High Fidelity Simulation of High Granularity Calorimeters with High Speed

ML4Jets meeting
07.07.2021

E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, D. Hundhausen, G. Kasieczka,
W. Korcari, A. Korol, K. Krüger, P. McKeown and L. Rustige

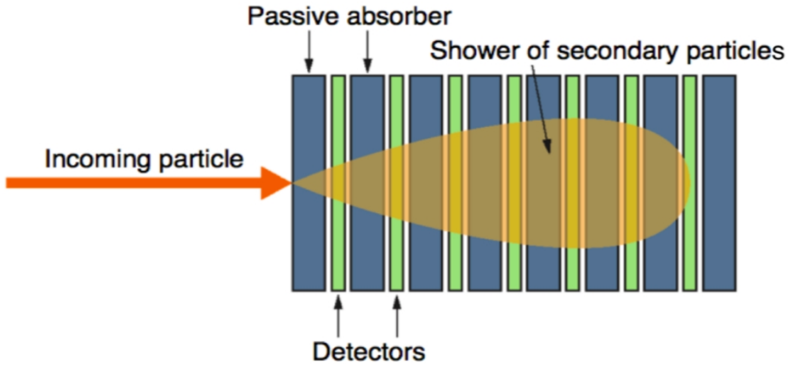
DESY.

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

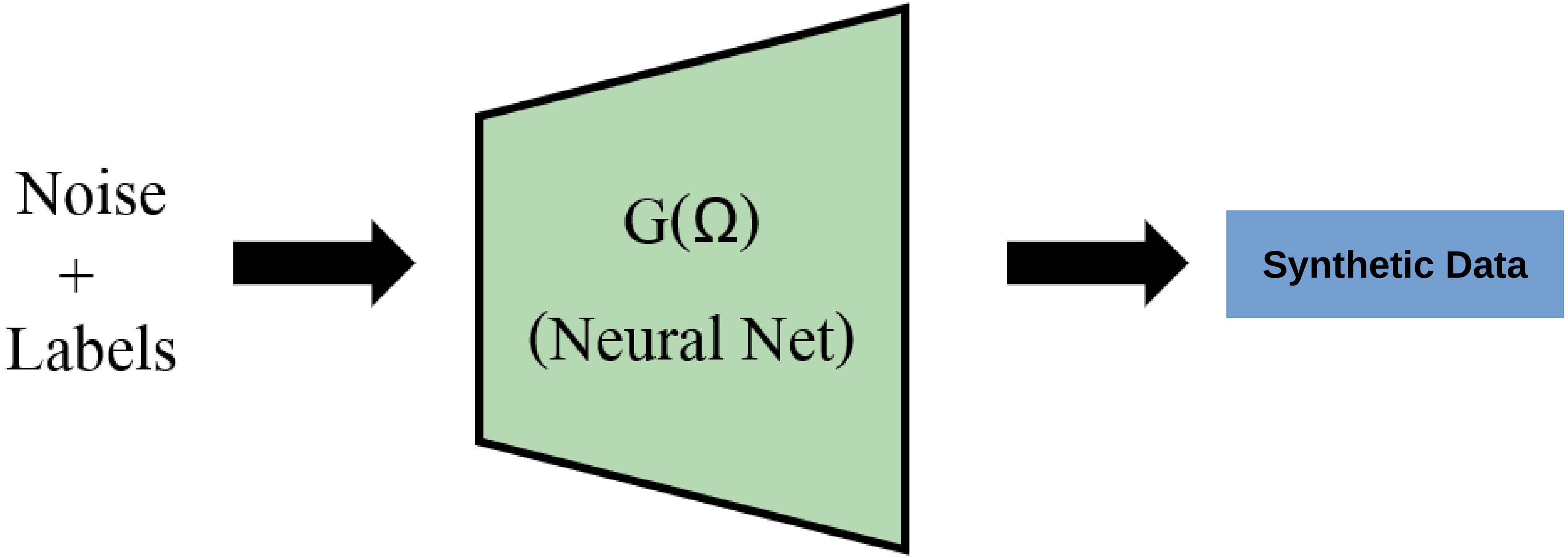
UH Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Deep Generative Models



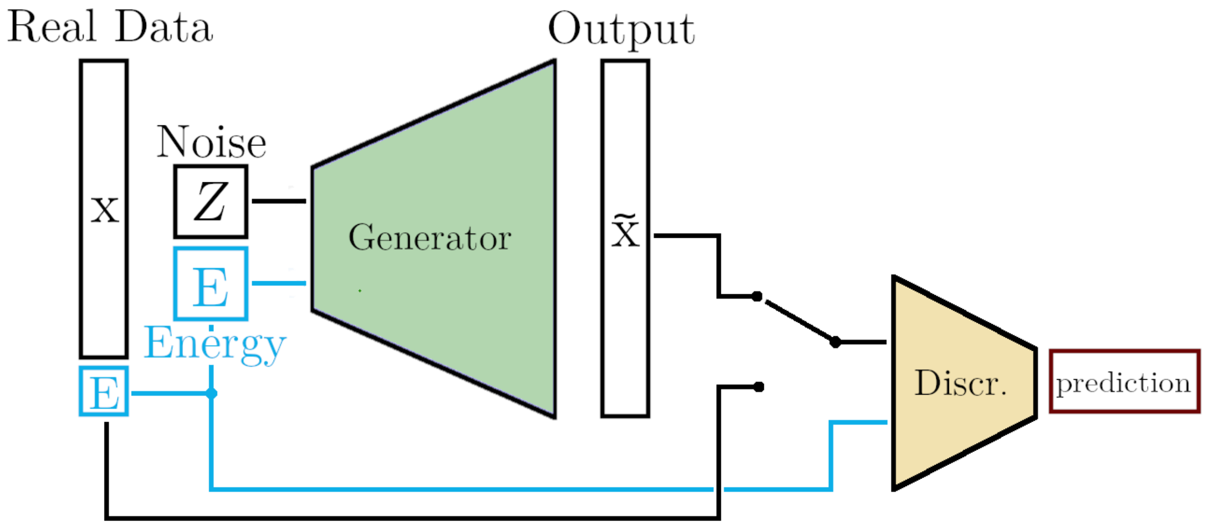
- Calorimeter simulation in HEP is CPU expensive!
- Promising solution for a **fast shower simulation**
  - Generate new samples by following the distribution of original data (i.e Geant4)
  - Map random noise to data
  - Conditioning



Recap: Generative Adversarial Neural Networks (GANs)

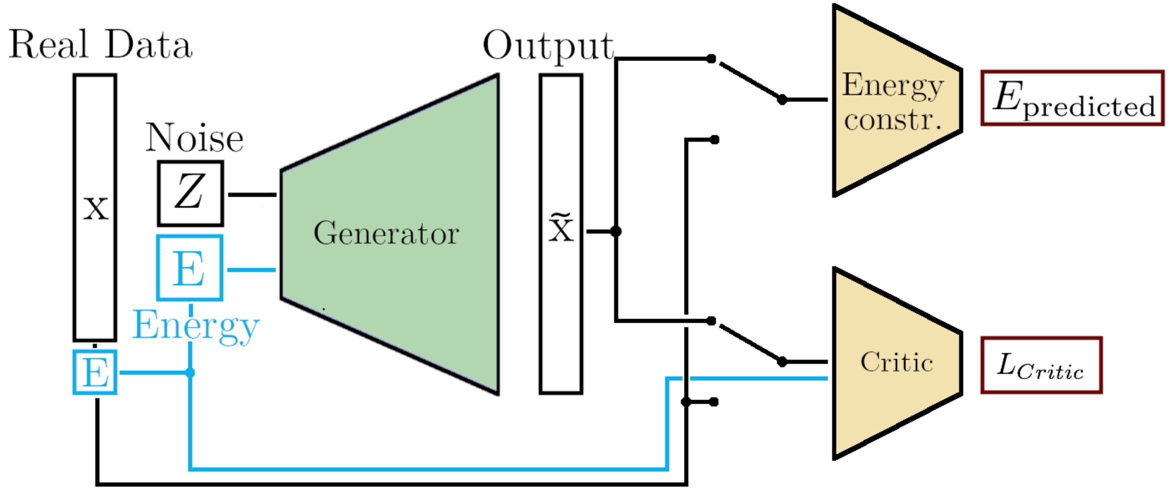
GAN

- First generative architecture used for simulating showers
- Discriminator tries to differentiate: Fake or Real ?
- Generator tries to fool the distriminator
- **New:** Apply mini-batch discrimination (pion)



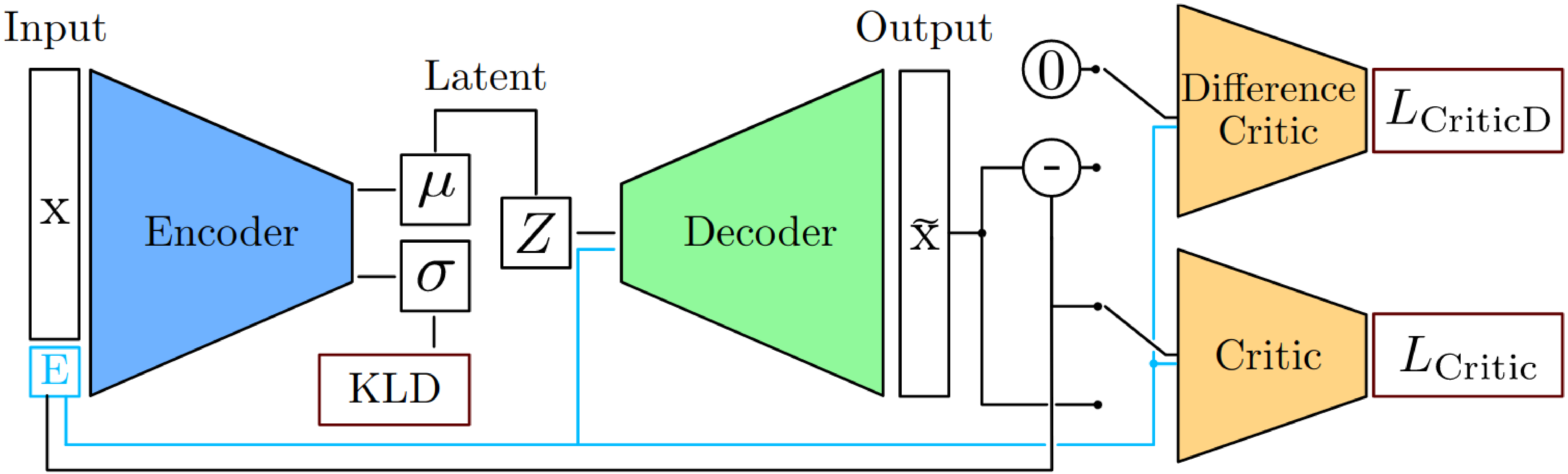
WGAN

- Alternative to classical GAN training:
  Helps improve the stability of the training
  Use Wasserstein-1 distance as a loss with gradient penalty
- Second network to constrain energy
- **New:** Latent optimization method (LO) is employed (pion)



Bounded-Information Bottleneck Autoencoder (BIB-AE)

- Unifies features of GANs and Autoencoders
- WGAN-like critics evaluate the quality of reconstructed images

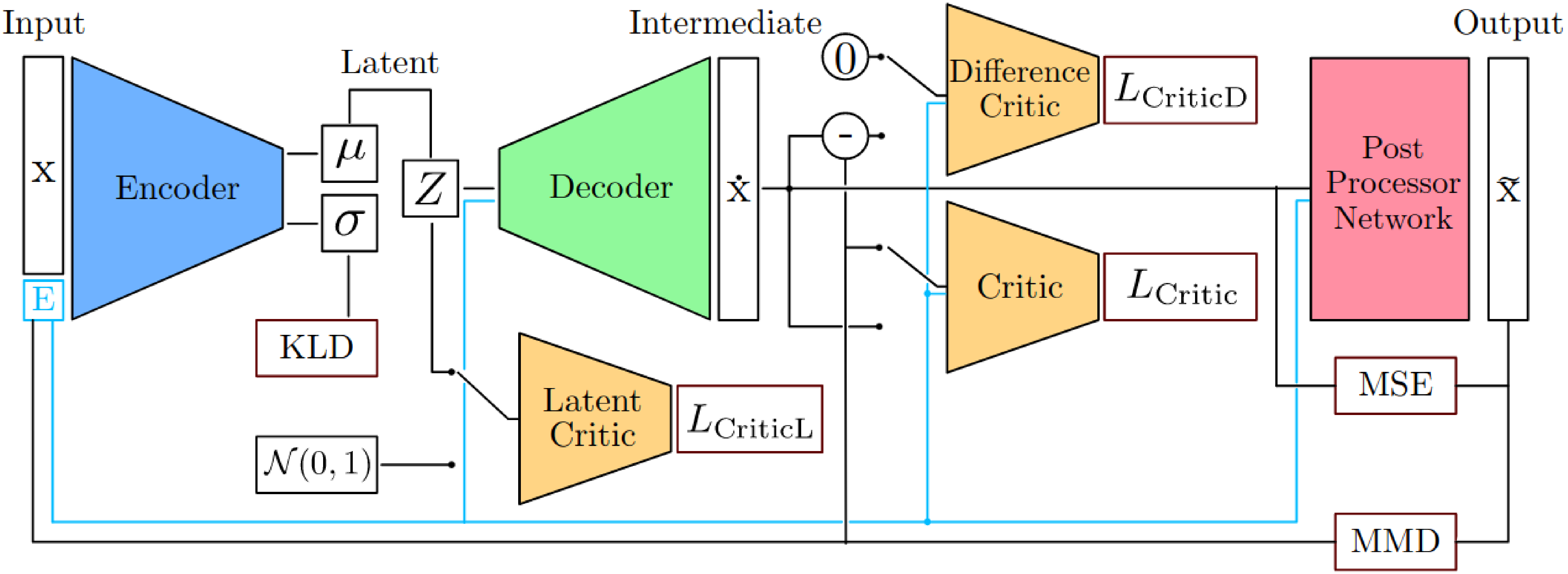


Buhmann, et al.: **Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed.** Comput Softw Big Sci 5, 13 (2021)

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

Bounded-Information Bottleneck Autoencoder (BIB-AE)

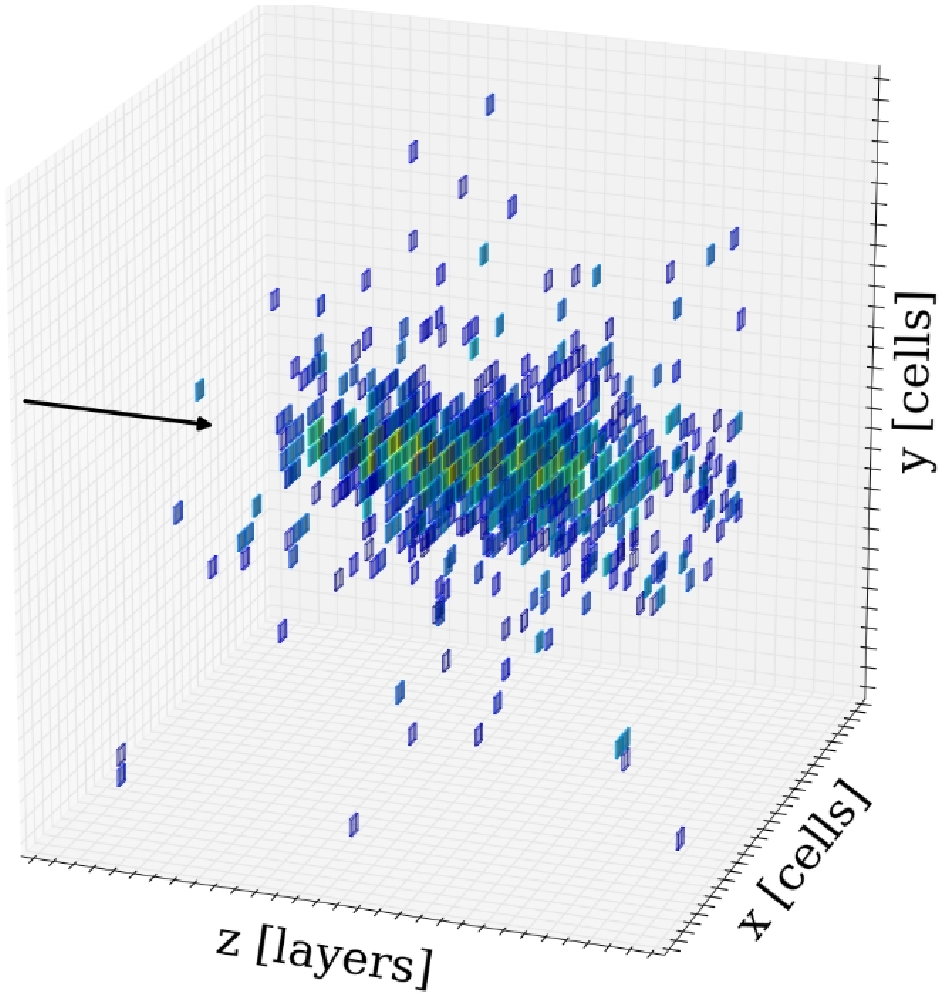
- Latent regularization is improved by an additional critic and a Maximum Mean Discrepancy (MMD) term
- Additional Post-Processor network, trained in a second step, is used to improved per-pixel energies
- **New**: Sampling from encoded latent space via multi-dimensional Kernel Density Estimation (KDE)
- **New**: Batch statistics involved in backward propagation



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

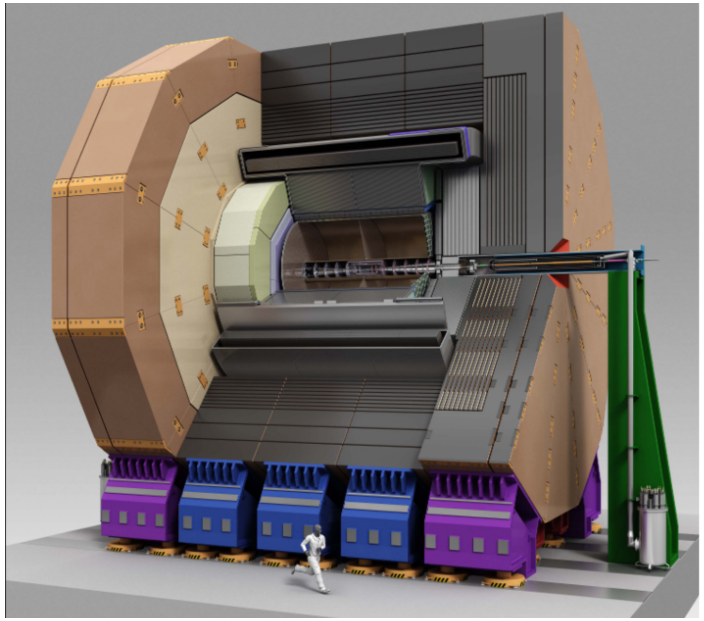
$$L_{\mathrm{Post}} = \mathrm{MMD} + \mathrm{MSE}$$

Photon Dataset



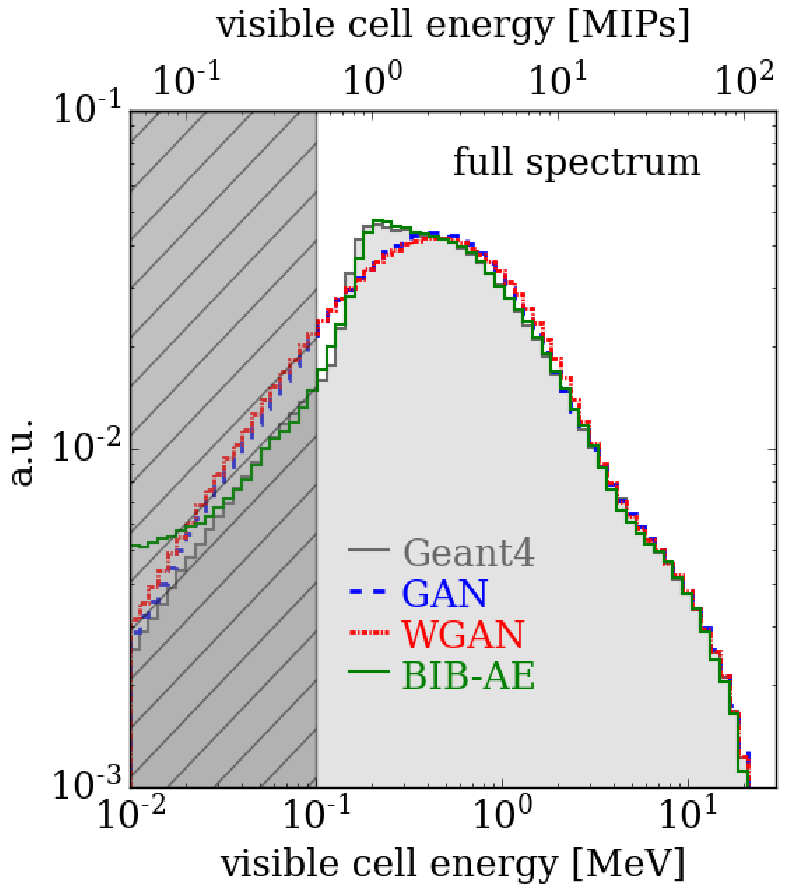
- 950k showers
- Fixed incident point and angle
- Uniform 10 GeV to 100 GeV
- 30x30x30 image

Highly Granular ECAL of ILD

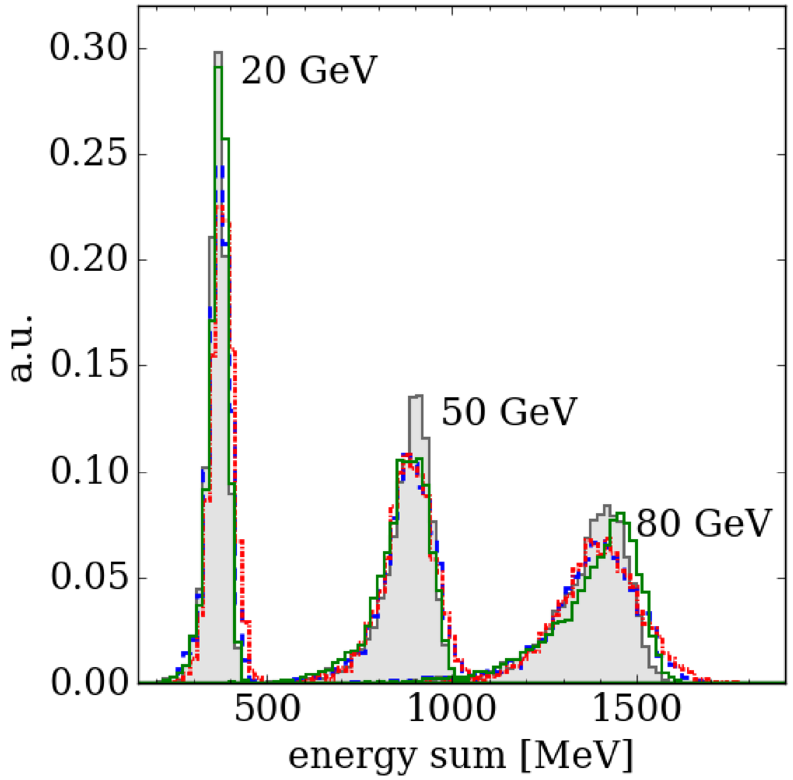


Zenodo Data Repository

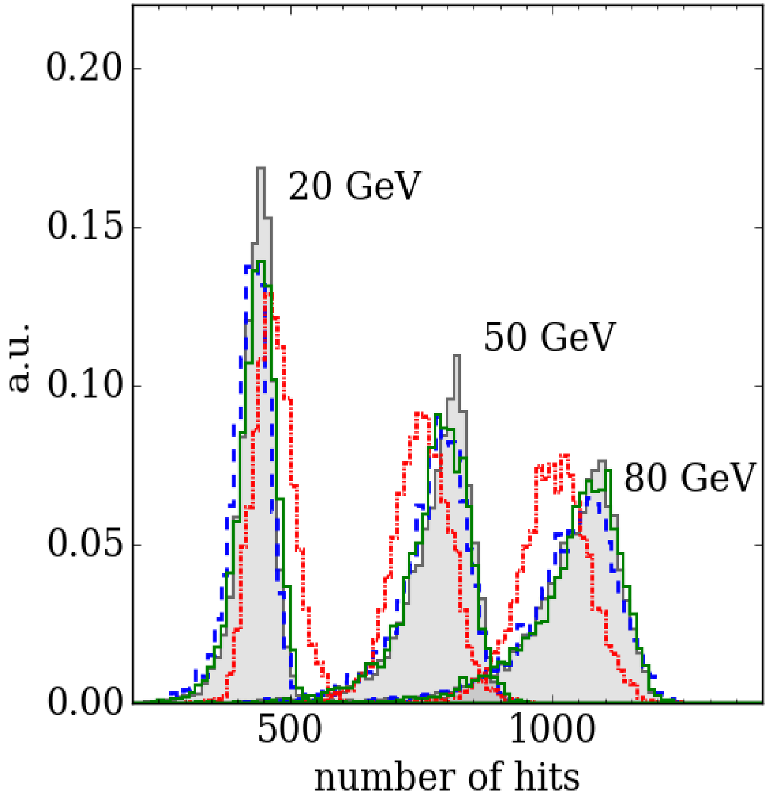
Photon Results



Very good agreement of MIP peak for BIB-AE with Post-Processing!

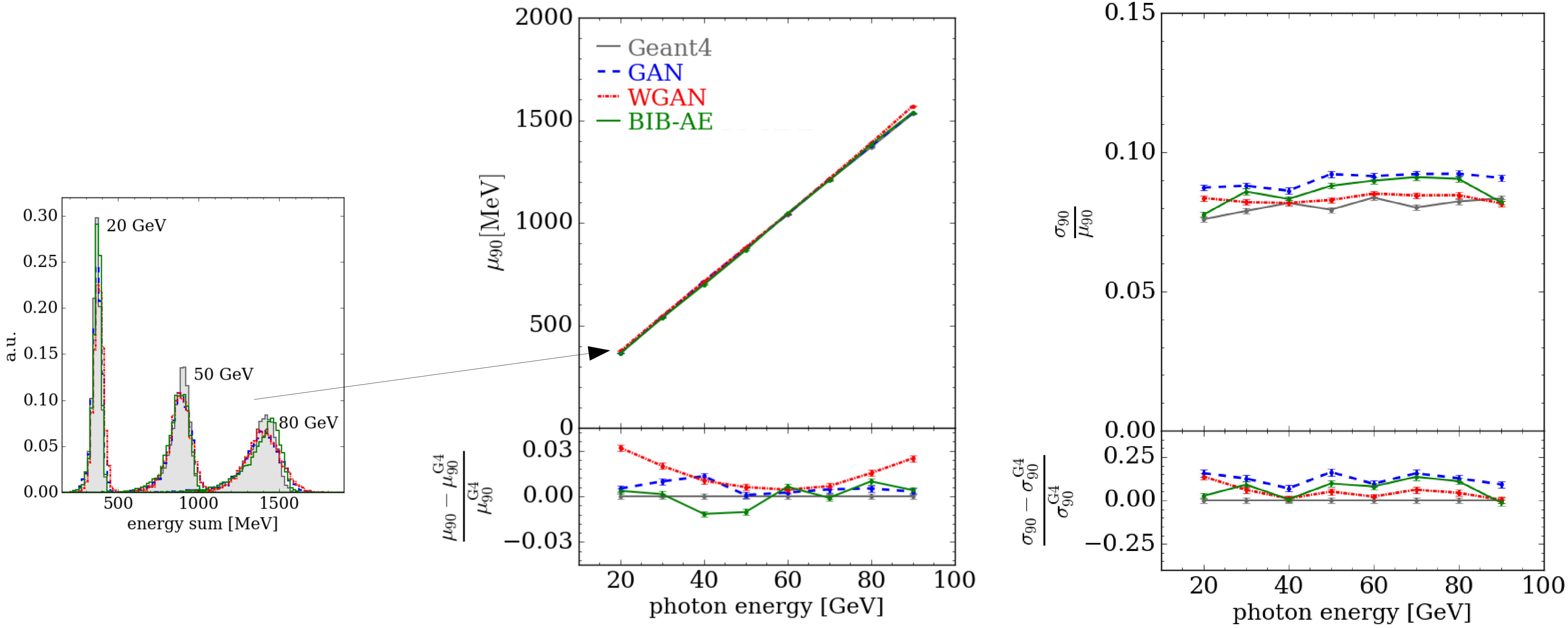


Peak and shape of the energy-sum is nicely reproduced by all models



BIB-AE and GAN correctly model number of hits

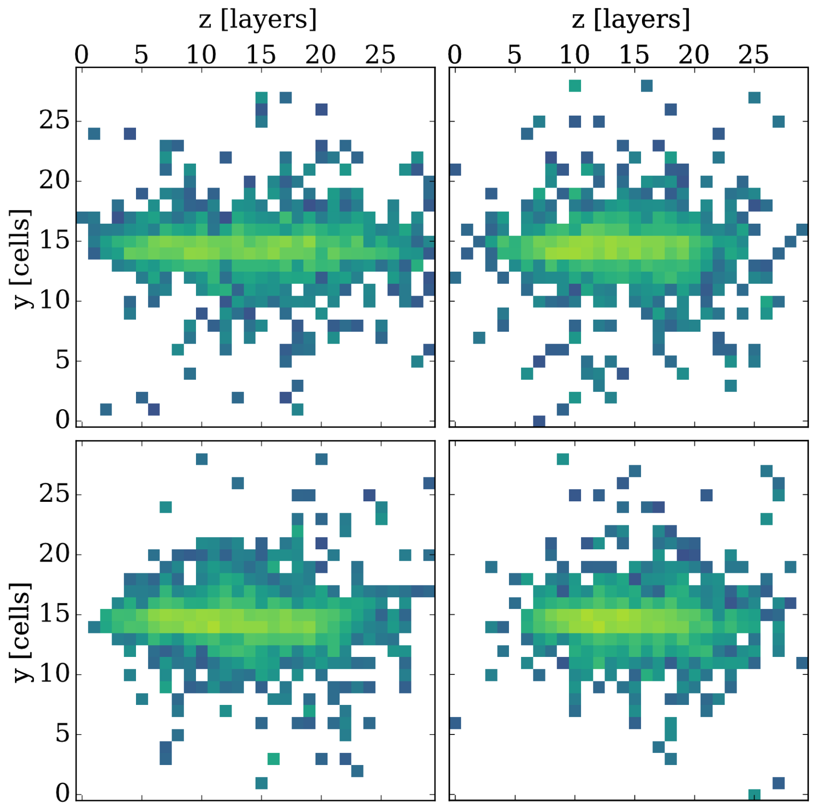
Linearity and Resolution*

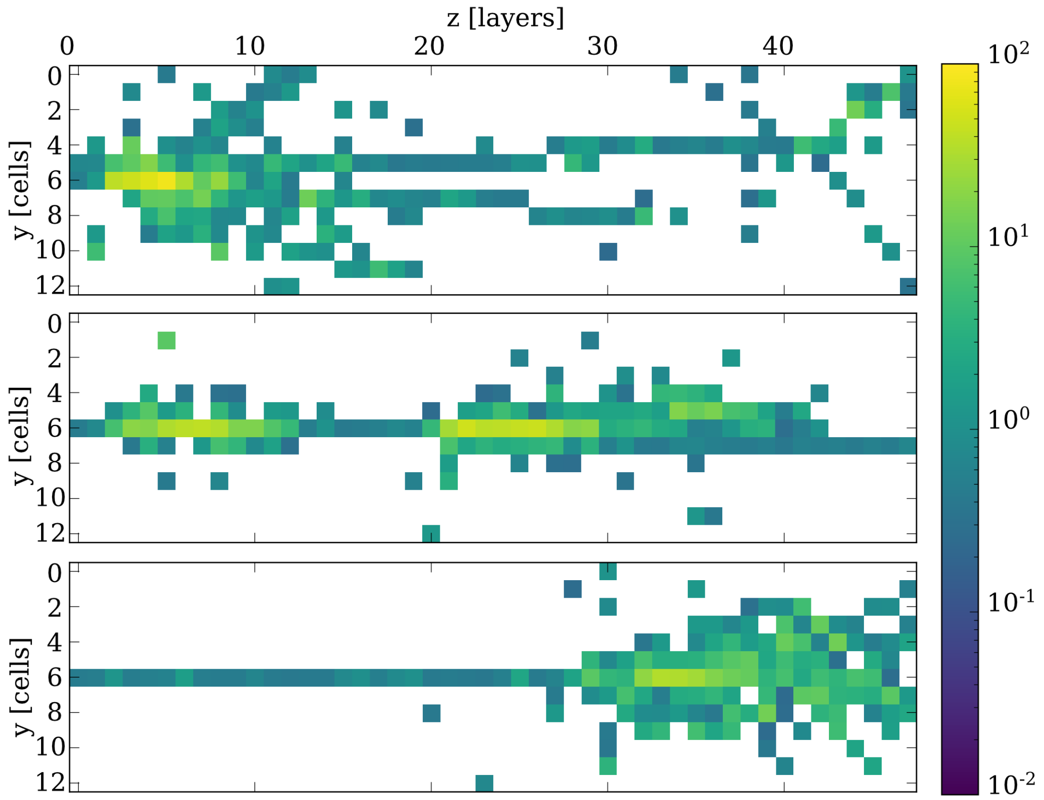


*not the actual ECAL resolution, no correction for sampling fraction variation performed

Hadron Showers

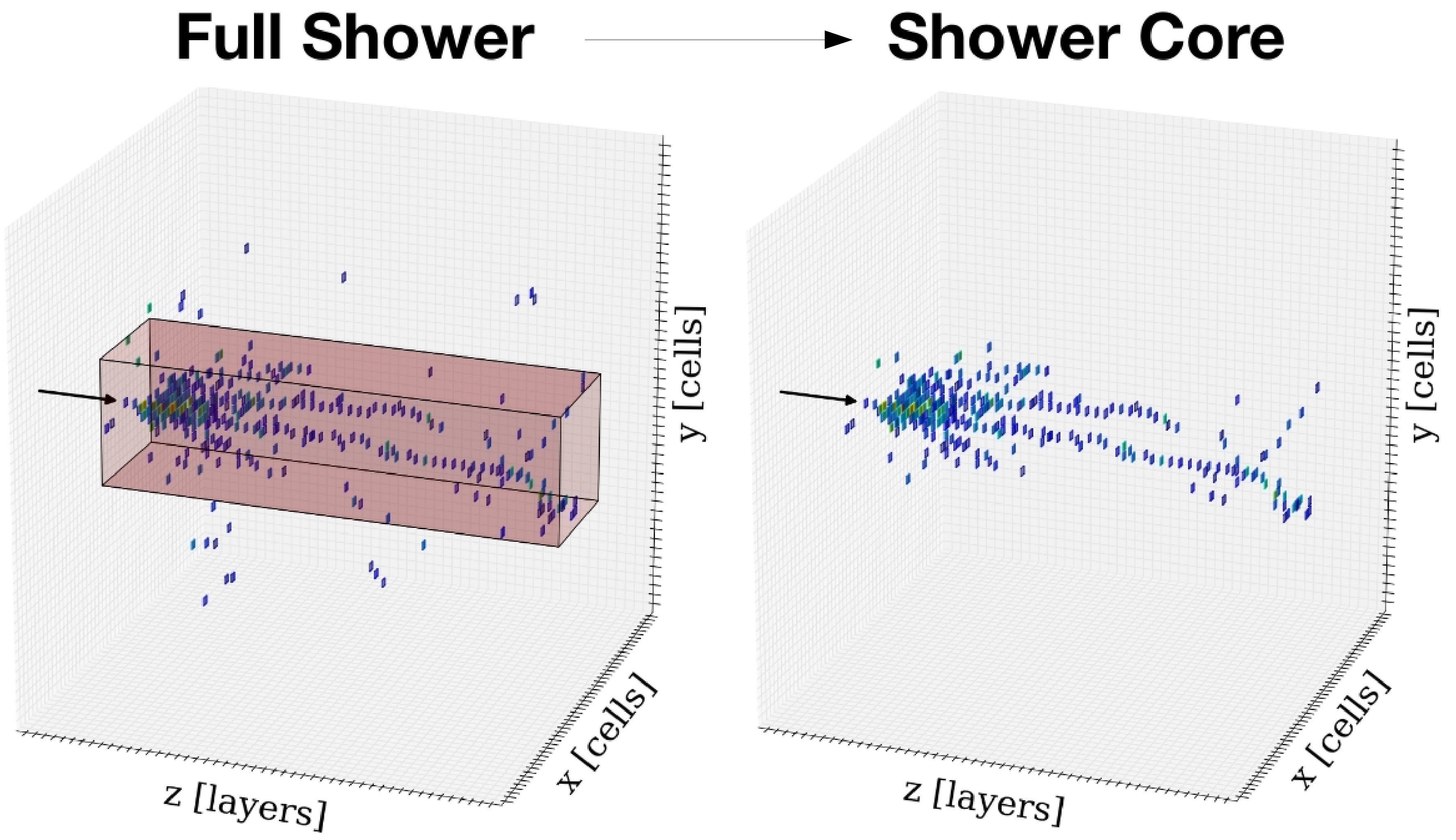
- Success for electromagnetic showers, now started to address hadronic (pion) showers:
  - Much more complex shower structure
  - Currently training with a smaller 3D image containing the active area (i.e shower core)
  - Started with GAN, WGAN, BIB-AE and alternatives





photon showers **vs.** **pion showers**

Pion Dataset



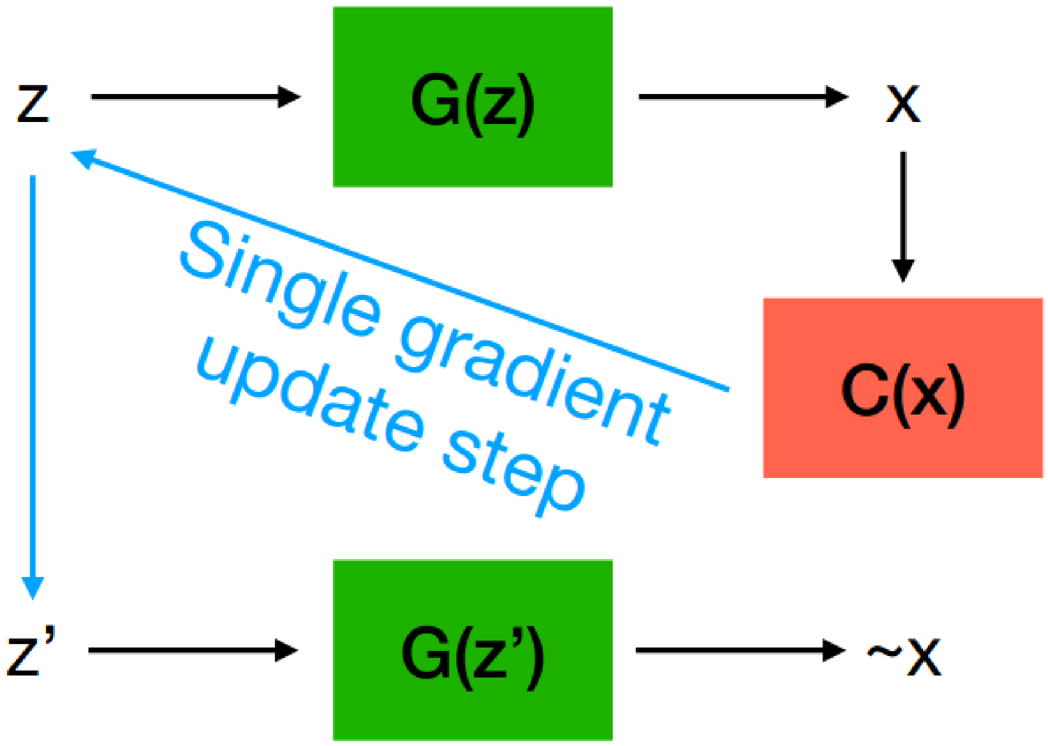
48x48x48 **48x13x13**

- 500k showers
- Fixed incident point and angle
- Uniform 10 GeV to 100 GeV

Latent Optimized WGAN

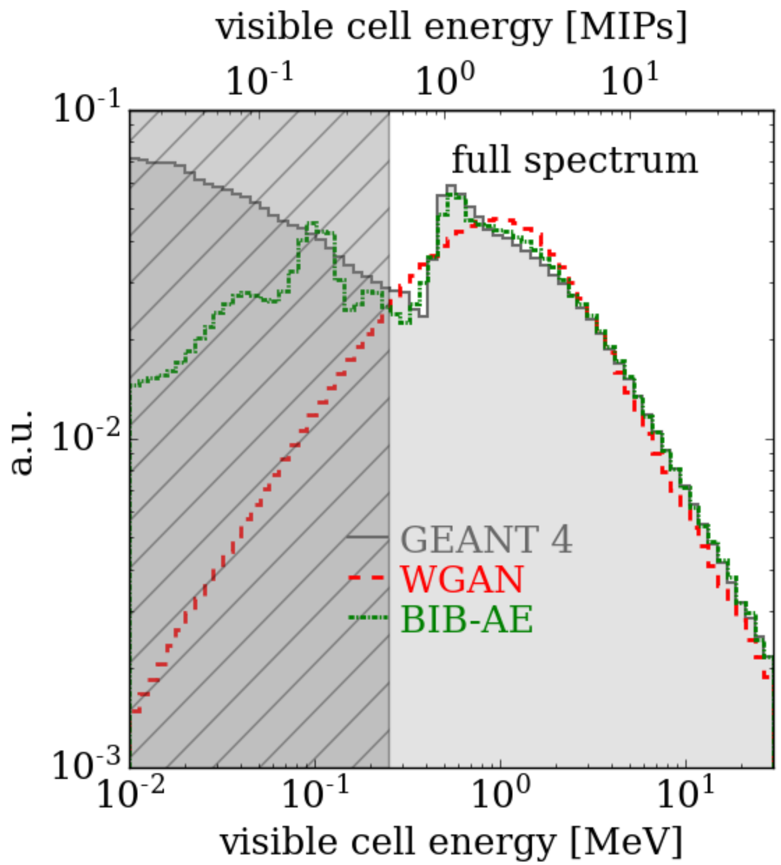
Standard WGAN

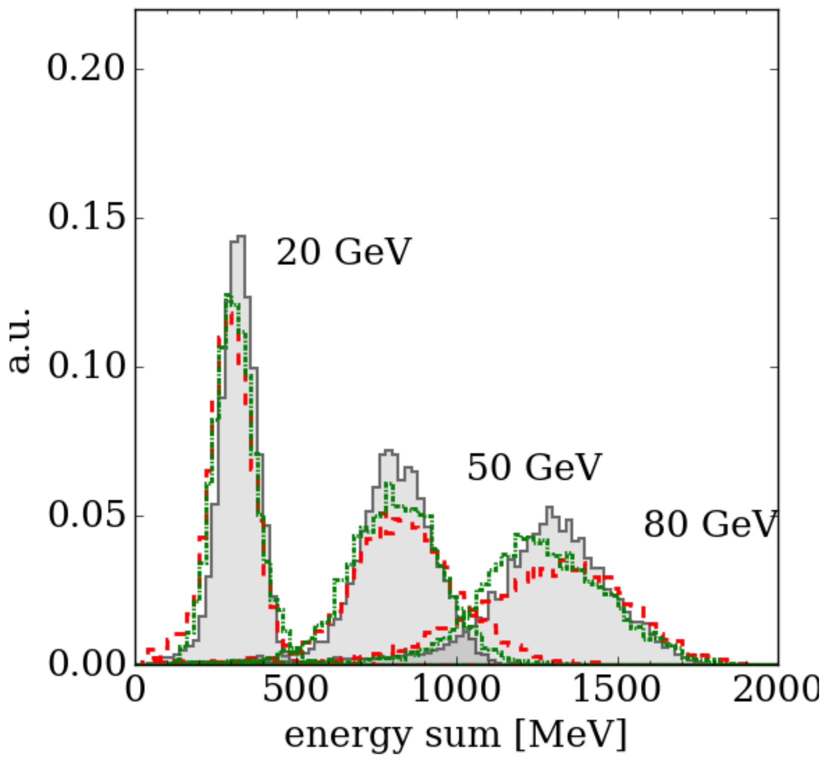
Latent Optimisation

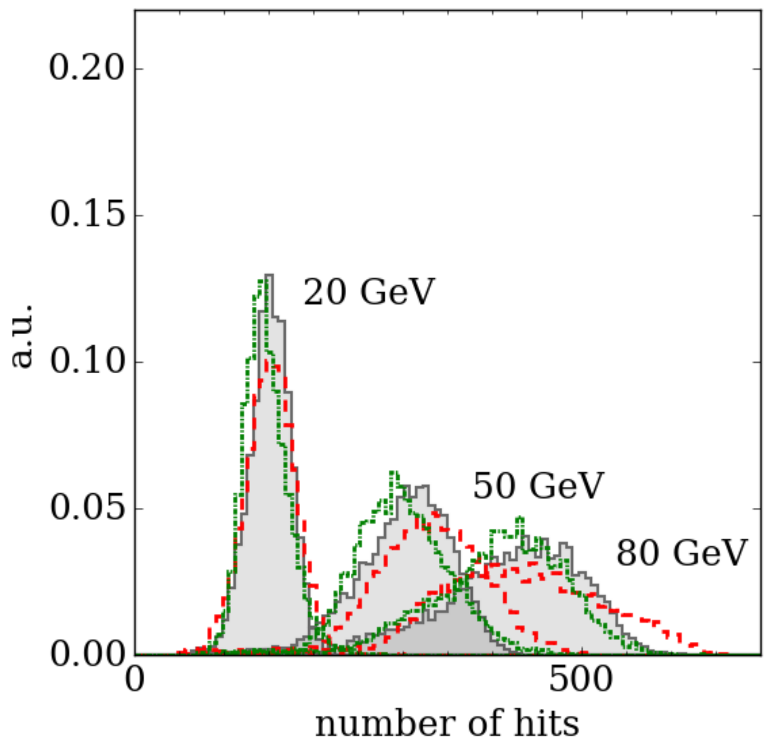


Wu et al.: **LOGAN: Latent Optimisation for Generative Adversarial Networks** 1912.00953

Pion Shower Results I



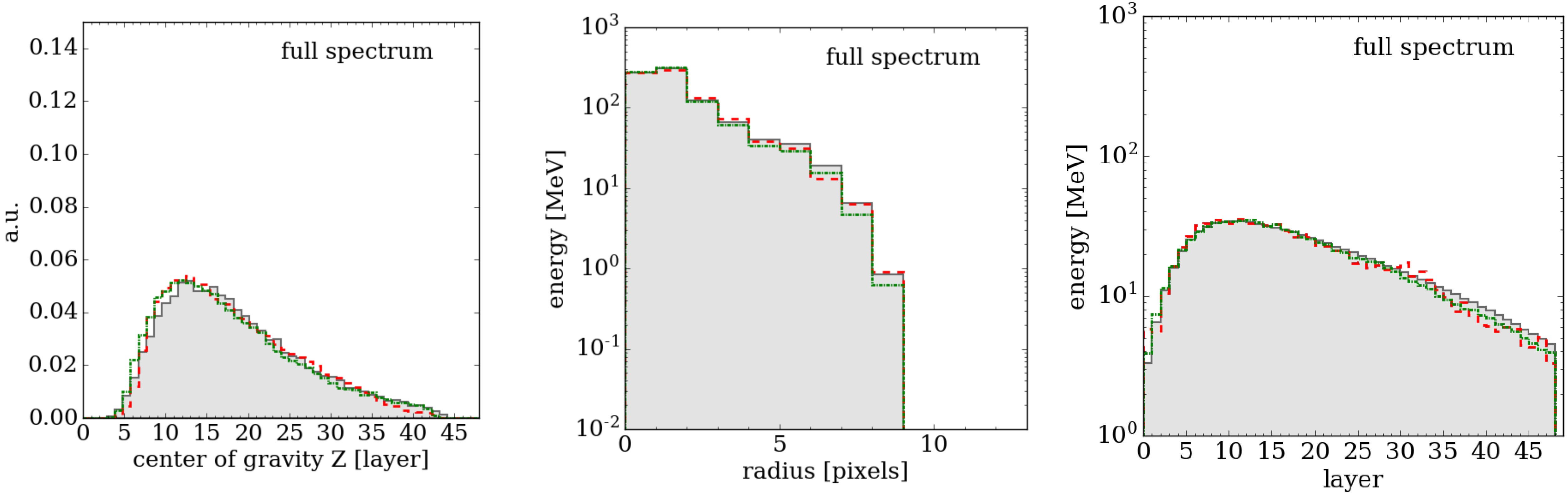




Very good agreement of MIP peak for BIB-AE with Post-Processing!

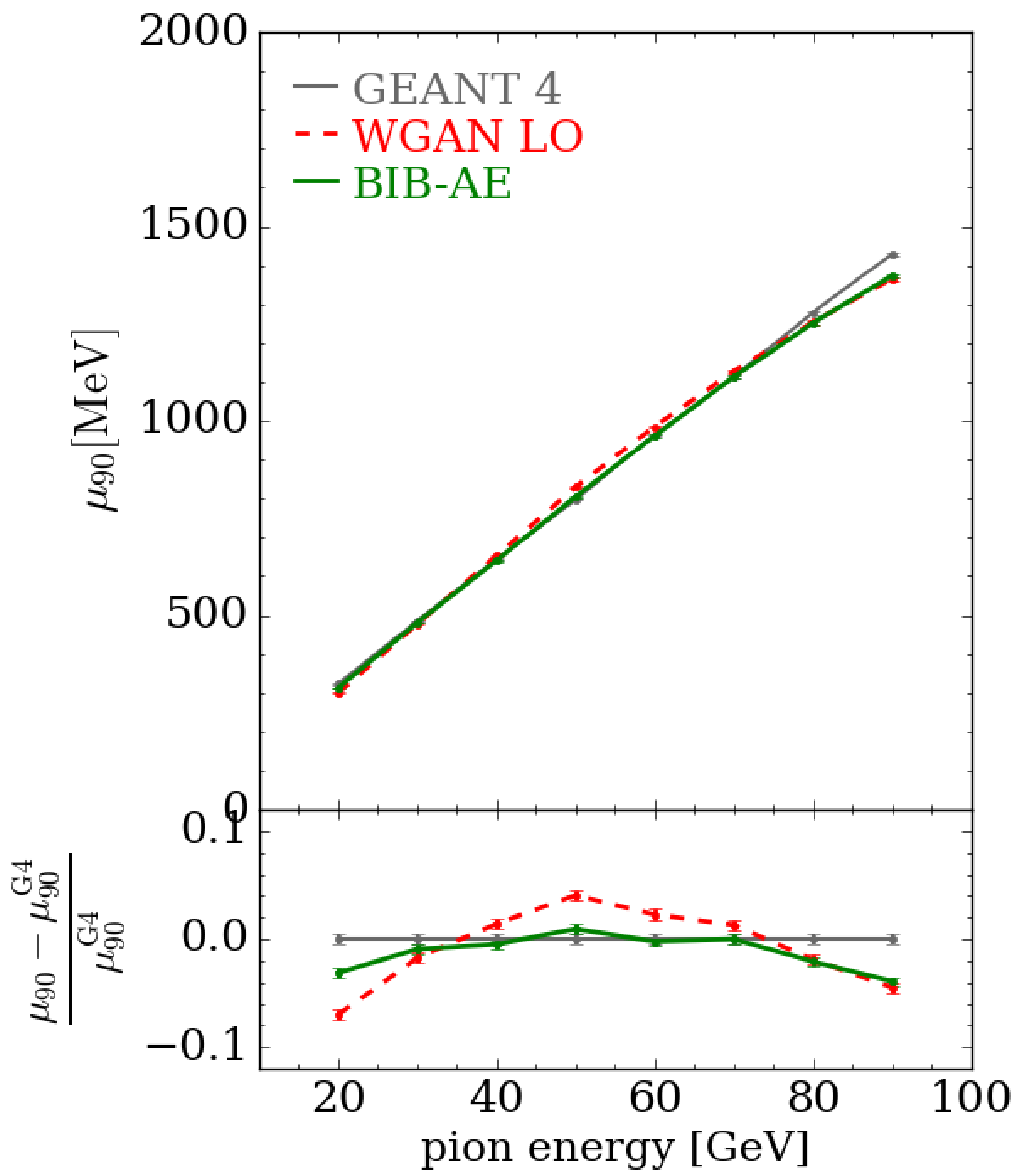
Overall good agreement with Geant4, still room for improvement

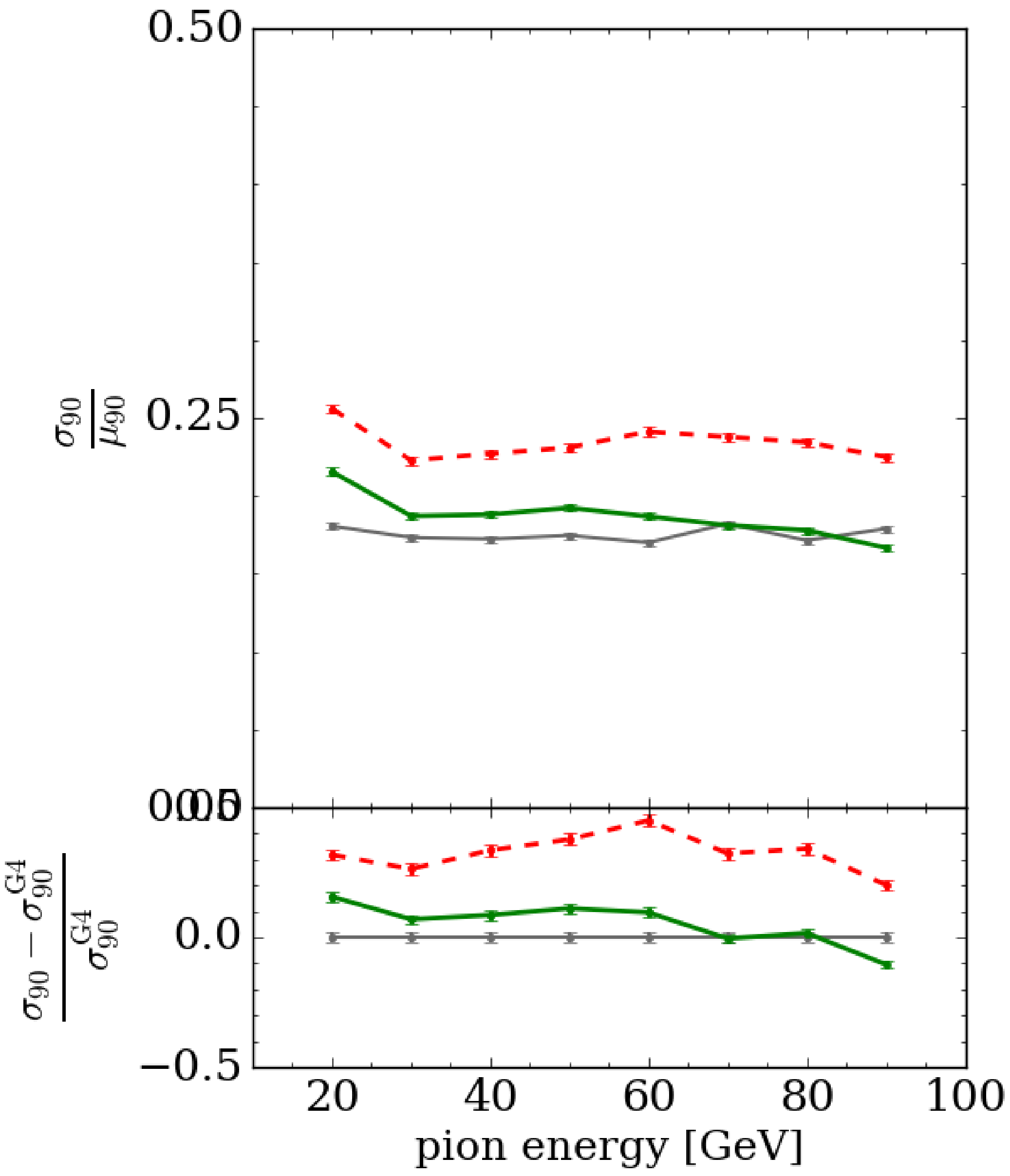
Pion Shower Results II



Both WGAN and BIB-AE fairly reproduces bulk of Geant4 distributions

Linearity and Resolution





Computation Time

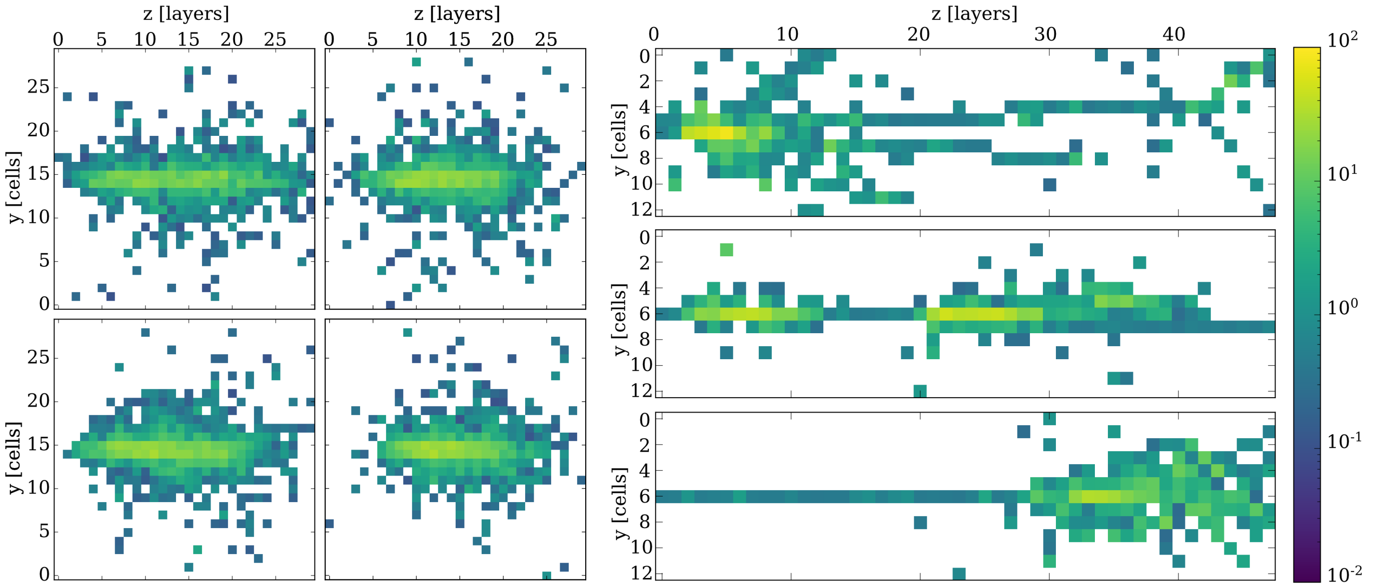
Hardware	Simulator	Photons		Pions	
		Time/shower[ms]	Speed-up	Time/shower[ms]	Speed-up
CPU	Geant4	4082±170	×1	2684±125	×1
	WGAN	61.44±0.03	×66	195.67±0.56	×14
	BIB-AE	95.98±0.08	×43	36.05±0.82	×74
GPU	WGAN	3.93±0.03	×1039	2.695±0.004	×996
	BIB-AE	1.60±0.03	×2551	1.101±0.004	×2438

We observe speed-ups of three orders of magnitude

Conclusions and Outlook

Application of generative models to high resolution EM and hadronic showers simulation

- ✔ Modelling of MIP peak and achieving high fidelity
- ✔ Speed-up: 3 orders of magnitude



Architectures and extensions:

- GAN (+ Mini-Batch discr.)
- WGAN (+ Latent Opt.)
- BIB-AE (+ KDE sampl. + Mini-Batch discr.)

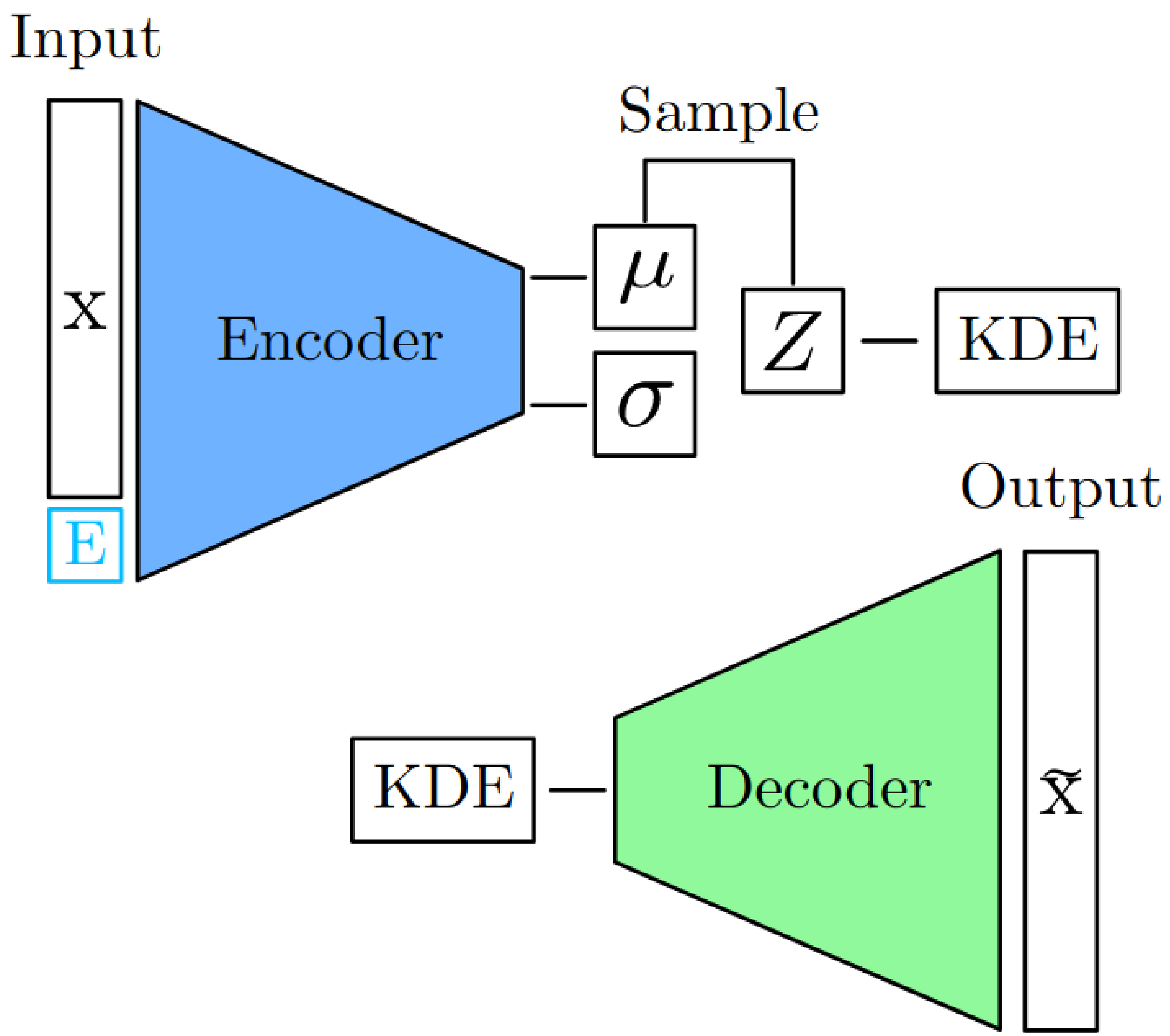
Future plans:

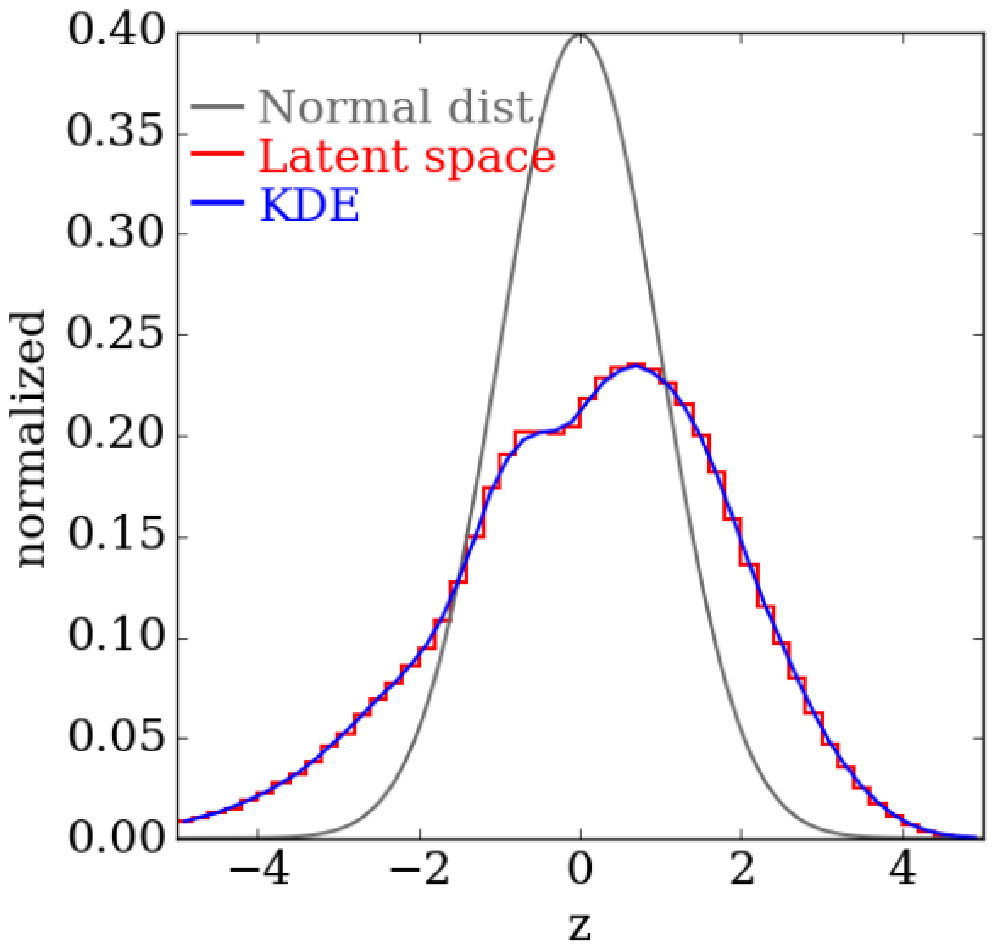
- Condition on incident position / angle
- More focus on hadron showers and full size-showers
- Integrate into existing tools / frameworks

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

Thank you

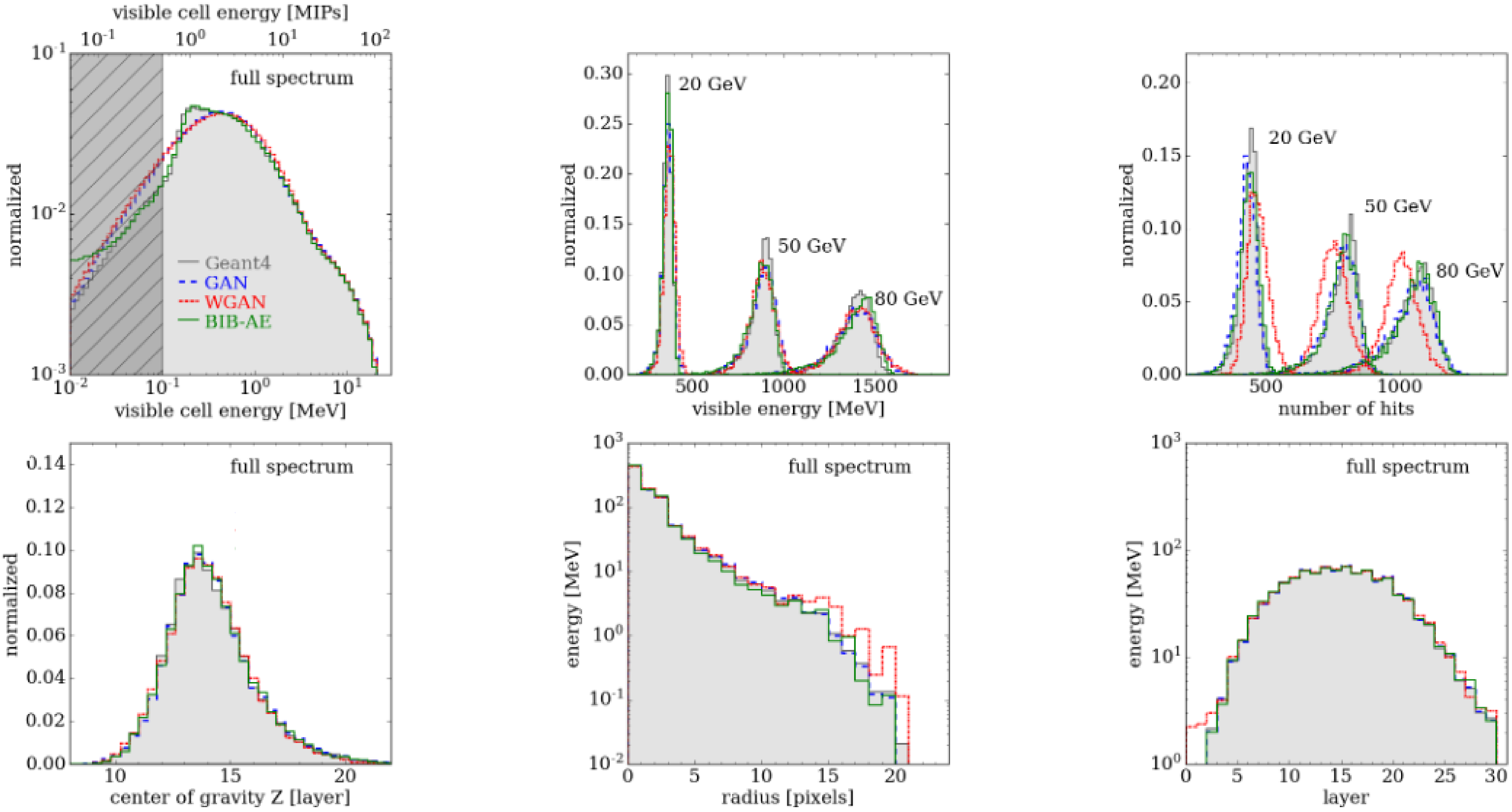
Kernel Density Estimation BIB-AE





Buhmann et. al. **Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network:** 2102.12491

Photon all



visible cell energy [MIPs]
full spectrum
Geant4
GAN
WGAN
BIB-AE
normalized
visible cell energy [MeV]
20 GeV
50 GeV
80 GeV
visible energy [MeV]
number of hits
center of gravity Z [layer]
energy [MeV]
radius [pixels]
layer

Correlations (Pion)

GEANT4 - BIB-AE PP KDE

	$m_{1,x}$	$m_{1,y}$	$m_{1,z}$	$m_{2,x}$	$m_{2,y}$	$m_{2,z}$	E_{vis}	E_{inc}	n_{hit}	E_1/E_{vis}	E_2/E_{vis}	E_3/E_{vis}
$m_{1,x}$	0.00											
$m_{1,y}$	-0.06	0.00										
$m_{1,z}$	0.00	-0.02	0.00									
$m_{2,x}$	-0.06	-0.00	-0.04	0.00								
$m_{2,y}$	-0.01	0.11	-0.03	-0.04	0.00							
$m_{2,z}$	0.00	0.04	-0.11	-0.04	-0.05	0.00						
E_{vis}	0.02	-0.02	-0.02	0.02	-0.00	-0.01	0.00					
E_{inc}	0.01	0.01	-0.02	0.08	0.04	-0.03	0.01	0.00				
n_{hit}	0.04	0.03	-0.02	0.08	0.02	-0.02	0.00	0.01	0.00			
E_1/E_{vis}	-0.00	0.04	0.00	0.02	0.01	0.09	0.02	0.01	0.02	0.00		
E_2/E_{vis}	-0.00	-0.05	-0.03	0.03	0.03	-0.03	-0.01	-0.01	-0.01	0.02	0.00	
E_3/E_{vis}	0.00	0.00	0.01	-0.06	-0.05	-0.11	-0.01	-0.01	-0.01	-0.01	-0.02	0.00

GEANT4 - WGAN LO

	$m_{1,x}$	$m_{1,y}$	$m_{1,z}$	$m_{2,x}$	$m_{2,y}$	$m_{2,z}$	E_{vis}	E_{inc}	n_{hit}	E_1/E_{vis}	E_2/E_{vis}	E_3/E_{vis}
$m_{1,x}$	0.00											
$m_{1,y}$	-0.17	0.00										
$m_{1,z}$	0.01	-0.12	0.00									
$m_{2,x}$	0.07	0.07	-0.14	0.00								
$m_{2,y}$	-0.07	0.05	-0.23	-0.02	0.00							
$m_{2,z}$	-0.07	0.04	-0.08	0.02	-0.06	0.00						
E_{vis}	-0.01	0.11	0.05	0.11	0.12	0.06	0.00					
E_{inc}	-0.00	0.19	0.04	0.03	0.01	0.03	0.06	0.00				
n_{hit}	-0.02	0.08	-0.01	0.23	0.23	0.14	-0.02	0.05	0.00			
E_1/E_{vis}	-0.01	0.12	0.01	0.12	0.17	0.06	-0.06	-0.06	0.00	0.00		
E_2/E_{vis}	-0.02	-0.08	-0.09	-0.02	-0.01	-0.03	0.06	0.05	0.00	0.05	0.00	
E_3/E_{vis}	0.02	-0.10	0.00	-0.14	-0.24	-0.09	0.02	0.01	-0.01	0.02	-0.10	0.00