Generative Models for Fast Simulation of Electromagnetic and Hadronic Showers in Highly Granular Calorimeters

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

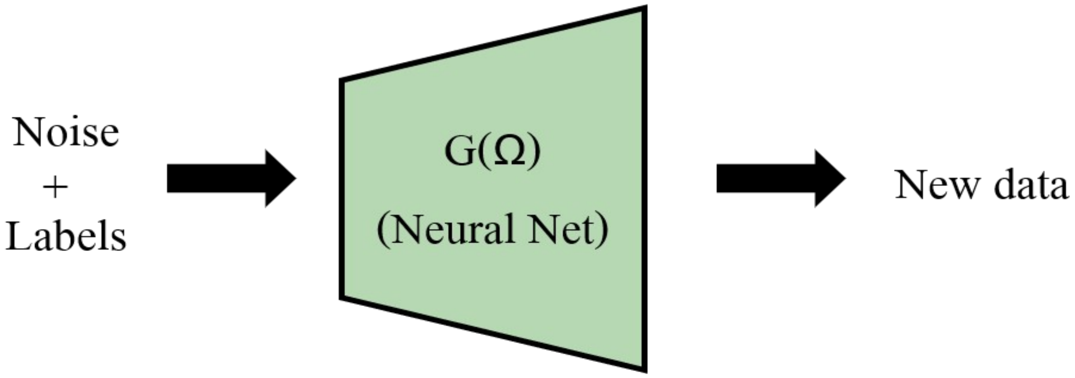
[1] **Deutsches Elektronen-Synchrotron**

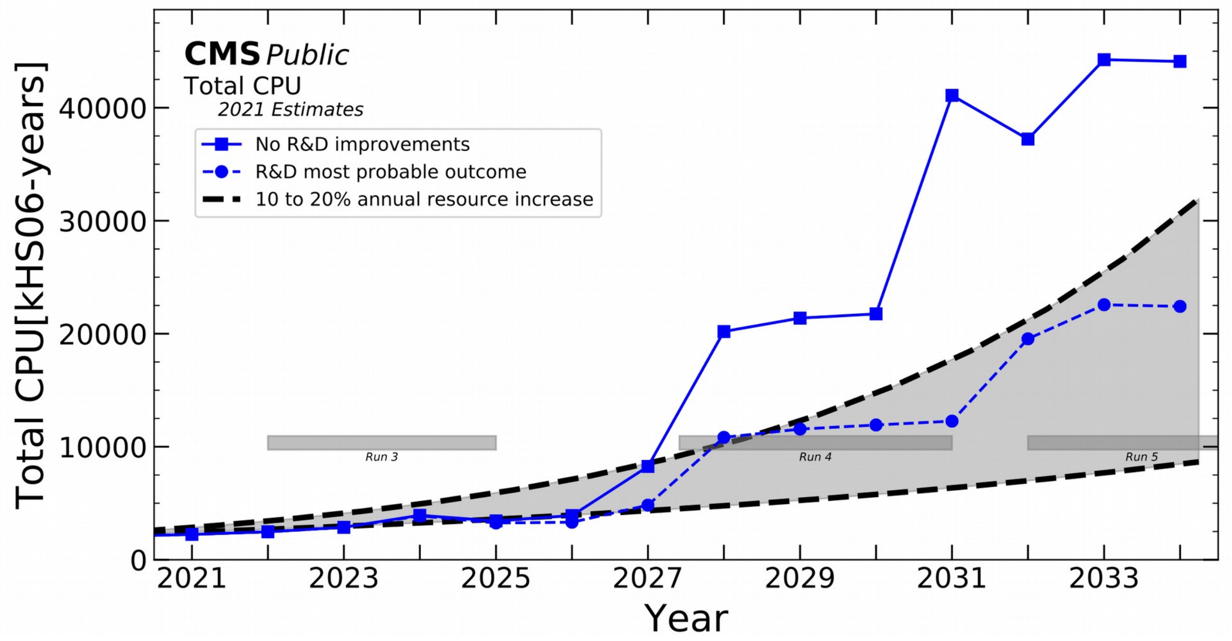
25.11.2022

peter.mckeown@desy.de

Reducing the Strain on HEP Computing Resources

- **MC simulation (Geant4)** is computationally **expensive**
  - **Calorimeters** most intensive part of detector simulation
- Major bottleneck e.g. **HL-LHC**
- **Generative models** potentially offer orders of magnitude speed up



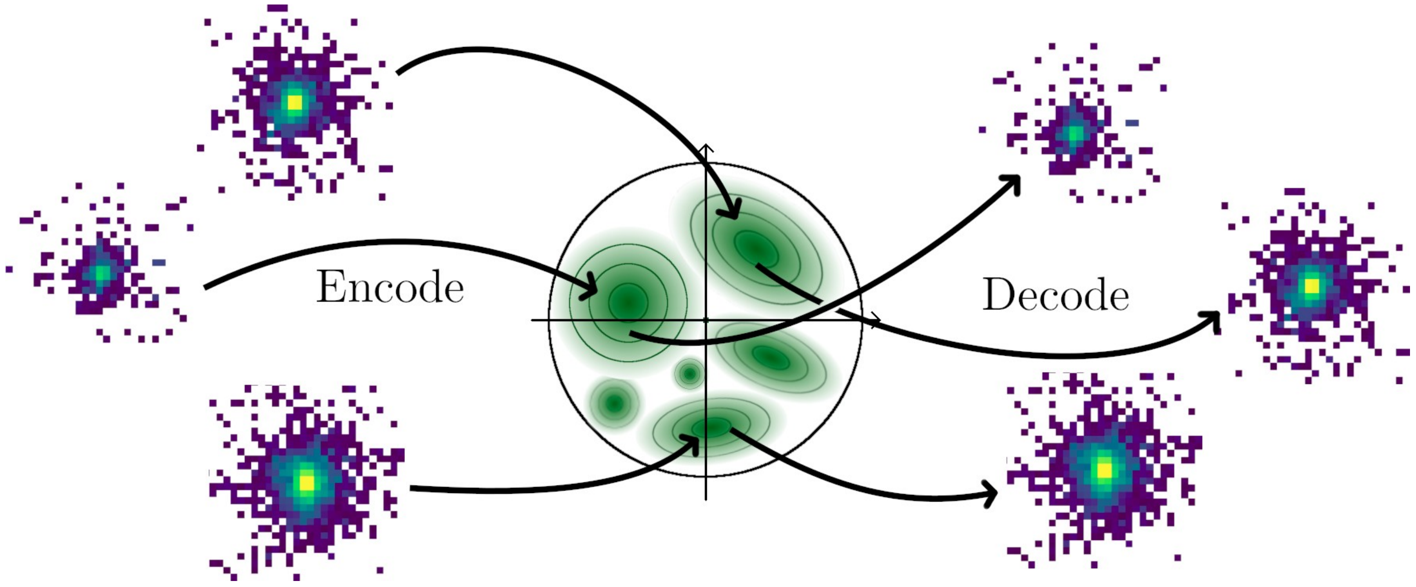


CMS Collaboration, Offline and Computing Public Results (2021), https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults

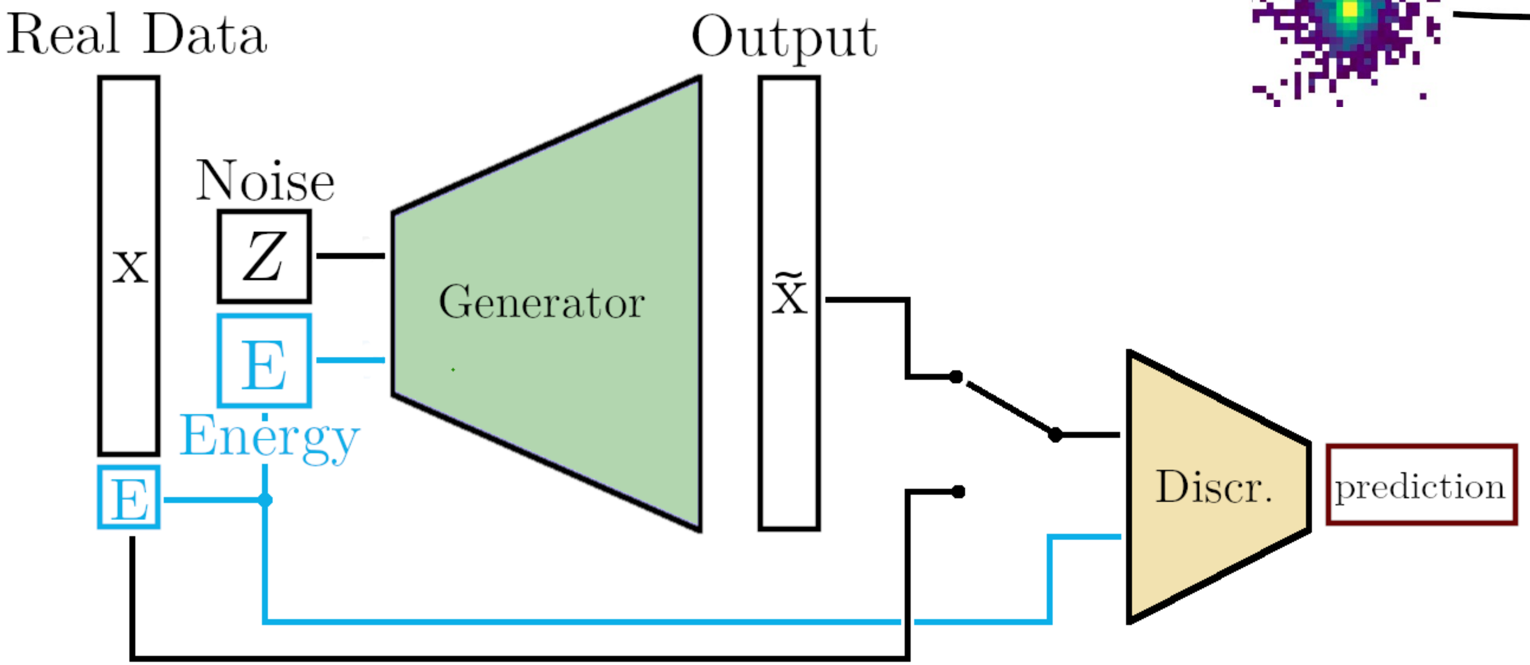
Common Generative Models

- **VAE**[1]**:** Encoder-decoder structure
- **GAN**[2]**:** Adversarial feedback from discriminator

Variational Autoencoder (**VAE**)



Generative Adversarial Network (**GAN**)



- Models studied:
  - Wasserstein GAN **(WGAN)**
  - Bounded Information Bottleneck Autoencoder **(BIB-AE)**

[1] D.P. Kingma, M. Welling. Auto-encoding Variational Bayes (2014), arXiv:1312.6114

[2] Ian Goodfellow et. al., Generative Adversarial Nets (2014), arXiv:1406.2661

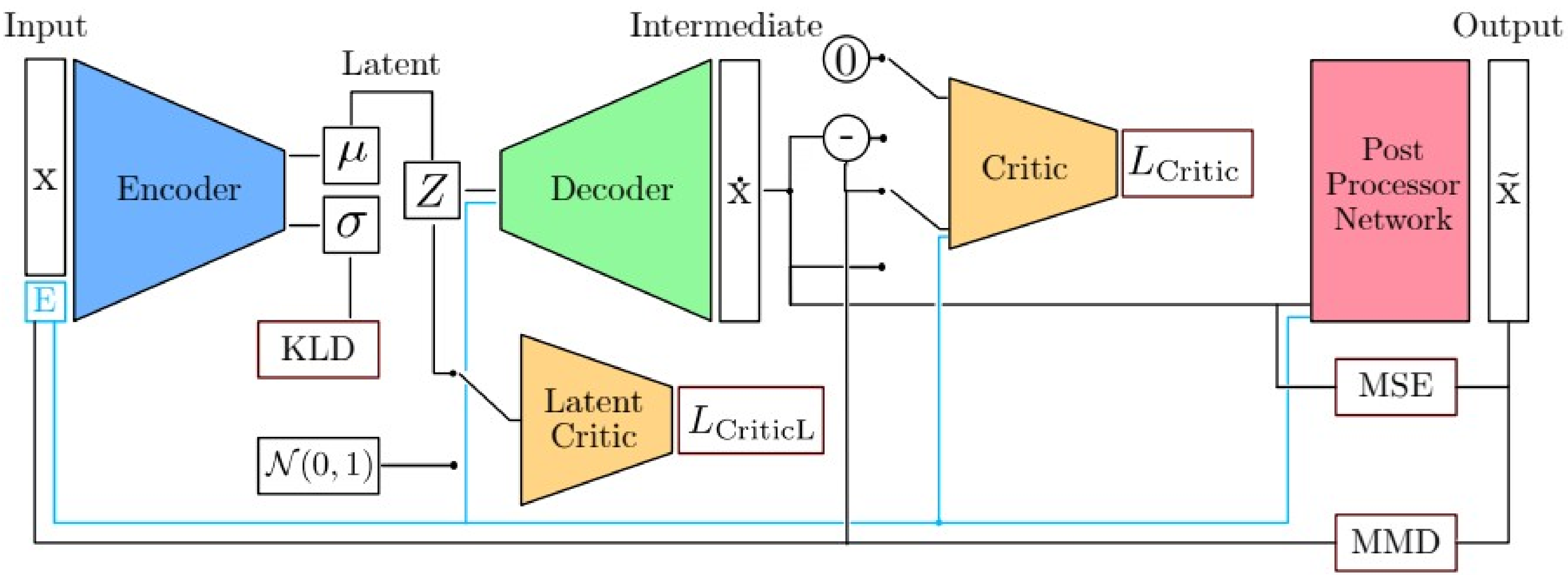
The BIB-AE

Bounded-Information Bottleneck Autoencoder

Voloshynovskiy et. al: Information bottleneck through variational glasses, arXiv:1912.00830 (2019)

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

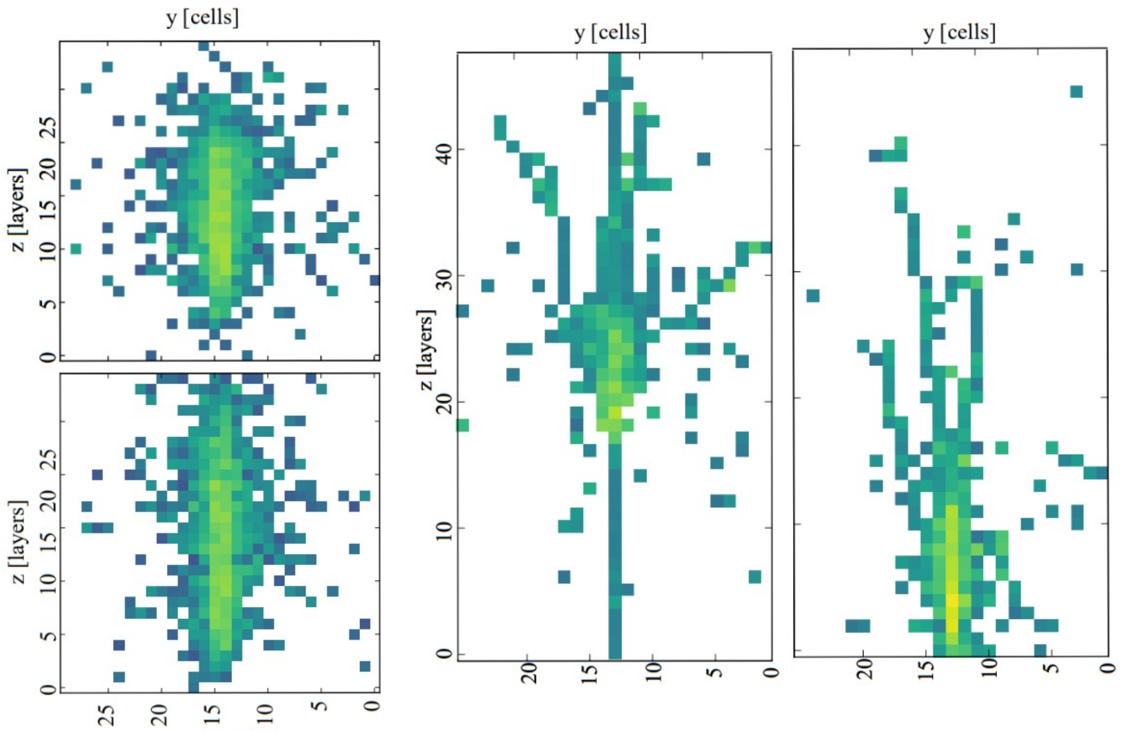
- **Unifies** features of both **GANs** and **VAEs**
- **Post-Processor** network: Improve per-pixel energies; second training
- Multi-dimensional KDE sampling: better modeling of latent space

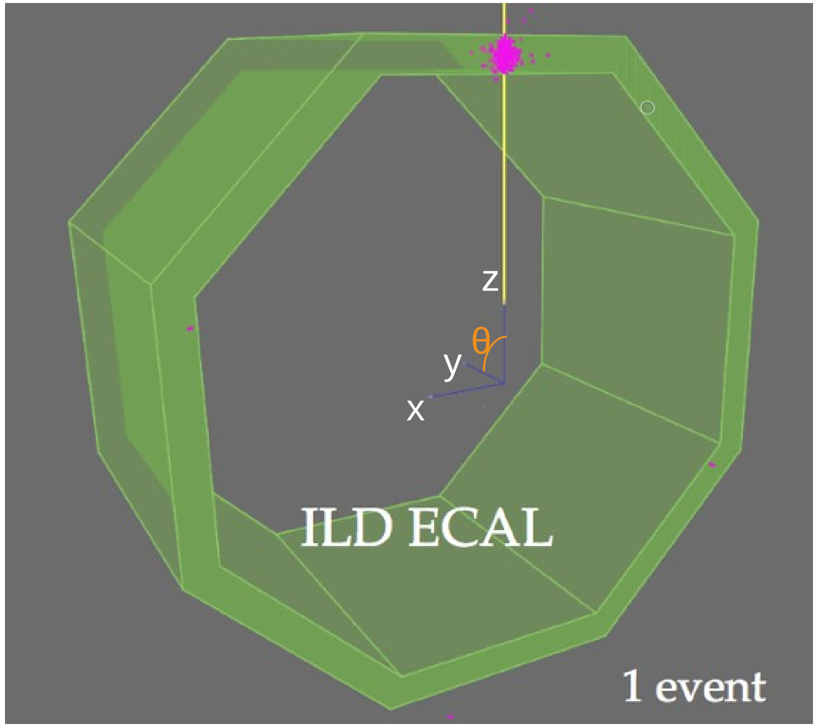


$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}}$

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

Challenges for Generative ML Calorimeter Simulations





From Photons to Pions

- **Hadronic** showers significantly harder to learn than electromagnetic showers
  - **Complex** topologies
  - Large event-to-event **fluctuations**

Multi-Parameter Conditioning

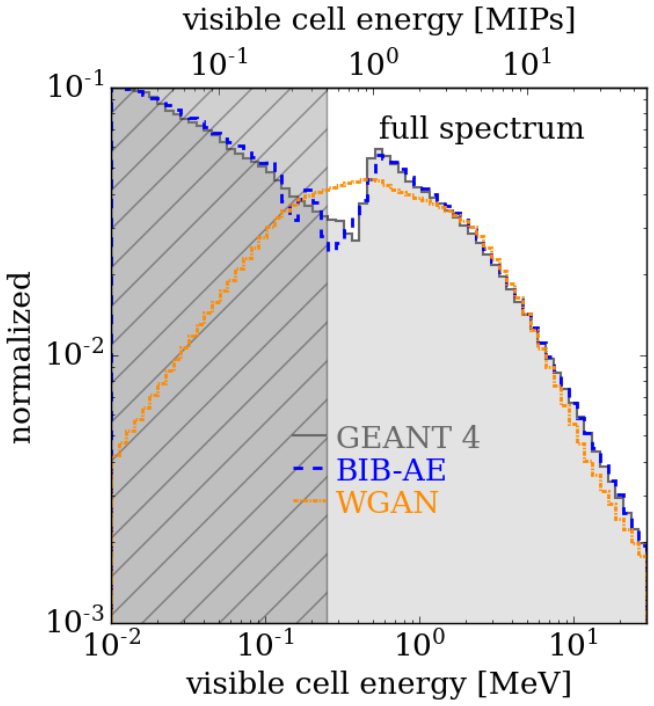
- **Simultaneous conditioning** on multiple parameters crucial for a general simulation tool
  - Start with **photons**
  - Vary incident **energy and angle**

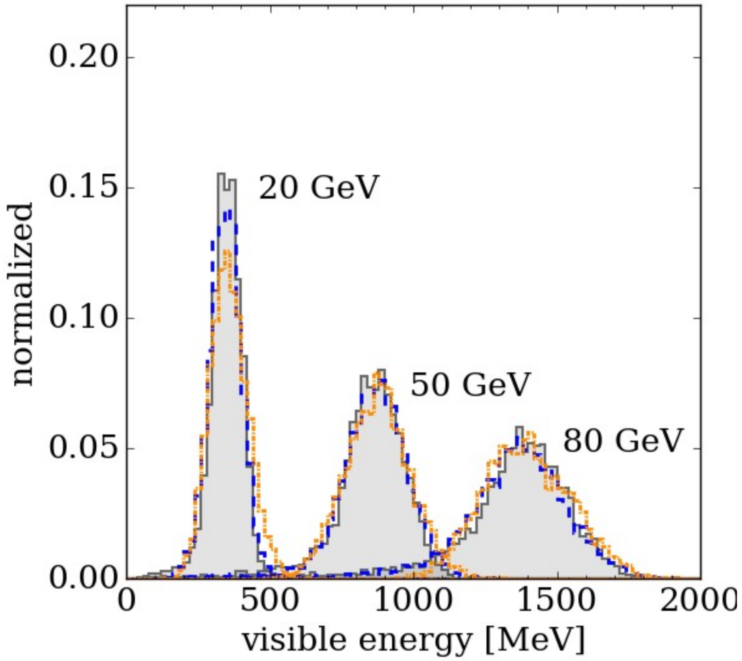
Latest Progress

Hadronic Calorimeter Showers

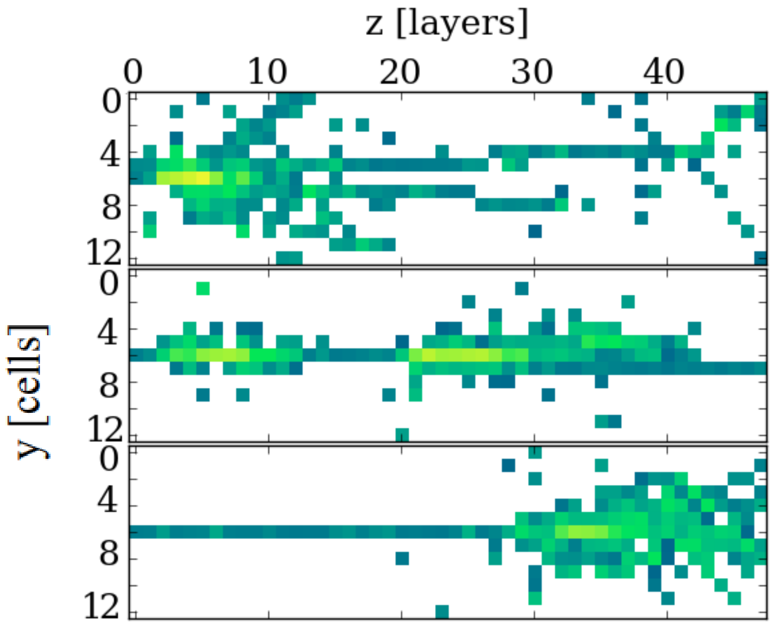
- Achieve **significant speedups** (CPU/GPU)
- Achieve **high degree of fidelity**

Hadrons, Better, Faster, Stronger,
E. Buhmann, et al.
MLST 3 025014 (2022)





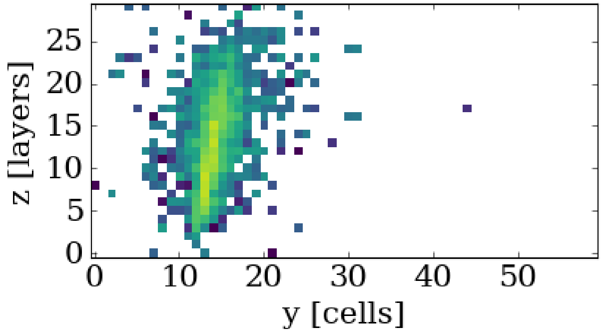
Hardware	Simulator	Time / Shower [ms]	Speed-up
CPU	GEANT4	2684 ± 125	×1
	WGAN	47.923 ± 0.089	×56
	BIB-AE	350.824 ± 0.574	×8
GPU	WGAN	0.264 ± 0.002	×10167
	BIB-AE	2.051 ± 0.005	×1309

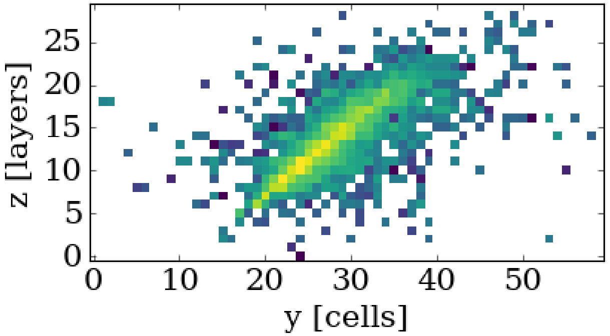


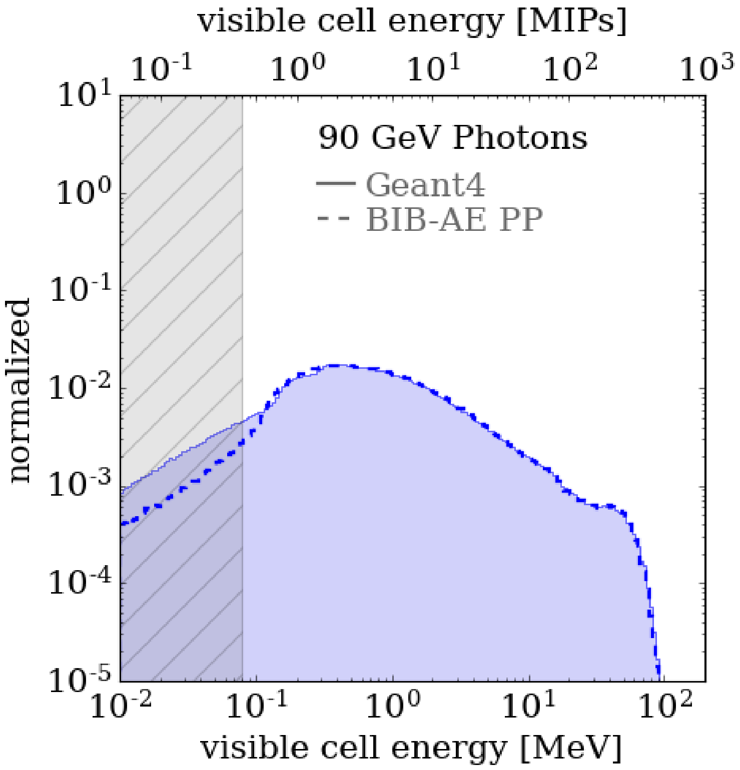
Latest Progress

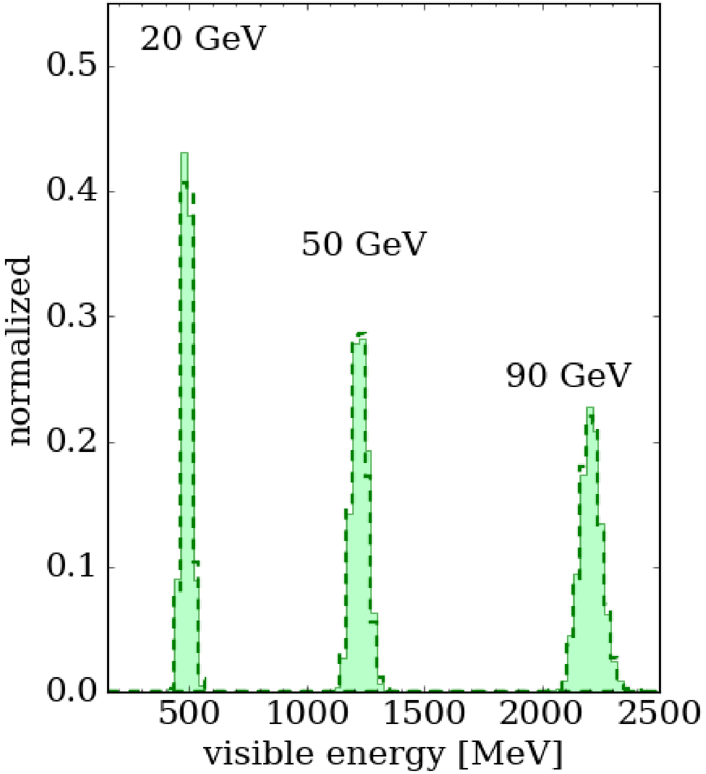
Angular Photon Showers

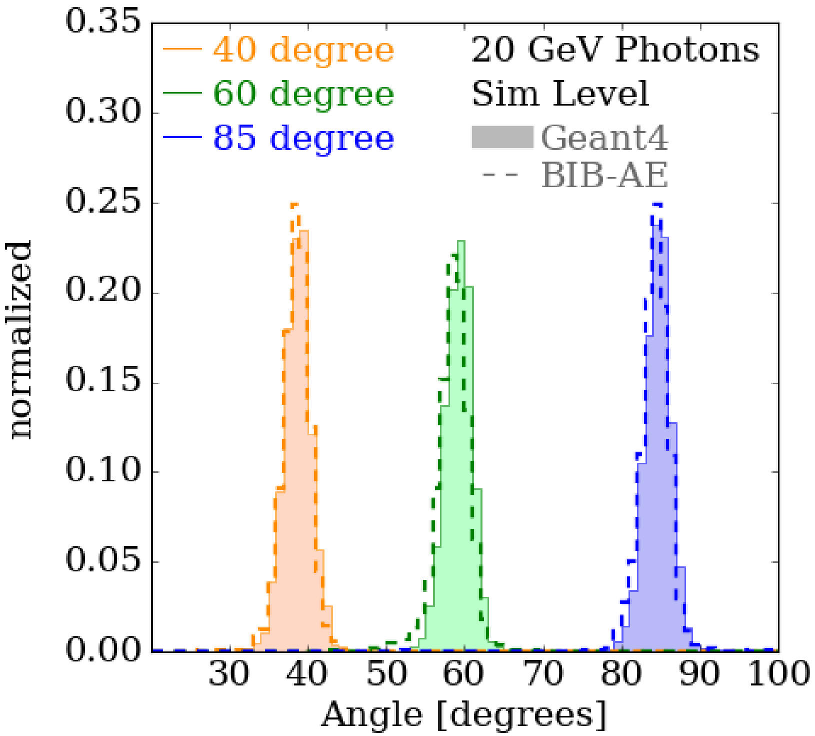
- Simultaneous **Energy and Angular conditioning** demonstrated while maintaining strong physics performance
  - Publication in preparation



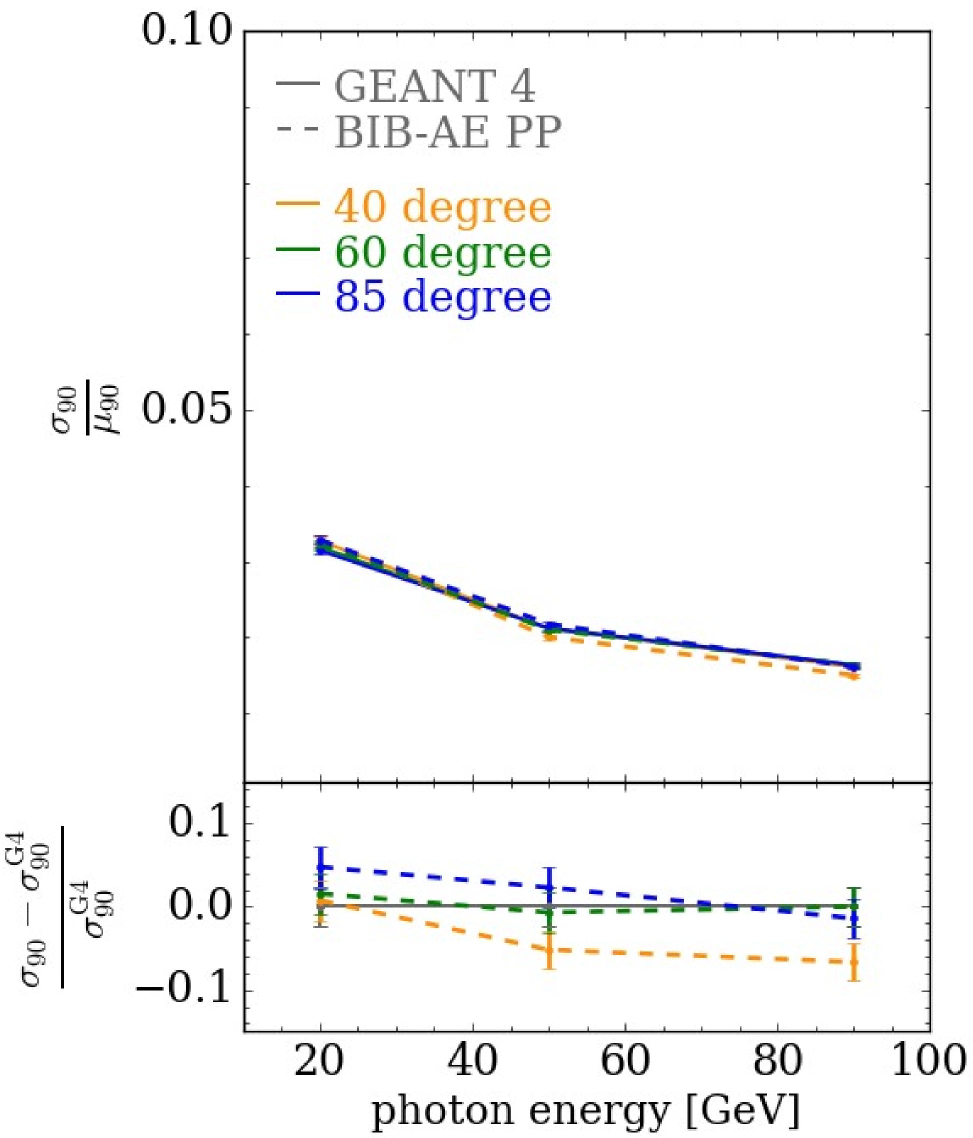


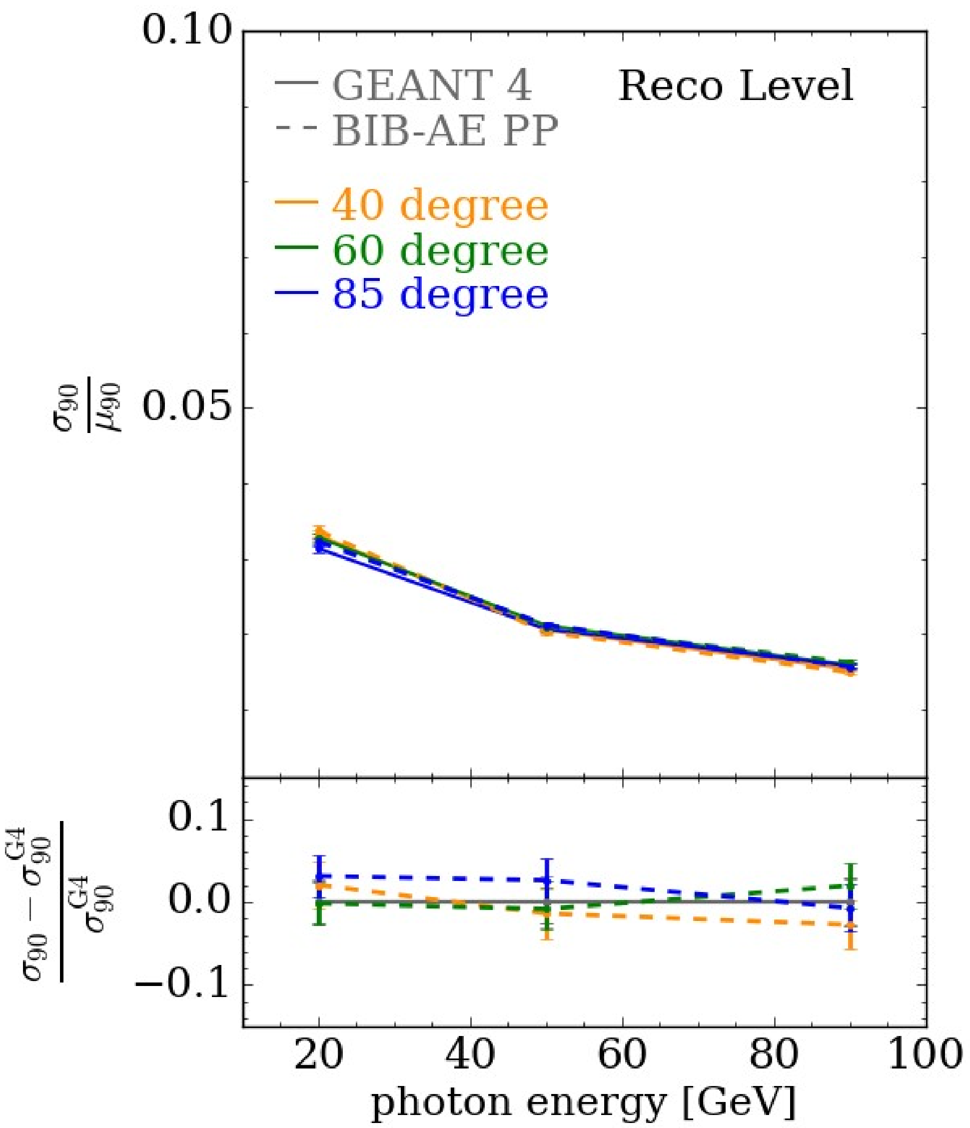






Results: Energy resolution Sim vs Rec

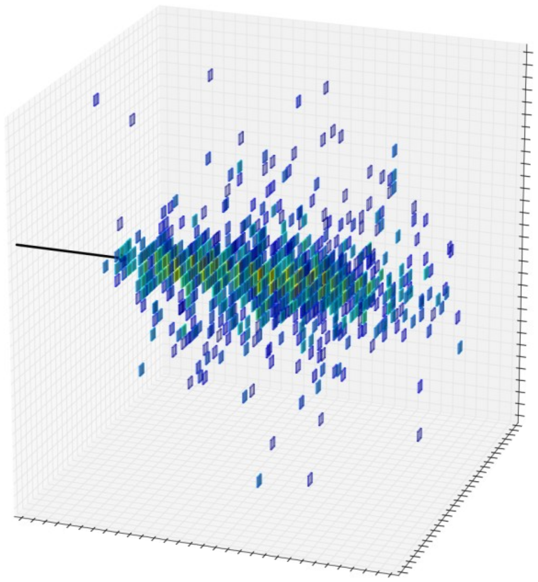




Summary

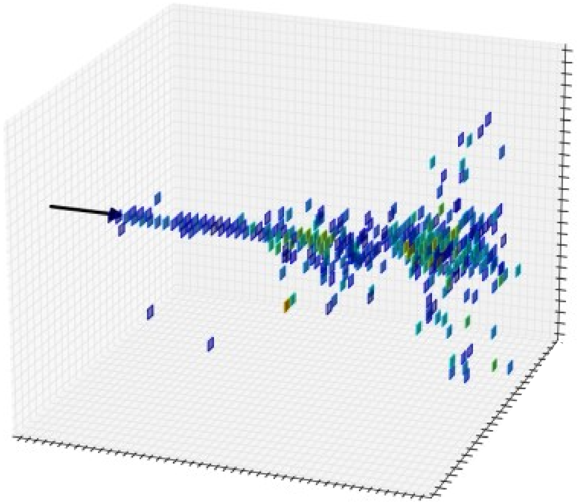
Achieved

- Generative models hold promise for **fast** simulation of showers in **high granularity** calorimeters with **high fidelity**
- Demonstrated high fidelity simulation of **hadronic** showers with generative models
- Demonstrated high fidelity simulation of **photon** showers with **angular and energy conditioning**



Next Steps

- Strategies for dealing with **complex** and **irregular geometries**
- **Integration** into the existing tools (Geant4)
- Full benchmark of **physics performance** after **reconstruction**

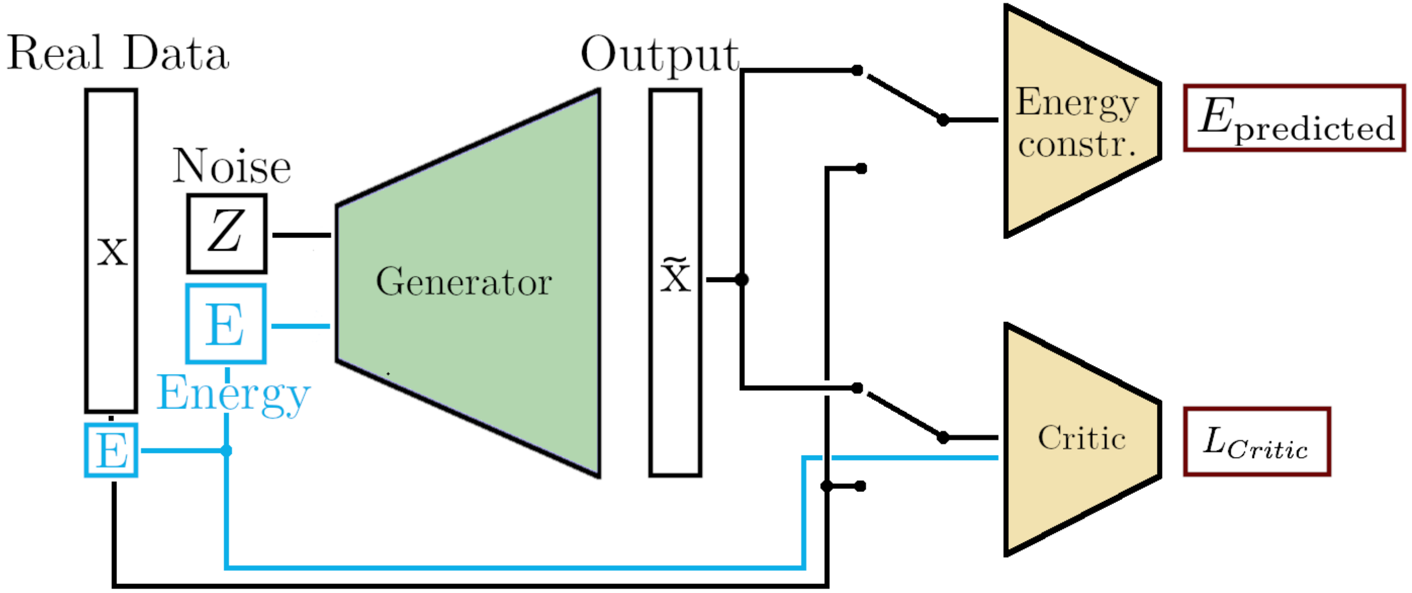


Backup

Architectures: WGAN

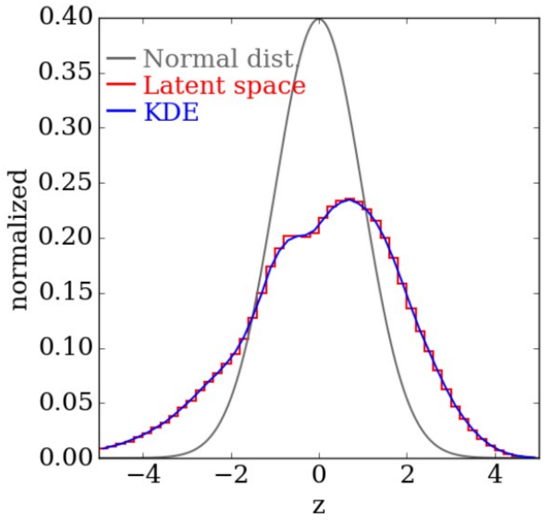
WGAN

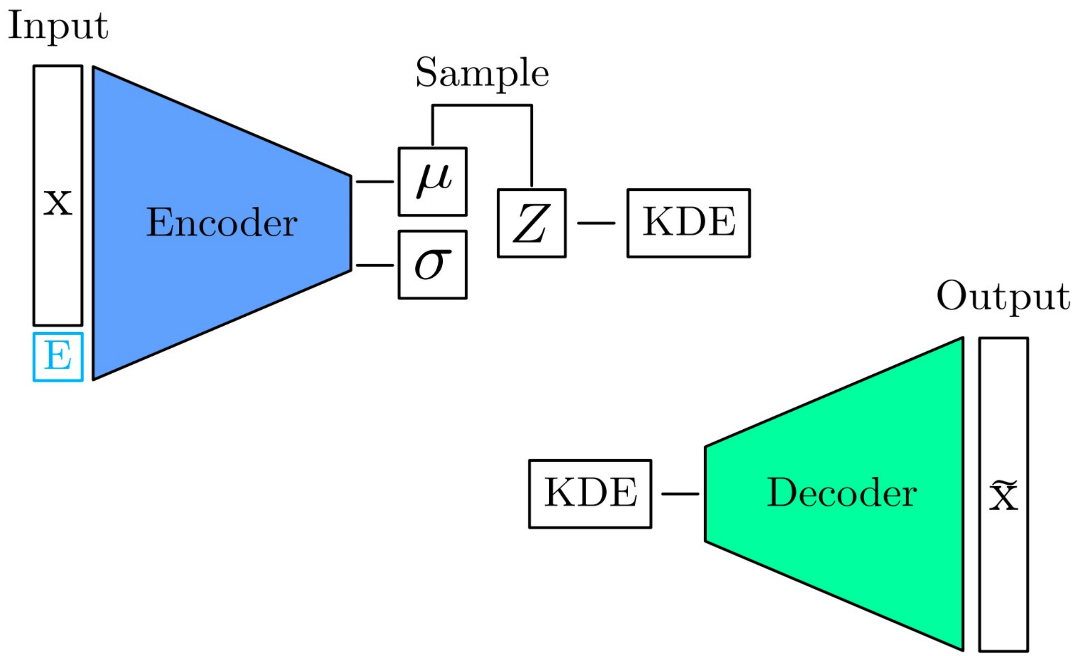
- Alternative to classical GAN training; Generator and Critic Networks
- Wasserstein-1 distance as loss with gradient penalty: **improve stability**
- **Addition of auxiliary constrainer network for improved conditioning performance**

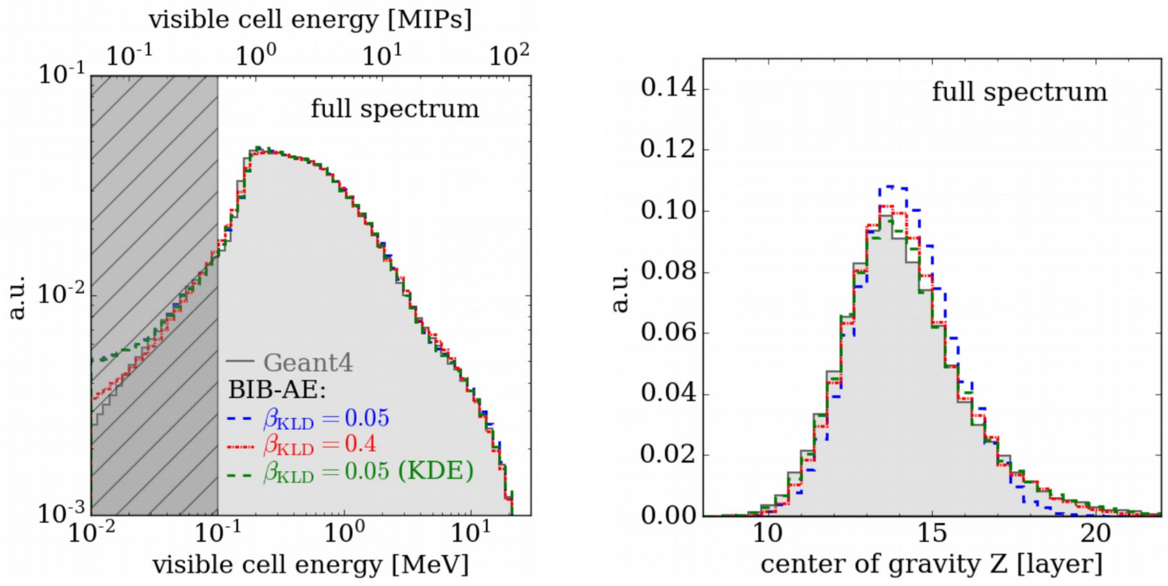


Latent Space sampling

- **Relaxing regularisation** of latent space allows more information to be stored
  - Latent space deviates from a Normal distribution
- Employ **density estimation** to produce latent sample (e.g. KDE)
- **Improve** modeling of **shower shape** (center of gravity)

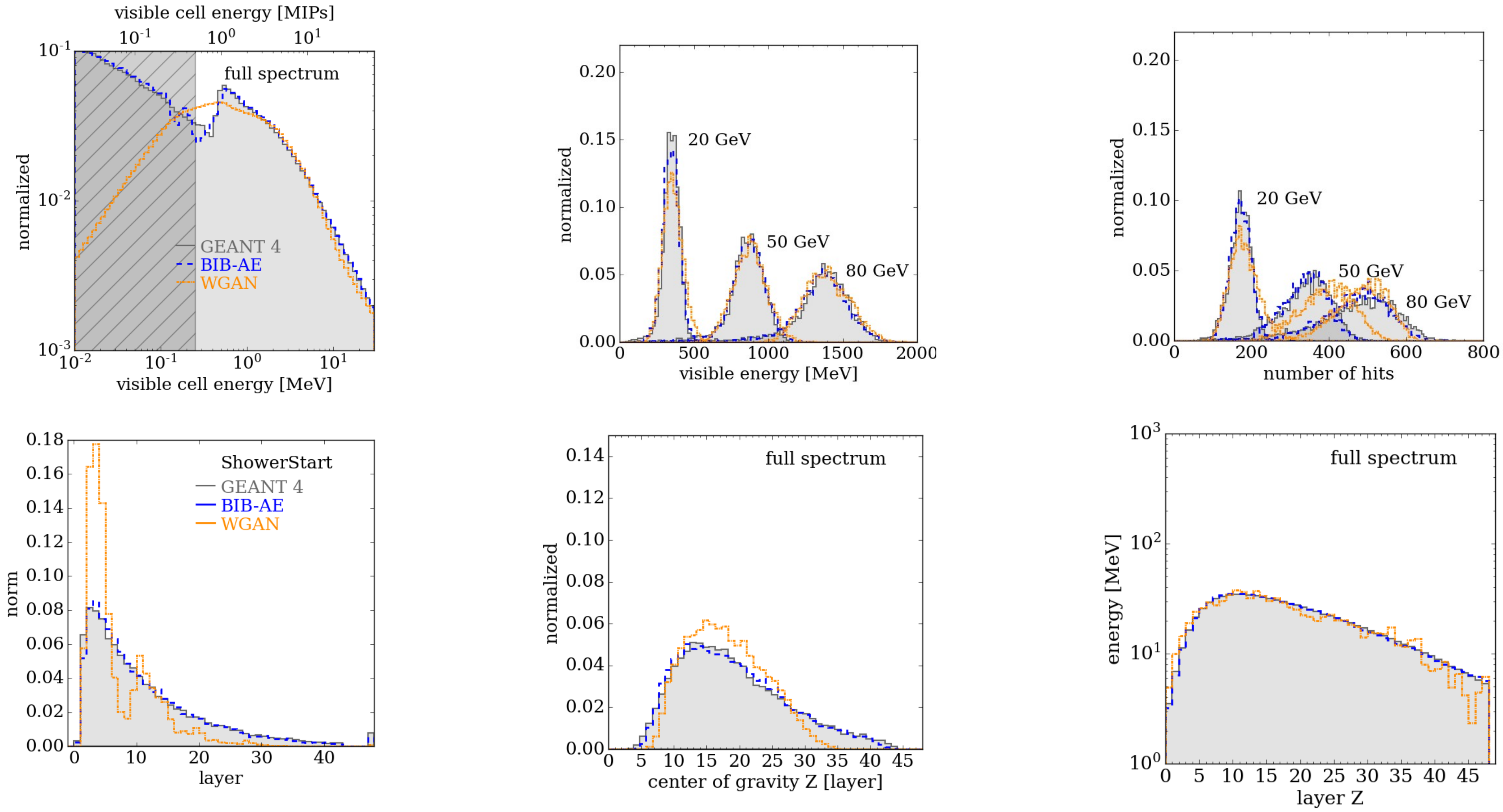






Buhmann et. al: **Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network**, EPJ Web of Conferences 251, 03003 (2021)

Pion Showers: Sim Level Results



Pion correlations

GEANT4 - BIB-AE

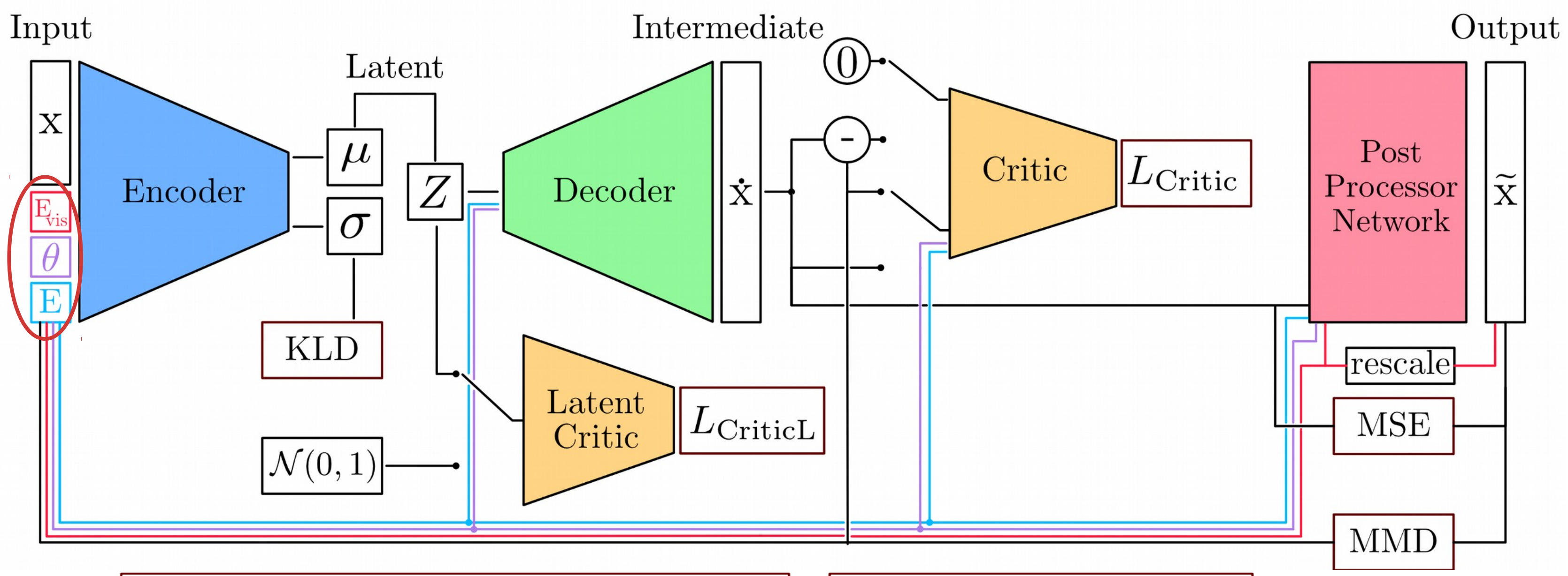
$m_{1,x}$	0.00											
$m_{1,y}$	-0.00	0.00										
$m_{1,z}$	-0.01	-0.04	0.00									
$m_{2,x}$	-0.08	-0.00	-0.06	0.00								
$m_{2,y}$	-0.10	-0.03	-0.05	0.01	0.00							
$m_{2,z}$	-0.06	0.01	-0.06	-0.08	-0.05	0.00						
E_{vis}	0.03	-0.02	-0.01	0.09	0.09	0.06	0.00					
E_{inc}	0.01	-0.03	-0.00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

GEANT4 - WGAN

$m_{1,x}$	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
	$m_{1,x}$	$m_{1,y}$	$m_{1,z}$	$m_{2,x}$	$m_{2,y}$	$m_{2,z}$	E_{vis}	E_{inc}	n_{hit}	E_1/E_{vis}	E_2/E_{vis}	E_3/E_{vis}

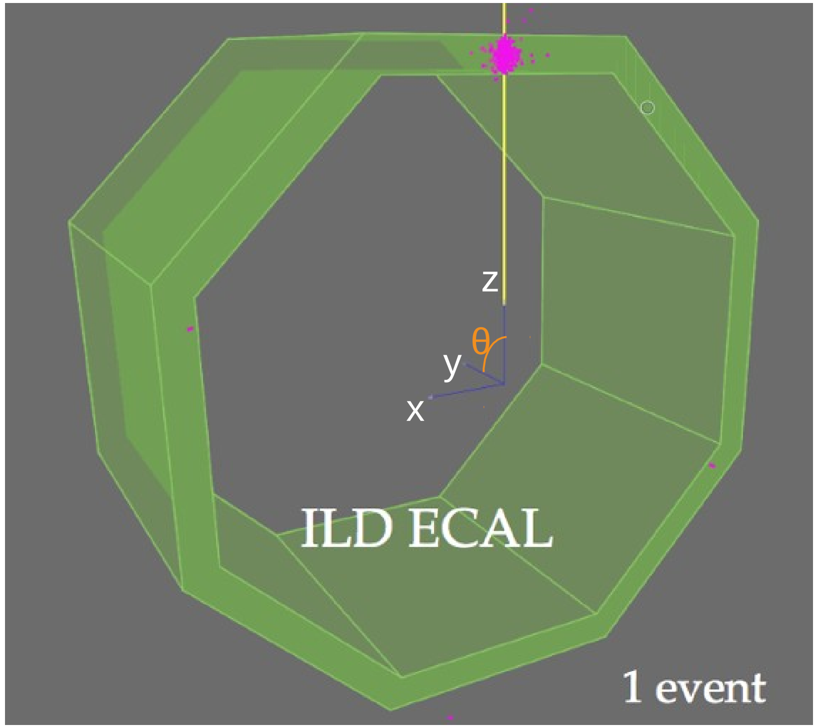
Angle and Energy Conditioning

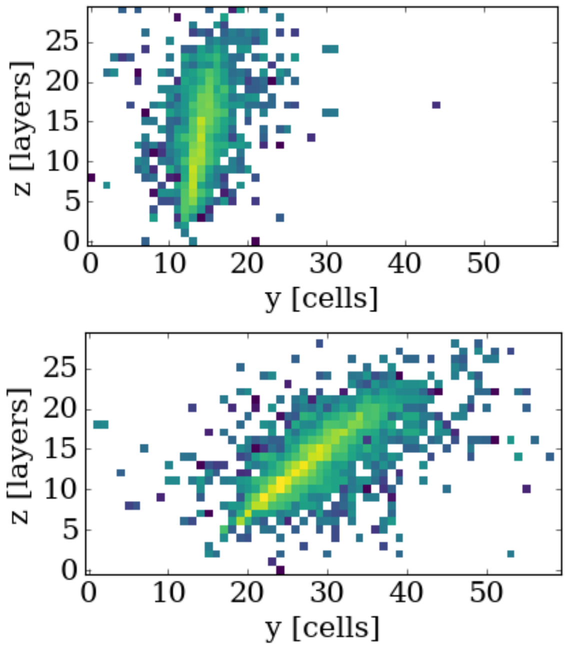
- **Multi-parameter conditioning** essential to **generalise** simulation tool
- **Normalising Flow** for latent sampling- **fast** sampling with multiple conditioning parameters
- Flow generates latent variables + Esum given angle and energy
- Additional **energy sum** conditioning in Post Processor- rescale per shower energy to pin down energy sum



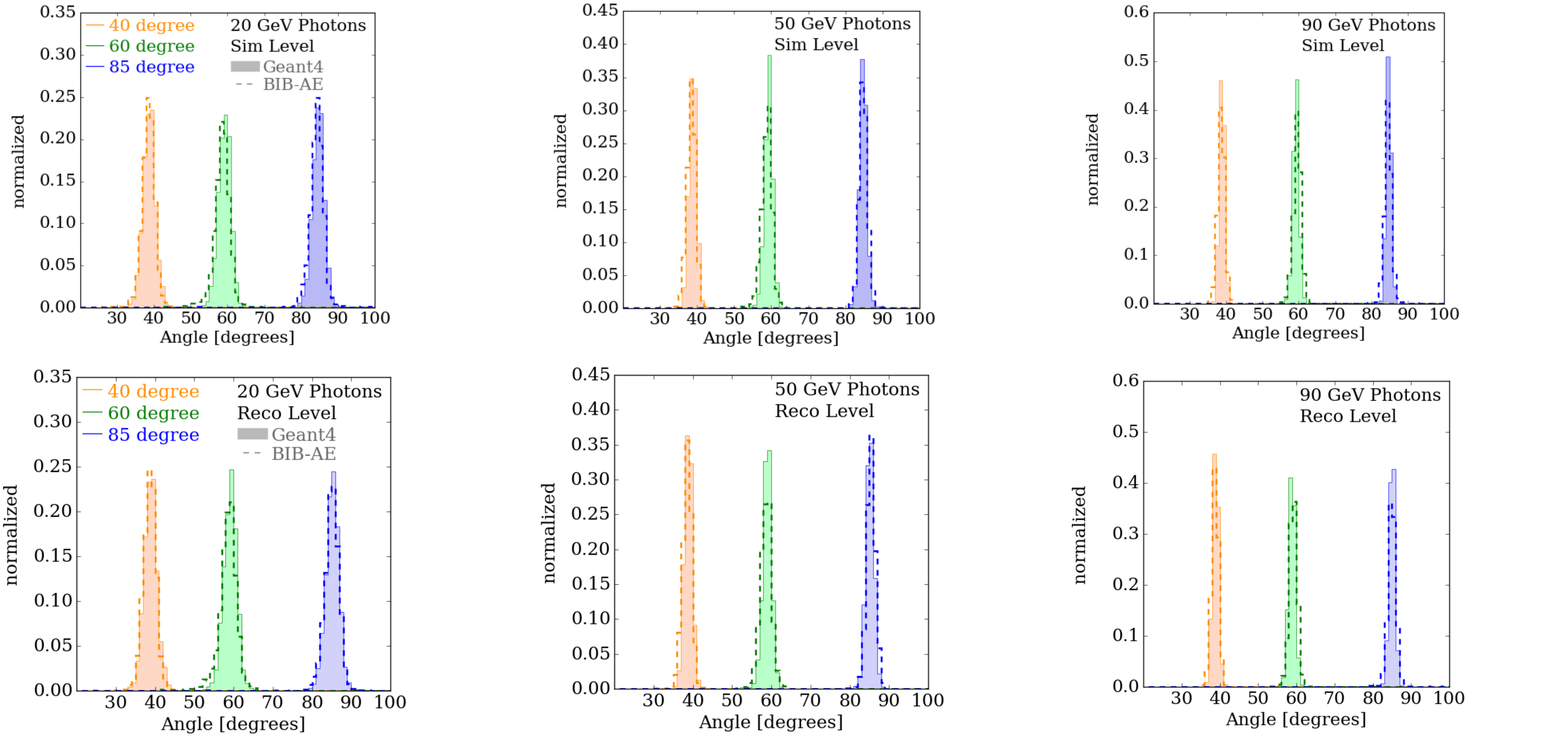
$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}}$

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

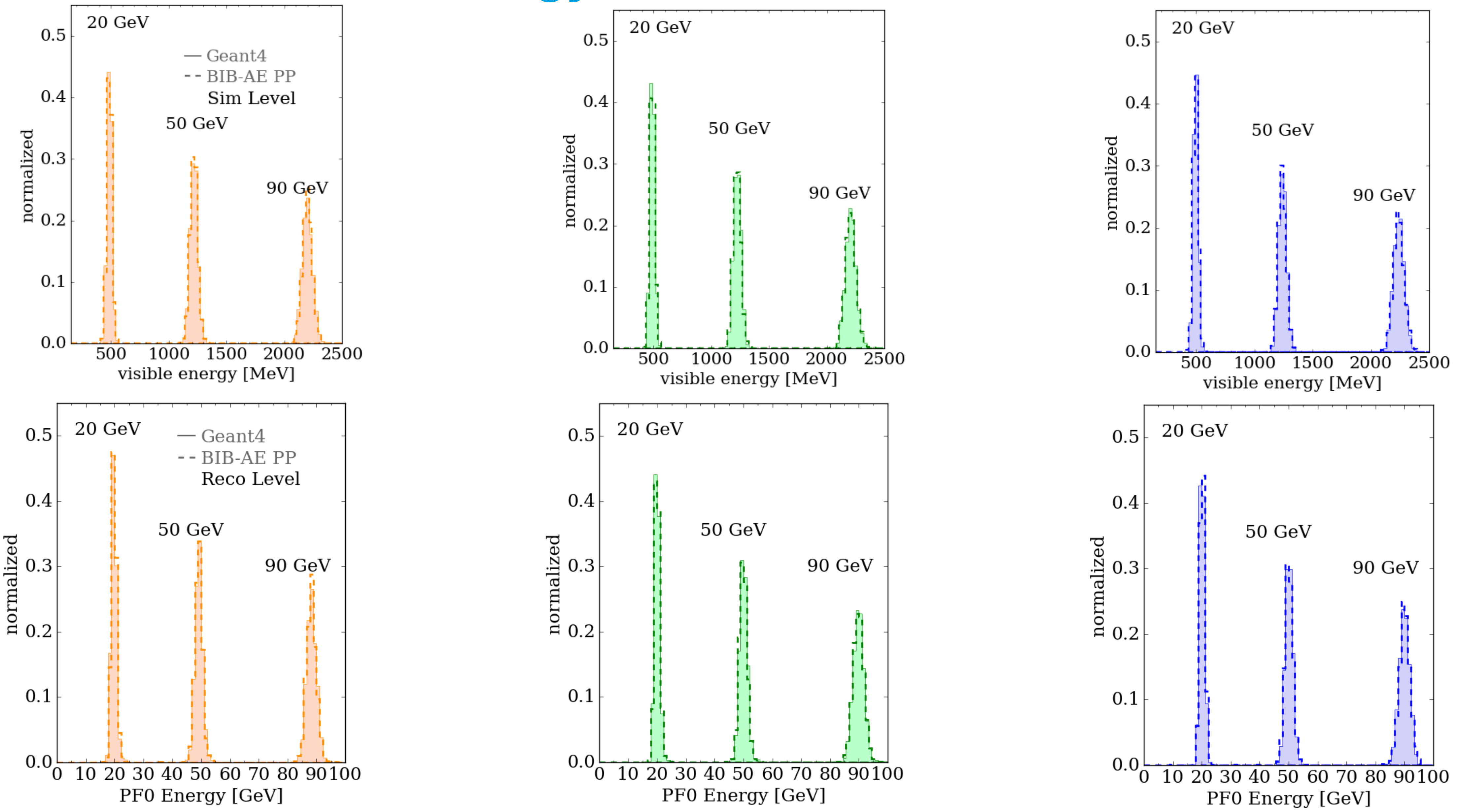




Results: Angular resolution- Sim vs Reco

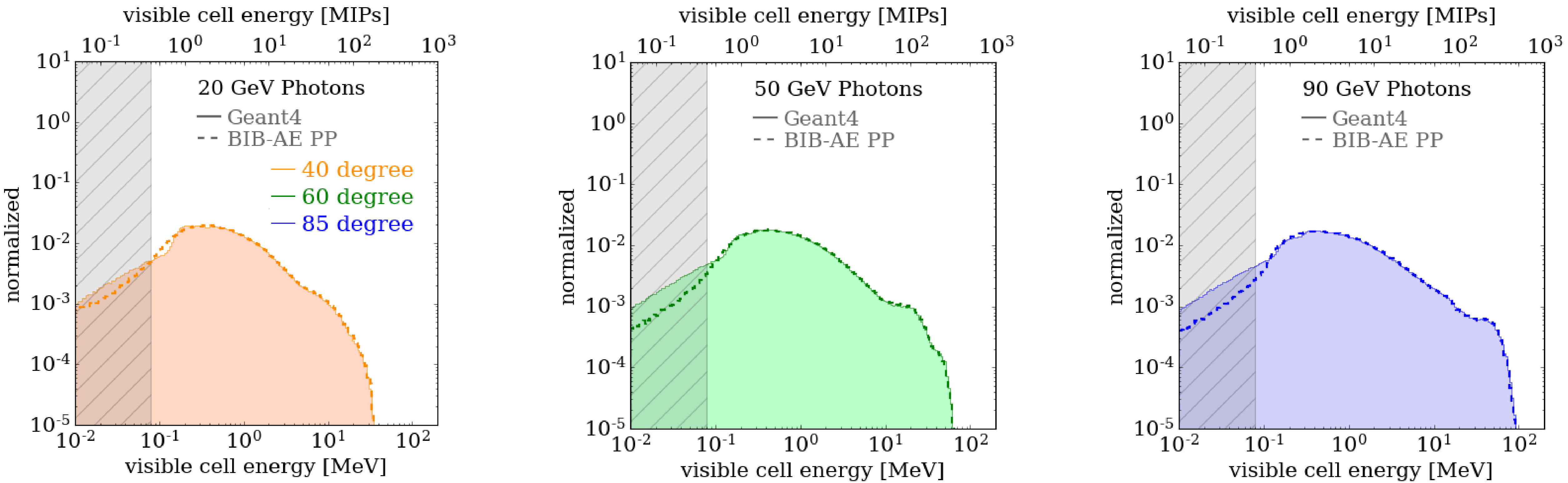


Results: Visible Energy Sum- Sim vs Reco

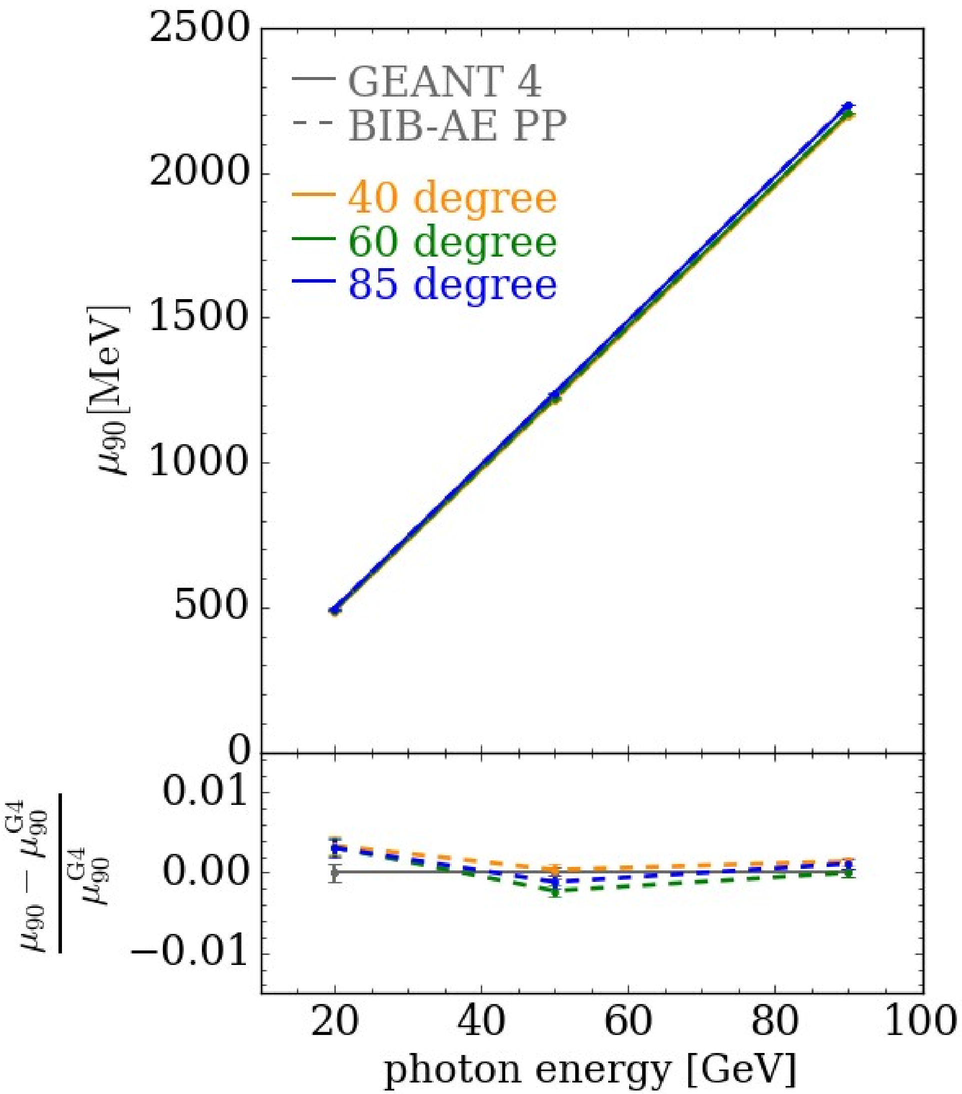


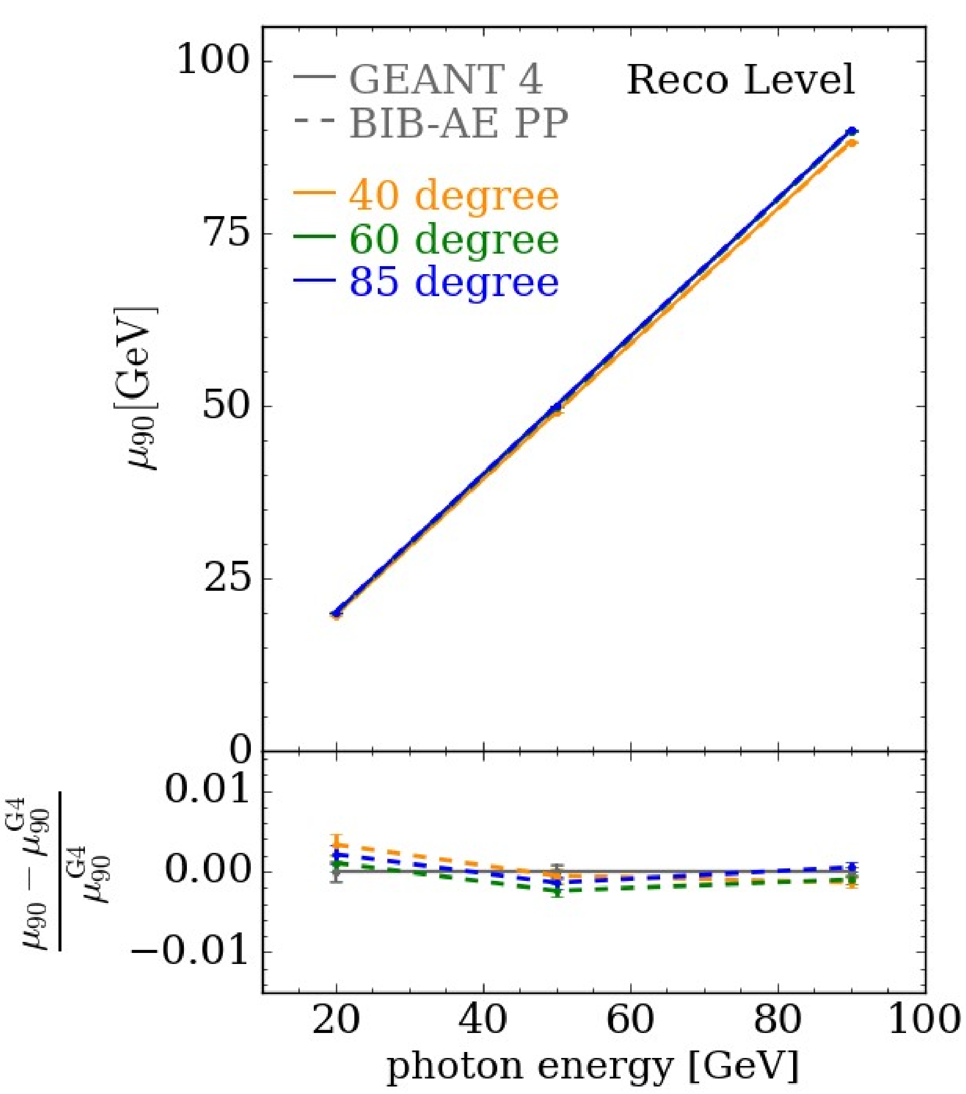
Results: Cell Energy Spectrum

- Post Processor Network retains its ability to correctly describe the cell energy distribution



Results: Energy linearity Sim vs Rec





Results: Energy resolution Sim vs Rec