Multi-parameter Conditioning of Generative Models for Fast Simulation of Highly Granular Calorimeter Showers

Sascha Diefenbacher, Engin Eren, Frank Gaede, Gregor Kasieczka,
Anatolii Korol, Katja Krüger, **Peter McKeown**[1], Lennart Rustige

[1] **Deutsches Elektronen-Synchrotron**

22.03.2023

peter.mckeown@desy.de

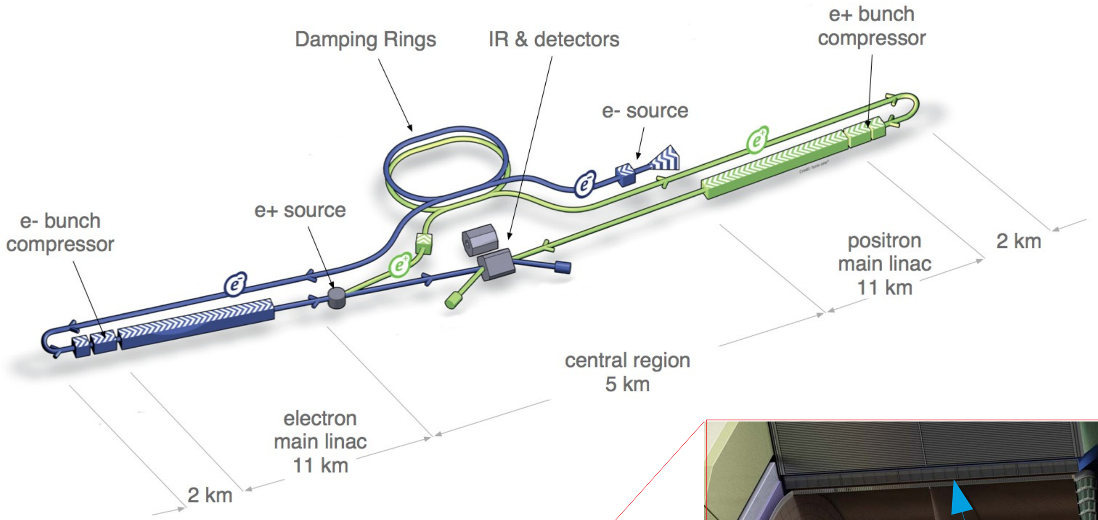
The ILD Concept

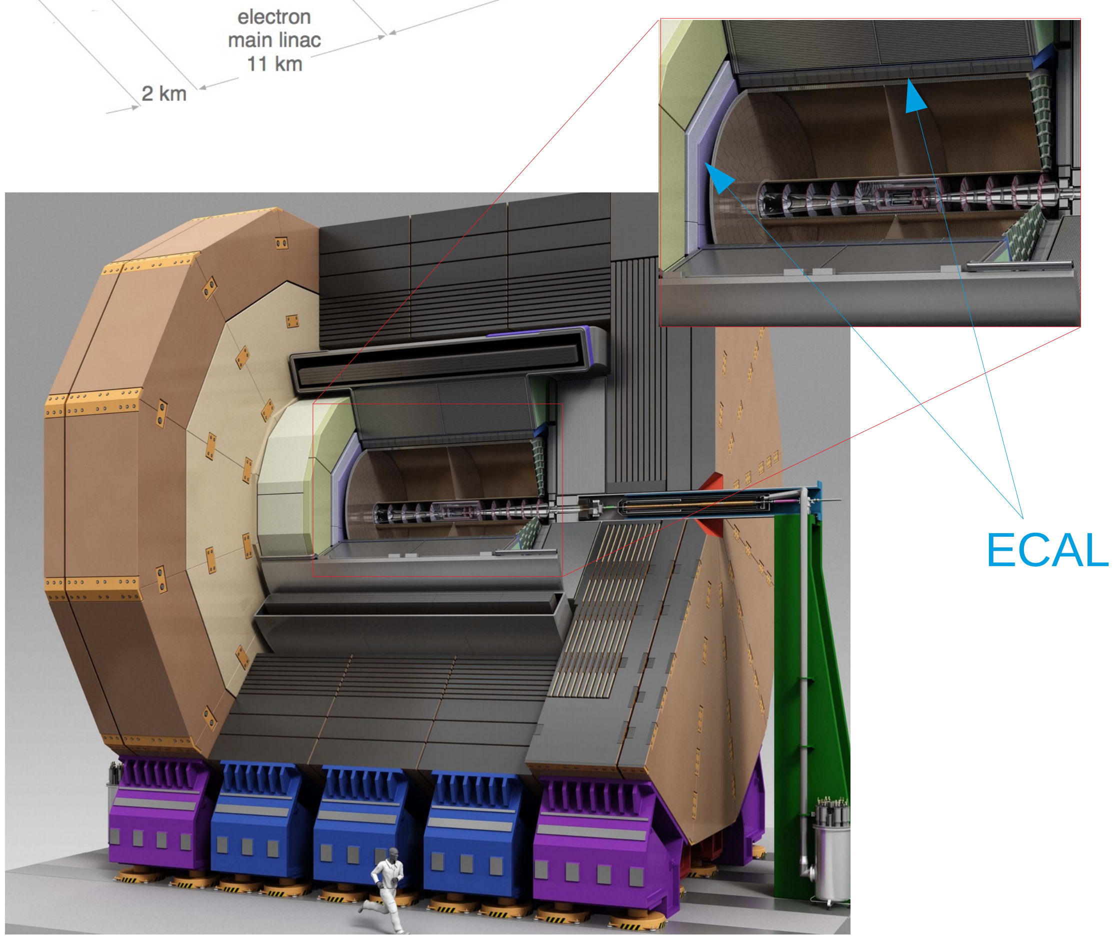
- Context: Future Higgs Factories

- Case Study: International Large Detector (ILD) concept for the International Linear Collider (ILC)

- 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**





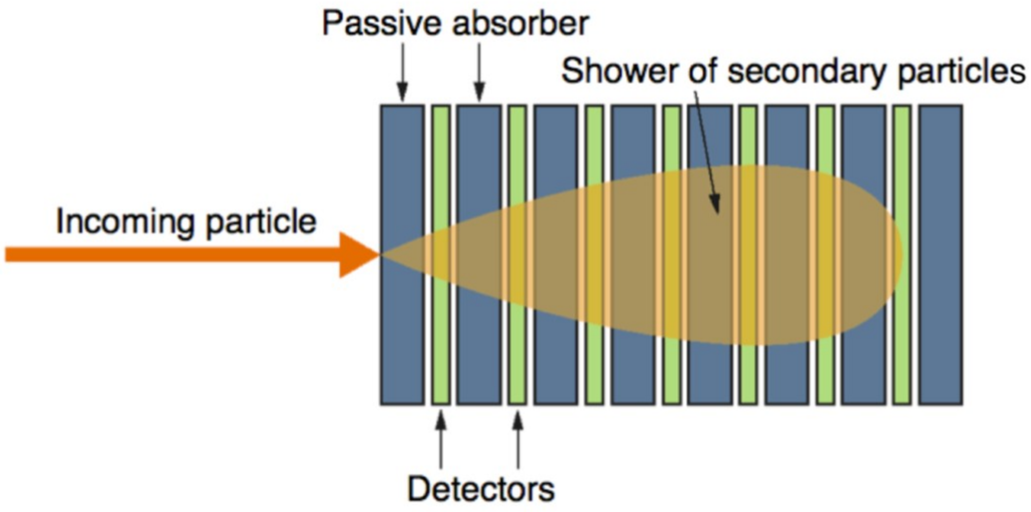
The ILD Electromagnetic Calorimeter

- 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 (Si) and passive (W) material
  - 5x5 mm^2 silicon cells
  - ~ 80 million channels for barrel

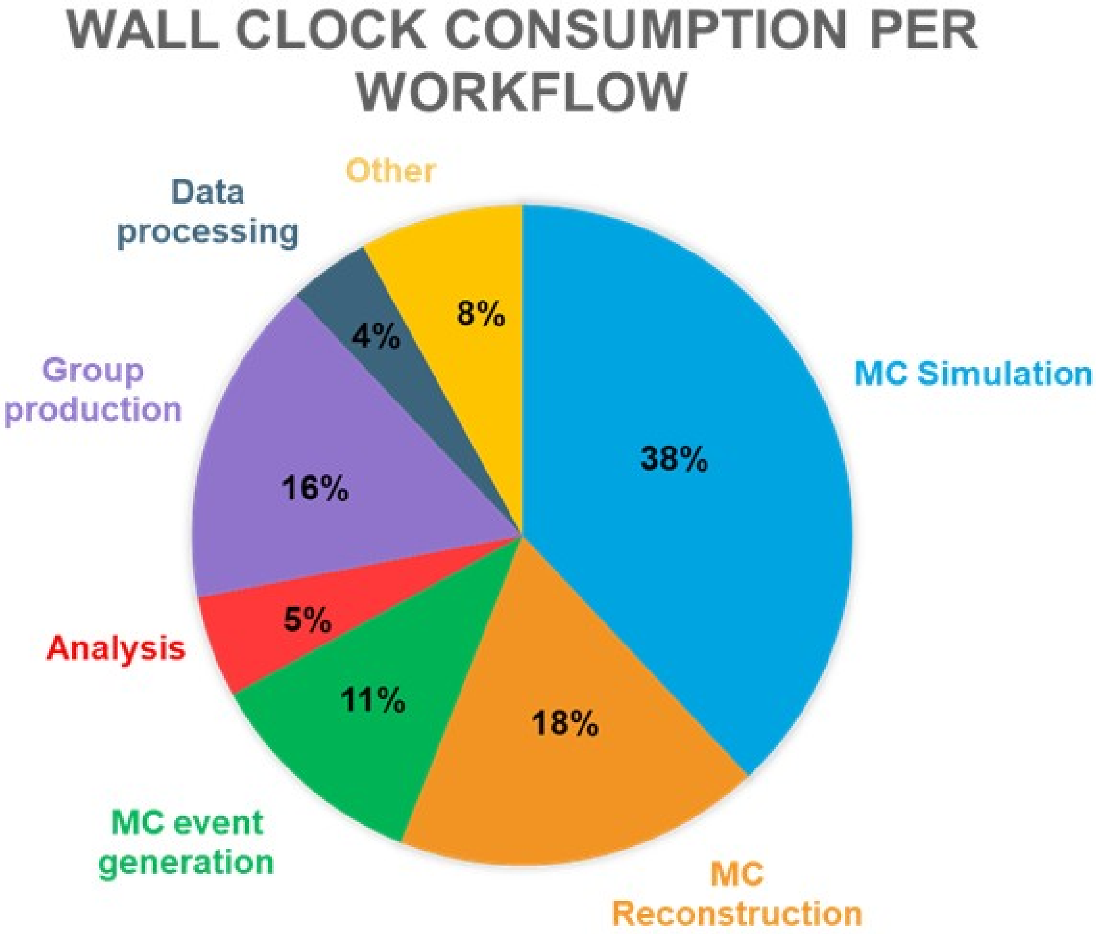
c.f. a few cm^2 for ATLAS/CMS ECAL (before High Lumi)



- High granularity → **High fidelity**

The Strain on HEP Computing Resources

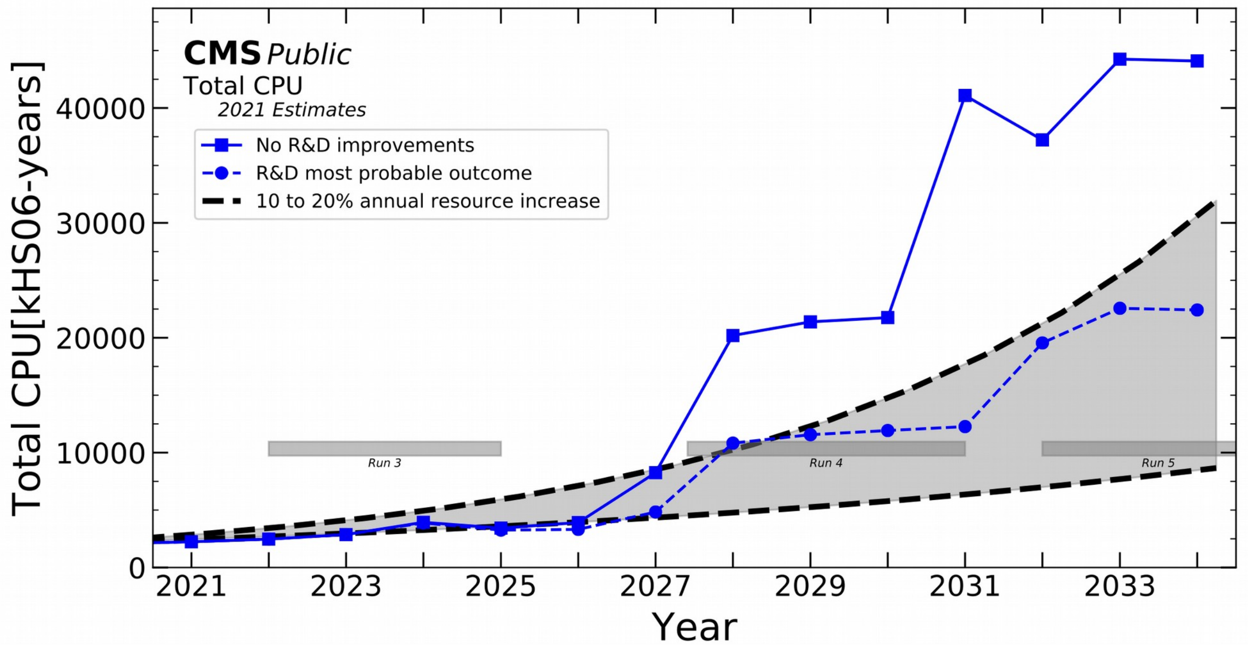
- Ever increasing demand for computing resources
  - MC simulation largest fraction
  - Calorimeters most intensive part of detector simulation



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

The Strain on HEP Computing Resources

- Ever increasing demand for computing resources
  - MC simulation largest fraction
  - Calorimeters most intensive part of detector simulation
- Projected computing resources required far outstrip what will be available
  - E.g HL-LHC
- Future lepton colliders also benefit from much faster MC
- **Generative models** potentially offer orders of magnitude speed up

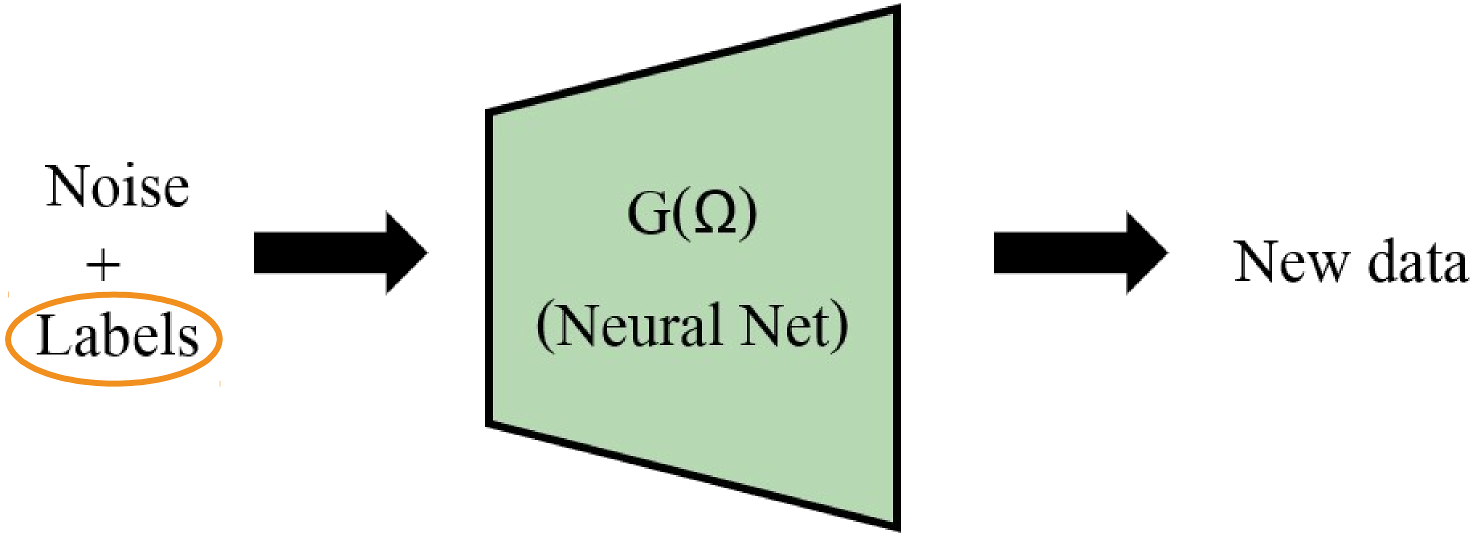


CMS Collaboration, Offline and Computing Public Results (2021), https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults

Generative Models

Butter et al., **GANplifying event samples**, [SciPost Phys. 10, 139, (2021)]

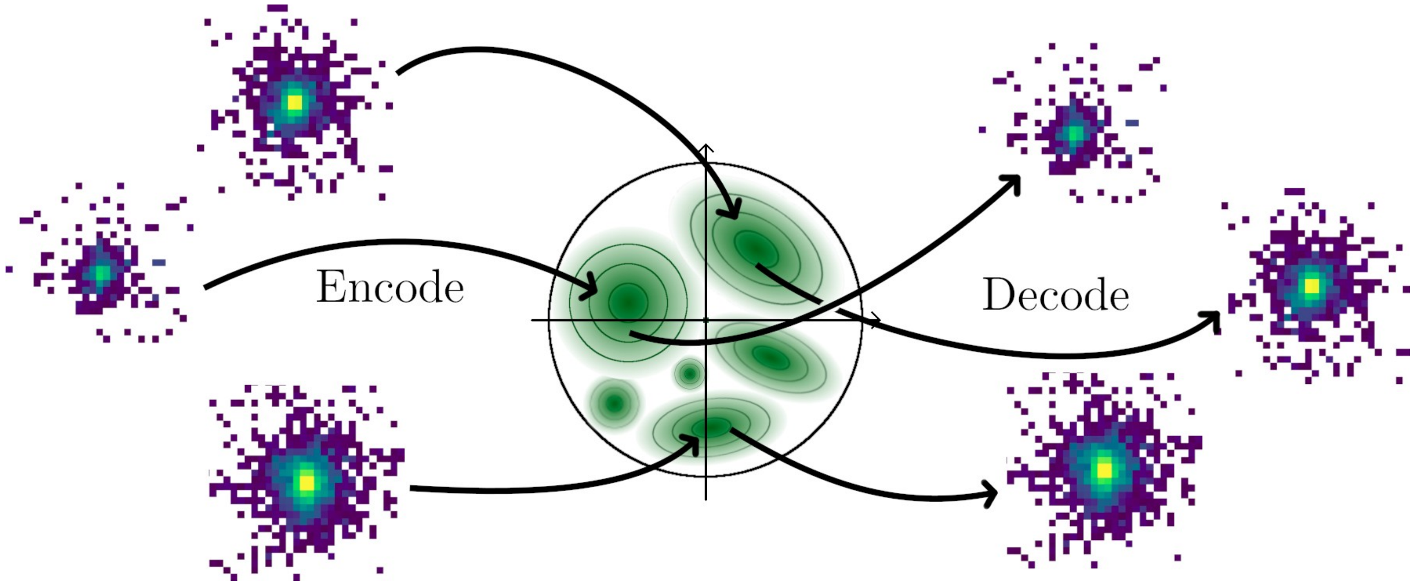
- Approach to fast simulation- **amplify statistics**
- Aim: augment full physics-based GEANT4 simulation by learning a transfer function
- Promising solution: **generative models**
  - Generate new samples following the distribution of original data
  - Map random noise to data
  - Highly parallelizable
  - **Conditioning (E.g., E, θ, φ)**



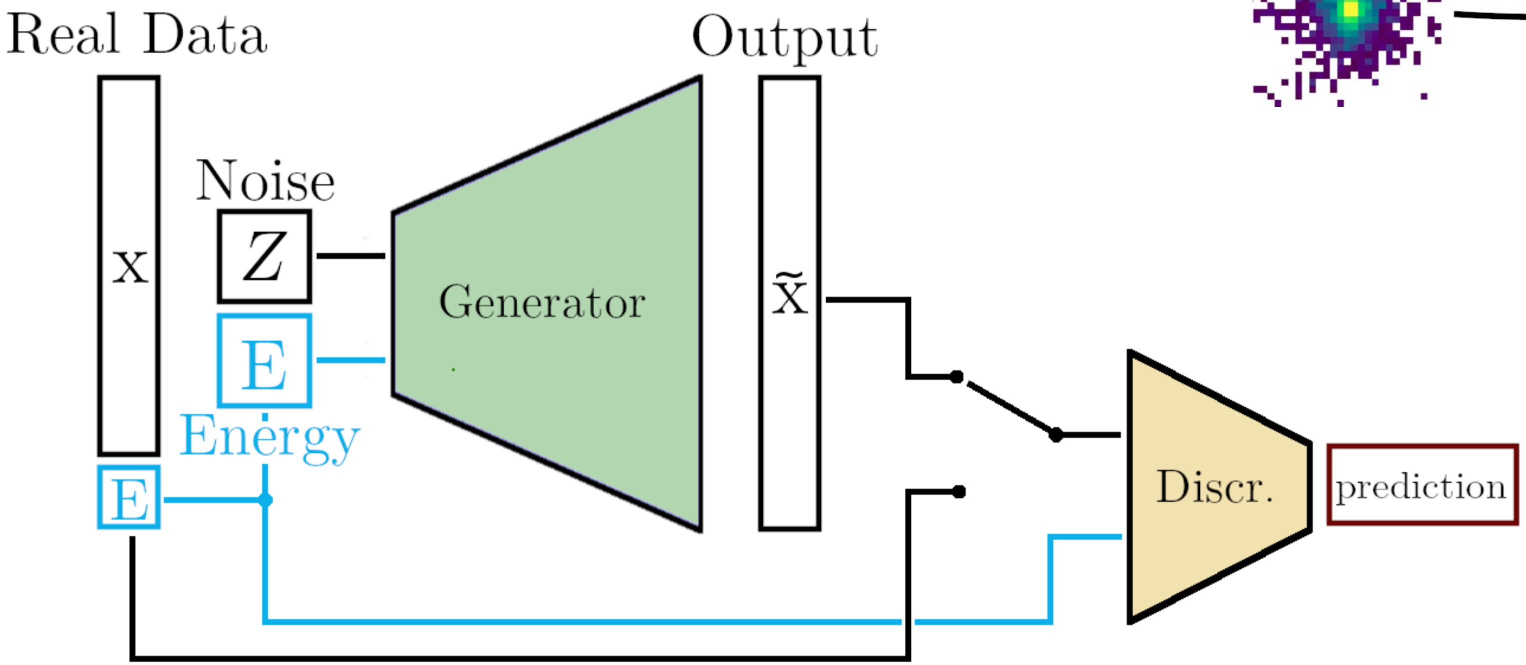
Common Generative Models

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

Variational Autoencoder (**VAE**)



Generative Adversarial Network (**GAN**)



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

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

Architectures: BIB-AE

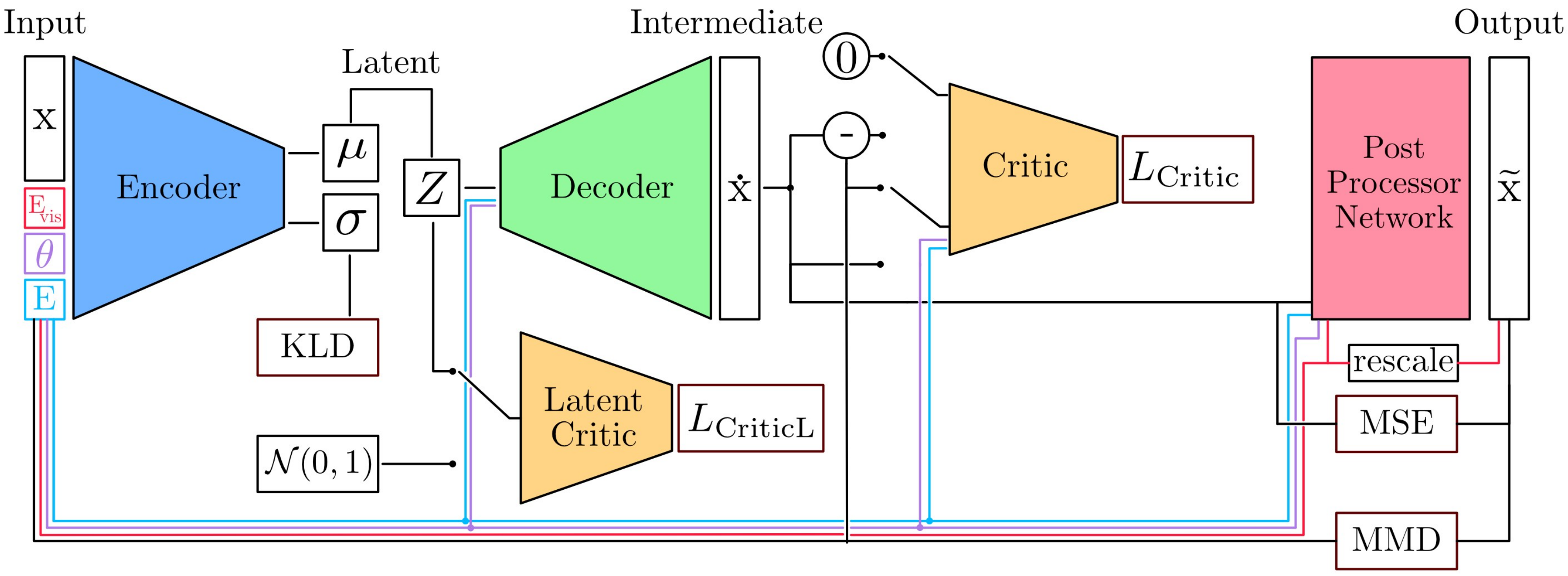
Voloshynovskiy et. al: Information bottleneck through variational glasses, arXiv:1912.00830, (2019)

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

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

Bounded-Information Bottleneck Autoencoder (BIB-AE)

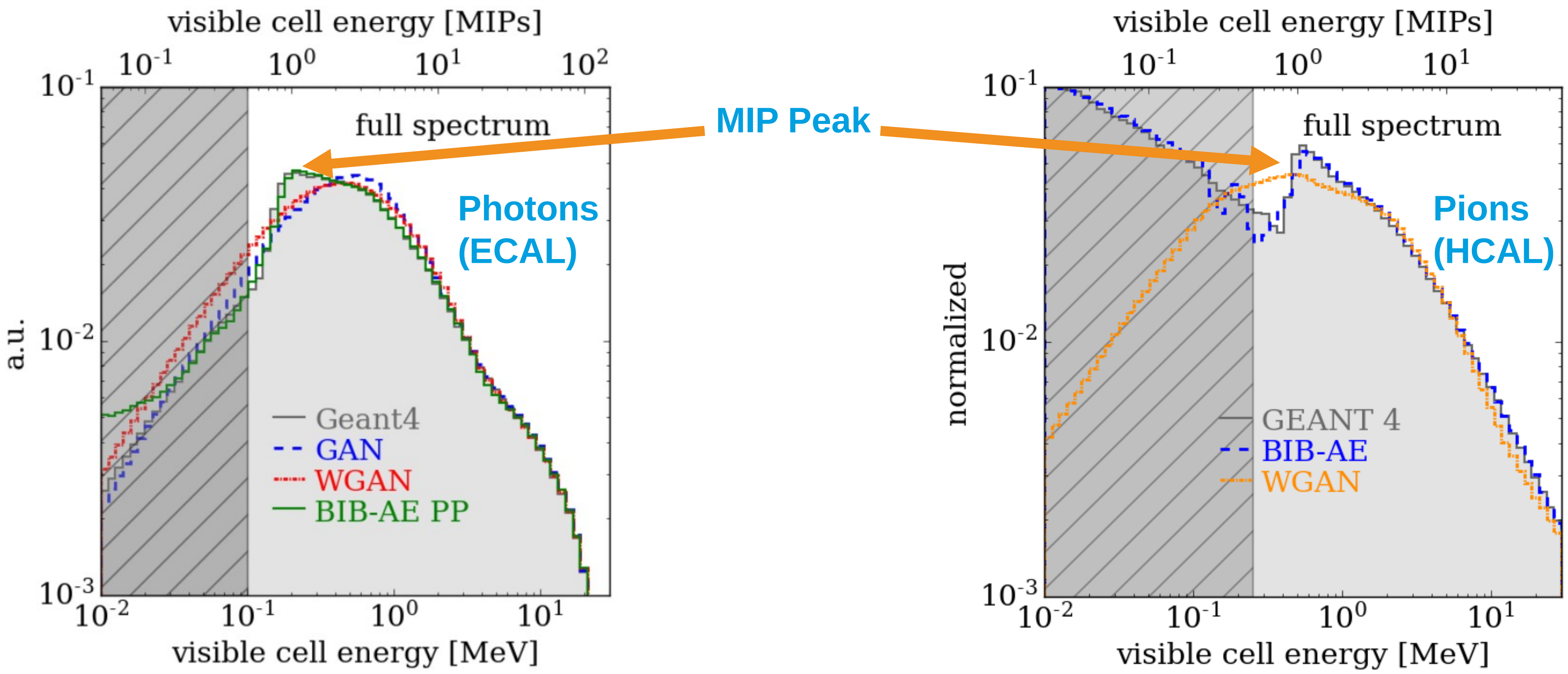
- **Unifies** features of both **GANs** and **VAEs**
- **Post-Processor** network: Improve per-pixel energies; second training



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

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

Previous work

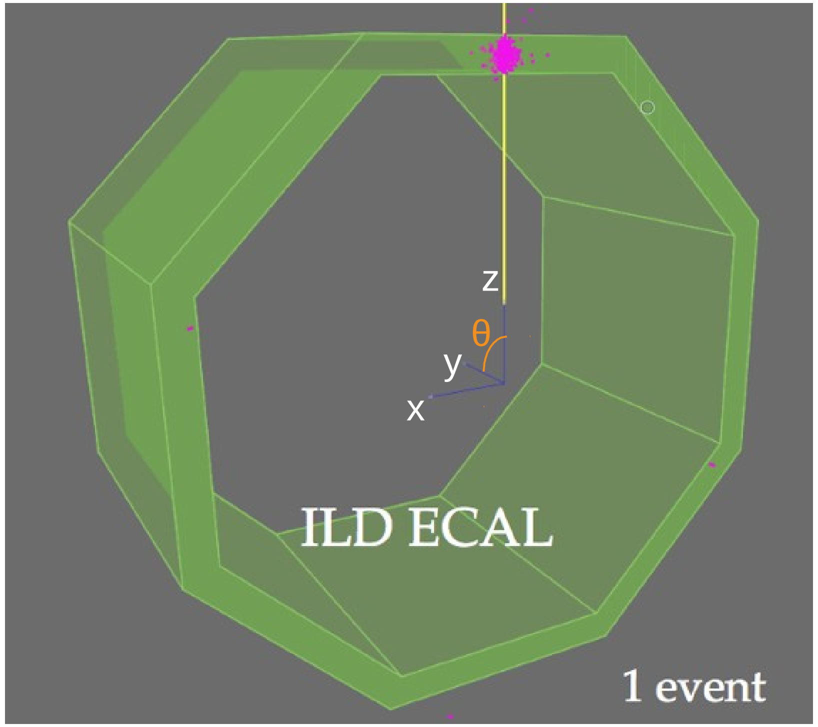


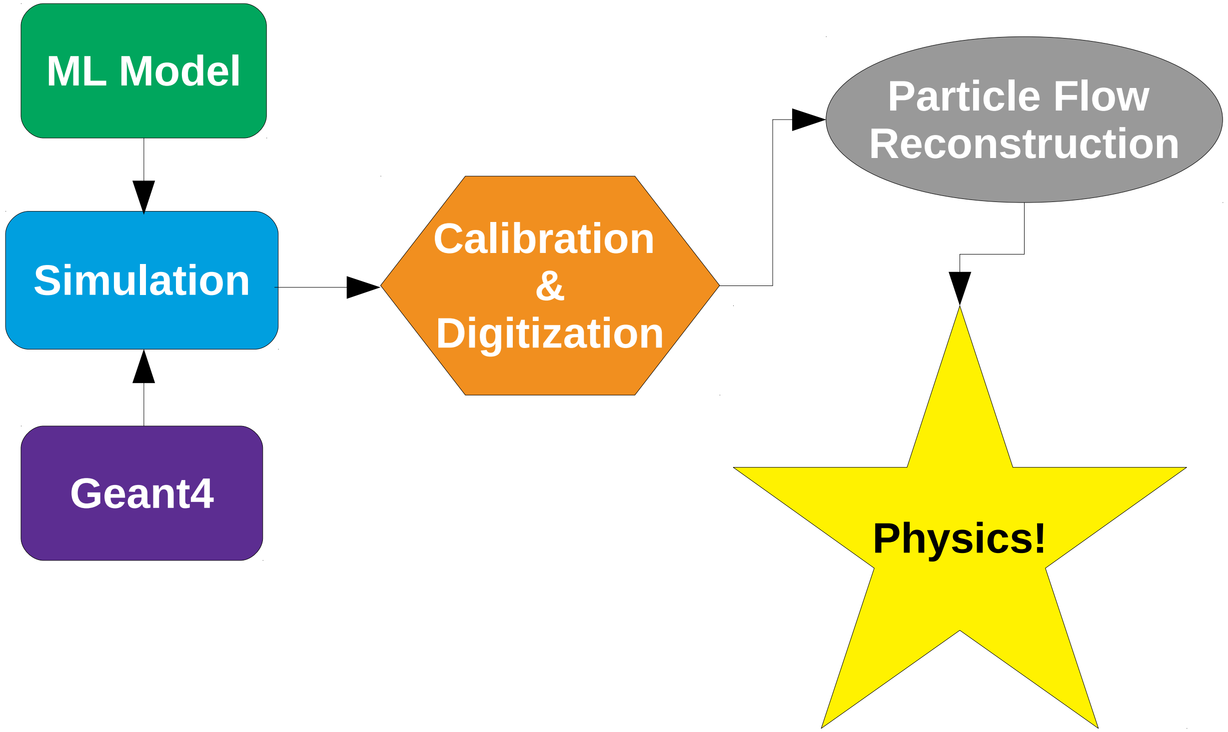
- **Cell energy spectrum** has a very steep rise (Minimum Ionizing Particle (MIP) peak- important for detector calibration)
  - Difficult to model with an adversarial approach
  - More sophisticated BIB-AE successfully reproduces MIP peak

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

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

Two New Angles





Multi-Parameter Conditioning

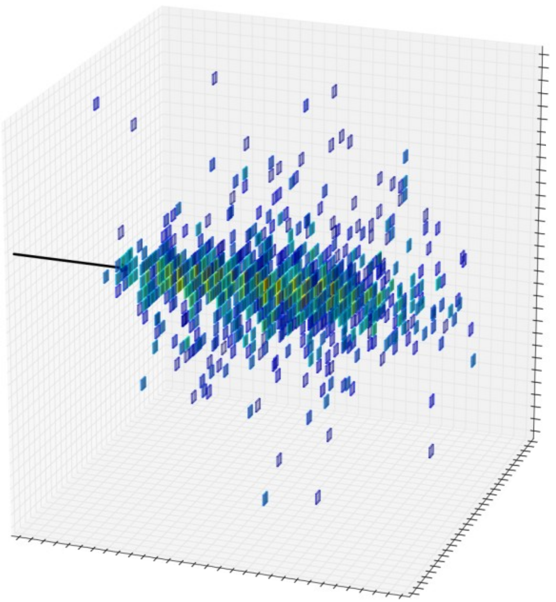
- **Simultaneous conditioning** on multiple parameters crucial for a general simulation tool
  - Start with **photons**
  - Vary incident **energy and angle**

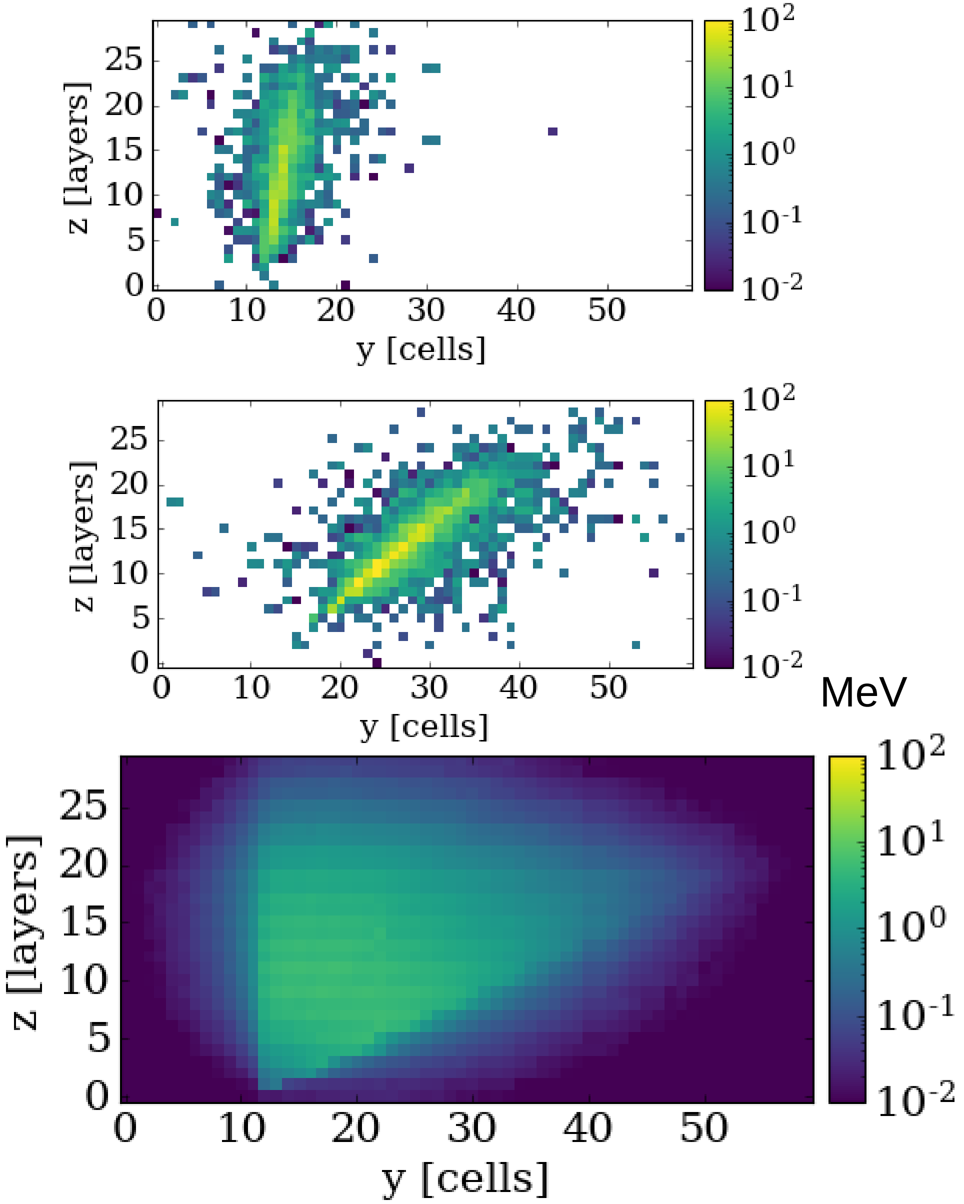
Study effects of reconstruction

- Ultimate goal is **physics performance** after **reconstruction**
- Want our models to **maintain performance** after reconstruction
  - **Compare observables** before and after

Training data

- 500,000 **photons** with fixed incident point
- Vary **energy**: **10-100 GeV**
- Vary **polar angle** in one direction: **90°-30°**
- Project to regular grid
  - Shape (30,60,30) (x,y,z)

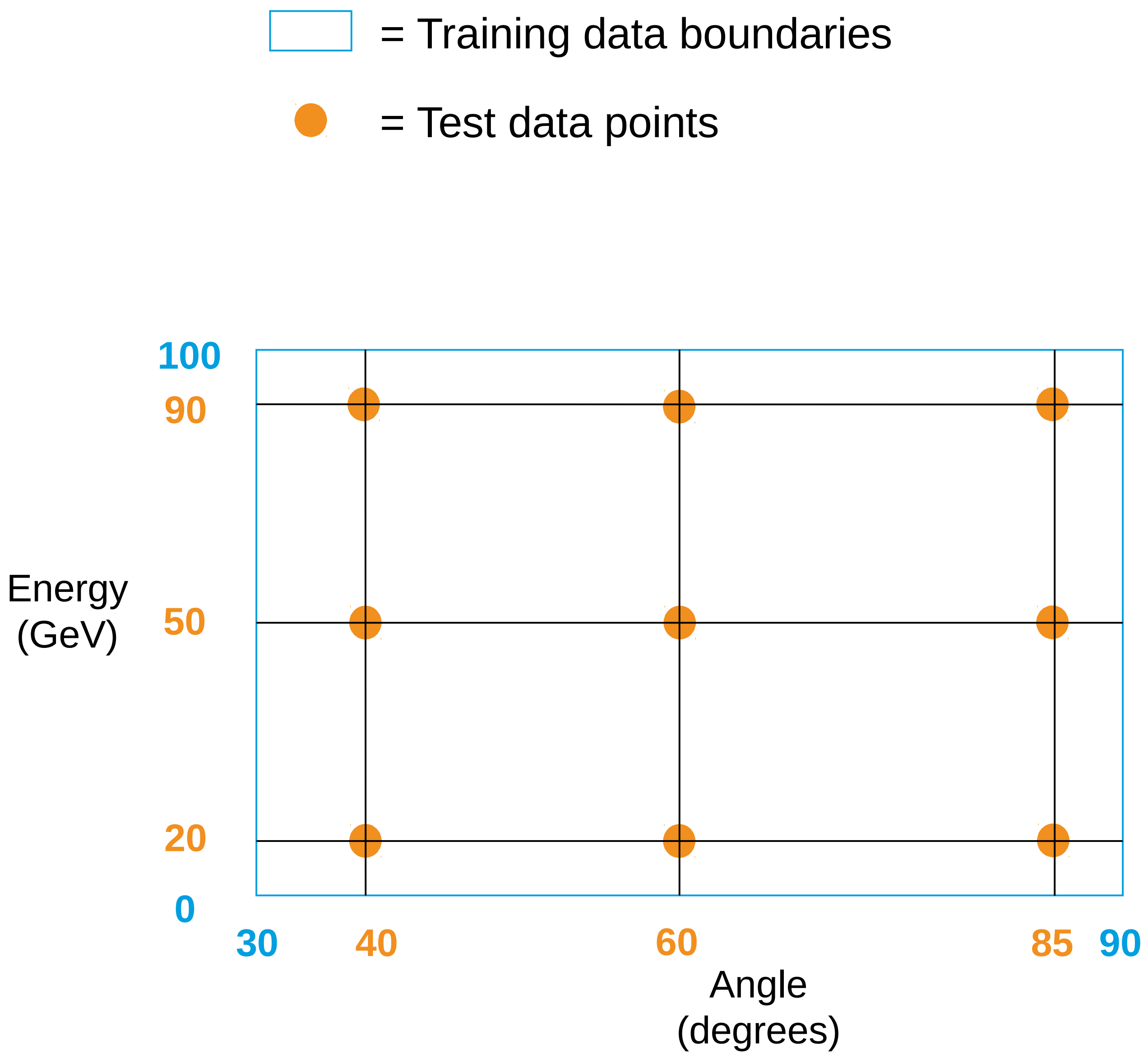




Test data

- Choose **fixed combinations** of energies and angles
- **9 points** in the phase space in total
- Avoid phase space boundaries

- Investigate how well generative model learns **key physics** properties of showers:
  - Visible energy sum
  - Number of hits and Hit energy spectrum
  - Shower profiles
  - Angular response

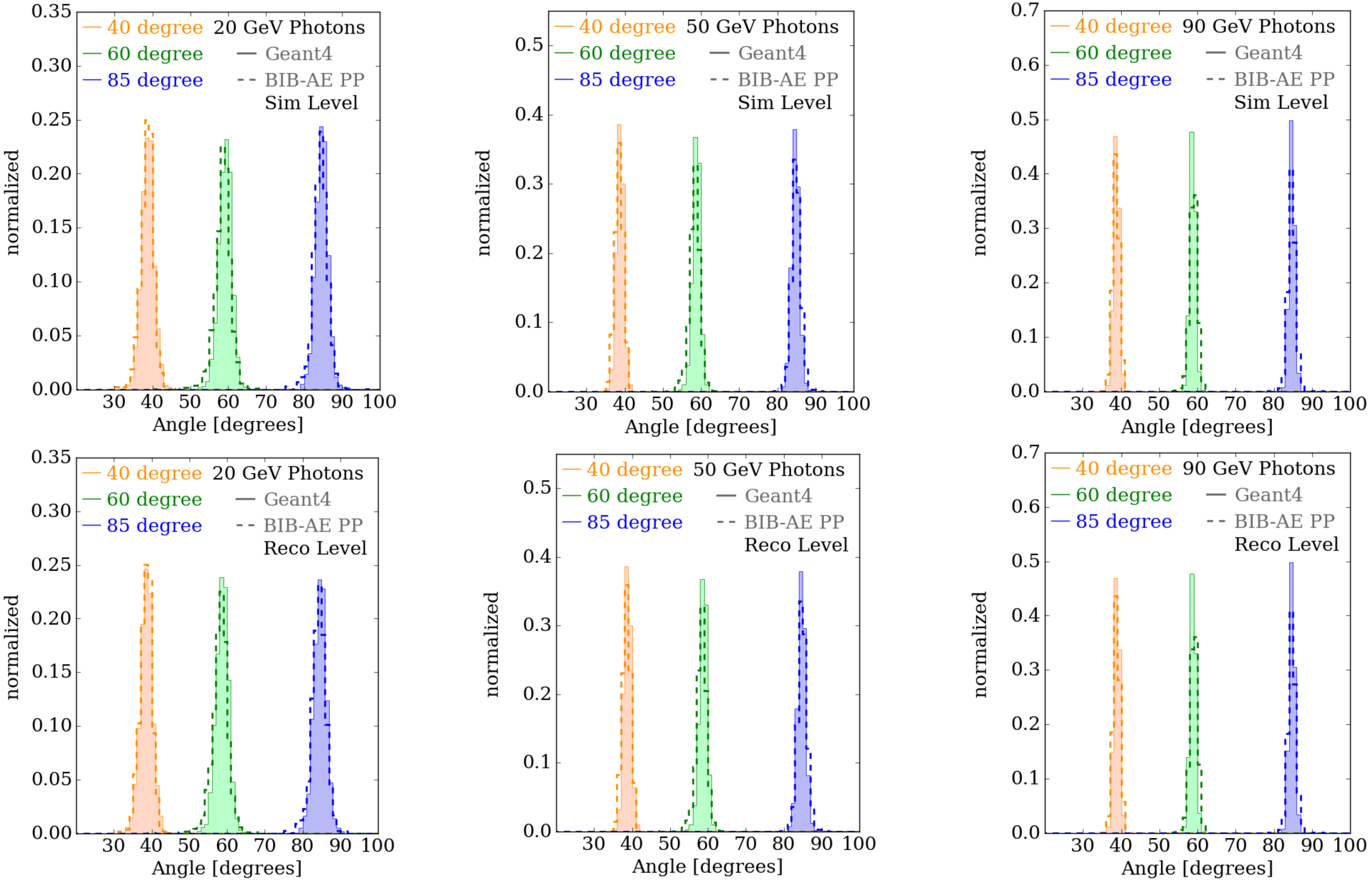


Results: Angular resolution- Sim vs Reco

20 GeV | 50 GeV | 90 GeV

Sim Level

Reco Level



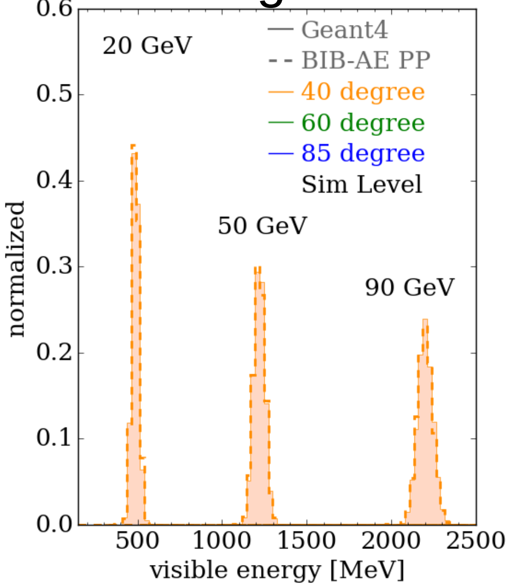
Results: Visible Energy Sum- Sim vs Reco

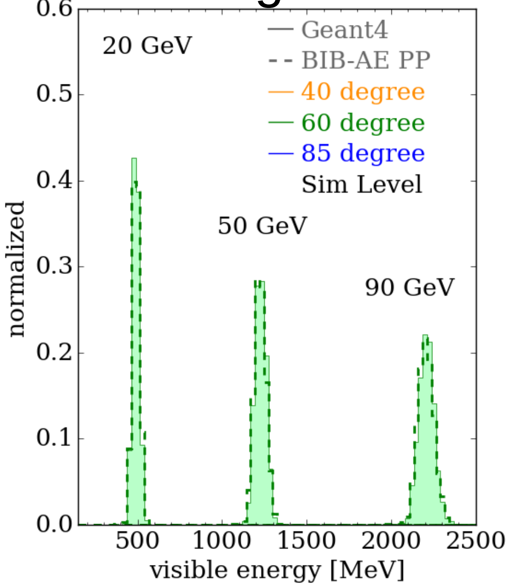
40 degrees

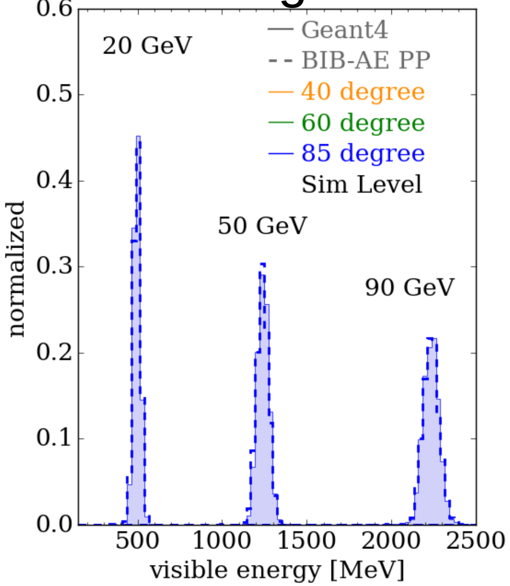
60 degrees

85 degrees

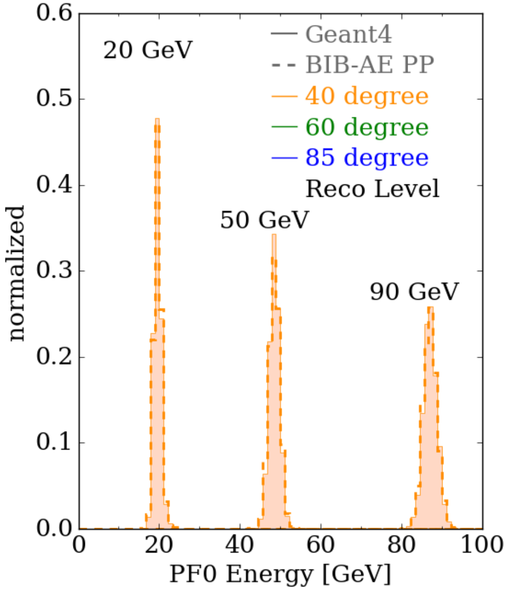
Sim Level

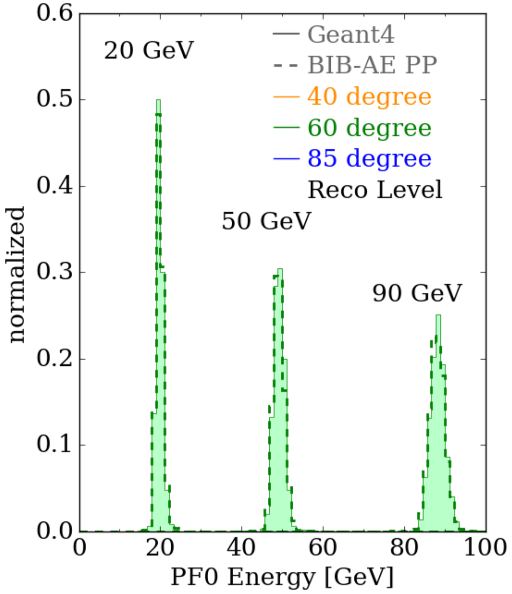


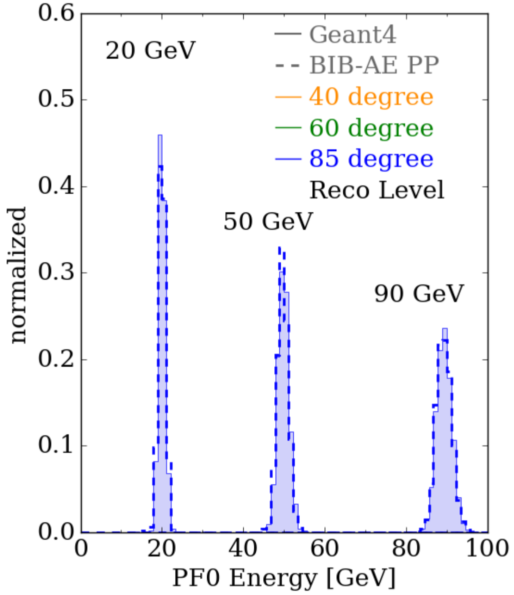




Reco Level



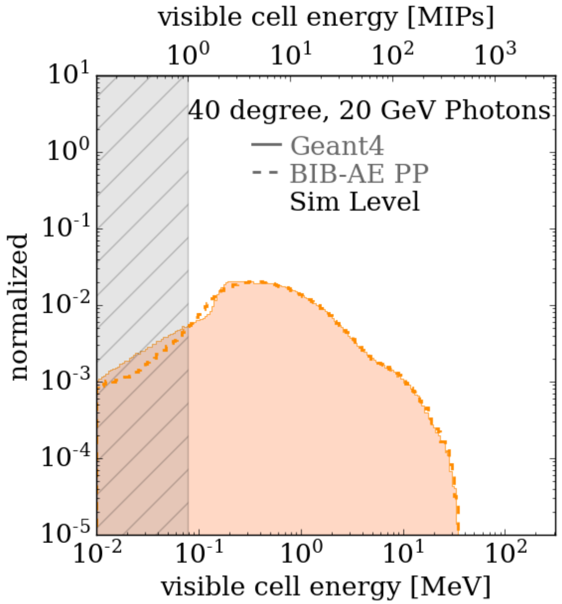


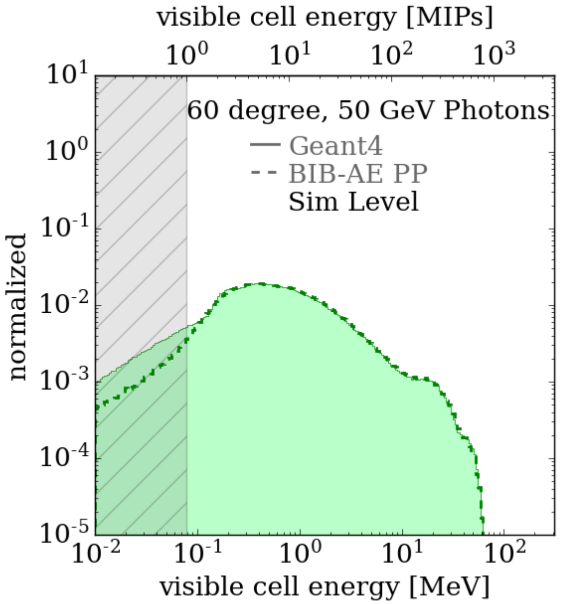


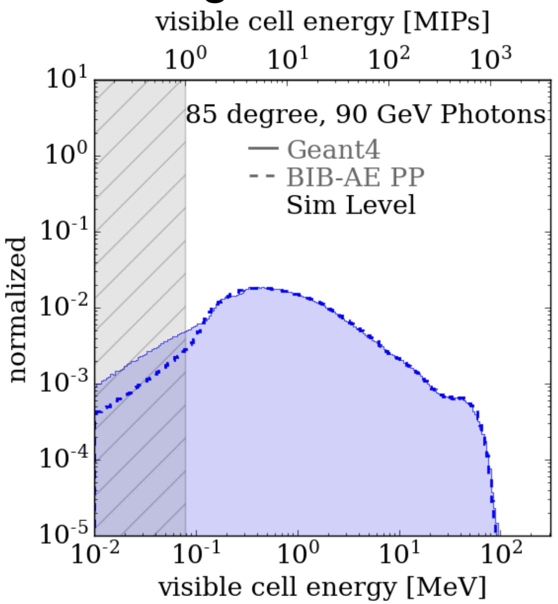
Results: Cell Energy Examples Sim vs Reco

40 degrees, 20 GeV | 60 degrees, 50 GeV | 85 degrees, 90 GeV

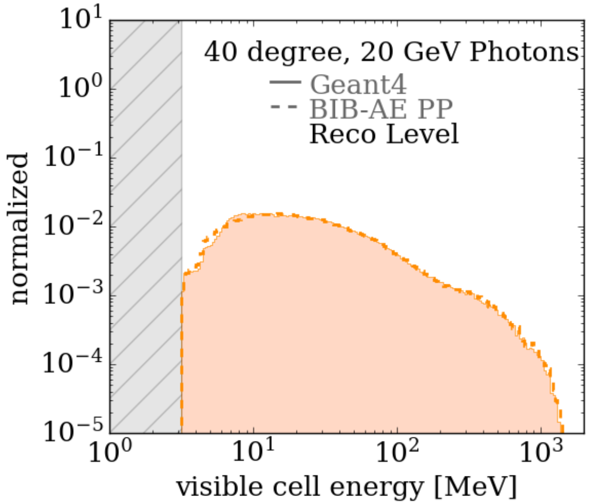
Sim Level

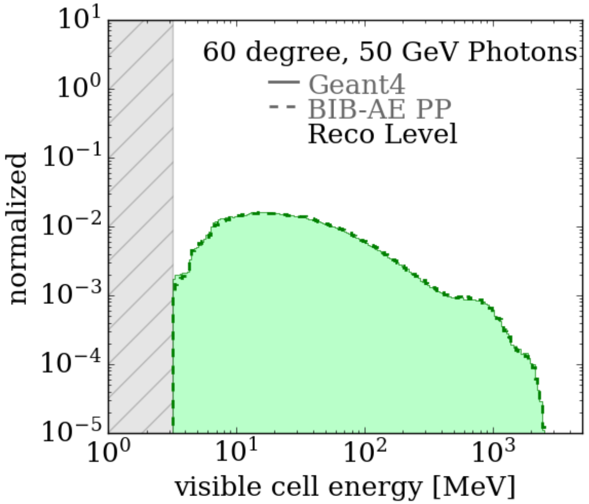


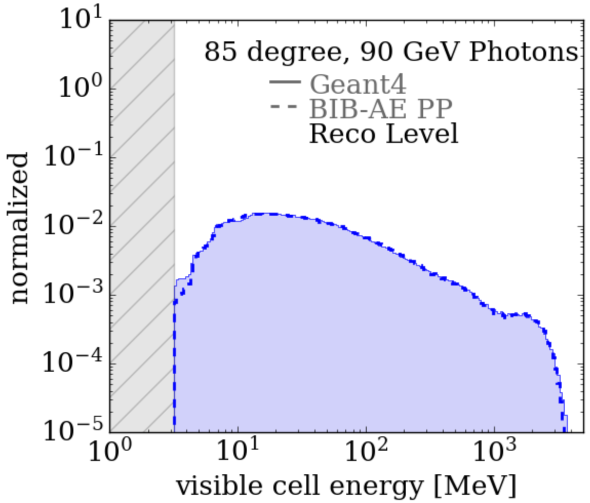




Reco Level







Conclusion

Achieved

- Generative models hold promise for **fast** simulation of calorimeter showers with **high fidelity**
- Demonstrated high fidelity simulation of **photon** showers with **angular and energy conditioning**
- Investigated generative model performance after **reconstruction**
    - **Performance remains high**

Paper out soon!

Next Steps

- Develop strategy for dealing with **arbitrary incident positions**
    - Dealing with **irregular geometries**- see talk by A. Korol (T 86.6) at 18:45
- Integration into **software chain** and study **physics performance**

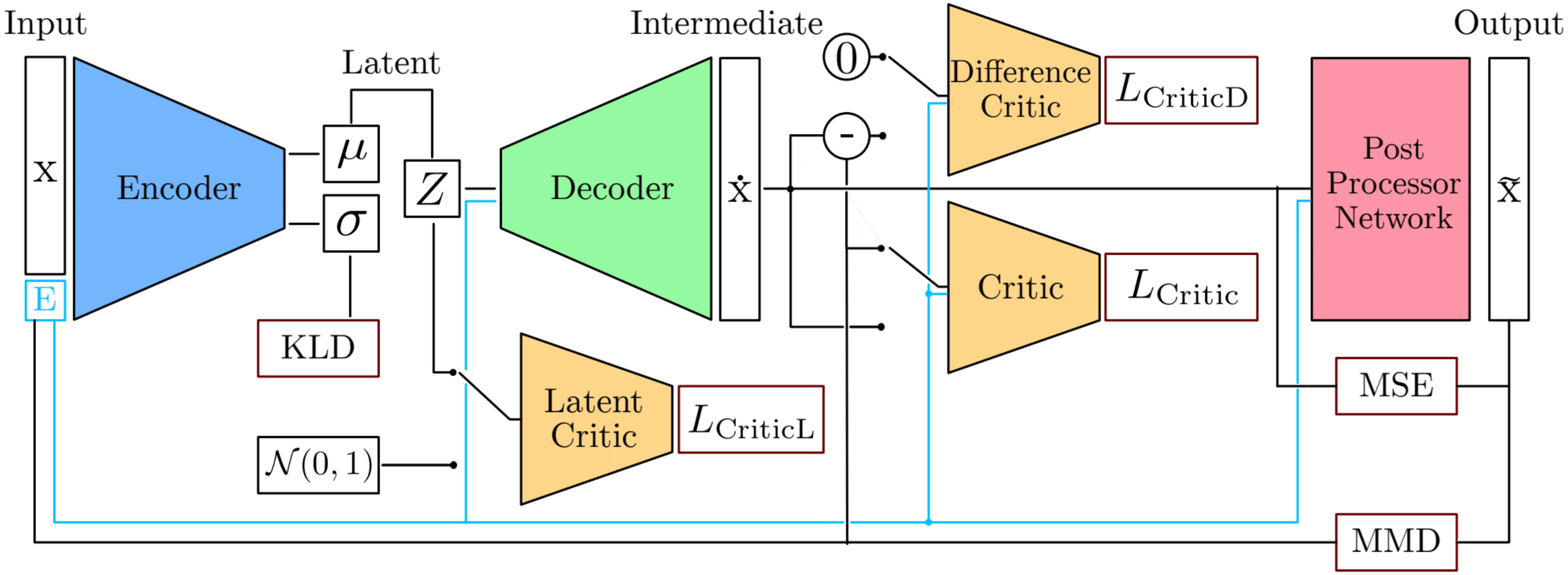
Backup

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



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

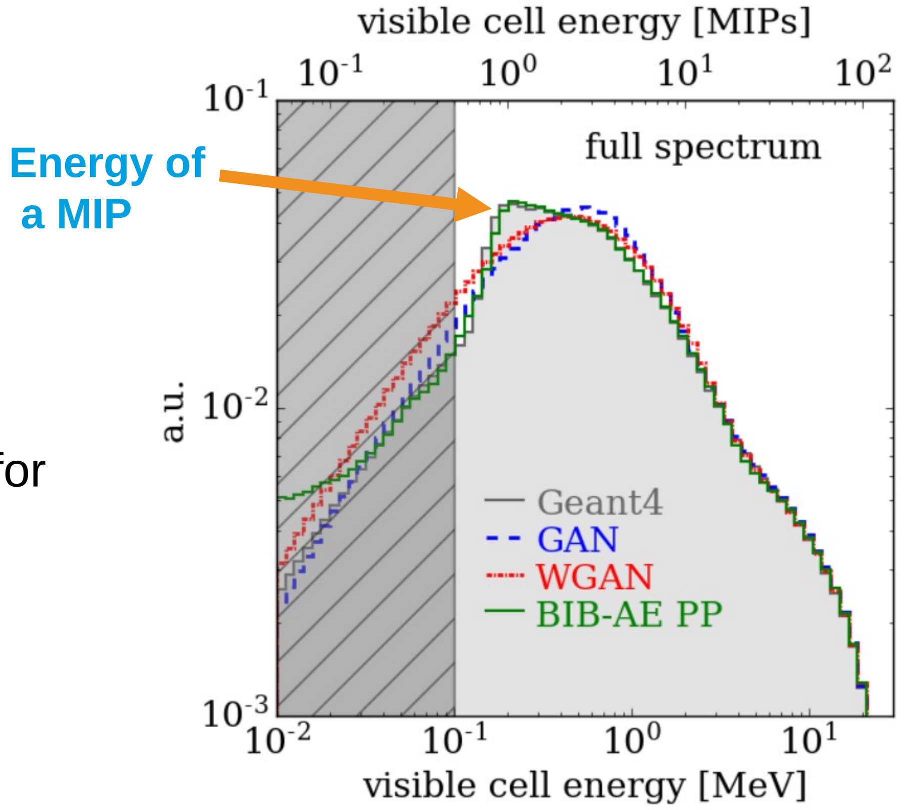
$$L_{\text{Post}} = \text{MMD} + \text{MSE}$$

Adaption to Highly Granular Calorimeter Shower Data

- **Highly granular** calorimeter data is very **sparse**
  - Causes problems for an MSE based loss
  - Switch to a discriminator based approach

- **Cell energy spectrum** has a very steep rise (MIP peak- important for calibration)
  - Difficult to model with an adversarial approach...

- Offload to separate **Post Processor** network:
  - 3D convolutions, kernel size 1
  - MSE loss and Sorted Kernel MMD loss
  - Encourage network to modify individual pixels



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

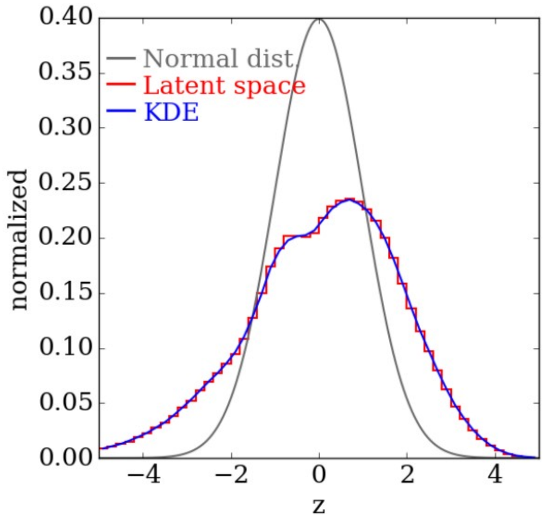
Computing Time for Inference

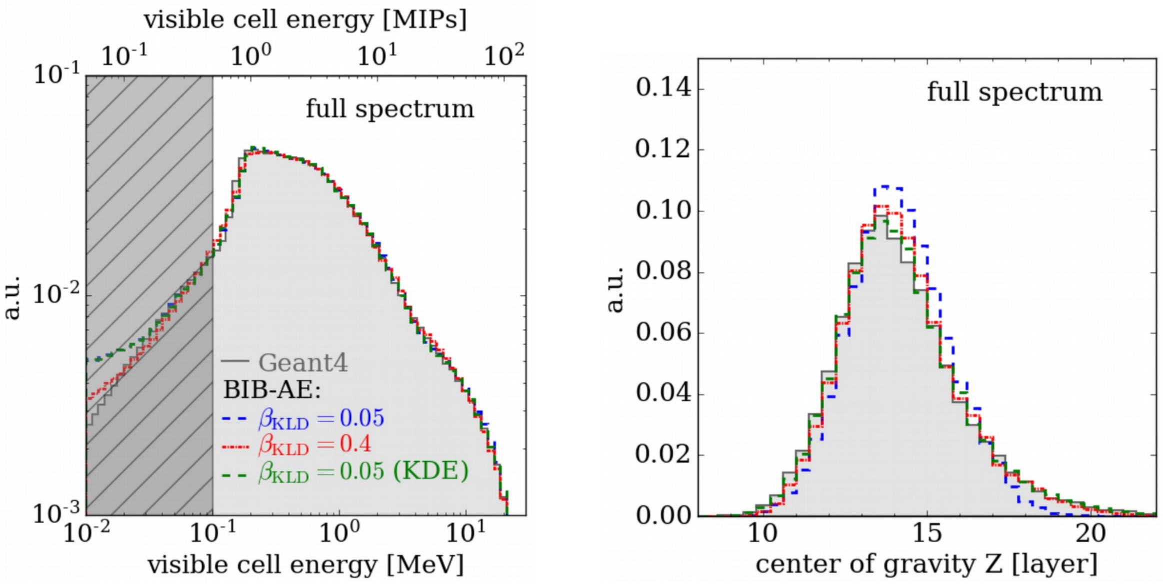
Hardware	Simulator	Time / Shower [ms]	Speed-up
CPU	GEANT4	4417 ± 83	$\times 1$
	BIB-AE	362 ± 2	$\times 12$
GPU	BIB-AE	4.32 ± 0.09	$\times 1022$

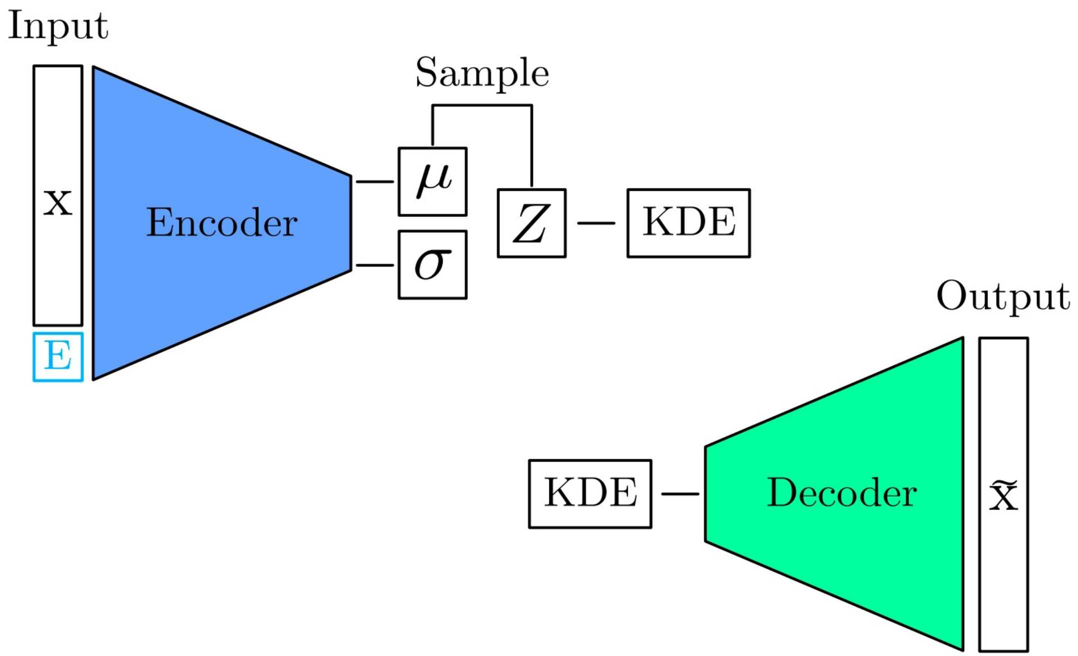
Speed-up of as much as three orders of magnitude on single core of Intel® Xeon® CPU E5-2640 v4 and NVIDIA® A100 for the best performing batch size

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 (**normalising flow**)
- **Improve** modeling of **shower shape** (center of gravity)

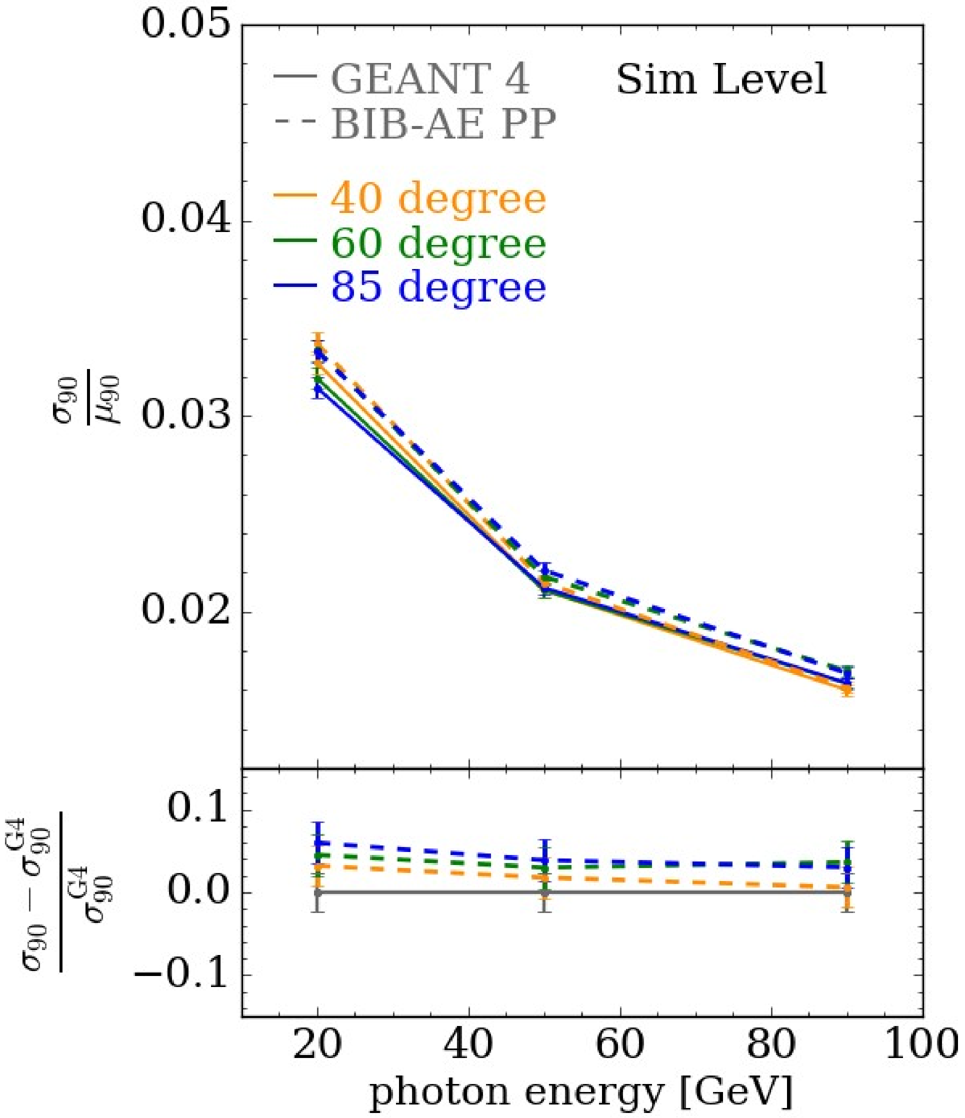


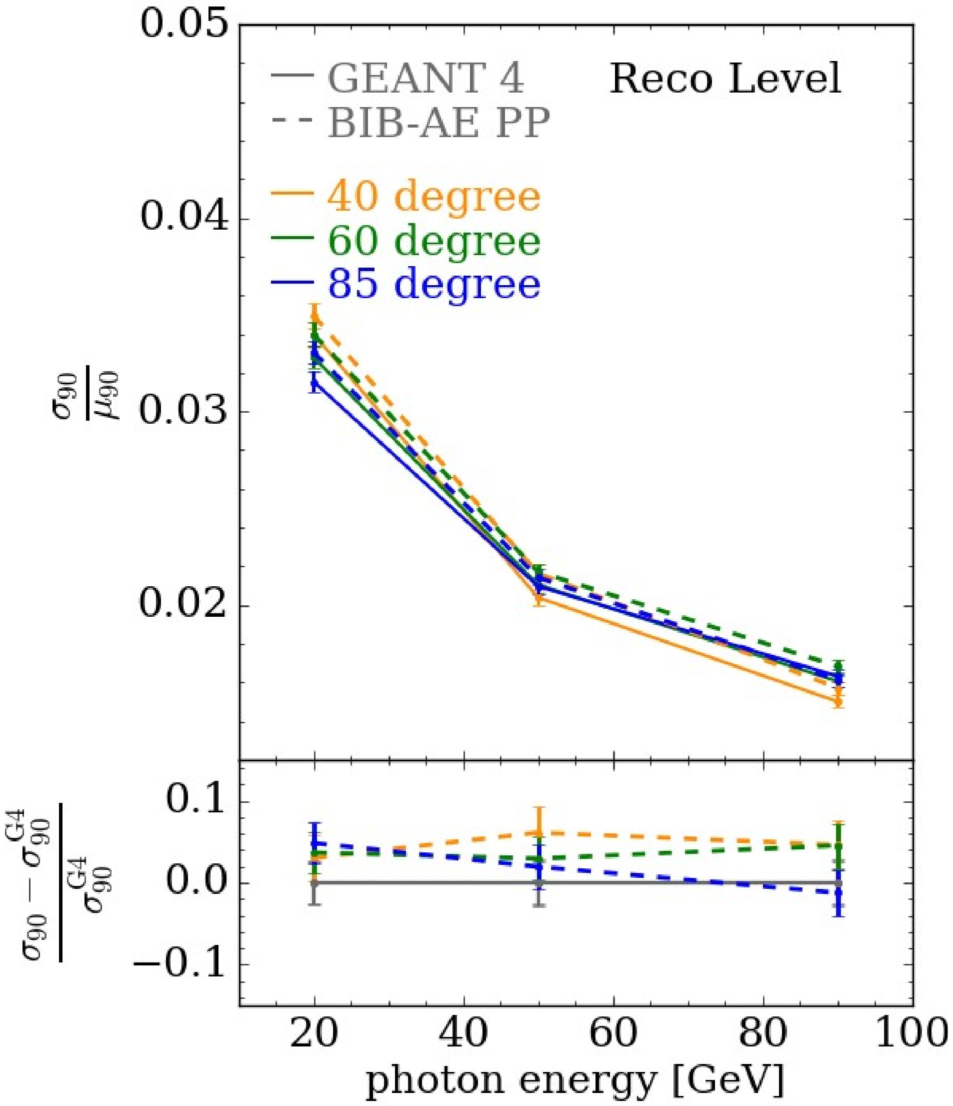




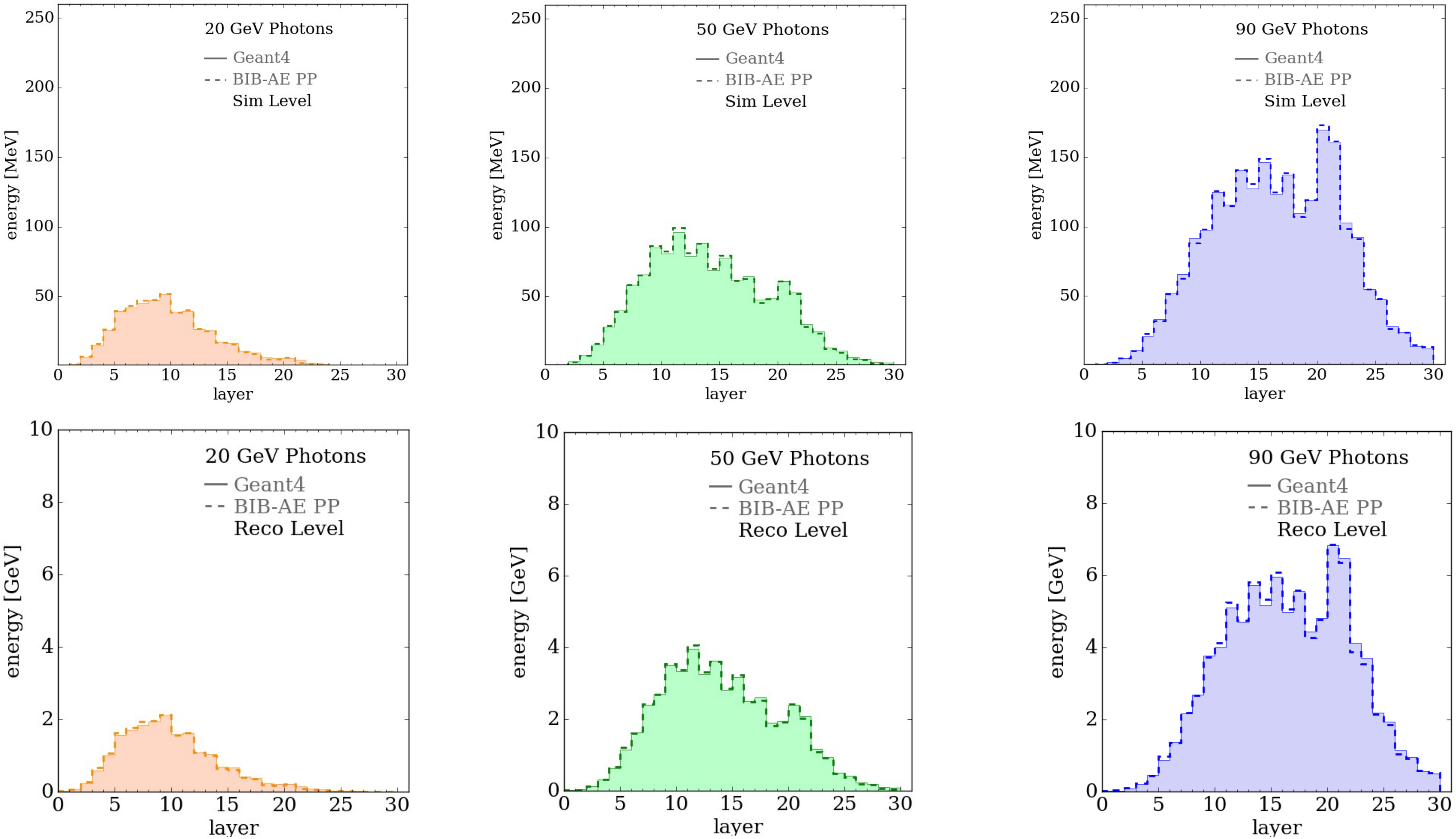
Buhmann et. al: **Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network**, [EPJ Web of Conferences 251, 03003](https://doi.org/) (2021)

Results: Energy resolution Sim vs Rec

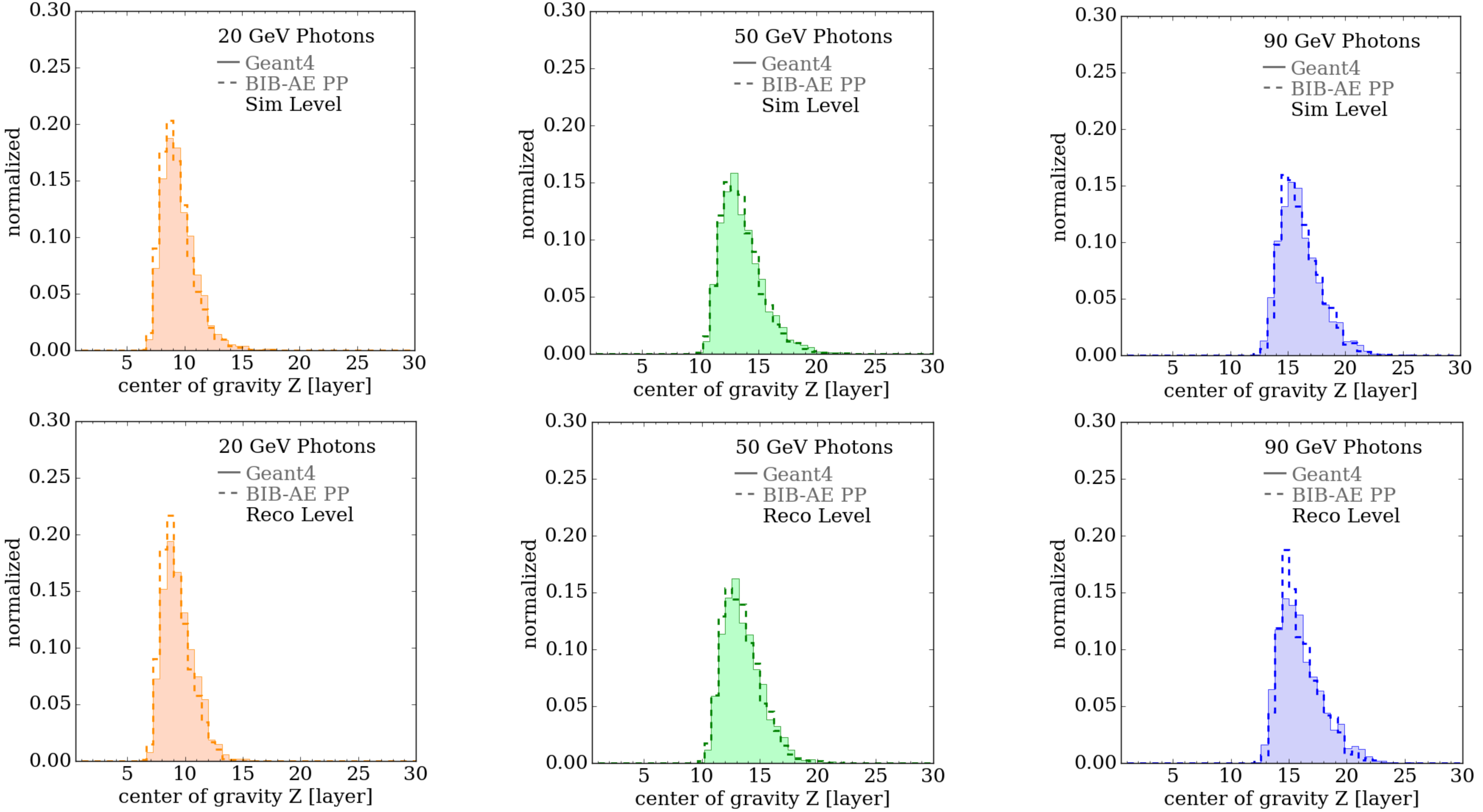




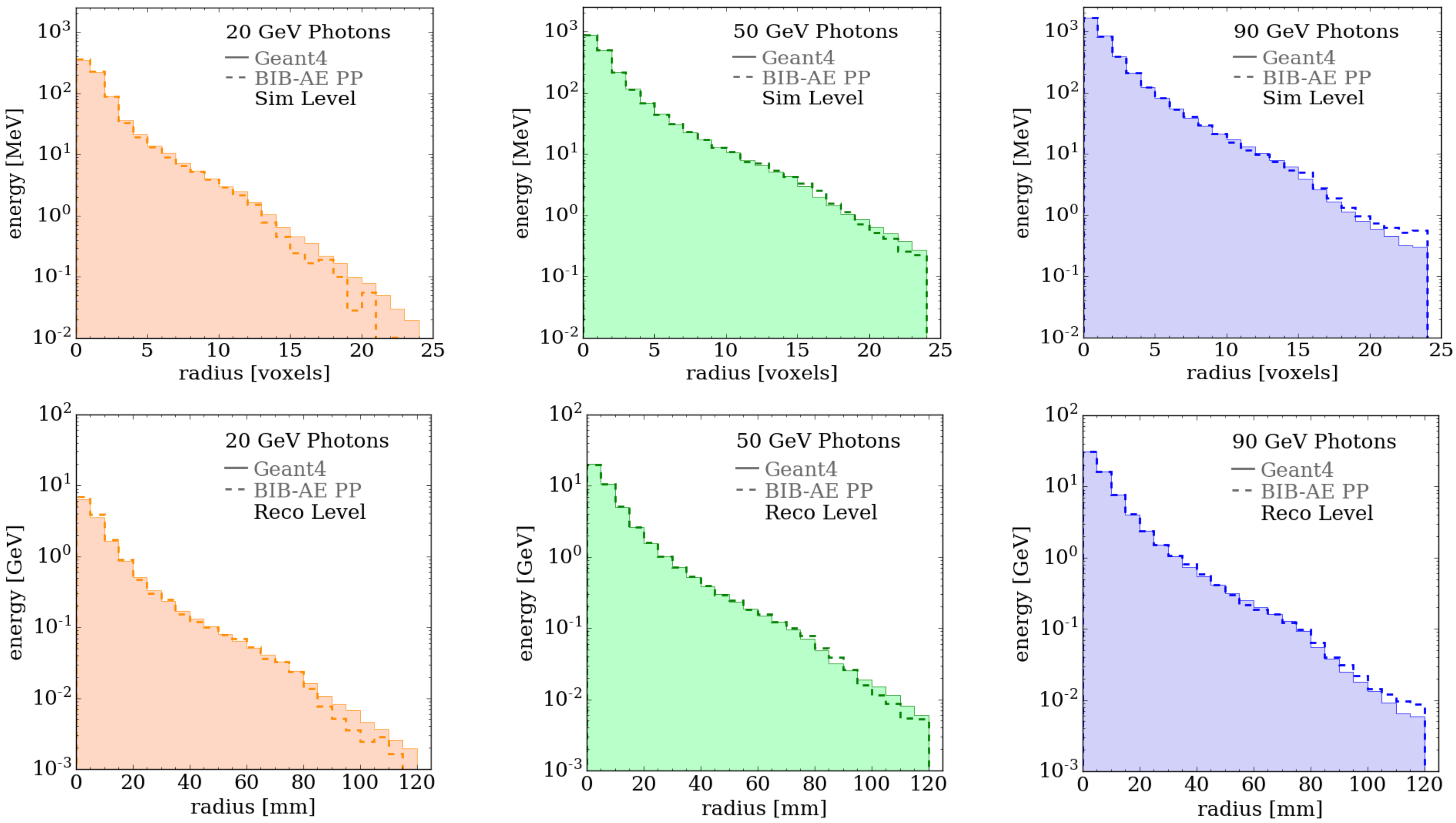
Results: Longitudinal Profile



Results: Center of Gravity



Results: Radial Profile



Results: Number of Hits