Generative Models for Fast Simulation of Showers in Highly Granular Calorimeters

ACCLAIM Meeting 2023, Jena

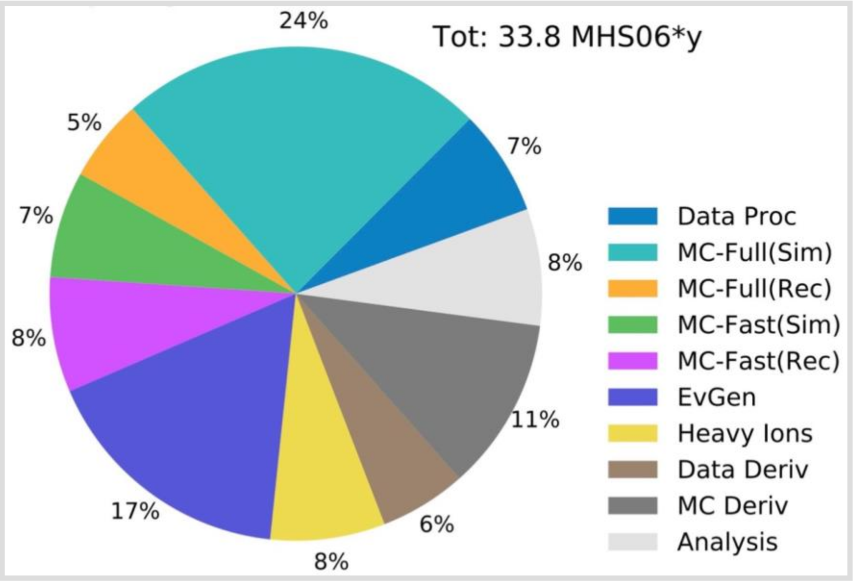
E. Buhmann[2], S. Diefenbacher[2,3], E. Eren[1], F. Gaede[1], G. Kasieczka[2], W. Korcari[2], A. Korol[1], K. Krüger[1], **P. McKeown[1*]**, L. Rustige[1]

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

* peter.mckeown@desy.de

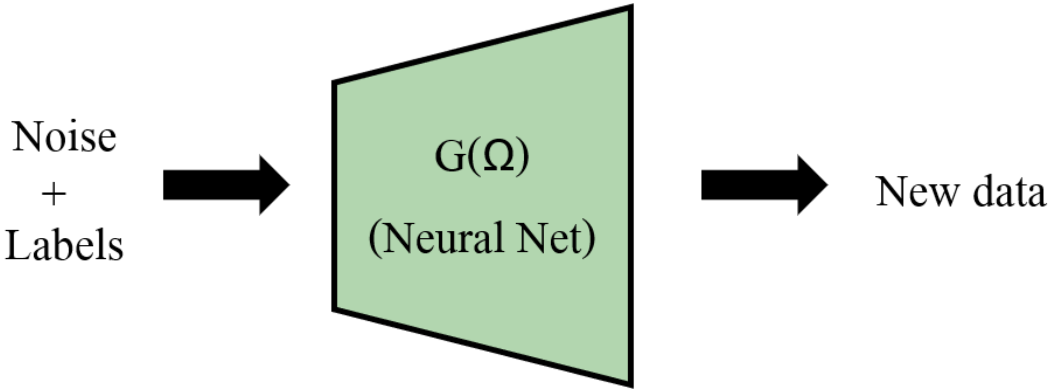
Introduction

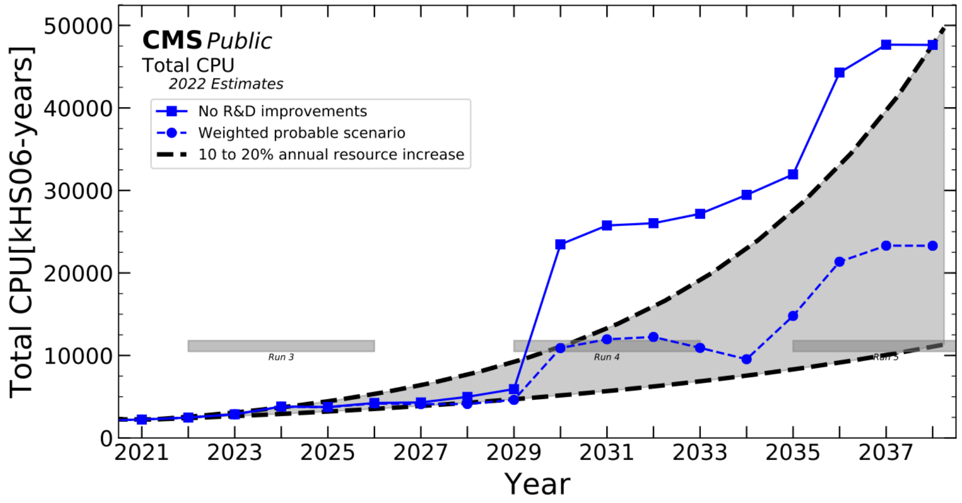
- **Full MC simulation (Geant4)** is computationally **expensive**
  - Calorimeters most intensive part of detector simulation



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

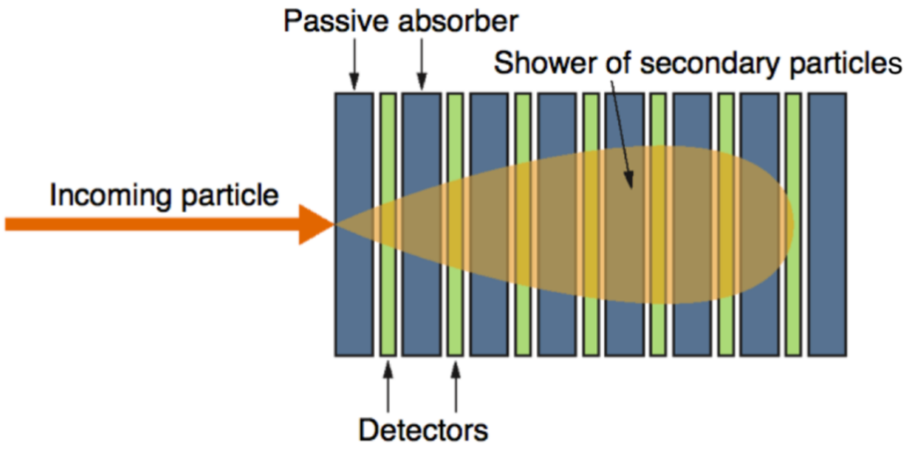
- **Generative models** potentially offer **high fidelity** simulation with significant **speed up:**
  - More sustainable computing





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

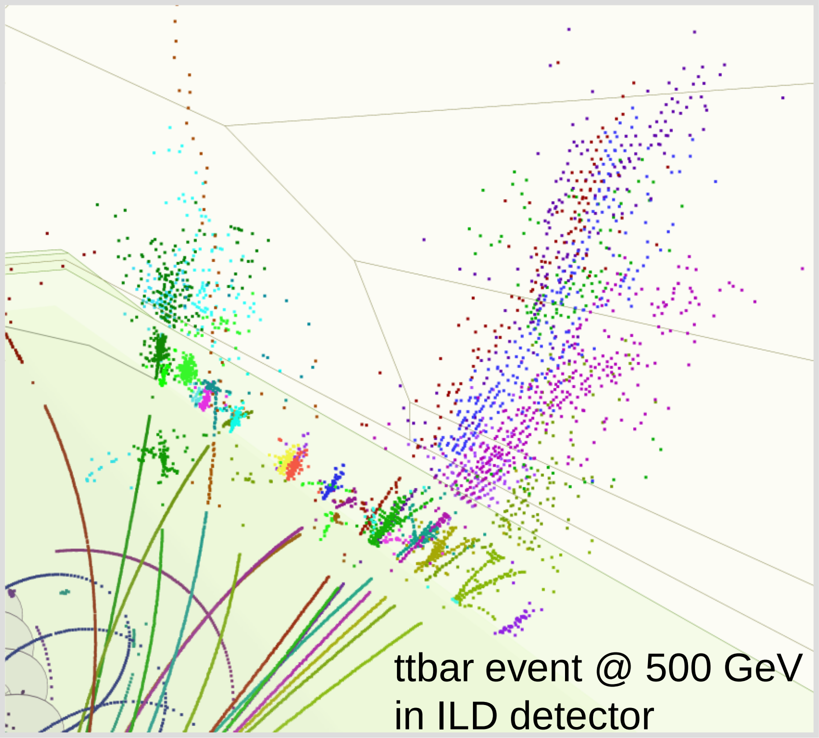
High Granularity Calorimeters



- Destructively measure particle’s energy
  - Produce shower of secondary particles until totally absorbed
- Sampling calorimeters- measure fraction of energy
- Future calorimeters feature ever increasing granularity (e.g. @ **HL-LHC**, future **Higgs Factories**)
- DESY-UHH group investigating generative ML architectures in context of **ILD** detector:
  - **ECAL**: Si-W - 5mm x 5mm
  - **HCAL**: Sci-Fe - 30mm x 30 mm

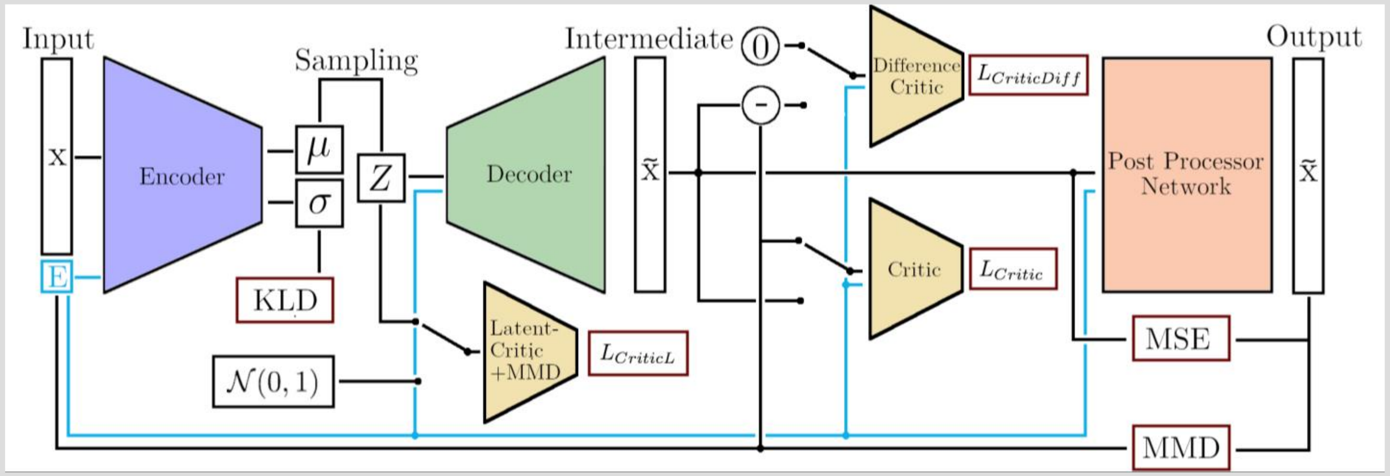
c.f. a few cm^2 for ATLAS/CMS ECAL (before High Lumi)

- High granularity ⟶ High fidelity



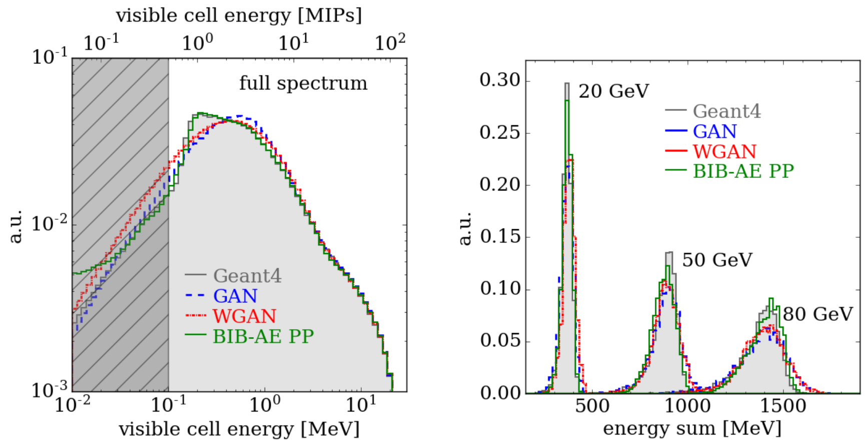
Previous Work

- 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)



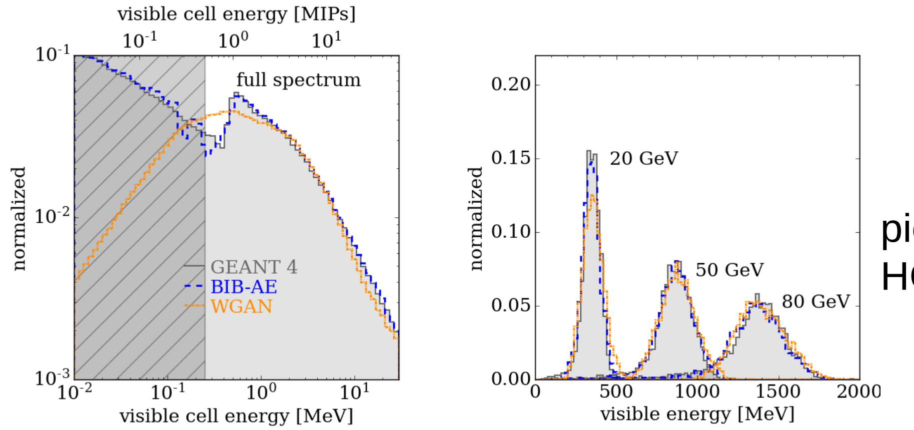
BIB-AE: Bounded Information Bottleneck Auto-Encoder

also comparison to GAN and WGAN …



photons
ECAL

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

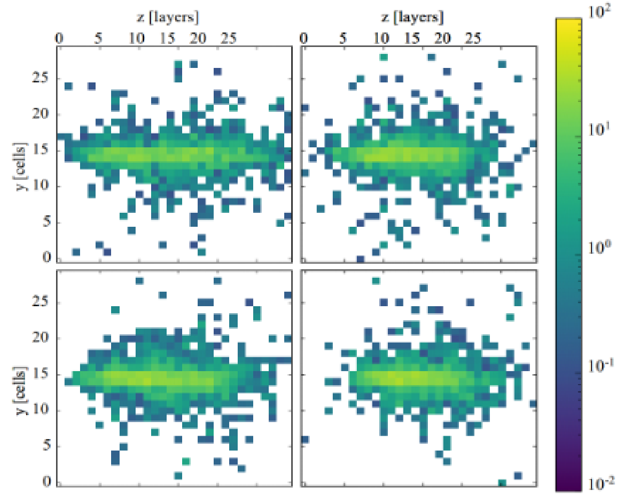


pions
HCAL

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

Towards An Application In Realistic Detector Simulation

From Photons to Pions

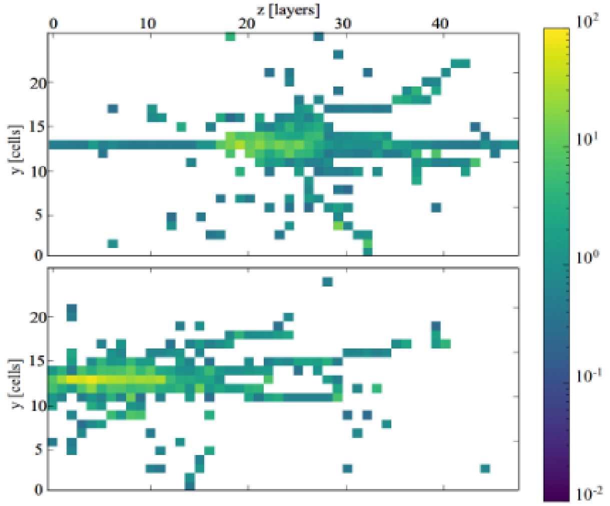


Photon showers

- 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

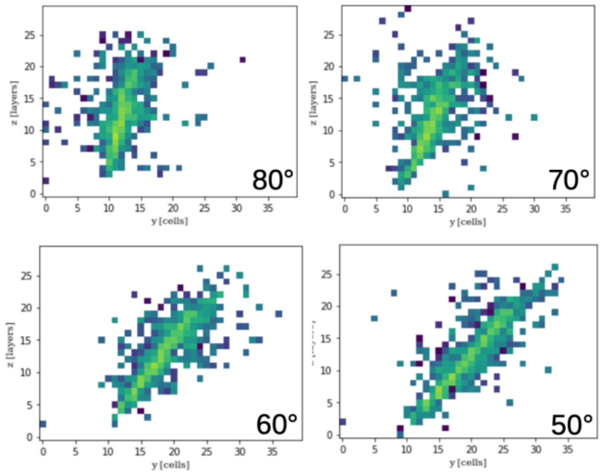
− = Partially Addressed　+ = Addressed here

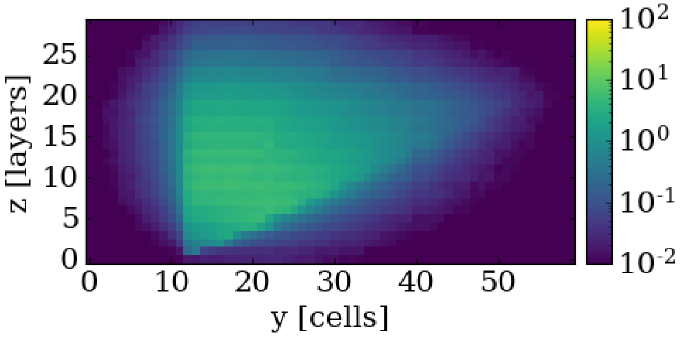
- For **realistic application** need to address several next steps:
  - **Angular conditioning**
  - Performance after **reconstruction**
  - …
- Continue to investigate other promising **ML architectures**

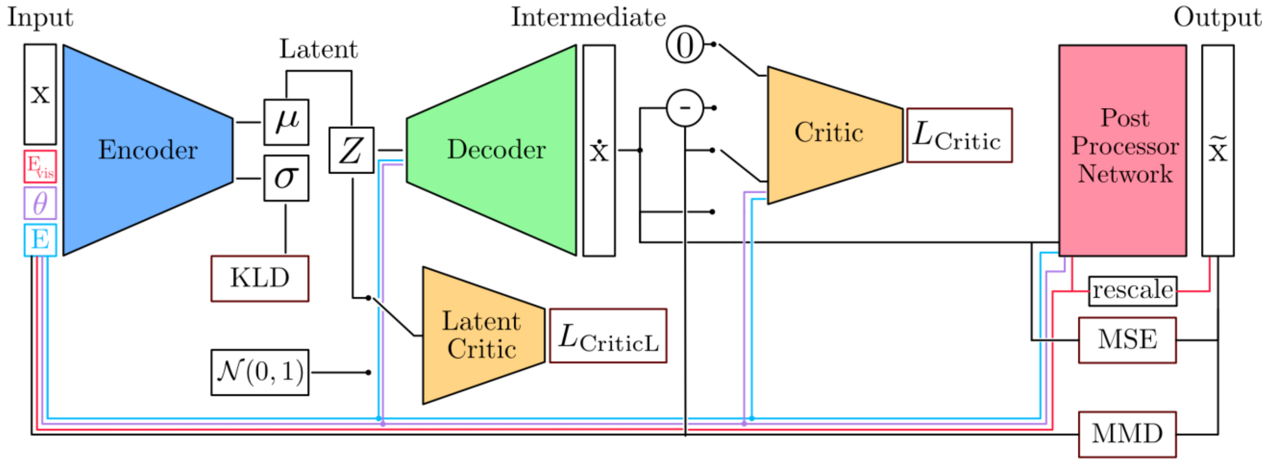
⟶ this talk

Energy and Angular Conditioning

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

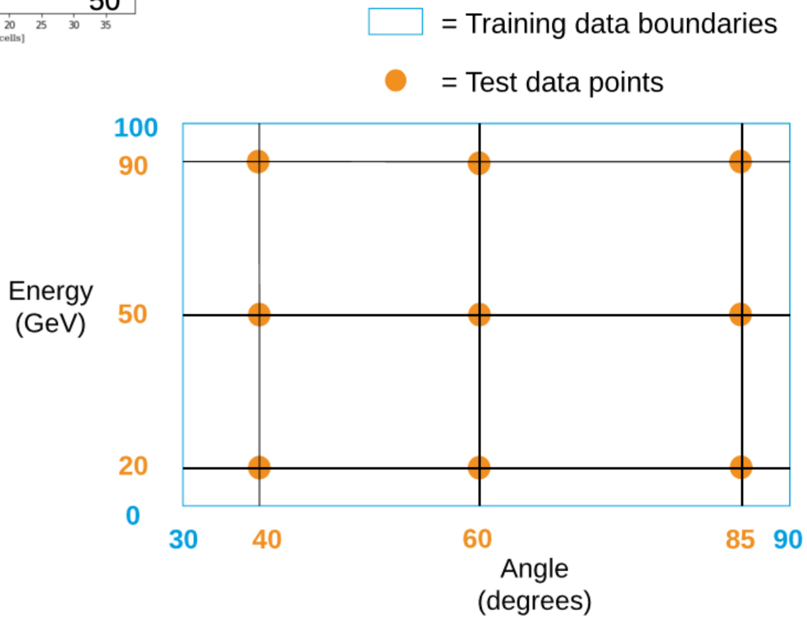






$$L_{\mathrm{BIB-AE}} = \mathrm{KLD} + L_{\mathrm{CriticL}} + L_{\mathrm{Critic}}$$

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

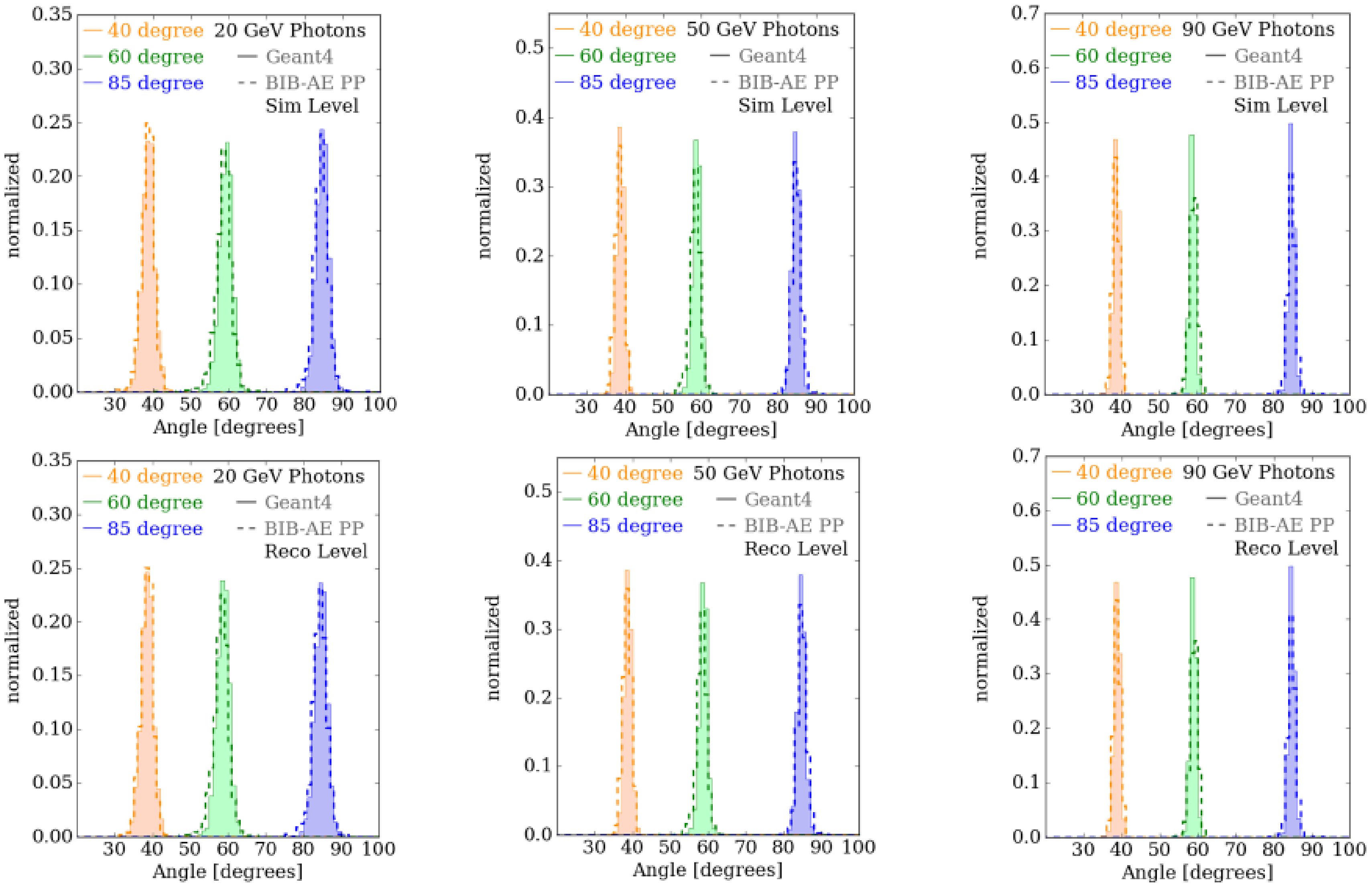


Training and validation samples - 30x60x30 grid

Angular Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, S.Diefenbacher et al.
arXiv: 2303.18150, submitted to MLST

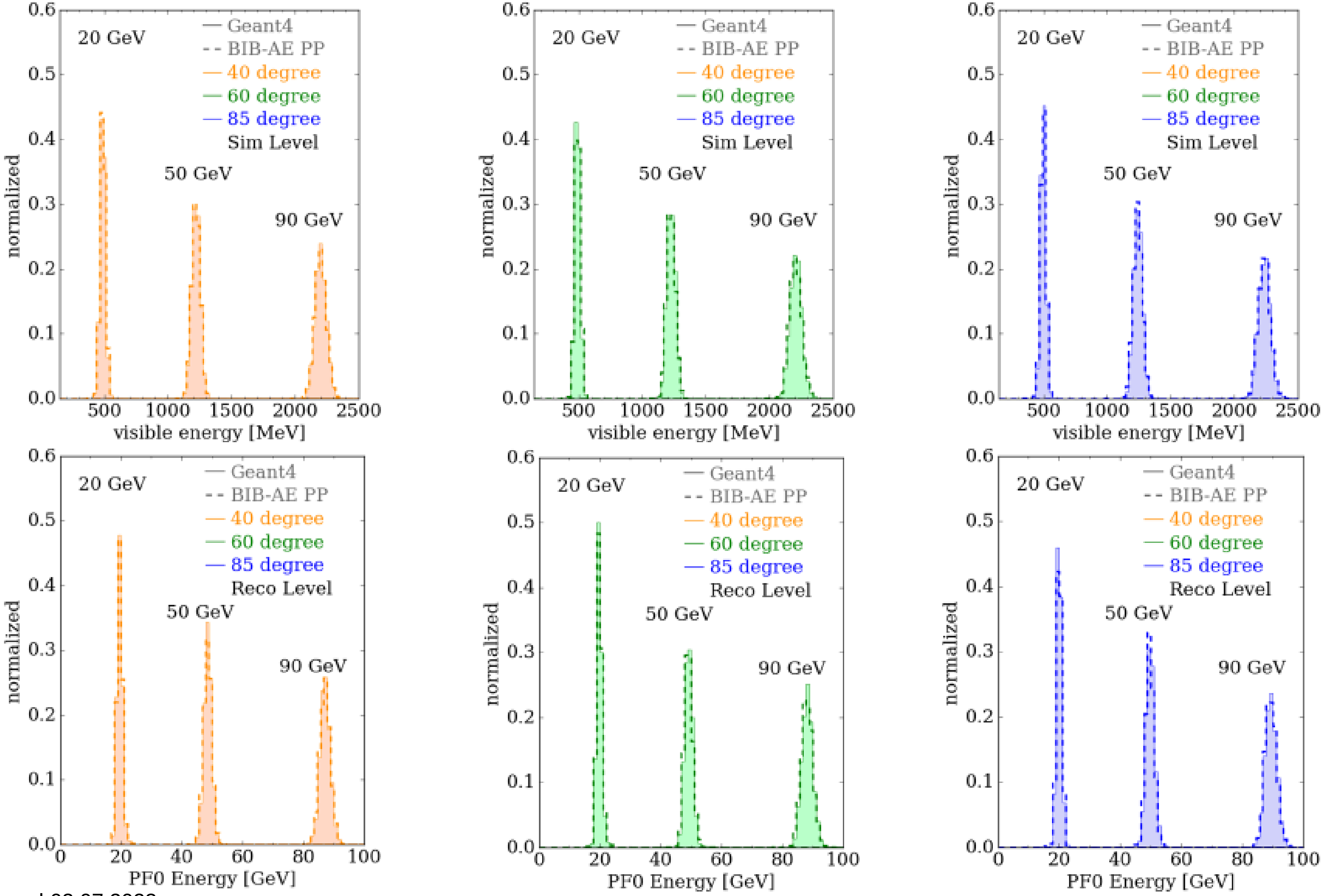
- **Sim** level PCA angle reconstruction
- **Rec** level PCA angle reconstruction
  - After full PandoraPFA reco



Energy Conditioning Performance

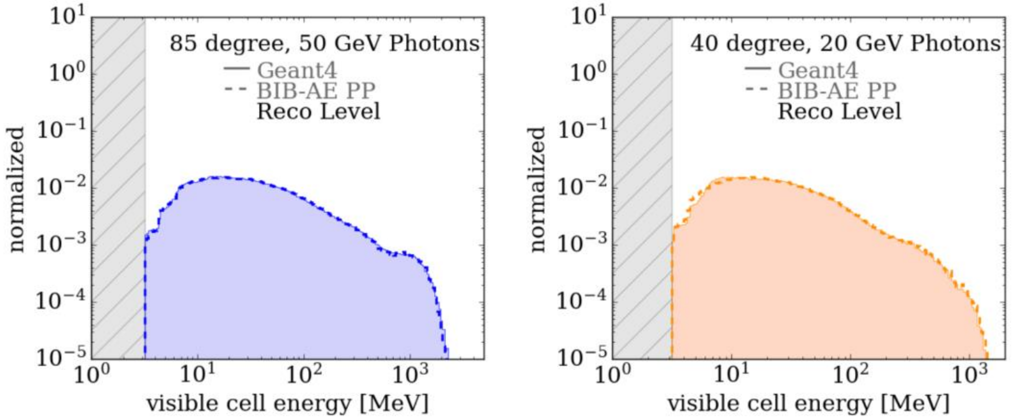
New Angles on Fast Calorimeter Shower Simulation, S.Diefenbacher et al.
arXiv: 2303.18150, submitted to MLST

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

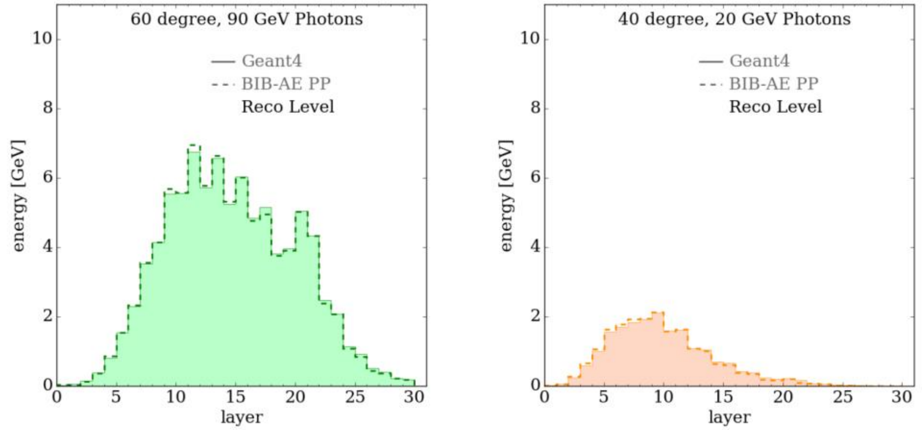


Performance After Reconstruction

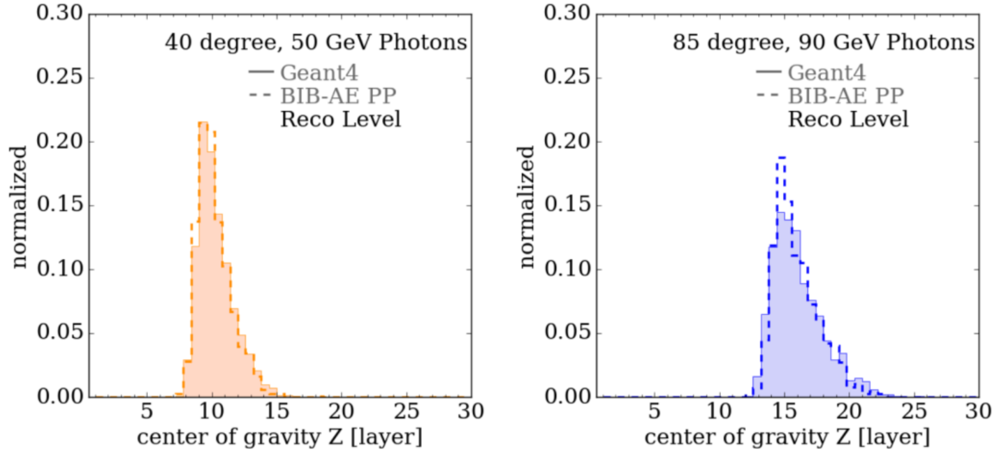
New Angles on Fast Calorimeter Shower Simulation, S.Diefenbacher et al.
arXiv: 2303.18150, submitted to MLST



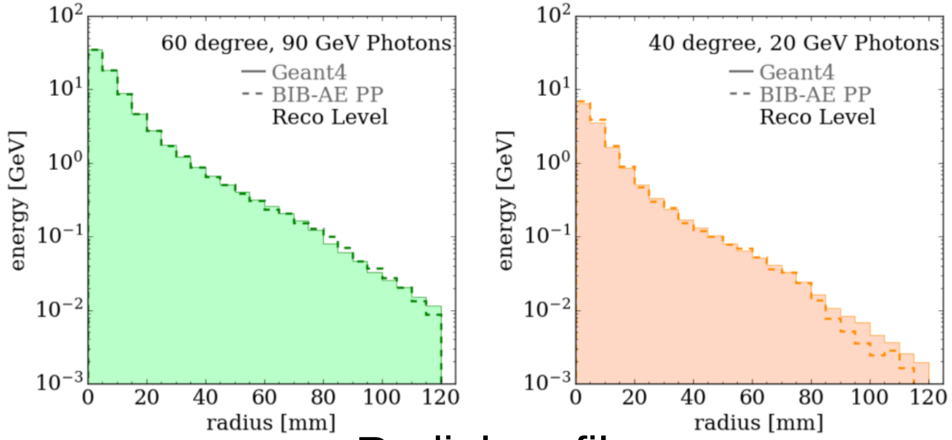
Hit energy spectrum



Longitudinal profile



Centre of gravity in z



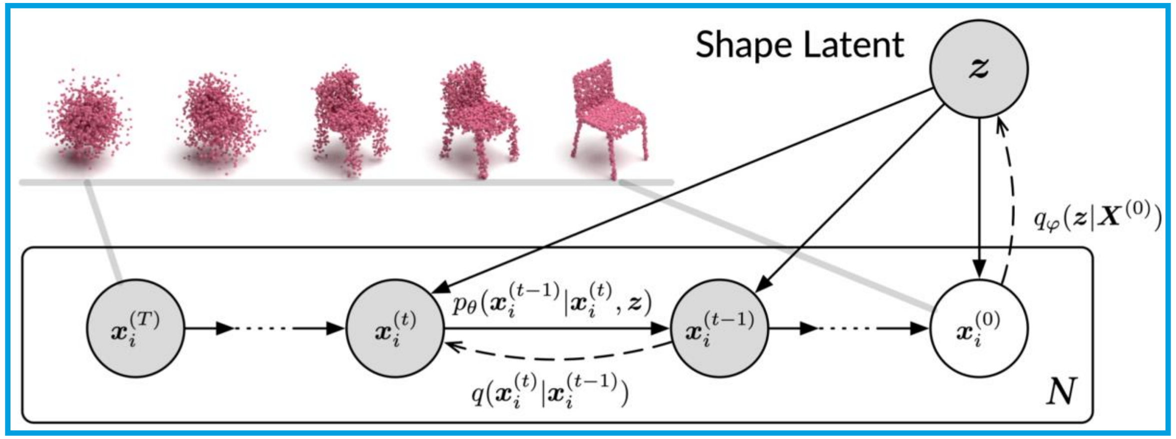
Radial profile

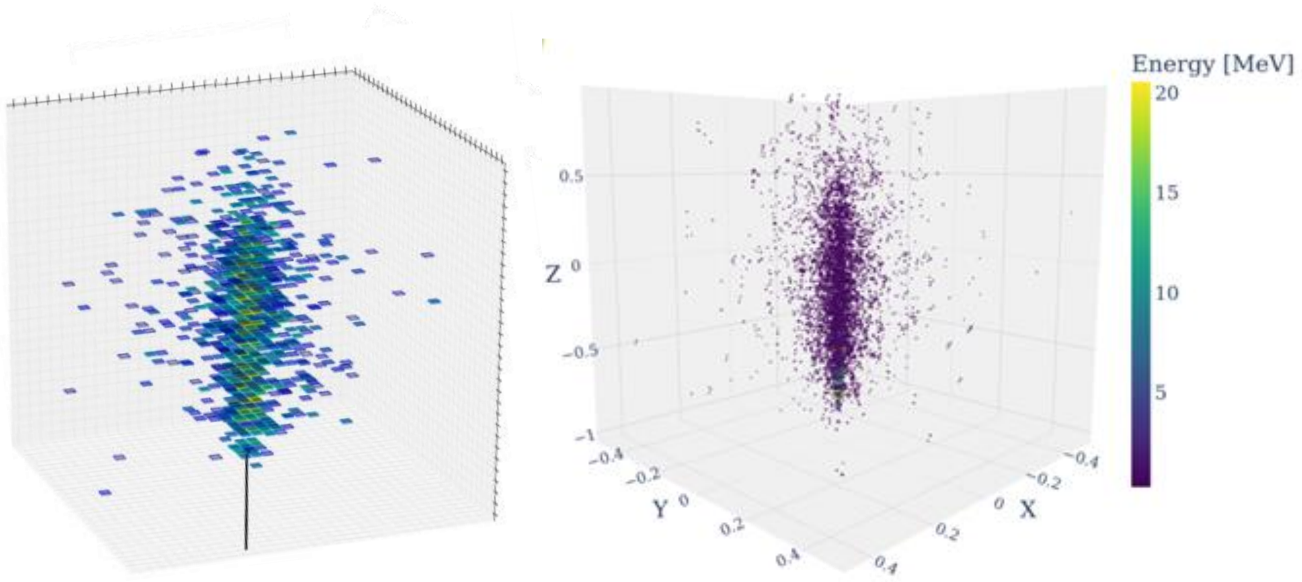
Best (left) and worst (right) test point - based on JSD $\longrightarrow$ **Excellent** physics fidelity

CaloClouds: Motivation

- **Regular grid models** show very high physics fidelity - yet they have **two drawbacks**:
  - **Low occupancy** ⟶ lots of superfluous compute
  - **Irregular** detector **geometries** ⟶ back projection creates **artifacts**

S.Luo, W.Hu: Diffusion Probabilistic Models for 3D Point Cloud Generation, arXiv:2103.01458

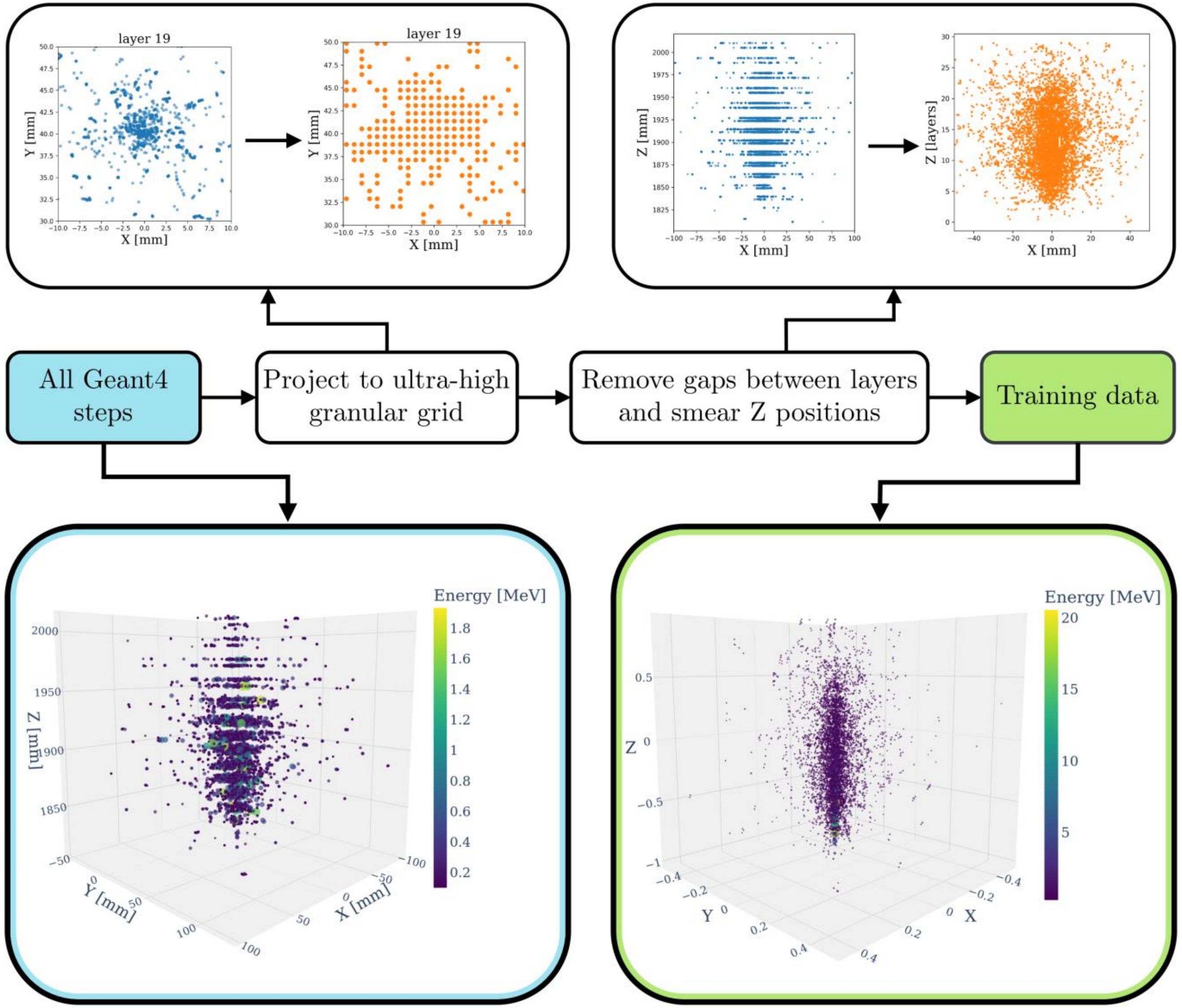




- Solution: **point cloud** based models
- Adapt models to HEP calorimeter use case
- Use **additional information** in the Geant4 **simulation**:
  - Much **higher granularity** than **physical geometry**
  - Gain **geometry independence**

CaloClouds: Data Preprocessing

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, E. Buhmann et al., arXiv:2305.04847

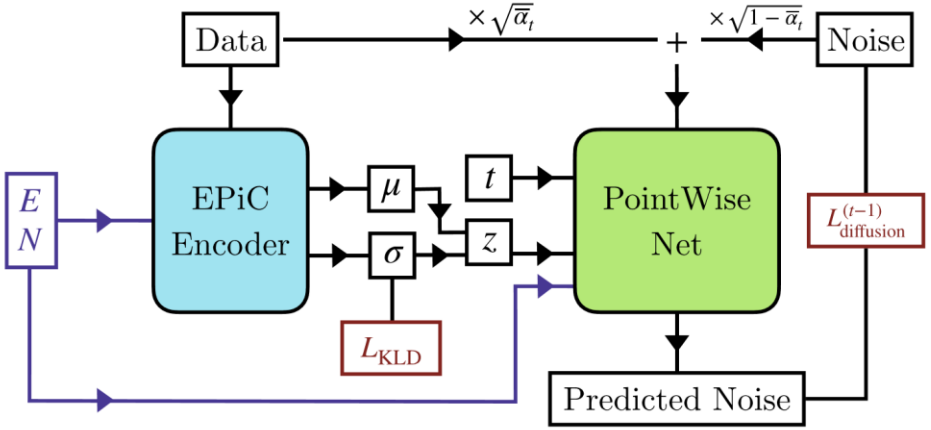


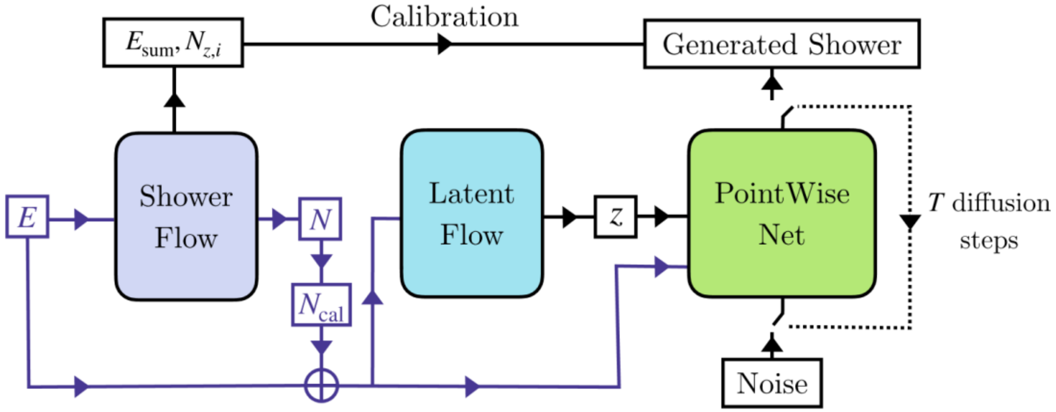
- Using all Geant4 steps directly is **computationally prohibitive**
  - 40k Geant4 steps at 90 GeV
- Apply preprocessing step:
  - Project Geant4 steps into **ultra-high granularity grid** (36 times more granular than ILD ECAL)
    - Reduce number of points by factor ~7:
    - Up to **6000 space points**
- Again study **photons** in ILD ECAL
  - 10-100 GeV, 90 deg impact
  - Additionally **check** the effects of **varying incident point** (hence geometry)

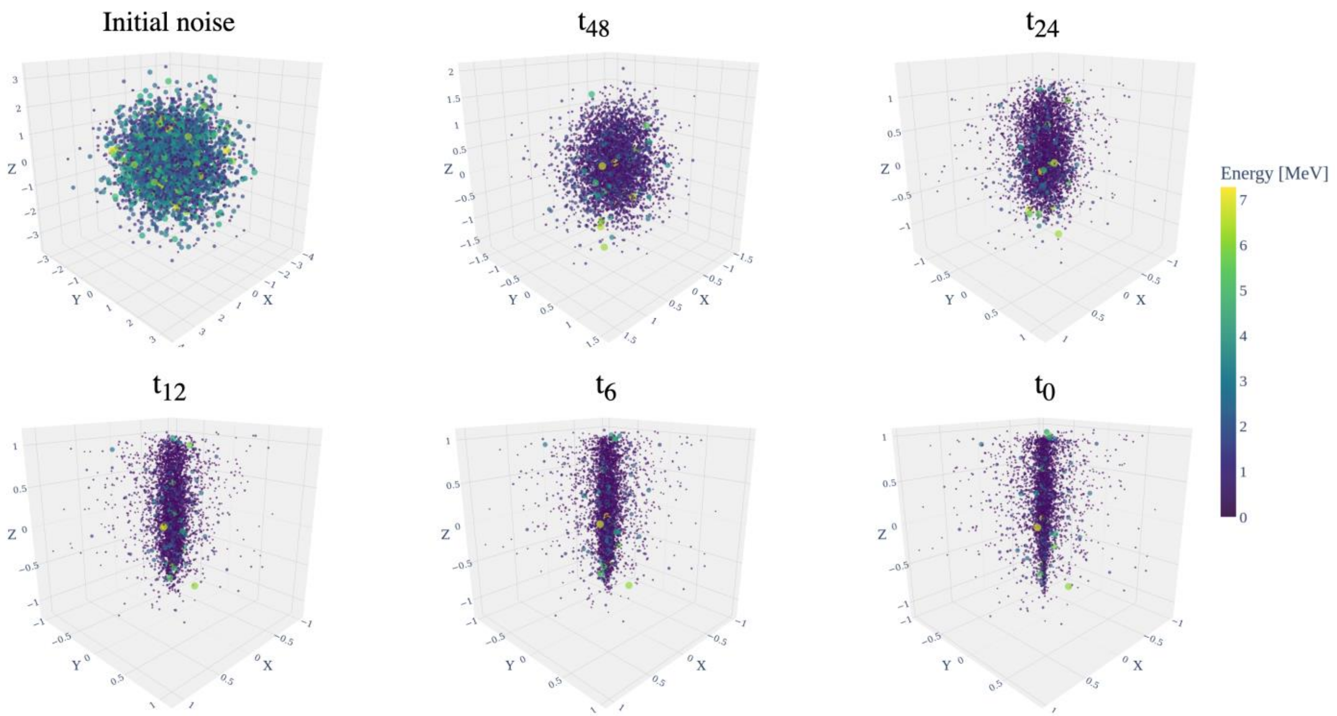
CaloClouds: Diffusion Model Architecture

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, E. Buhmann et al., arXiv:2305.04847

- Training of PointWise Net with EPiC Encoder (e-Print: 2301.08128)
- Inference uses two additional flows for number of space points, calibration and latent space

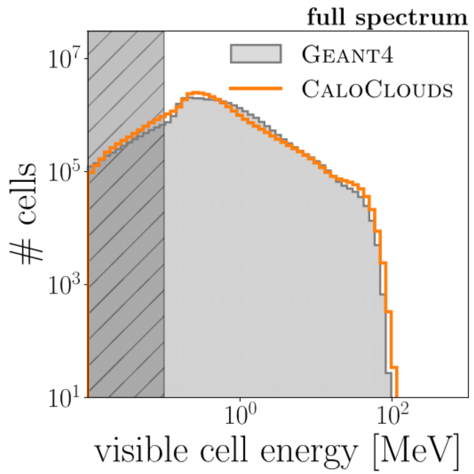


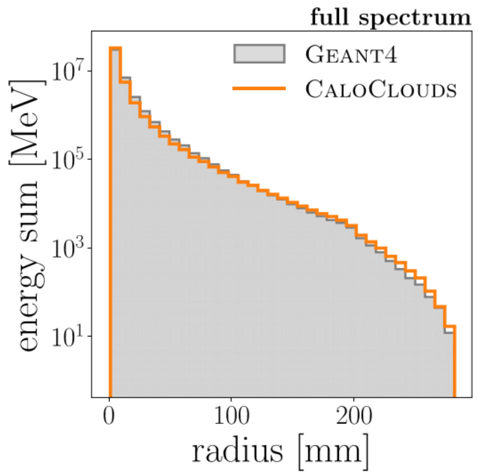


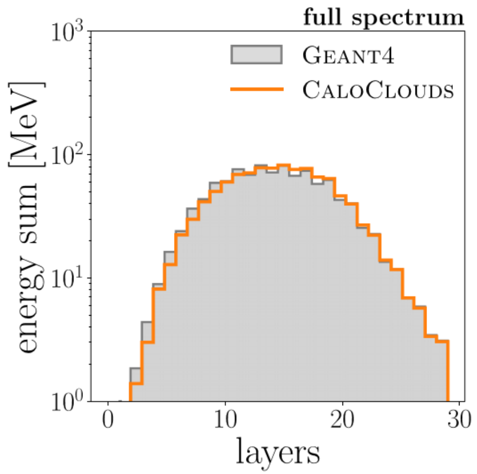


CaloClouds: Results

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, E. Buhmann et al., arXiv:2305.04847



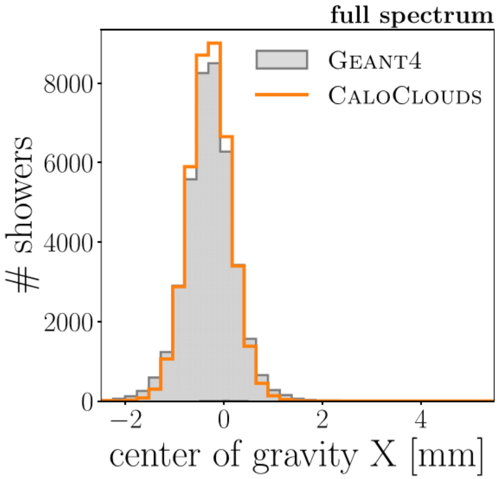


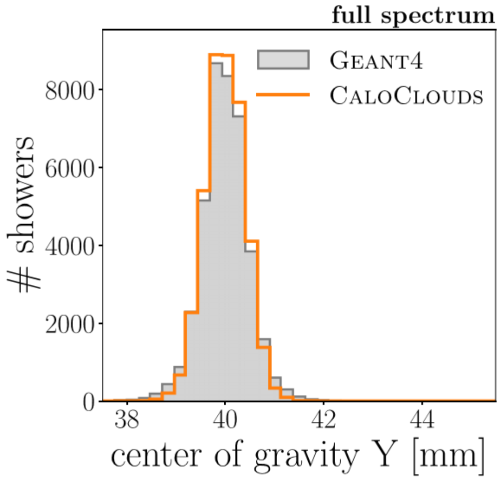


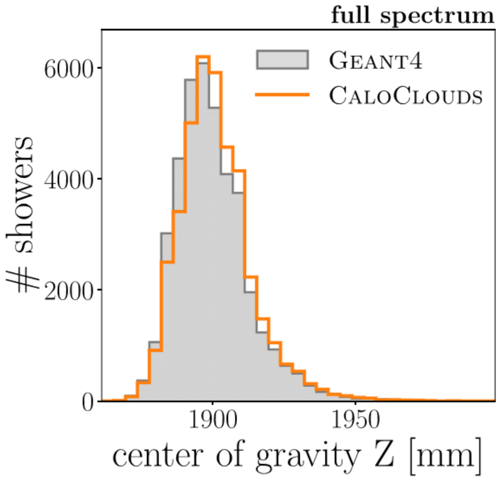
- Overall observe very good physics fidelity
- **First successful application of diffusion models to (high granularity) calorimeter simulation using point clouds**

Hardware	Simulator	Time / Shower [ms]	Speed-up
CPU	GEANT4	4082 ± 170	$\times 1$
	CALOCLOUDS	3509 ± 220	$\times 1.2$
GPU	CALOCLOUDS	38 ± 3	$\times 107$

Future work: focus on improving timing







Conclusion

Achieved

- **Energy and angular conditioning** for EM showers with high physics fidelity
  - **Strong performance** after **reconstruction** with PandoraPFA

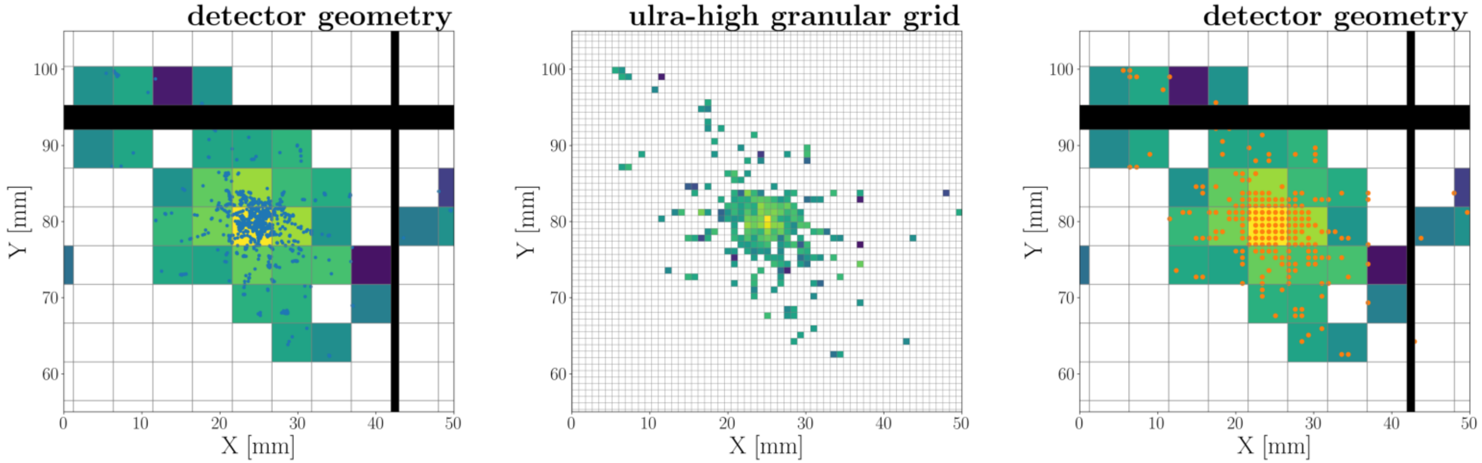
- **CaloCloud diffusion model** for EM showers
  - **First** successful generation of calorimeter showers as a **point cloud**
  - Achieve a high degree of **geometry independence**

Next Steps

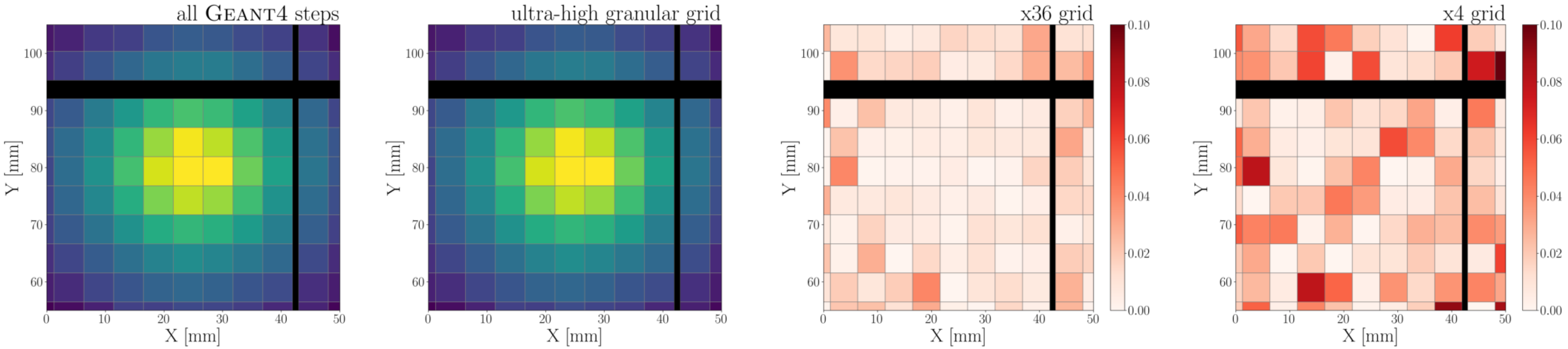
- Continue to **improve** simulation **speed** with models and **extend functionality**
  - **Integrate** models into **full** detector **simulation pipelines** and study full **physics benchmarks**
  - Investigate methods to **speed up CaloCloud** diffusion model

Backup

Effects Of Pre-clustering To Ultra-high Granularity



example:
Geant4 90 GeV shower in layer 21 with full round-trip pre-clustering and back-projection

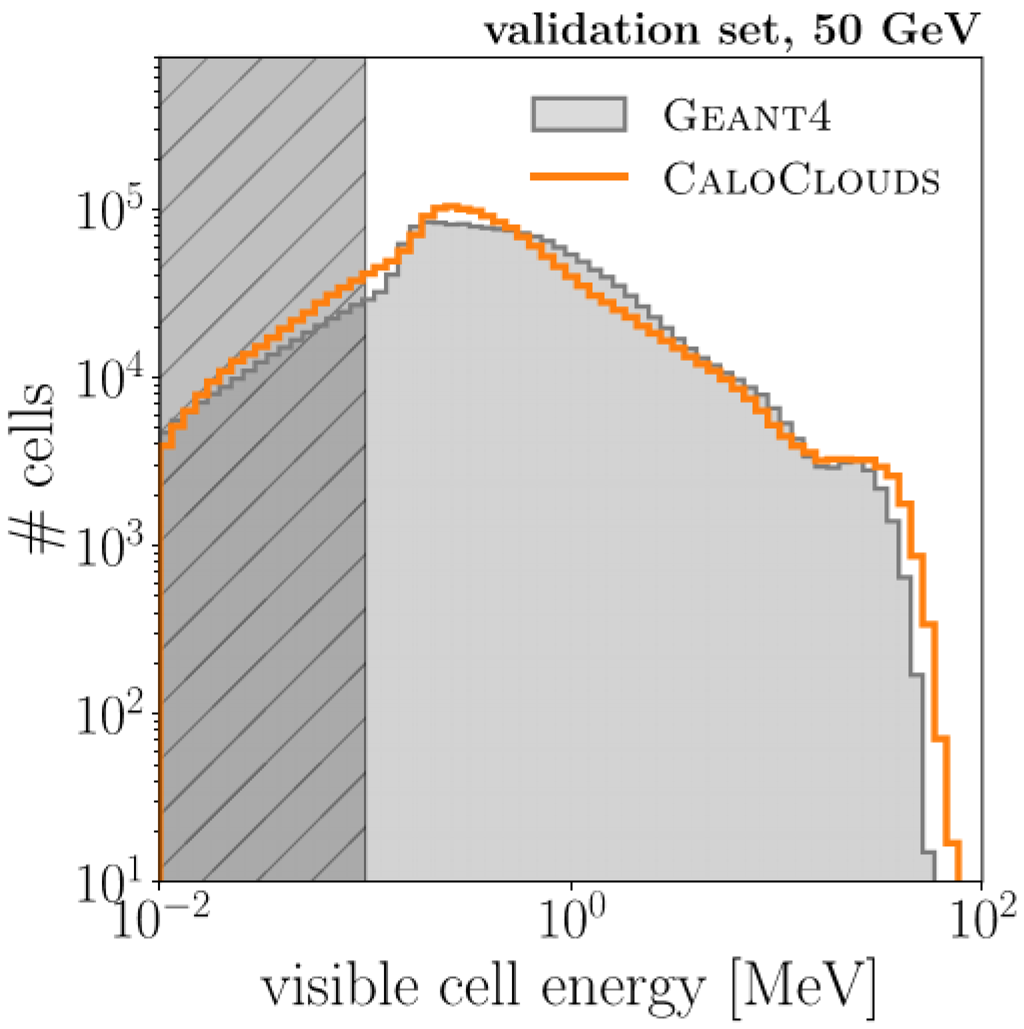


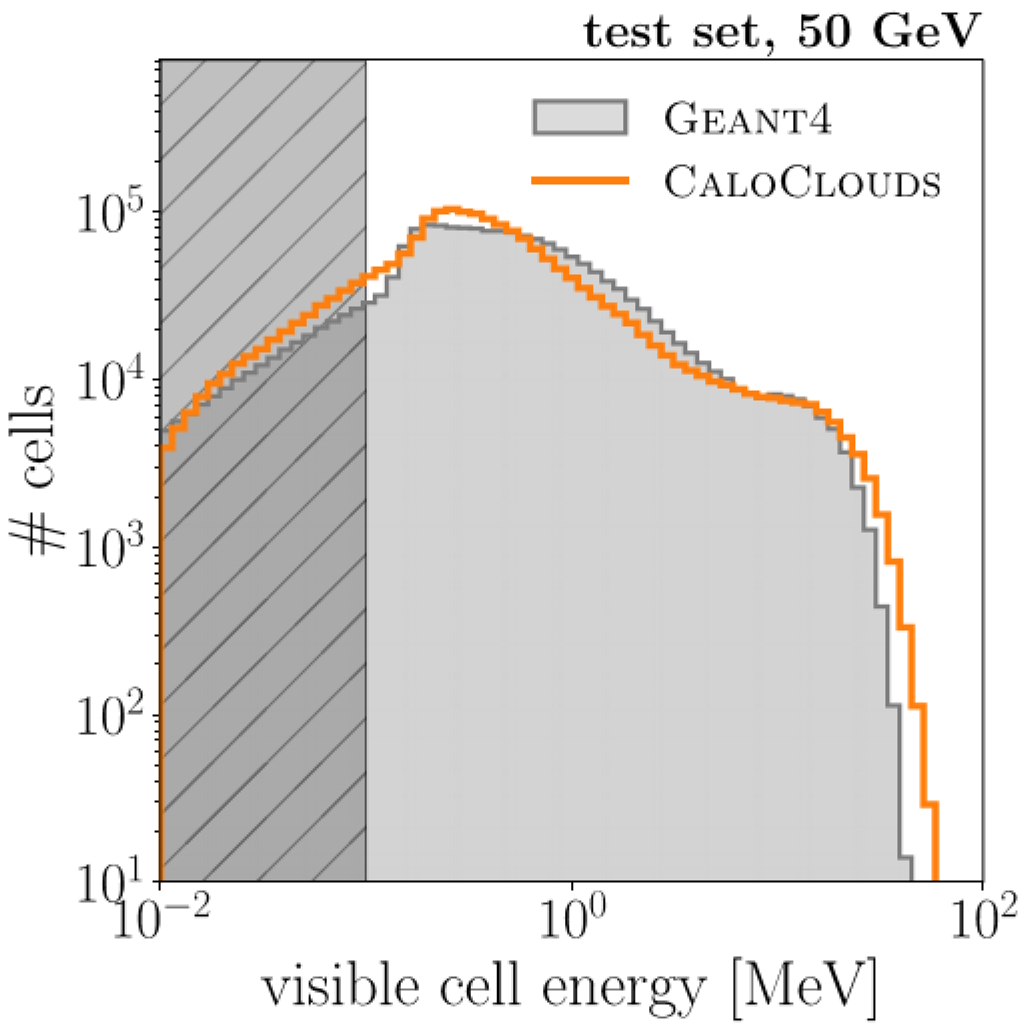
overlay of 2k Geant4 90 Gev showers in layer 21 with full round-trip pre-clustering and back-projection

relative difference per cell < 2% in core of shower

relative difference with lower granularity

CaloClouds: Effects Of Varying Geometry





Timing Of Generative ML Methods

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

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

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

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

Hardware	Simulator	Time / Shower [ms]	Speed-up
CPU	GEANT4	4082 $\pm$ 170	$\times$1
	CALOCLOUDS	3509 $\pm$ 220	$\times$1.2
GPU	CALOCLOUDS	38 $\pm$ 3	$\times$107

CaloClouds, photon showers 10-100 GeV uniform