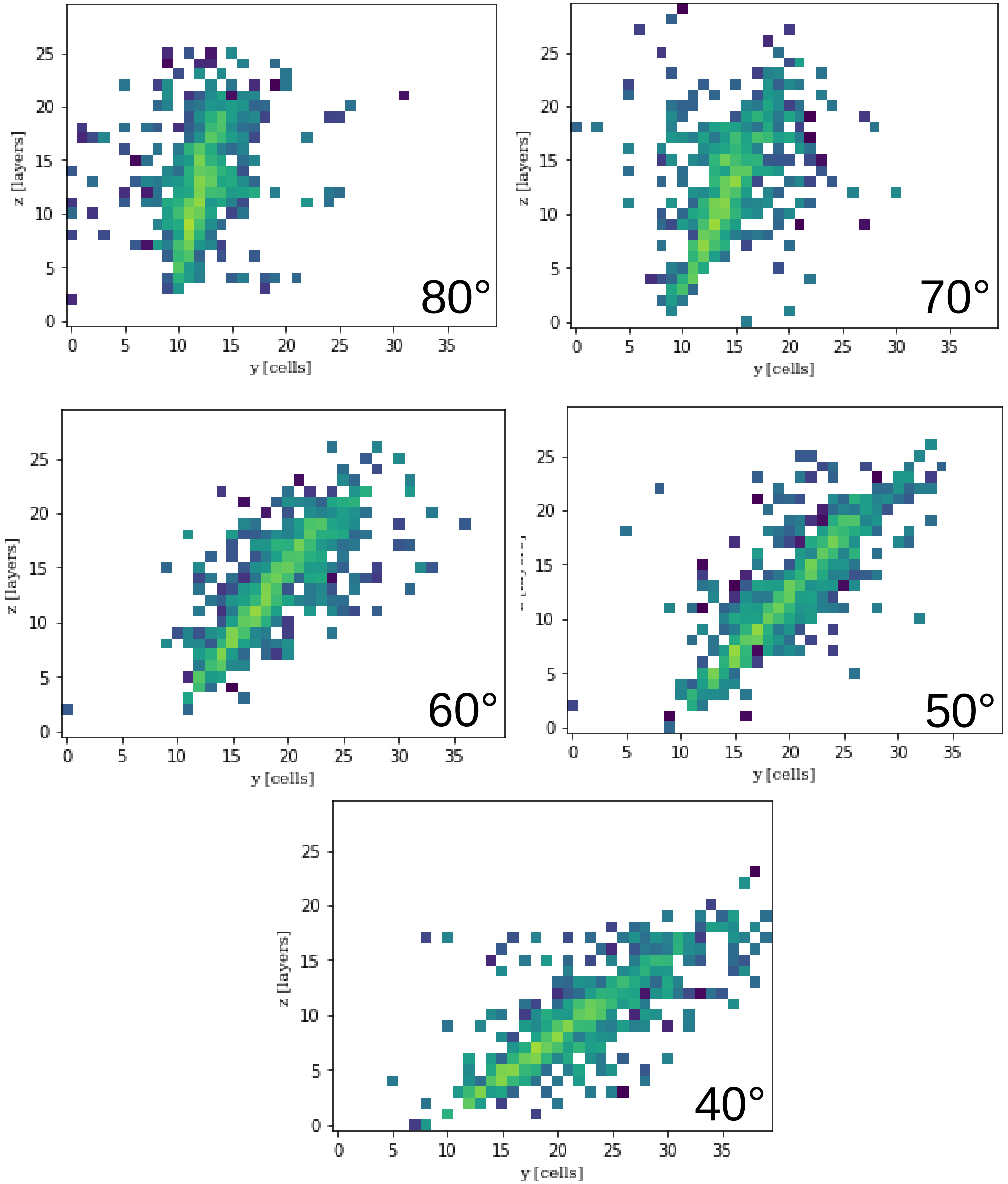


Angular conditioning- Preliminary results

GEANT4

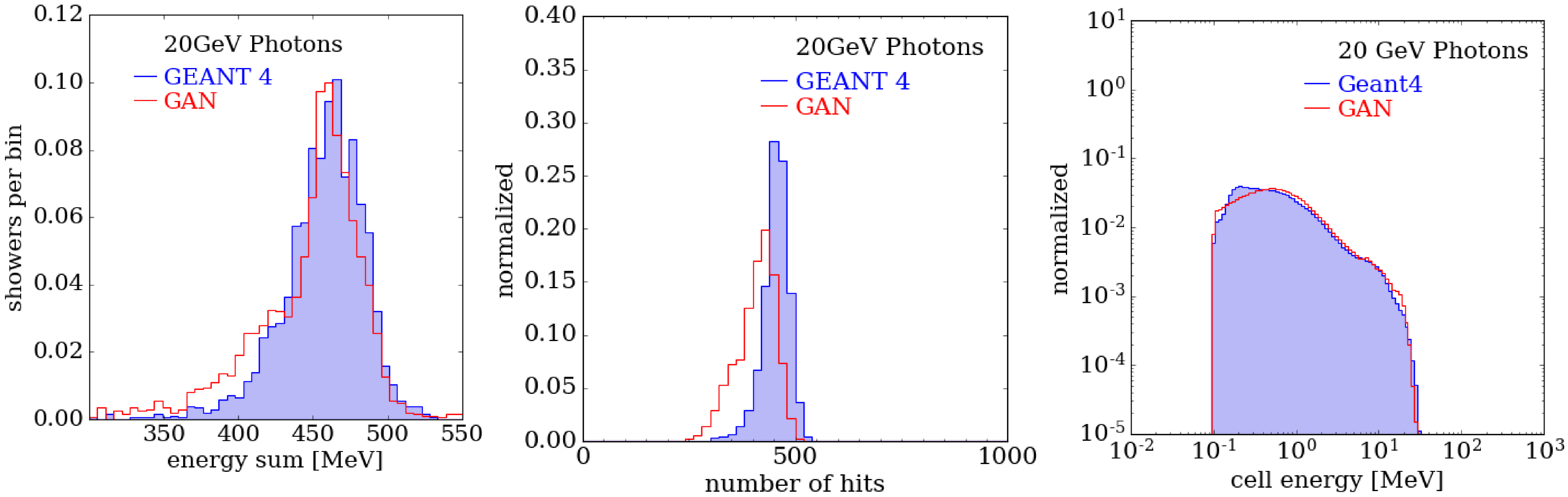


GAN

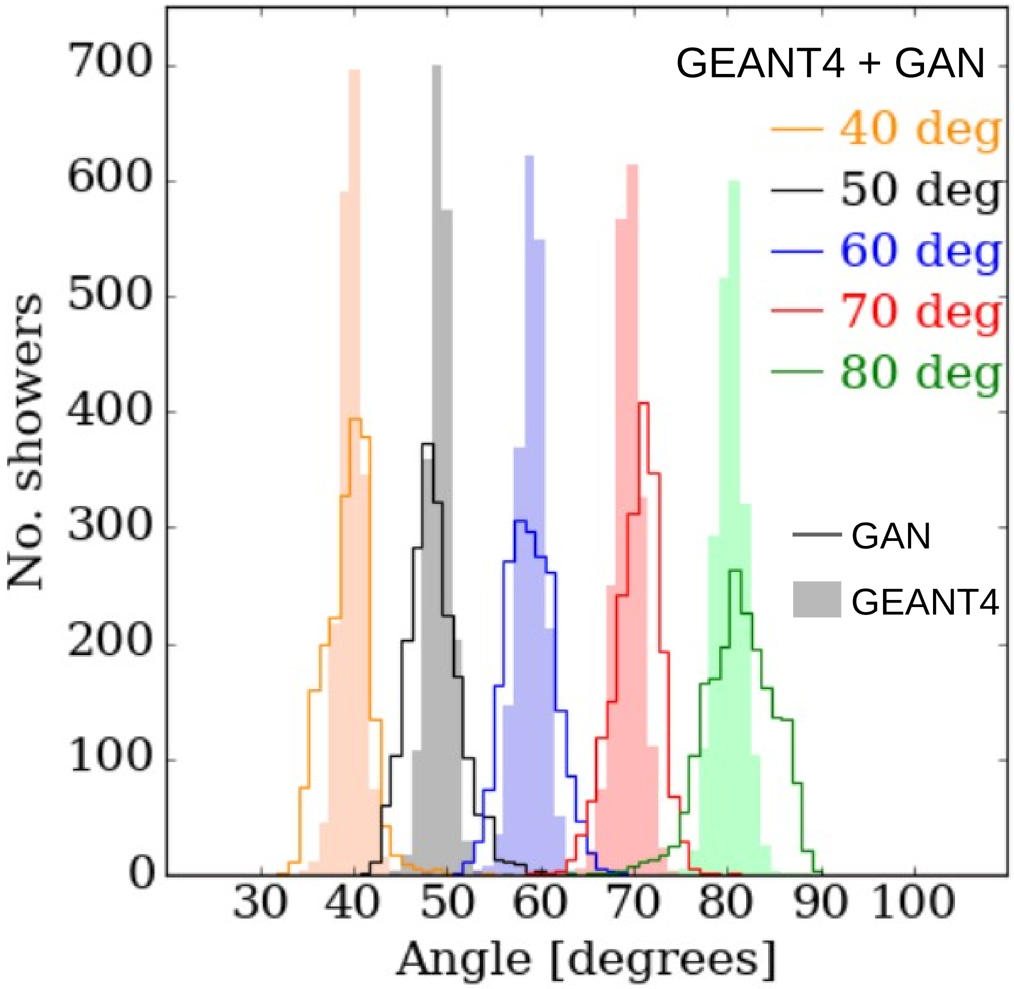


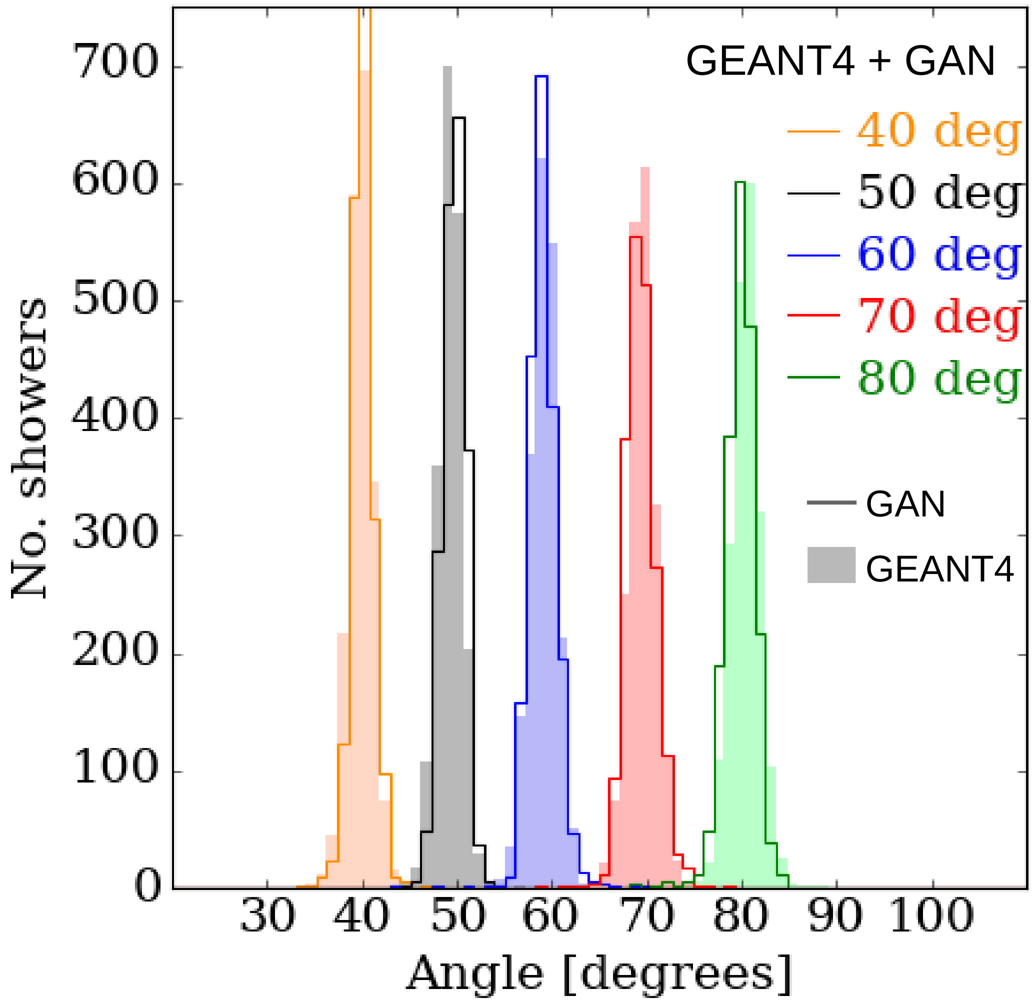
Angular conditioning- Some physics distributions

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



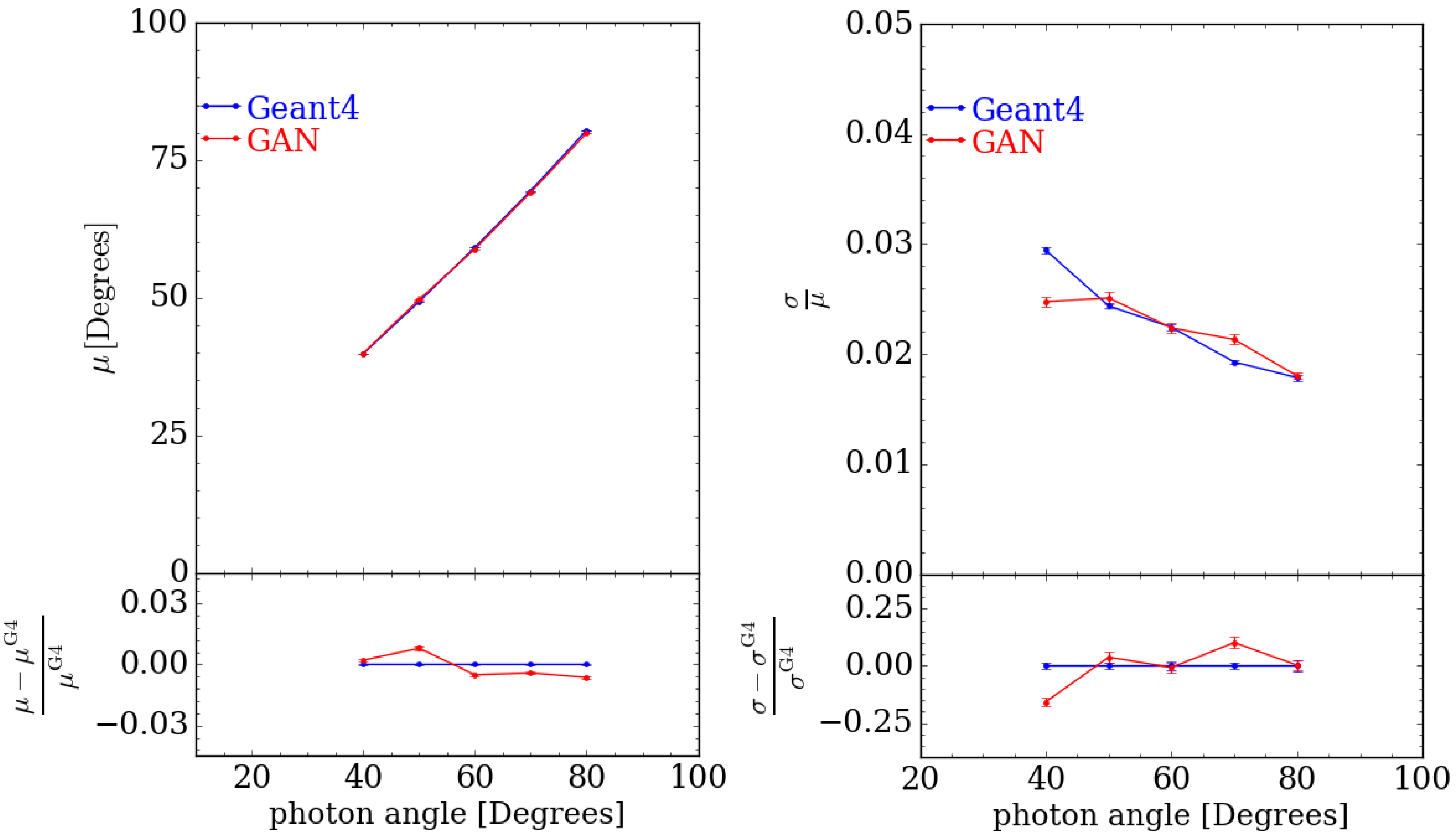
Angular conditioning- With a Constrainer Network





Angular linearity and resolution

- Good overall agreement



Conclusion

Achieved

- Generative models hold promise for fast simulation of calorimeter showers with high fidelity
- Demonstrated high fidelity simulation of hadronic showers with generative models
- Demonstrated angular and energy conditioning in a GAN architecture

Ongoing Work

- Vary energy and angle simultaneously and study effect on performance
- Incorporate angular conditioning in more sophisticated architectures e.g. BIB-AE

Next Steps

- Simulation of hadronic showers including HCAL and ECAL

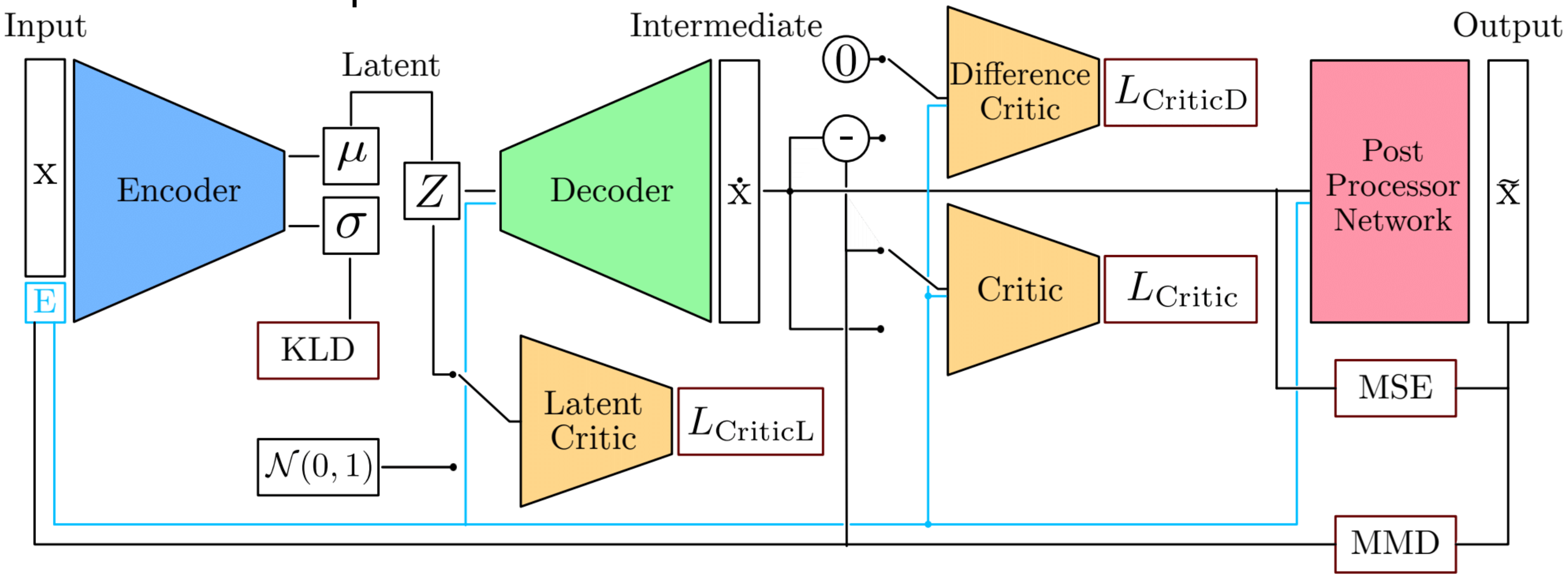
Backup

Architectures: BIB-AE

More Details

- Unifies features of both GANs and VAEs
- Adversarial critic networks rather than pixel-wise difference a la VAEs
- Improved latent regularisation: additional critic and MMD term
- Post-Processor network: Improve per-pixel energies; second training

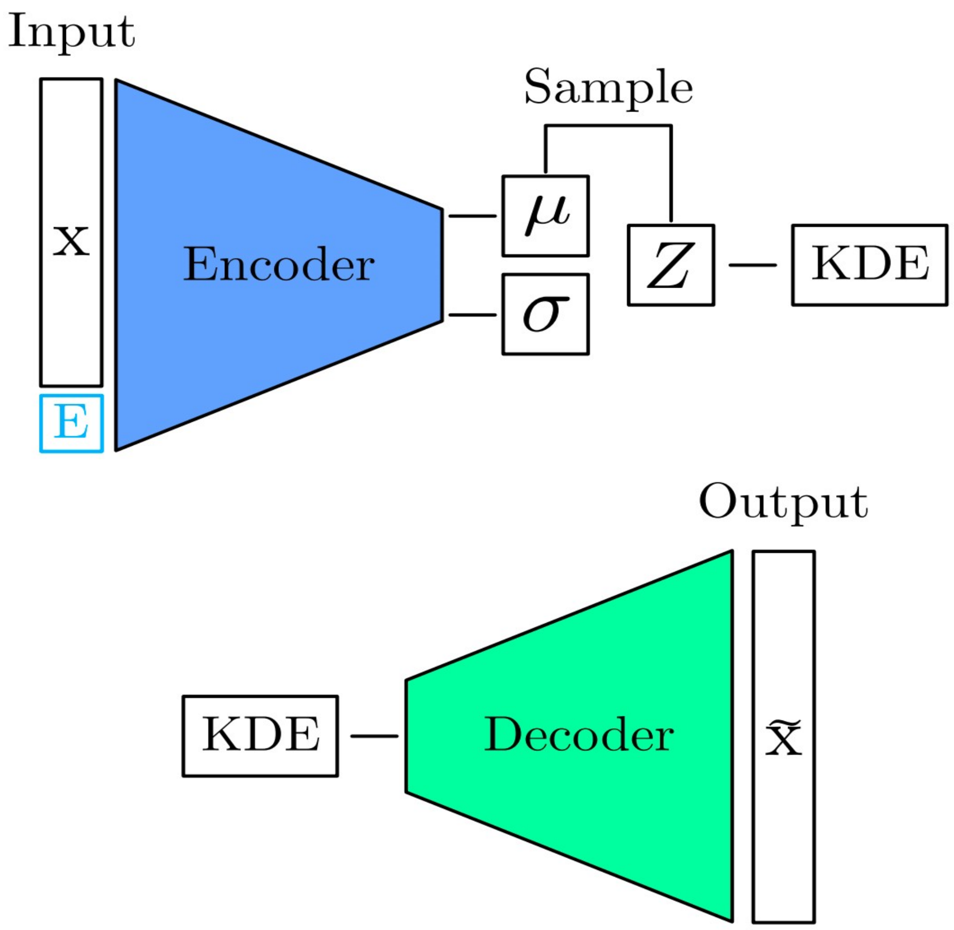
- Updates and improvements:
  - Dual and resetting critics: prevent artifacts caused by sparsity
  - Batch Statistics: prevent outliers/ mode collapse
  - Multi-dimensional KDE sampling: better modeling of latent space

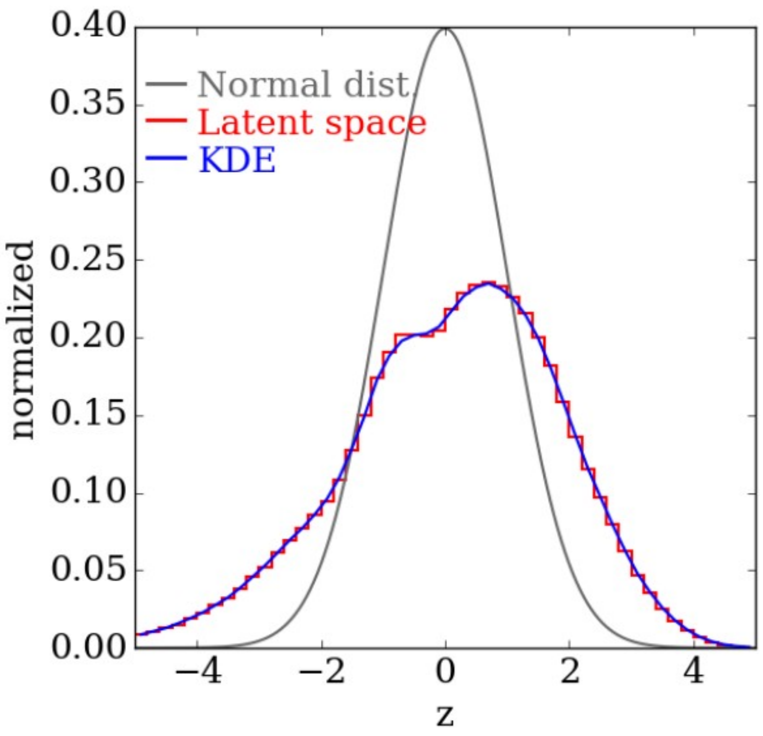


$$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}} + L_{\mathrm{CriticD}}$$

$$L_{\mathrm{Post}} = \mathrm{MMD} + \mathrm{MSE}$$

Kernel Density Estimation: BIB-AE





Buhmann et. al: **Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network**, EPJ Web of Conferences 251, 03003 (2021)

Pion correlations

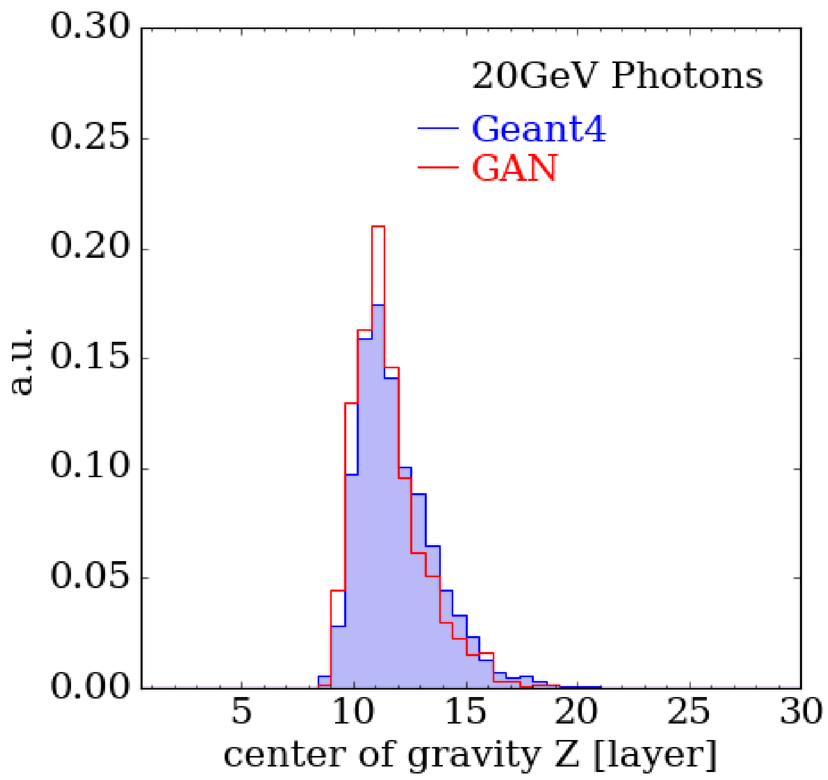
GEANT4 - BIB-AE

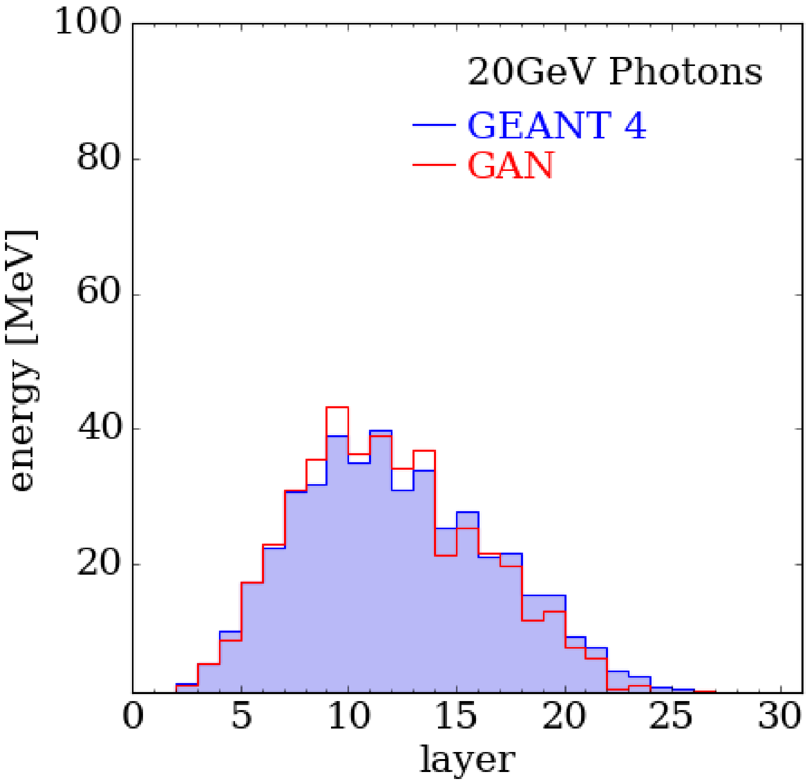
	$m_{1,x}$	$m_{1,y}$	$m_{1,z}$	$m_{2,x}$	$m_{2,y}$	$m_{2,z}$	E_{vis}	E_{inc}	n_{hit}	E_1/E_{vis}	E_2/E_{vis}	E_3/E_{vis}
$m_{1,x}$	0.00											
$m_{1,y}$	-0.00	0.00										
$m_{1,z}$	-0.01	-0.04	0.00									
$m_{2,x}$	-0.08	-0.00	-0.06	0.00								
$m_{2,y}$	-0.10	-0.03	-0.05	0.01	0.00							
$m_{2,z}$	-0.06	0.01	-0.06	-0.08	-0.05	0.00						
E_{vis}	0.03	-0.02	-0.01	0.09	0.09	0.06	0.00					
E_{inc}	0.01	-0.03	-0.00	0.08	0.09	0.06	-0.01	0.00				
n_{hit}	0.03	-0.02	-0.02	0.13	0.14	0.06	0.00	-0.01	0.00			
E_1/E_{vis}	0.00	0.03	0.00	0.04	0.04	0.04	0.01	0.00	0.02	0.00		
E_2/E_{vis}	-0.01	-0.00	-0.03	0.02	-0.02	0.01	-0.02	-0.02	-0.01	0.02	0.00	
E_3/E_{vis}	-0.01	-0.04	0.00	-0.07	-0.04	-0.07	0.00	0.01	-0.01	-0.00	-0.03	0.00

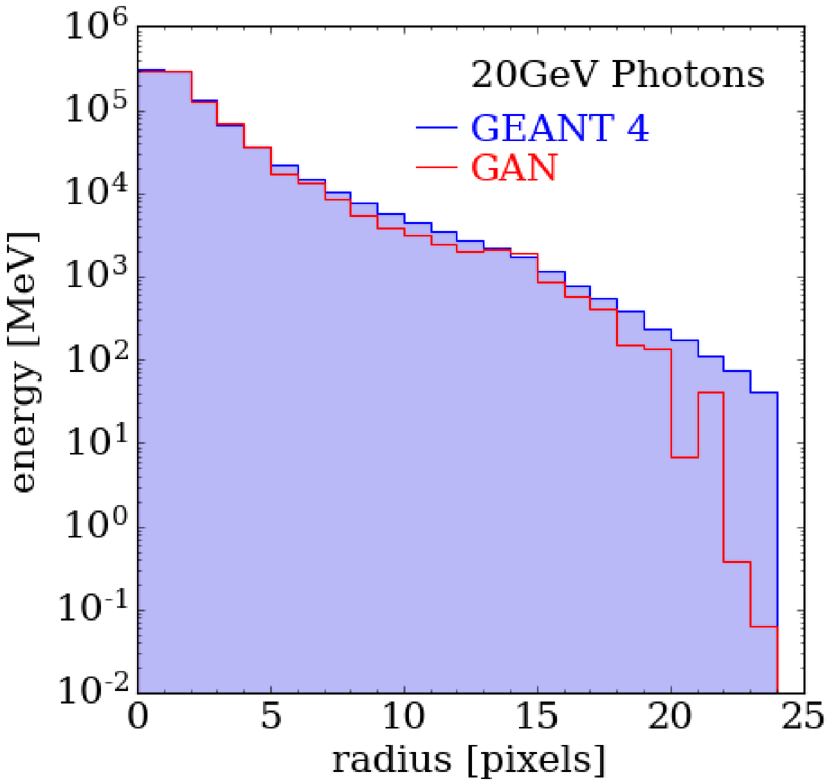
GEANT4 - WGAN

	$m_{1,x}$	$m_{1,y}$	$m_{1,z}$	$m_{2,x}$	$m_{2,y}$	$m_{2,z}$	E_{vis}	E_{inc}	n_{hit}	E_1/E_{vis}	E_2/E_{vis}	E_3/E_{vis}
$m_{1,x}$	0.00											
$m_{1,y}$	-0.21	0.00										
$m_{1,z}$	0.20	-0.10	0.00									
$m_{2,x}$	-0.14	-0.17	-0.20	0.00								
$m_{2,y}$	-0.20	-0.28	-0.21	-0.18	0.00							
$m_{2,z}$	-0.14	-0.24	-0.17	-0.20	-0.23	0.00						
E_{vis}	0.14	0.19	-0.21	0.11	0.18	0.06	0.00					
E_{inc}	0.11	0.17	-0.13	0.01	0.09	0.03	-0.03	0.00				
n_{hit}	0.10	0.07	-0.26	-0.02	0.05	-0.09	0.04	-0.00	0.00			
E_1/E_{vis}	-0.17	0.05	0.02	0.16	0.15	0.14	0.14	0.10	0.15	0.00		
E_2/E_{vis}	0.11	0.09	-0.28	-0.00	0.03	-0.01	-0.06	-0.12	-0.01	0.15	0.00	
E_3/E_{vis}	0.09	-0.22	0.03	-0.26	-0.29	-0.31	-0.15	-0.06	-0.27	0.03	-0.28	0.00

Angular conditioning- 60 degree shower shape distributions







Angular conditioning- 80 degree other distributions