New Angles on Fast Calorimeter Shower Simulation

ML4Jets 2023, Hamburg

S. Diefenbacher[3], E. Eren[1], F. Gaede[1], G. Kasieczka[2], A. Korol[1], K. Krüger[1], T. Madlener[1],
P. McKeown[1*], L. Rustige[1]

[1] Deutsches Elektronen Synchrotron, DESY
[2] University of Hamburg, UHH
[3] Lawrence Berkeley National Laboratory, LBNL

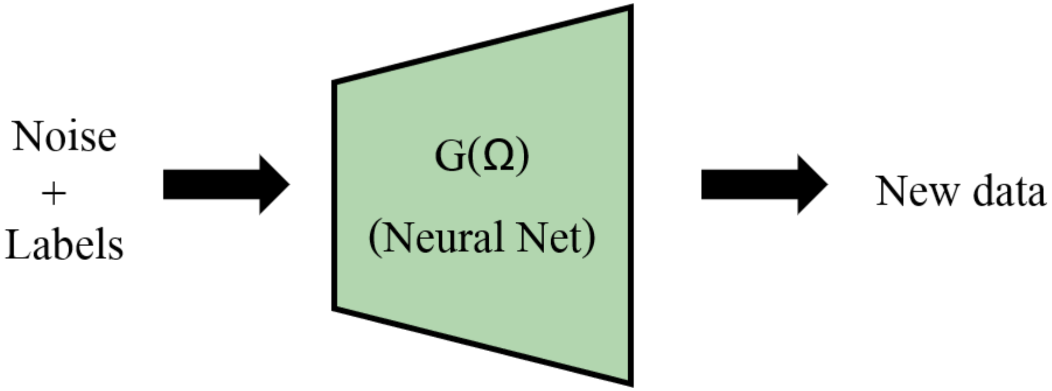
* peter.mckeown@desy.de

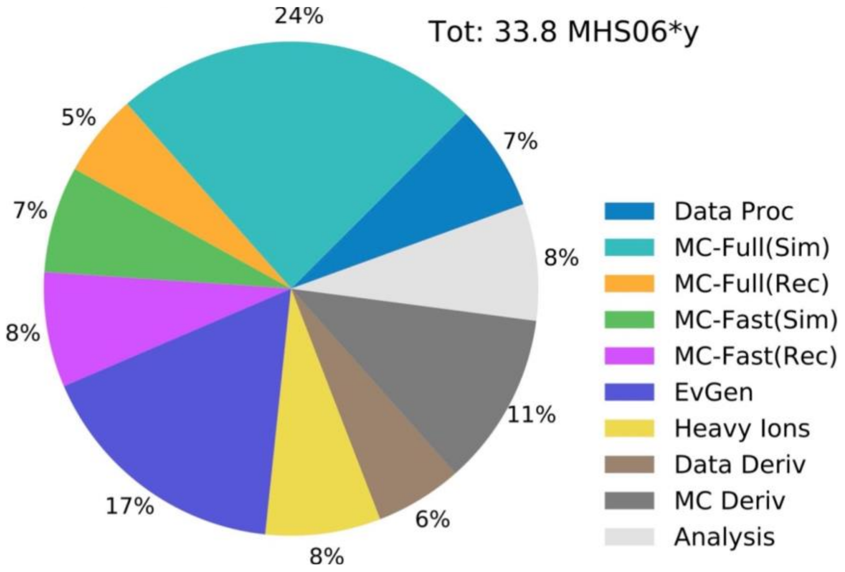
AIDA innova HELMHOLTZ KISS FTX UHH CLUSTER OF EXCELLENCE QUANTUM UNIVERSE DESY.

Introduction

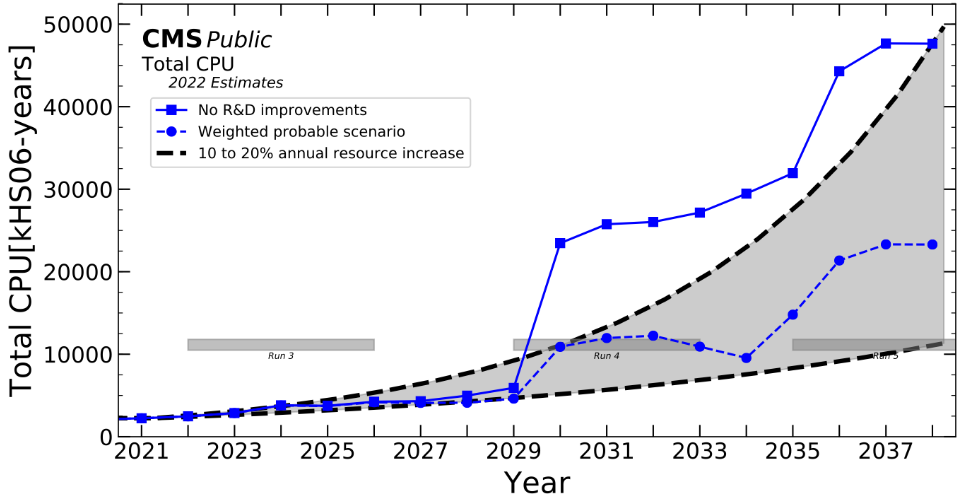
- **Full MC simulation (Geant4)** is computationally **expensive**
  - Calorimeters most intensive part of detector simulation

- **Generative models** potentially offer **high fidelity** simulation with significant **speed up:**
  - More sustainable computing





ATLAS Software and Computing HL-LHC Roadmap, CERN-LHCC-2022-005



CMS Collaboration, Offline and Computing Public Results (2022), https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults

Highly Granular Calorimeters for Future Experiments

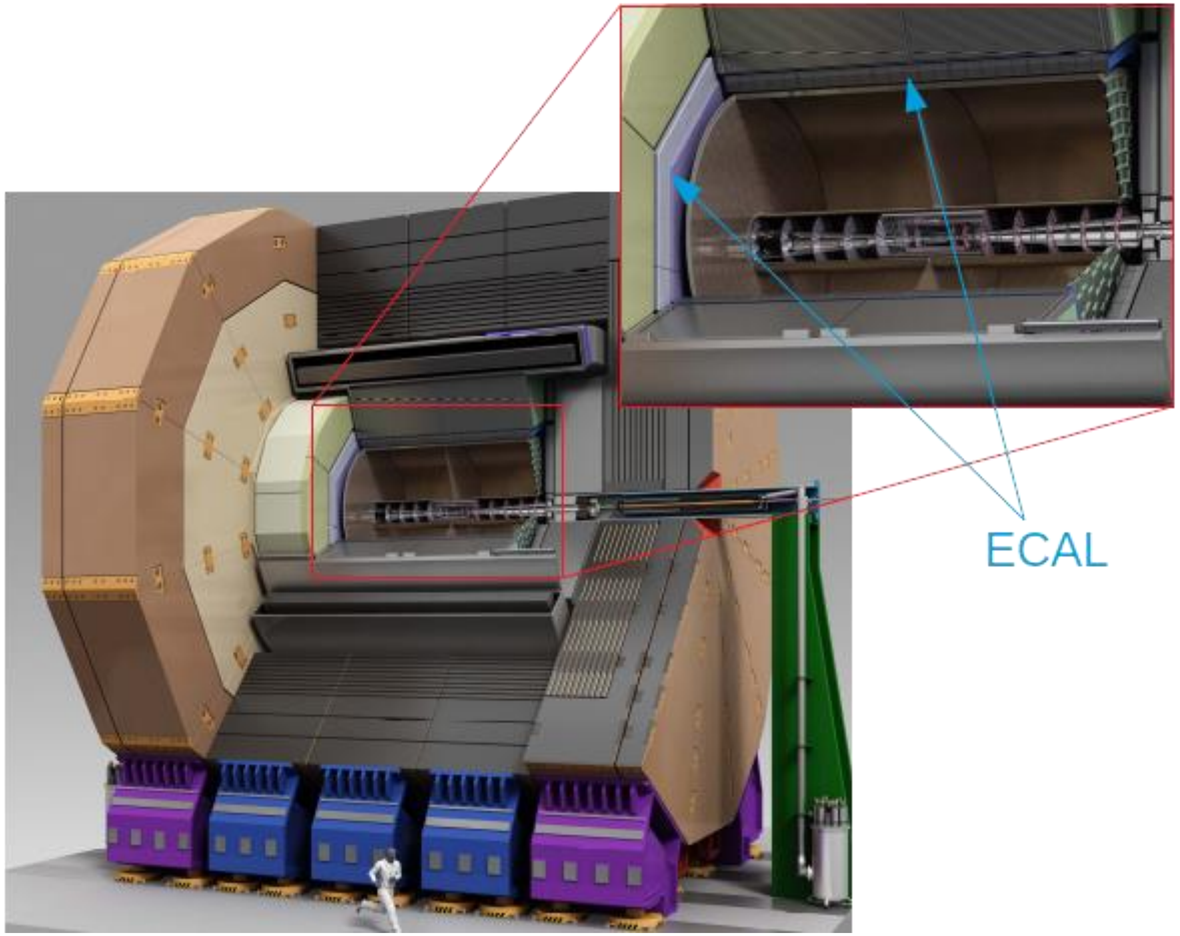
- Widely planned for future experiments: e.g. HL-LHC, e+e- Higgs Factories

- 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

- High granularity calorimeters:
  - **ECAL**: Si-W - 5mm x 5mm
  - **HCAL**: Sci-Fe - 30mm x 30 mm

- High granularity → **Need for high fidelity simulation**



Highly Granular Calorimeters for Future Experiments

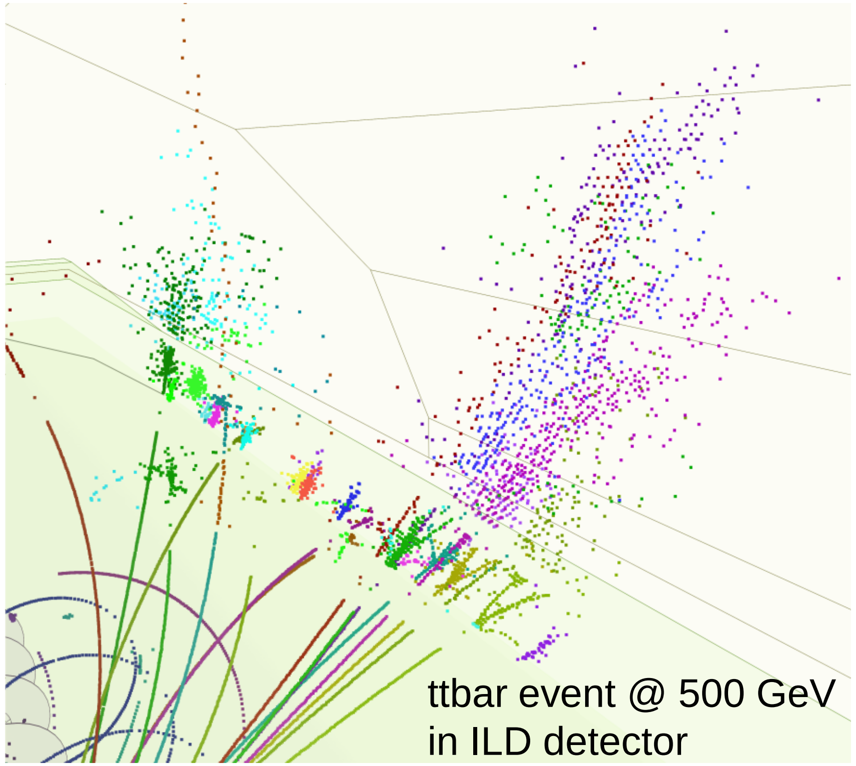
- Widely planned for future experiments: e.g. HL-LHC, e+e- Higgs Factories

- 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

- High granularity calorimeters:
  - **ECAL**: Si-W - 5mm x 5mm - ~ 80 million channels
  - **HCAL**: Sci-Fe - 30mm x 30 mm - ~ 8 million channels

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

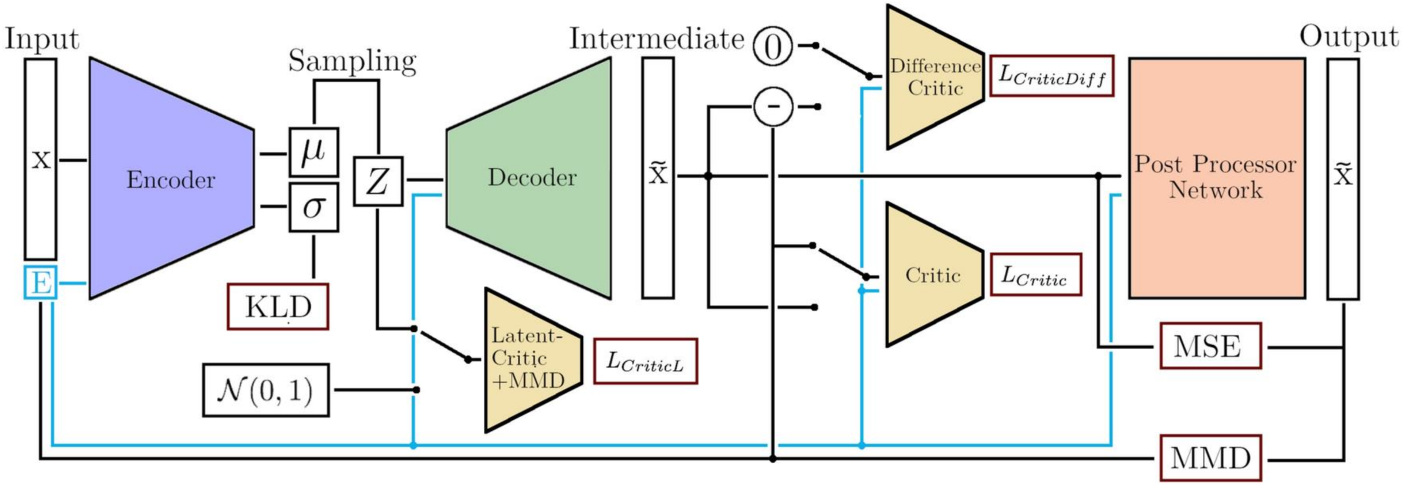


ttbar event @ 500 GeV in ILD detector

- High granularity → **Need for high fidelity simulation**

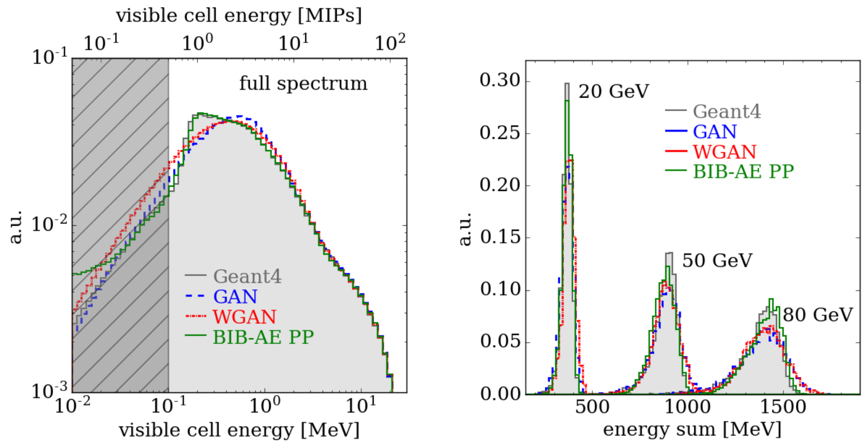
Initial Progress: Photons and Pions

- Achieved **high fidelity** generation of **photon** and **pion** showers with **BIB-AE** architecture (and post processing)
  - 90 deg impact angle, fixed position in calorimeter
  - Fixed regular 3D grid geometry (O(10-100k) voxels)



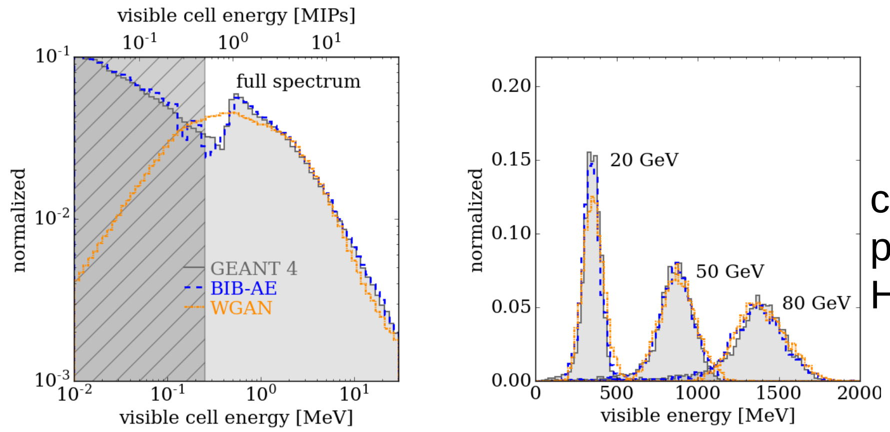
BIB-AE: Bounded Information Bottleneck Auto-Encoder

as well as comparison to GAN and WGAN …



photons
ECAL

Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed, Buhmann et al., arXiv:2005.05334, Comput Softw Big Sci 5, 13 (2021)

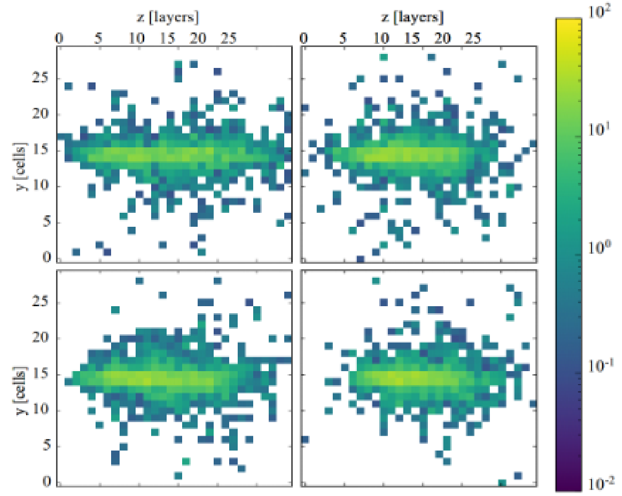


charged
pions
HCAL

Hadrons, Better, Faster, Stronger
Buhmann, P.M. et al, arXiv:2112.09709,
MLST 3 2, 025014 (2022),

Towards An Application In Realistic Detector Simulation

From Photons to Pions

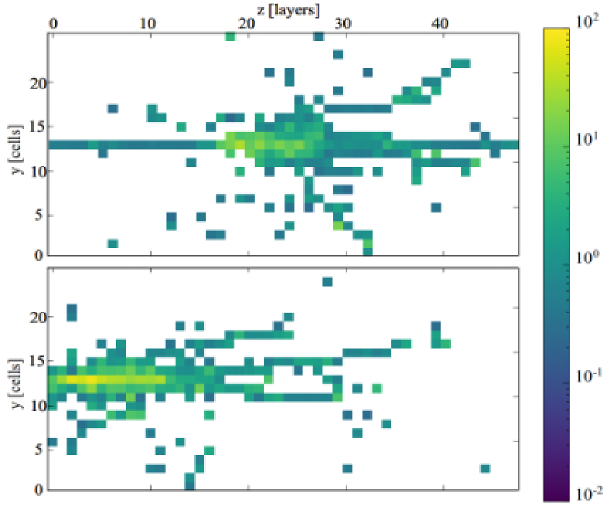


Photon showers

- Predominantly governed by EM interactions
- Compact structure

↓

Relatively easy to generalise



Pion showers

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

↓

Hard to learn

	Energy	Angles	ECAL +HCAL	Reco
Photon showers	✓	+	N/A	+
Pion showers	✓	X	X	−

✓ = Achieved X = Yet to be done

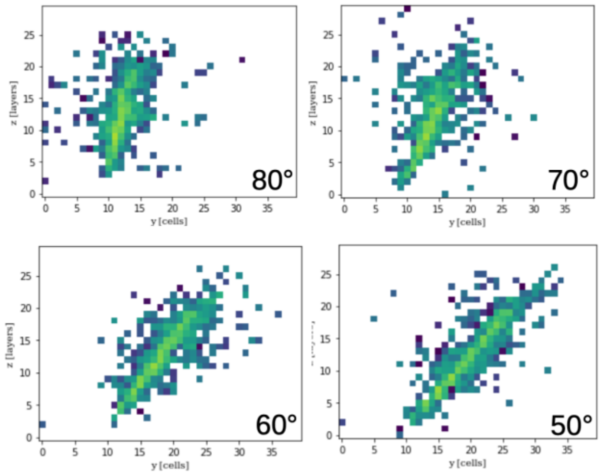
− = Partially Addressed + = Addressed here

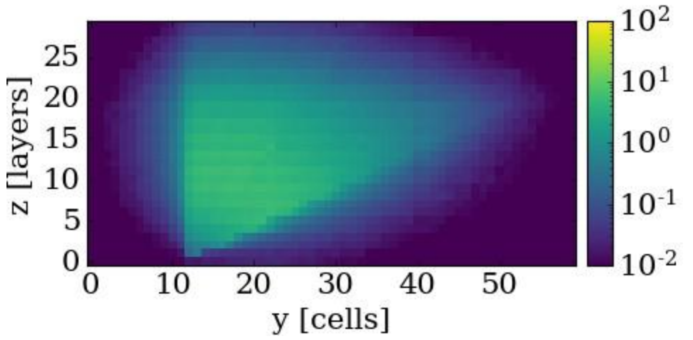
For **realistic application** also need:

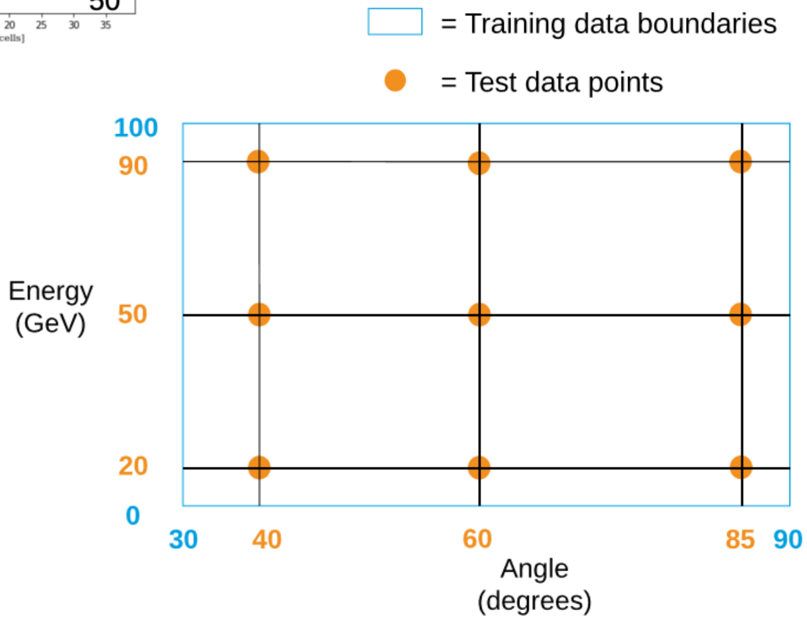
- **Angular conditioning**
- High performance after **reconstruction**
- **Integration** into existing software frameworks
- …

Energy and Angular Conditioning

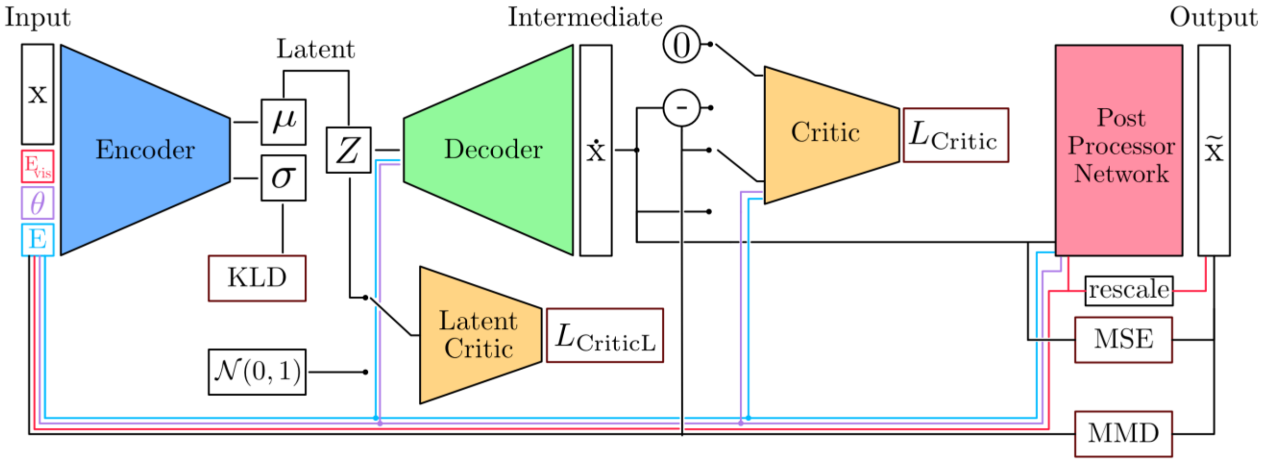
- Photons incident at fixed position
- Extend **BIB-AE** architecture
- Normalising flow for latent space sampling
- **Vary incident energy and polar angle**
  - Large training sample - 500k showers
    - Uniform in [10-100 GeV, 30-90 deg]
  - Test/validation samples at fixed energies and angles







Training and validation samples - 30x60x30 grid



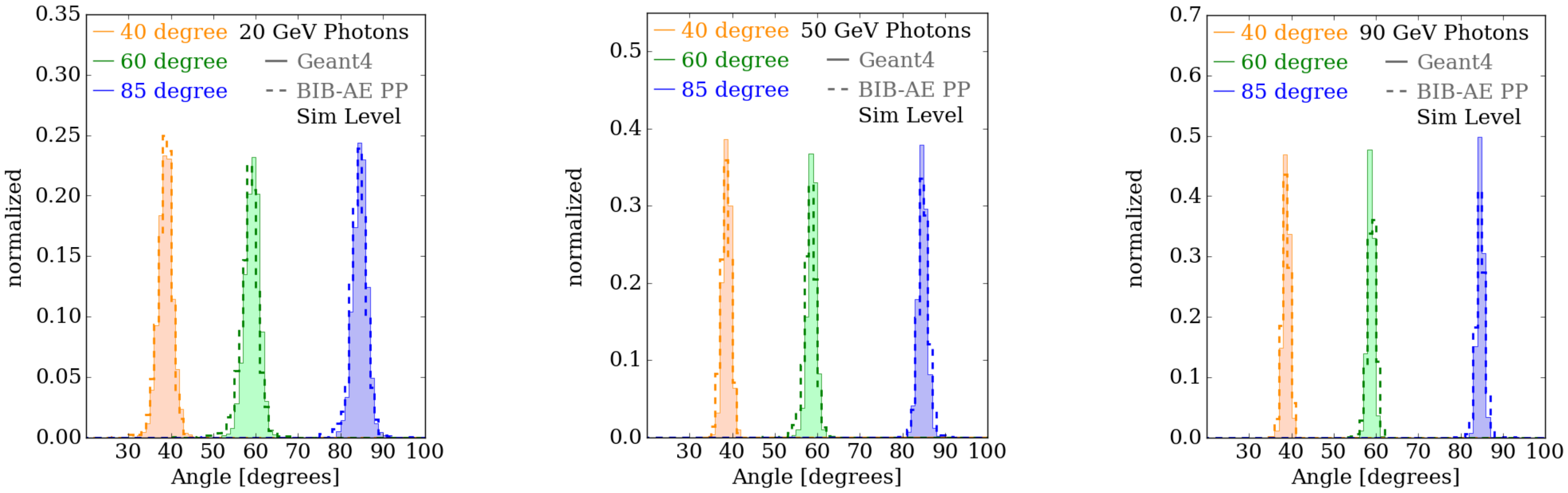
$$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}}$$

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

Angular Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, P.M. et al. 2023 MLST 4 035044
DOI 10.1088/2632-2153/acefa9, arXiv: 2303.18150

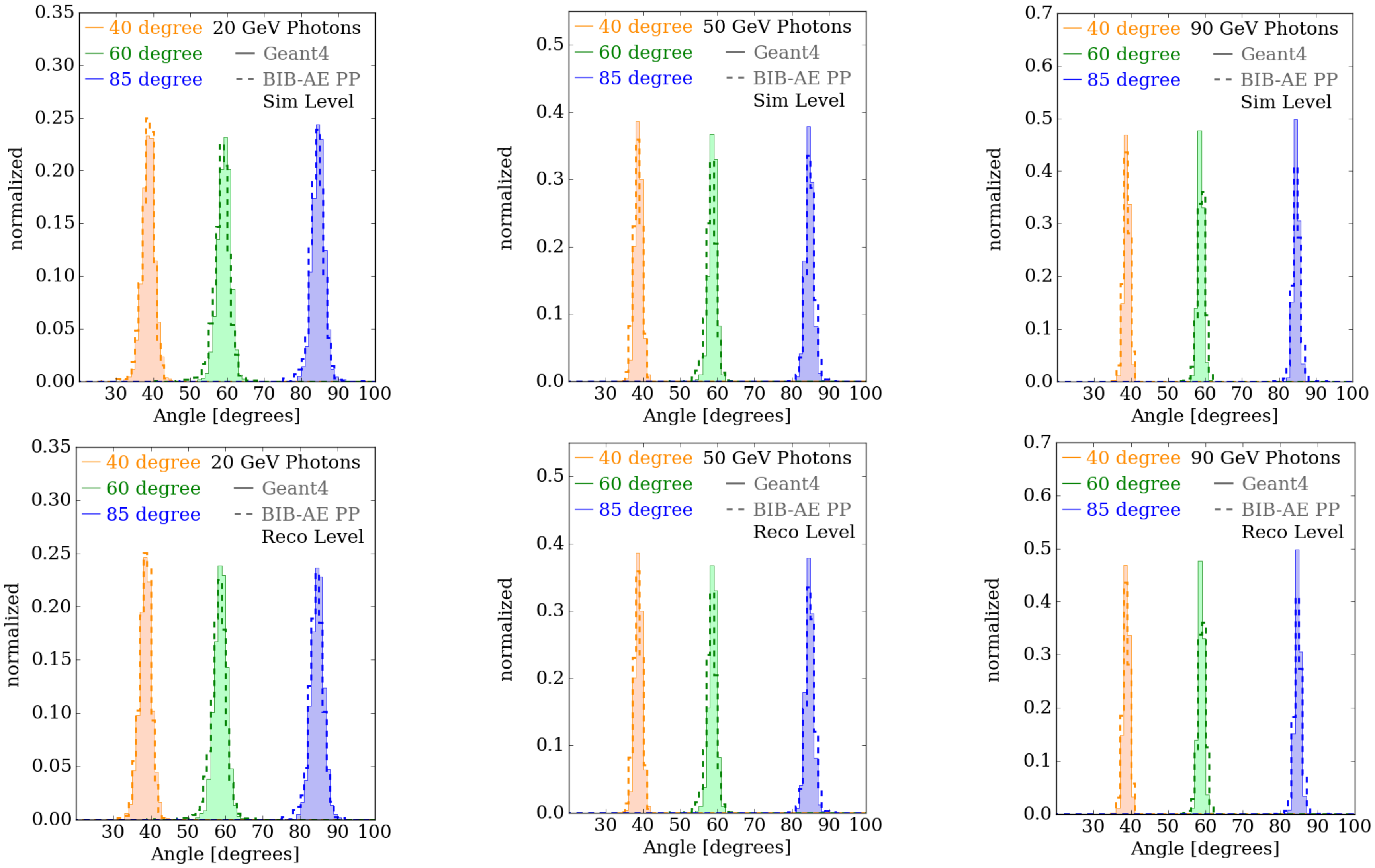
- **Sim** level angle reconstruction



Angular Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, P.M. et al. 2023 MLST 4 035044
DOI 10.1088/2632-2153/acefa9, arXiv: 2303.18150

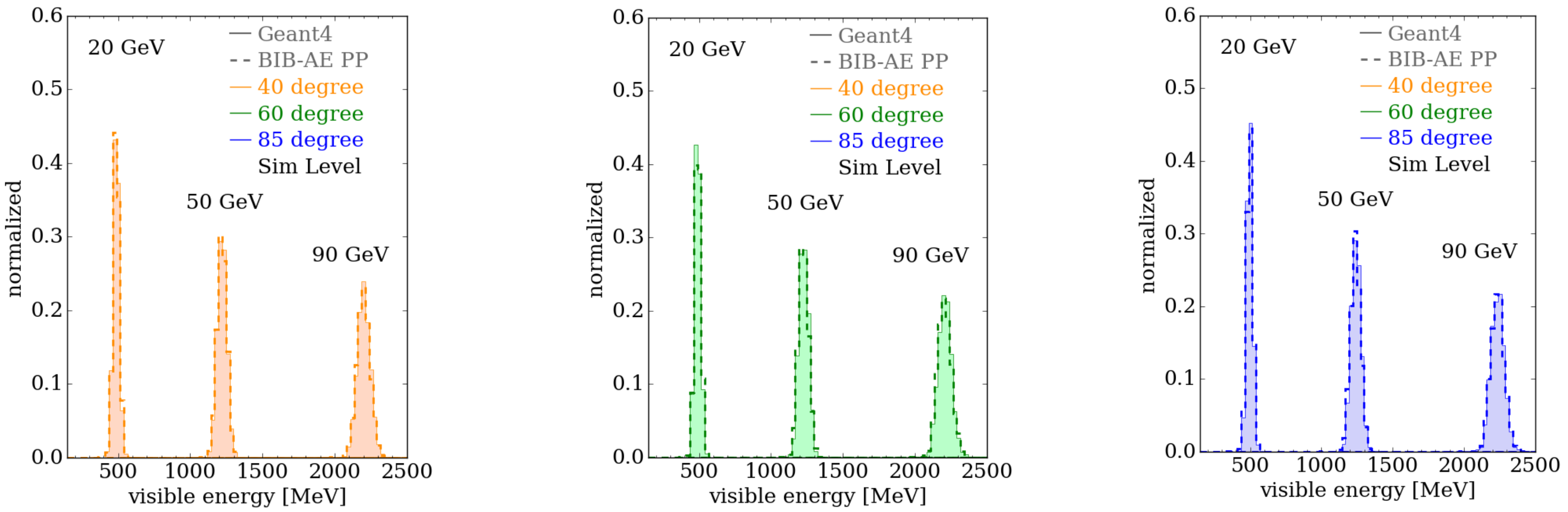
- **Sim** level angle reconstruction
- **Rec** level angle reconstruction
    - After full reconstruction with PandoraPFA



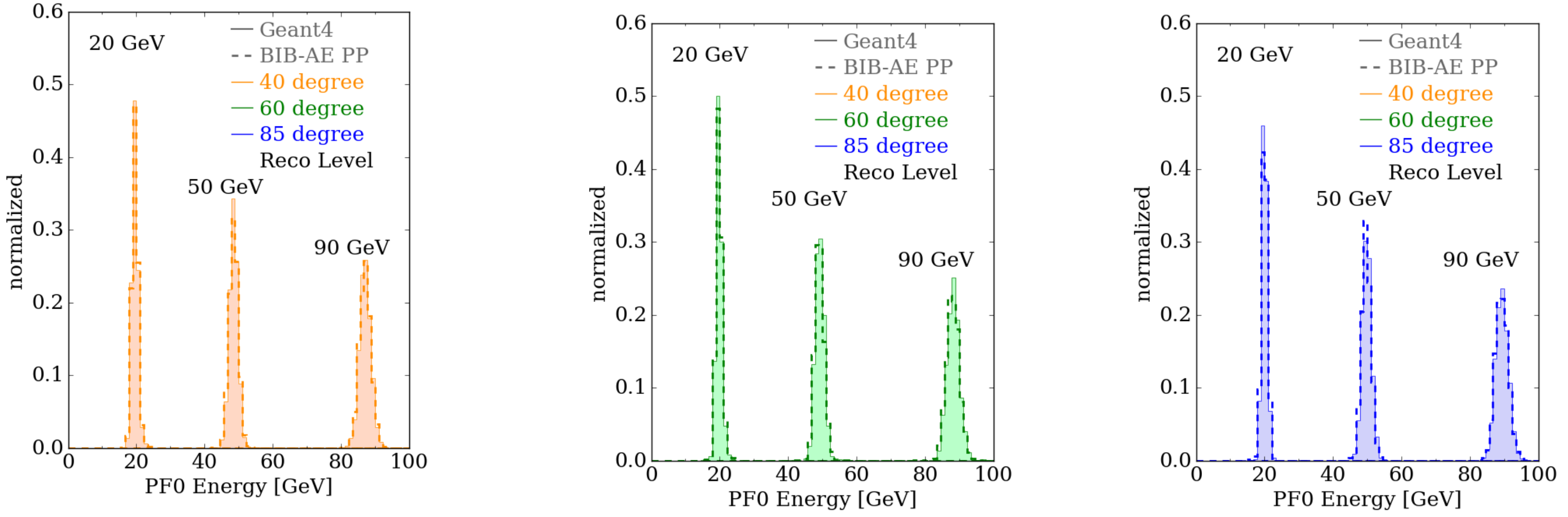
Energy Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, P.M. et al. 2023 MLST 4 035044
DOI 10.1088/2632-2153/acefa9, arXiv: 2303.18150

- **Sim** level **visible** energy

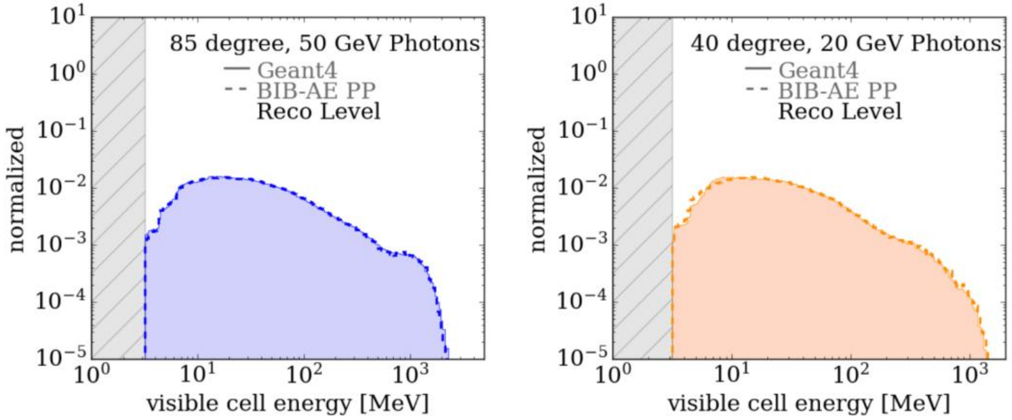


- **Rec** level **calibrated** energy
  - After full PandoraPFA reco

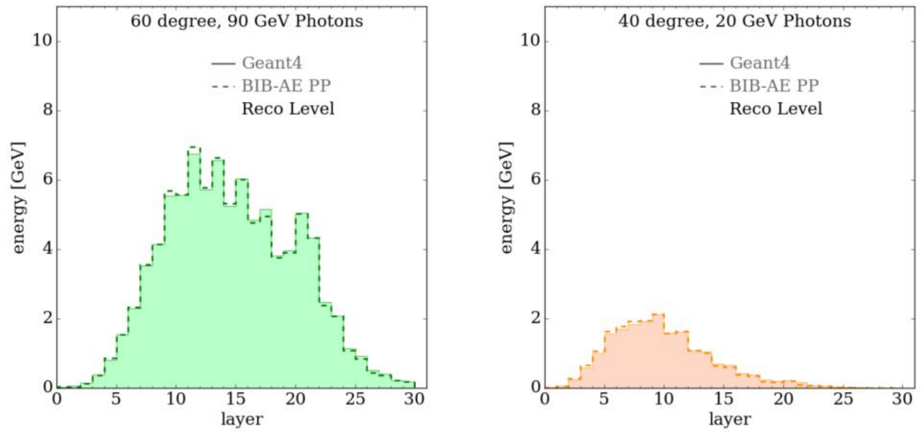


Performance After Reconstruction

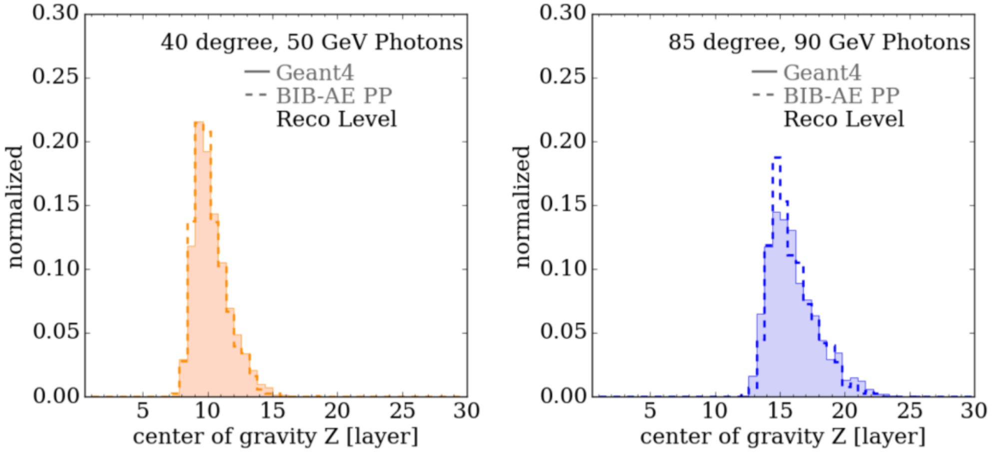
New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, P.M. et al. 2023 MLST 4 035044
DOI 10.1088/2632-2153/acefa9, arXiv: 2303.18150



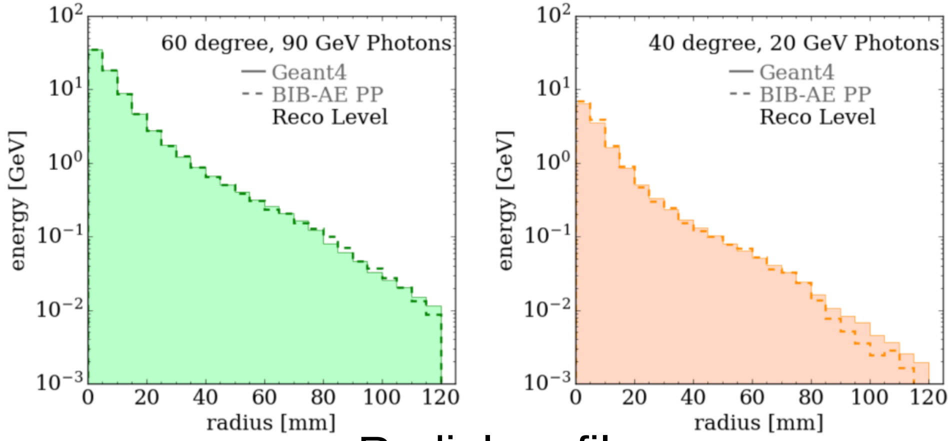
Hit energy spectrum



Longitudinal profile



Centre of gravity in z

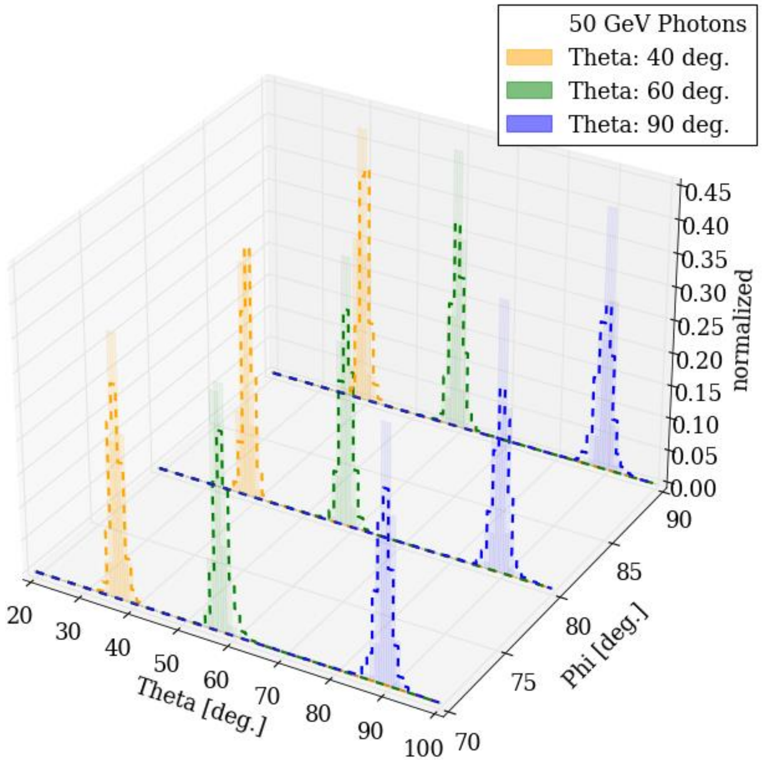


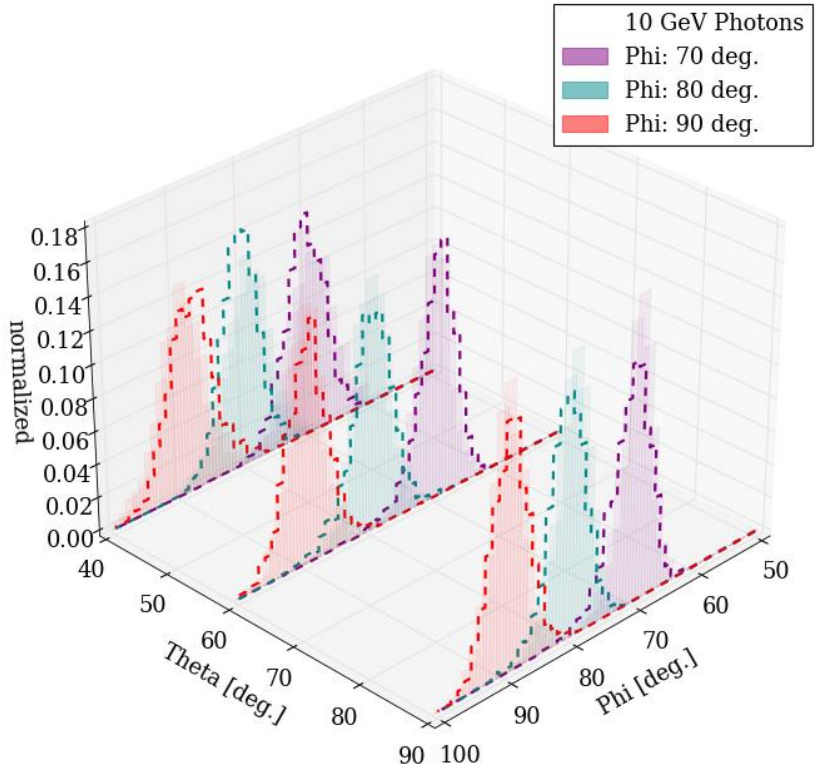
Radial profile

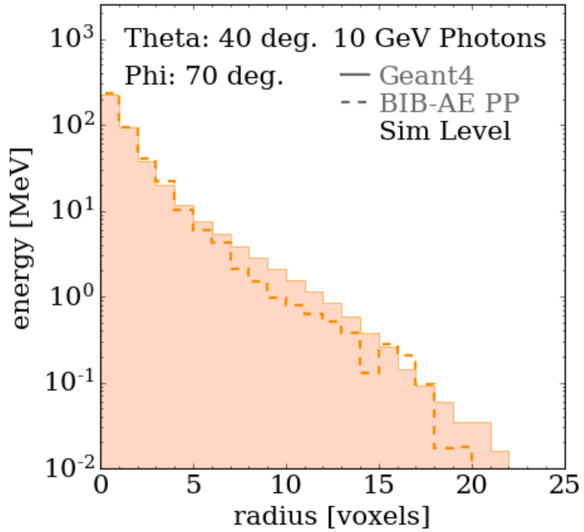
Best (left) and worst (right) test point → **Excellent** physics fidelity

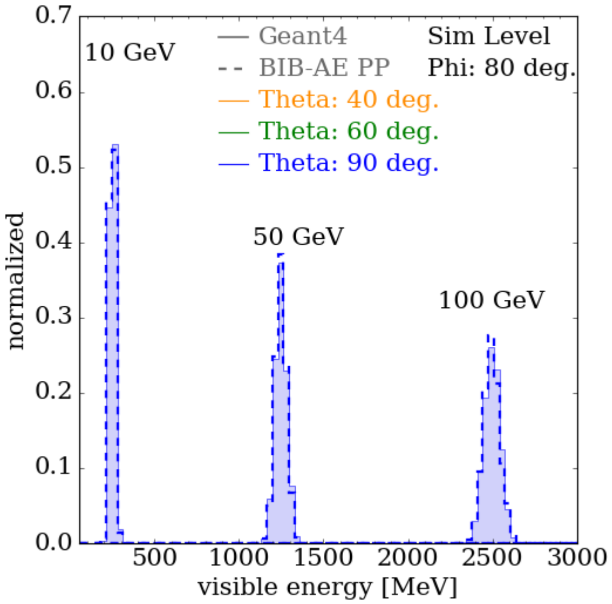
Adding Another Angle

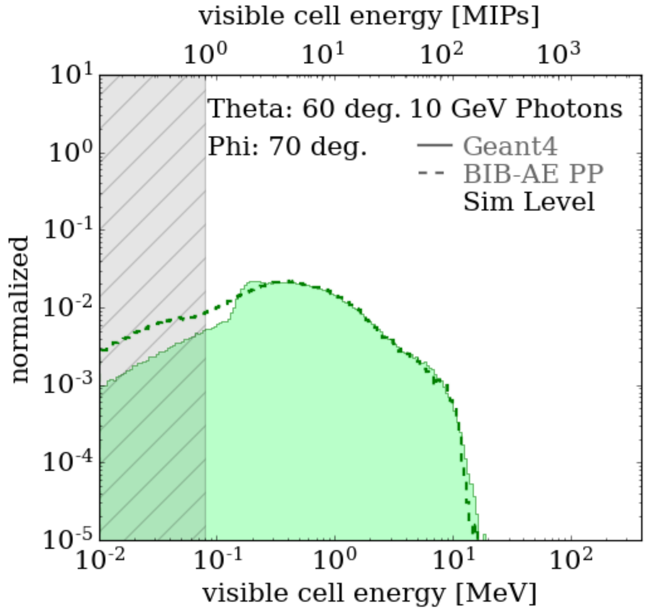
- Need to condition on **energy, theta and phi** for full application
- **Extending phase space** can be challenging



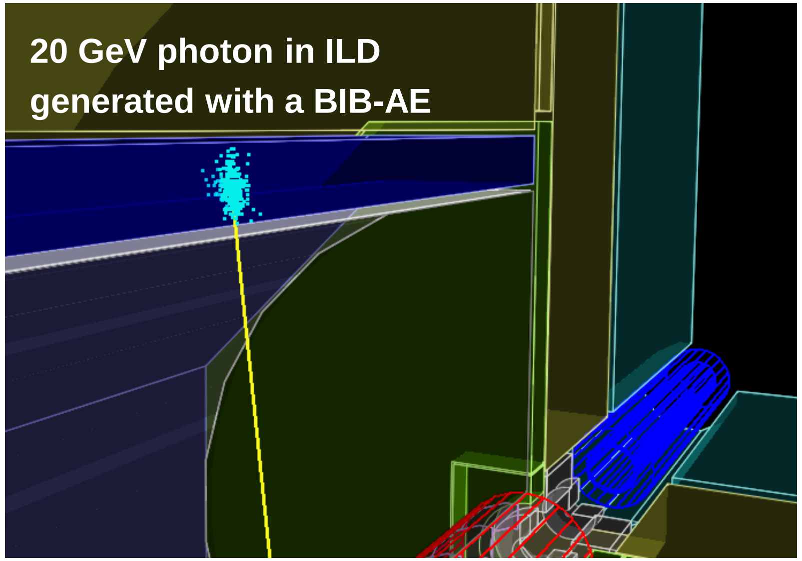








Integration into the Full Simulation Chain



- Prototype library for running ML-based fast sim models: ***DDFastShowerML***
  https://gitlab.desy.de/ilcsoft/ddfastshowerml
  - Use fast sim hooks in DDG4/Geant4
  - Use realistic, detailed detector models
  - Currently only supports CPU
  - Development ongoing

- Aim to have an easy to use library which can be adapted for all types of ML architectures in DD4hep

- **Essential** step to be able to study performance of model with **full physics benchmarks**

Trigger

- Fast Sim trigger
  - e.g. particle type, energy, geometry

Model

- Model-specific implementation of ML architecture
  - e.g. BIB-AE, Flow, Diffusion model

Inference

- Concrete inference in C++
  - ONNX, LibTorch etc…

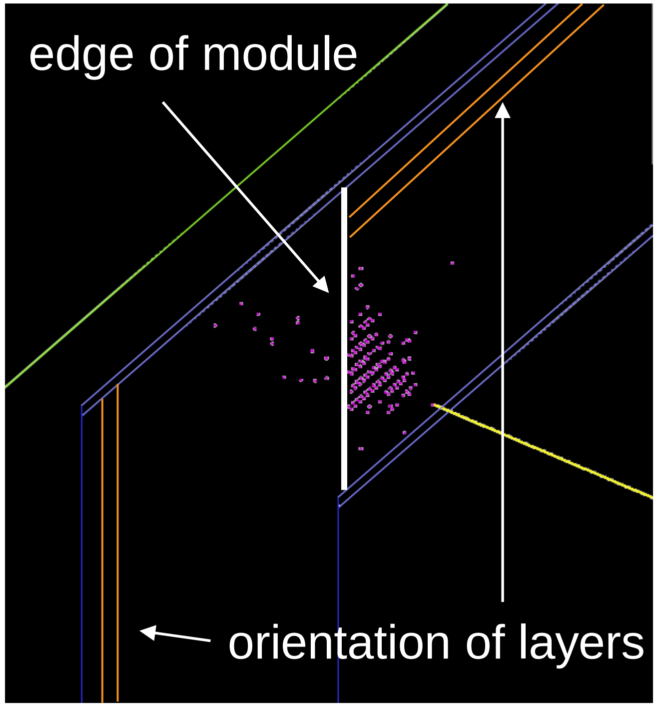
Geometry

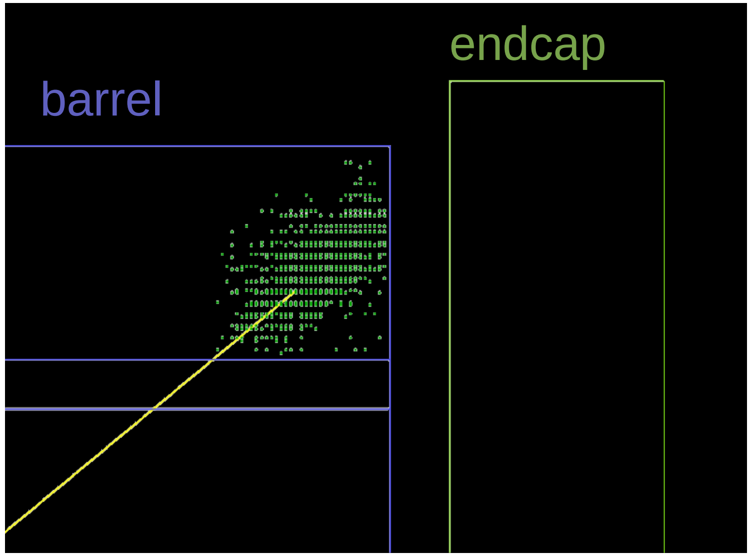
- Concrete placement in detector geometry
  - Endcap, barrel etc…

BIB-AE Integration Into Realistic Geometry

- **BIB-AE** model with full conditioning now **integrated** into ILD detector simulation chain

- **Exclude regions** of detector where model cannot be applied to geometry
  - **Corners** of octagonal barrel
  - **Transition** between barrel and endcap



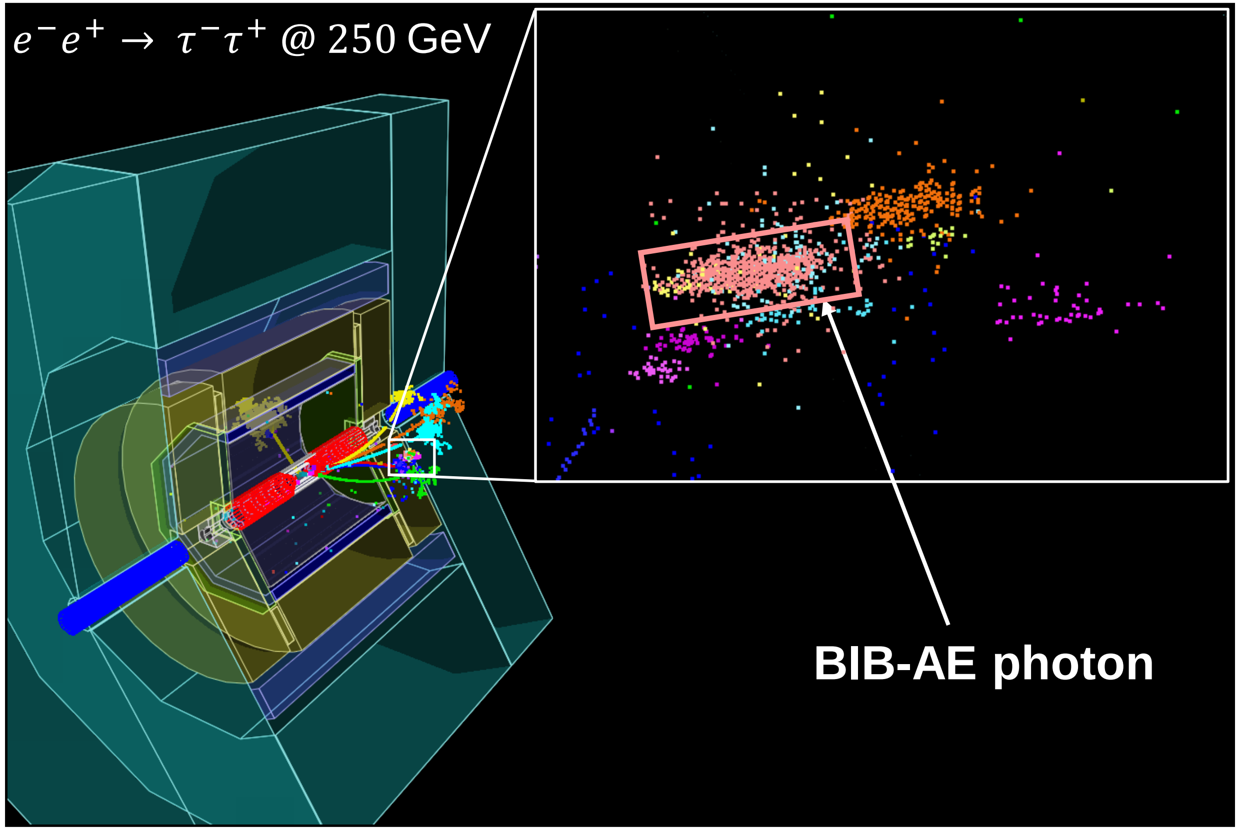


BIB-AE Integration Into Realistic Geometry

- **BIB-AE** model with full conditioning now **integrated** into ILD detector simulation chain

- **Exclude regions** of detector where model cannot be applied to geometry
  - **Corners** of octagonal barrel
  - **Transition** between barrel and endcap

- Now possible to run ML model in **full physics simulation**
  - e.g. physics benchmark on π^0 photons from tau pairs



Conclusion

Achieved

- **Energy and angular** conditioning for EM showers with high physics fidelity
  - **Strong performance** after **reconstruction** with PandoraPFA
- **Additional angle** added in conditioning
- An initial implementation of a **prototype library** for interfacing with the full simulation chain

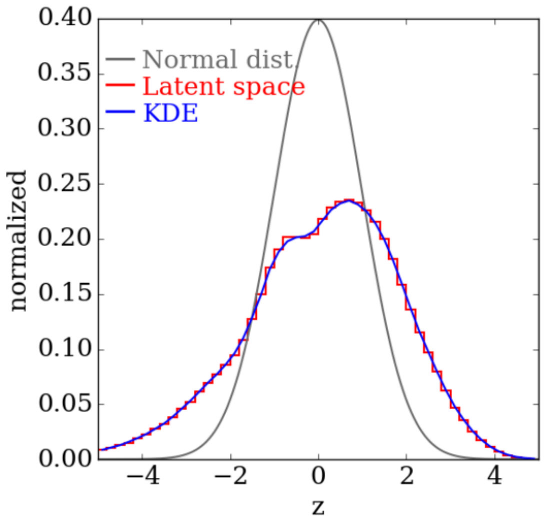
Next Steps

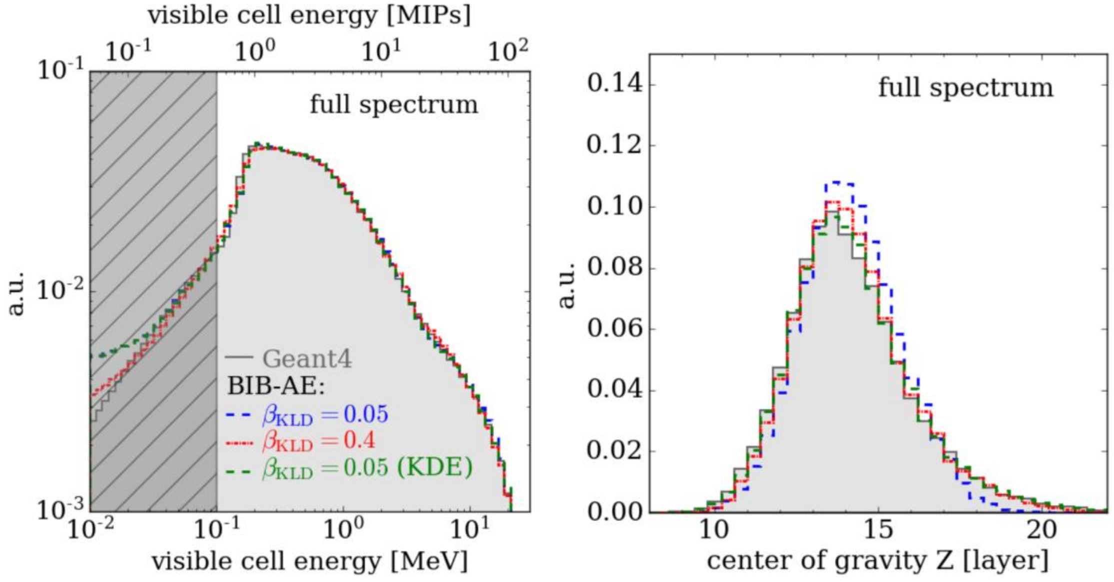
- Study full **physics benchmarks**
- **Extend** functionality of library (batching, GPU support etc.)

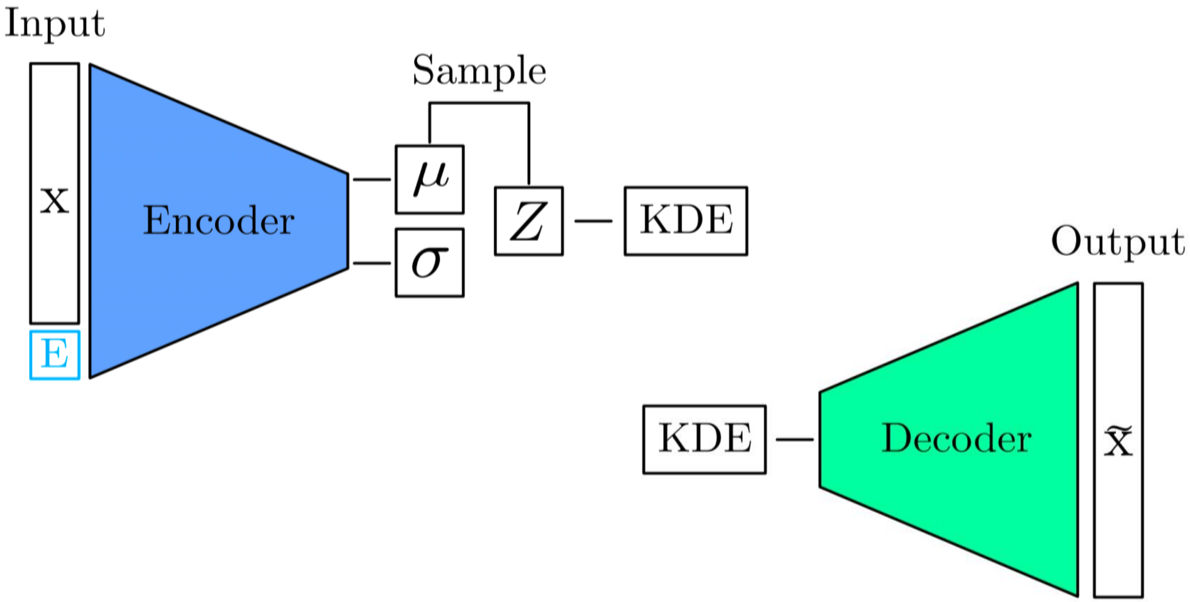
Backup

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)







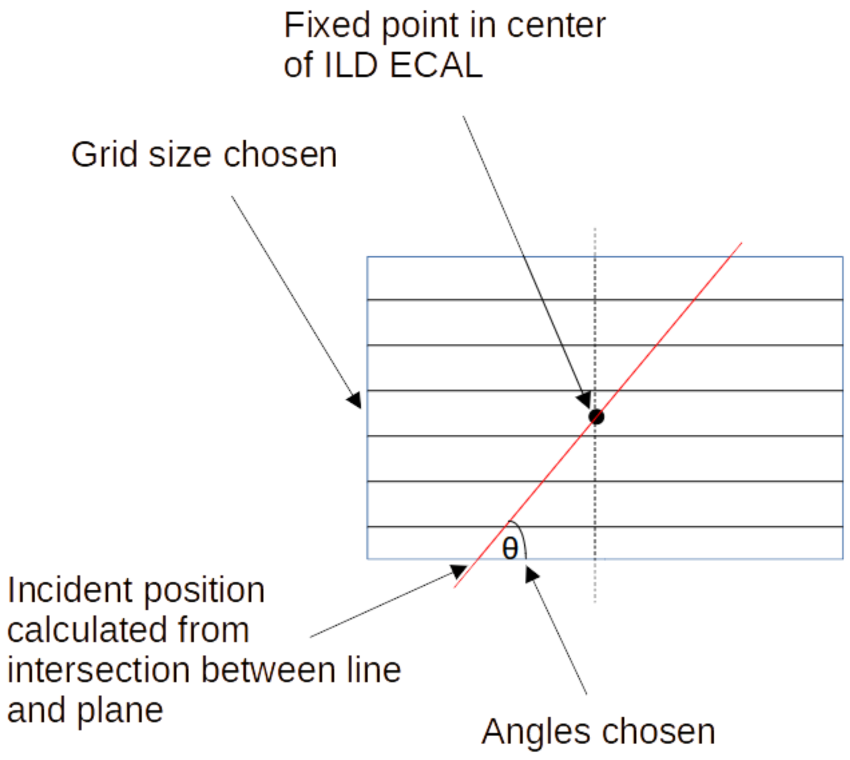
Buhmann et. al: **Decoding Photons: Physics in the Latent Space of a BIB-AE Generative Network**, EPJ Web of Conferences 251, 03003 (2021)

Two Angle Training Data

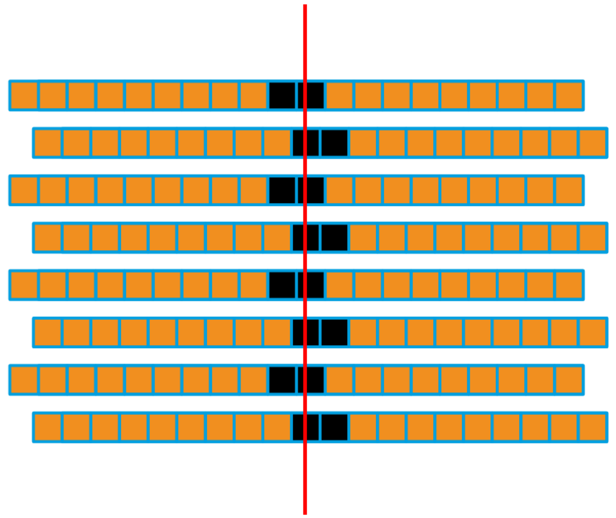
- Create ILD ECAL with **regular** structure for **training**
  - Exactly the **same layer wise** material **composition**
  - **Purely sensitive** material in active layers (remove dead material)

- Vary angles to **minimise box size**, but retain information about **incident position**

- During **simulation with realistic detector geometry**, hits in dead material are dropped by Geant4

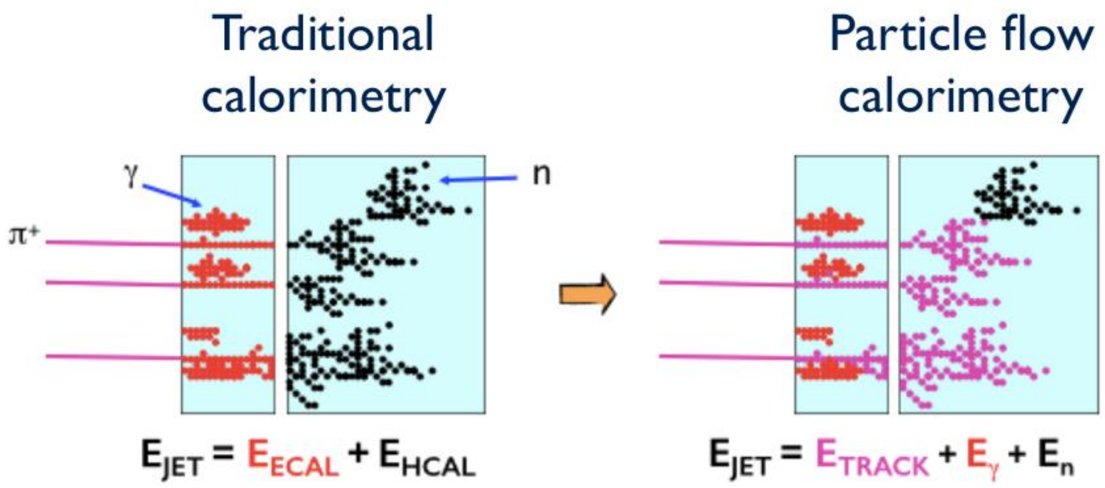


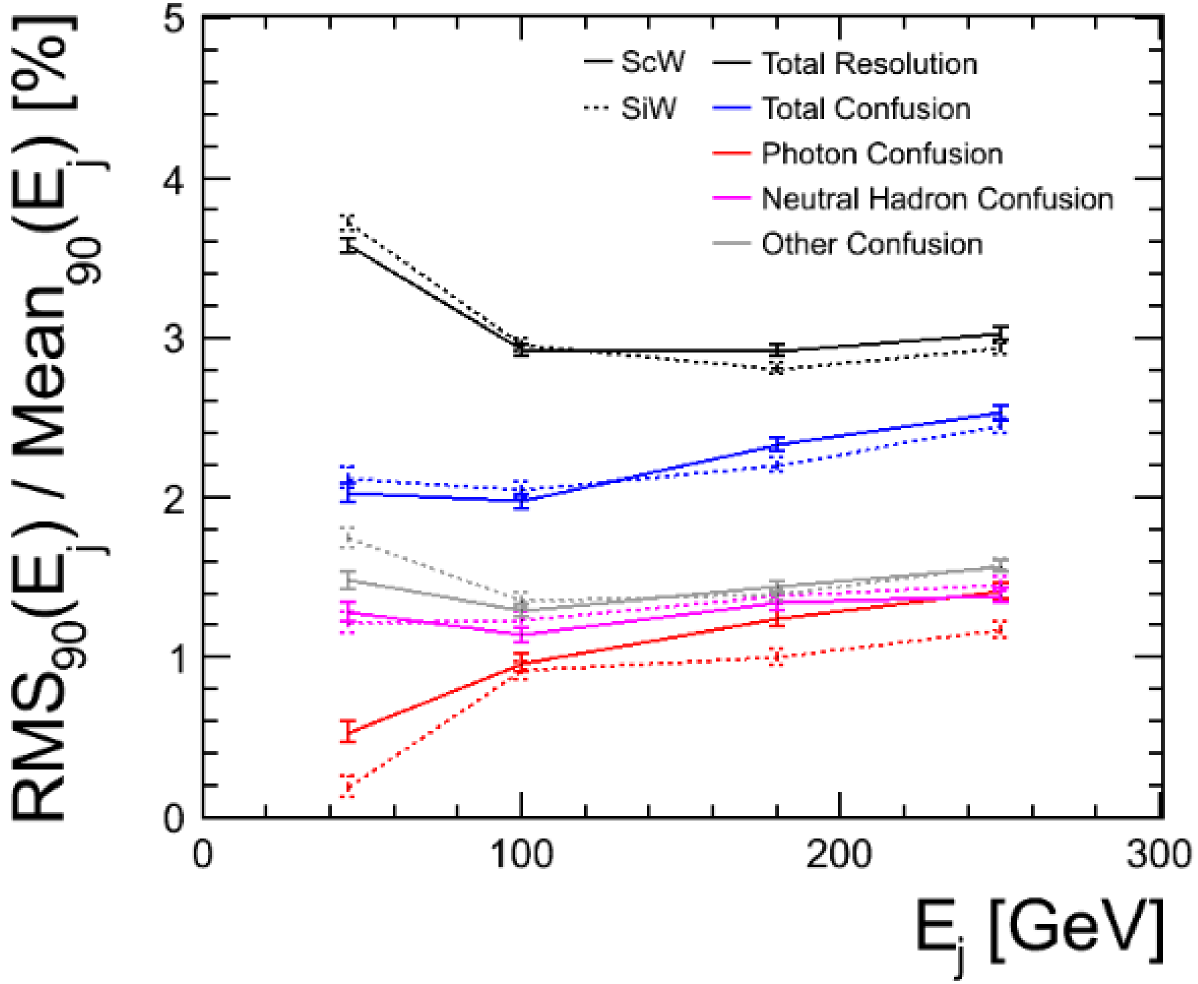
Staggering present in real ILD ECAL due to irregular structure



Particle Flow at a Future Lepton Collider

Particle Flow Calorimetry and the PandoraPFA Algorithm, Thomson, arXiv:0907.3577, Nucl.Instrum.Meth.A611:25-40,2009





Slide credits:
https://indico.cern.ch/event/342026/contributions/799531/attachments/672094/923645/AIDA_Pandora_10.12.14.pdf

Timing Of Generative ML Methods

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

BIB-AE/WGAN, pion showers 10-100 GeV uniform

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

BIB-AE, photon showers 10-100 GeV - 30-90 deg uniform