AIDAinnova 12.3 FastSim Activities at DESY

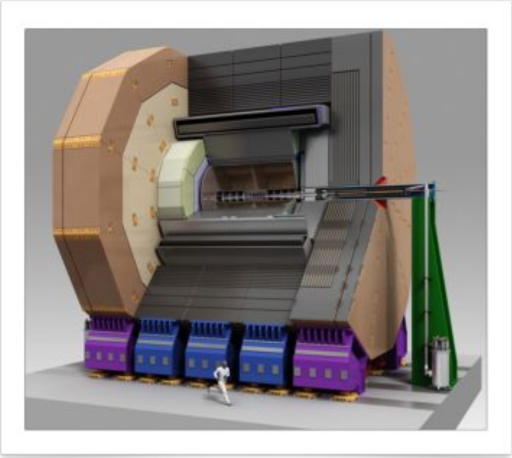
Annual Meeting 2024, Canatia

Aanatolii Korol*
*anatolii.korol@desy.de

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE
HELMHOLTZ

Calo-ML Working Group at DESY

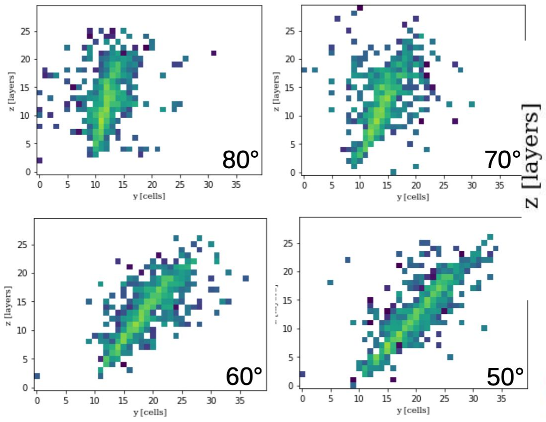
- goal: study the suitability of using generative networks for the simulation of calorimeter showers in highly granular calorimeters with high fidelity
  - ideally such that they can replace G4 in full simulation for physics analyses for Future Higgs Factories
  - work carried out in context of the ILD detector concept for the ILC and with the CALICE collaboration
- generative ML working group:
  - part of FTX-SFT at DESY
    - 1 post doc, 2 PhD students
- joined with UHH (G.Kasieczka) in Quantum Universe cluster of excellence
- AIDAinnova, ACCLAIM (Helmholtz Innovation Pool), CDCS,…

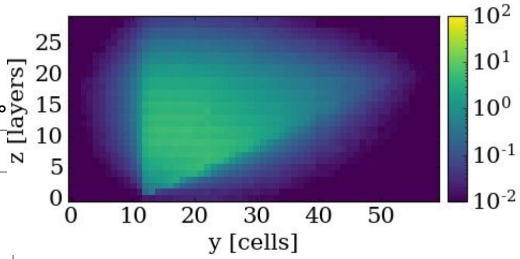


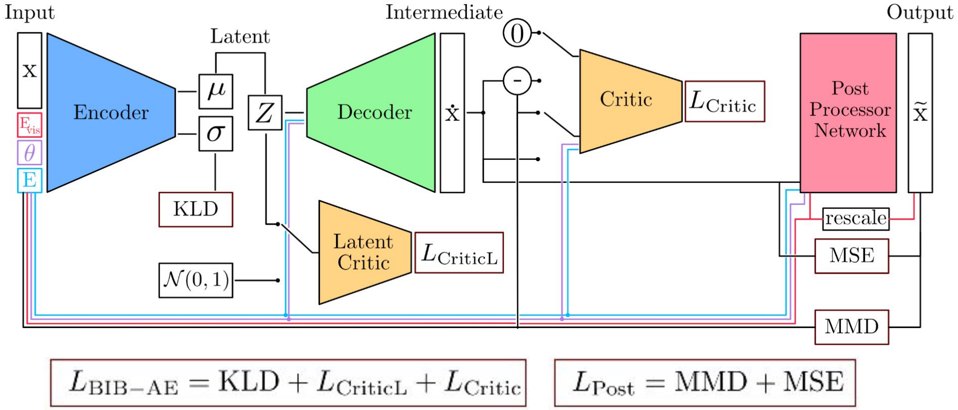
Progress on 3D Regular Grid

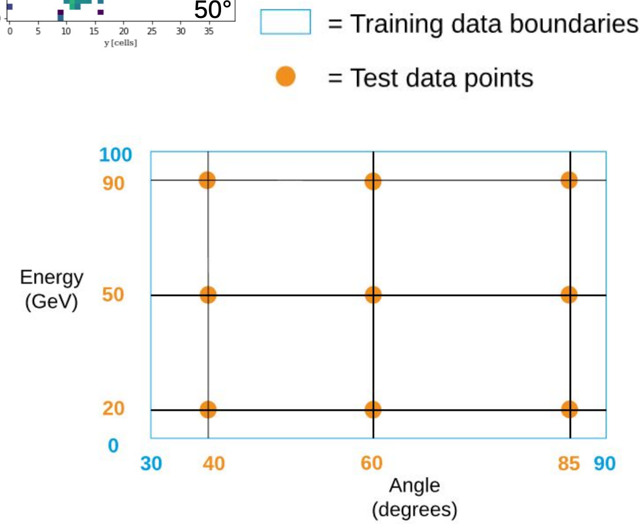
Energy and Angular Conditioning

- 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





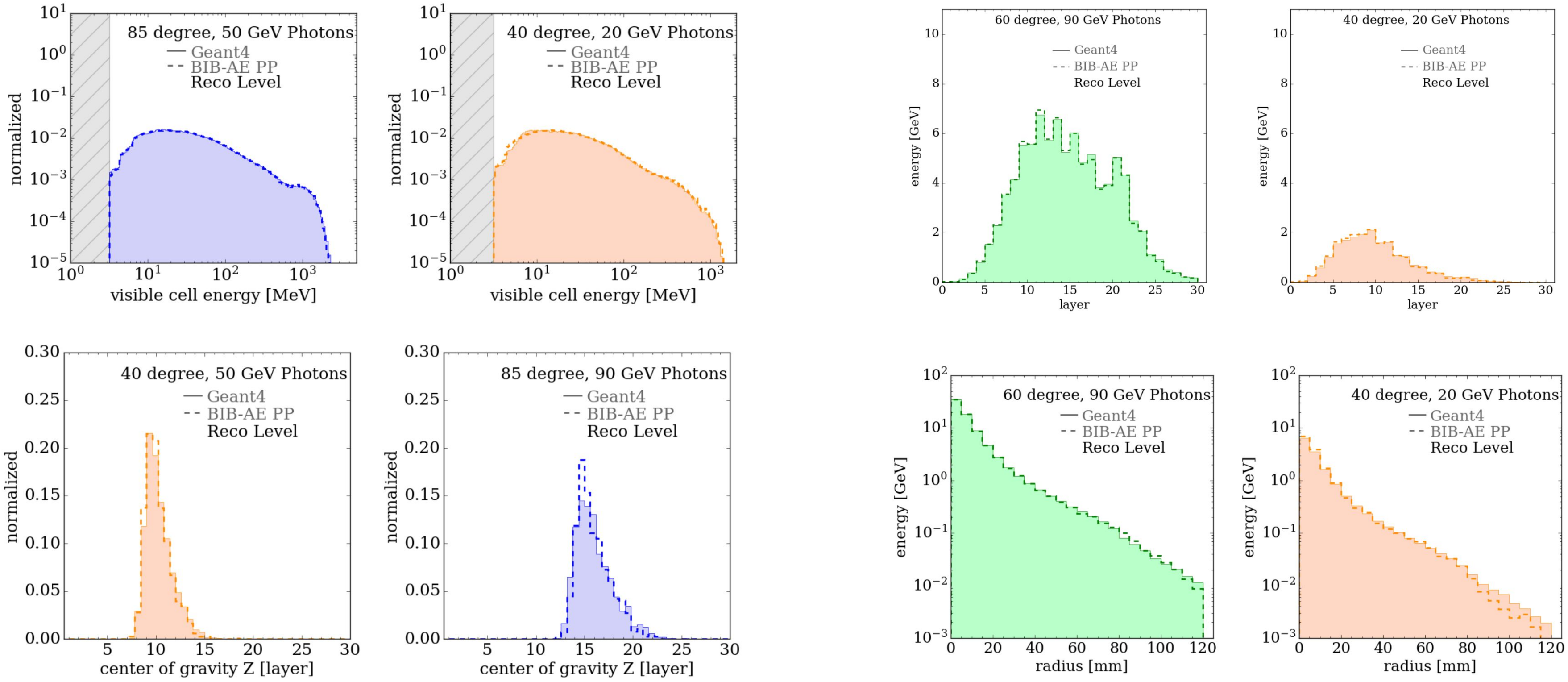




Progress on 3D Regular Grid

Energy and Angular Conditioning, Performance After Reconstruction

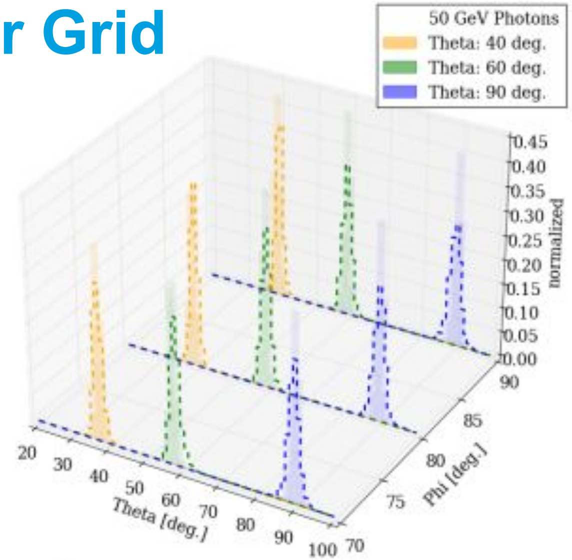
New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, et al., 2023 MLST in press
DOI 10.1088/2632-2153/acefa9, **arXiv: 2303.18150**

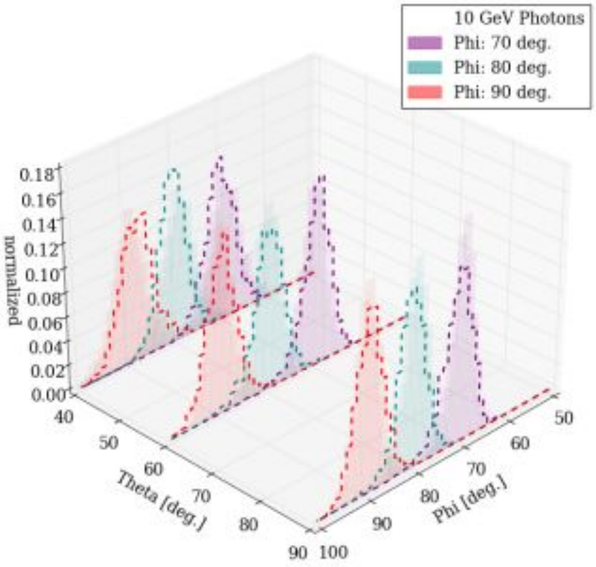


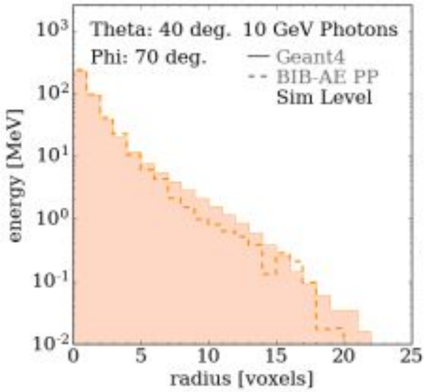
Progress on 3D Regular Grid

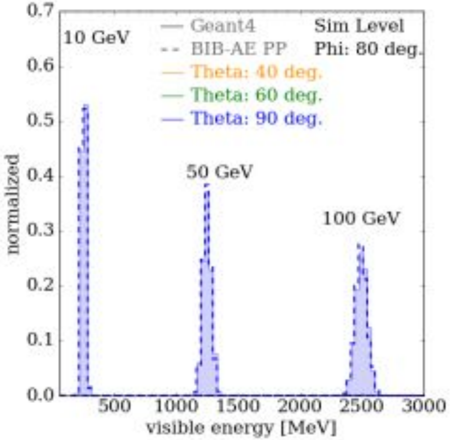
Adding Another Angle

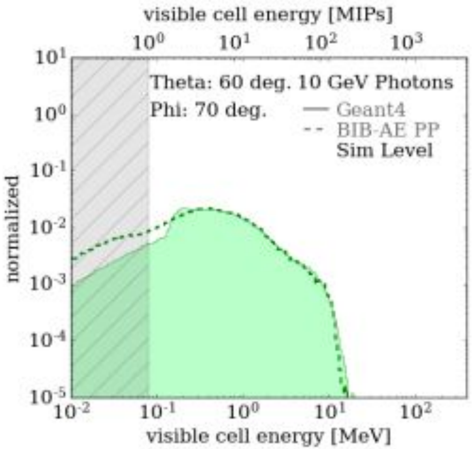
- Need to condition on **energy, theta and phi** for full application
- **Extending phase space** can be challenging







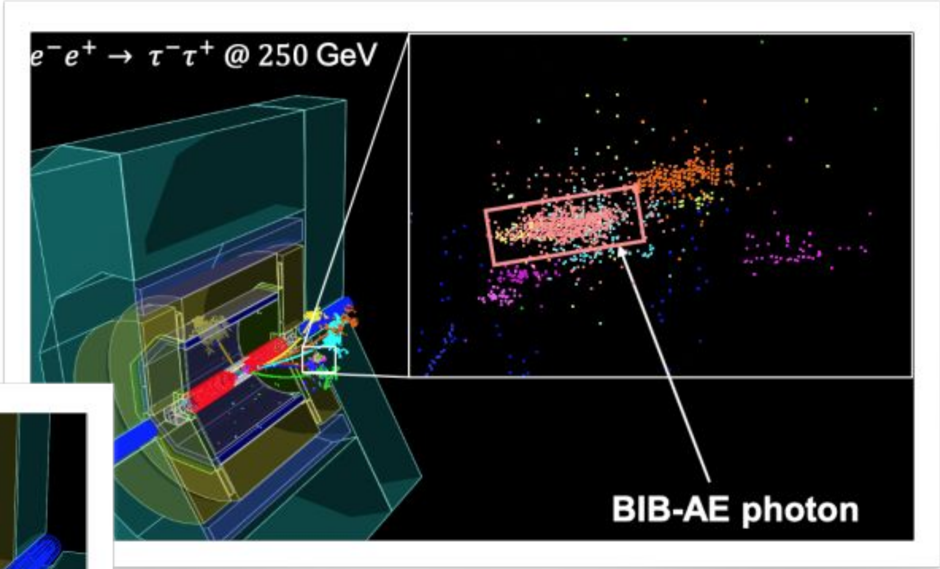




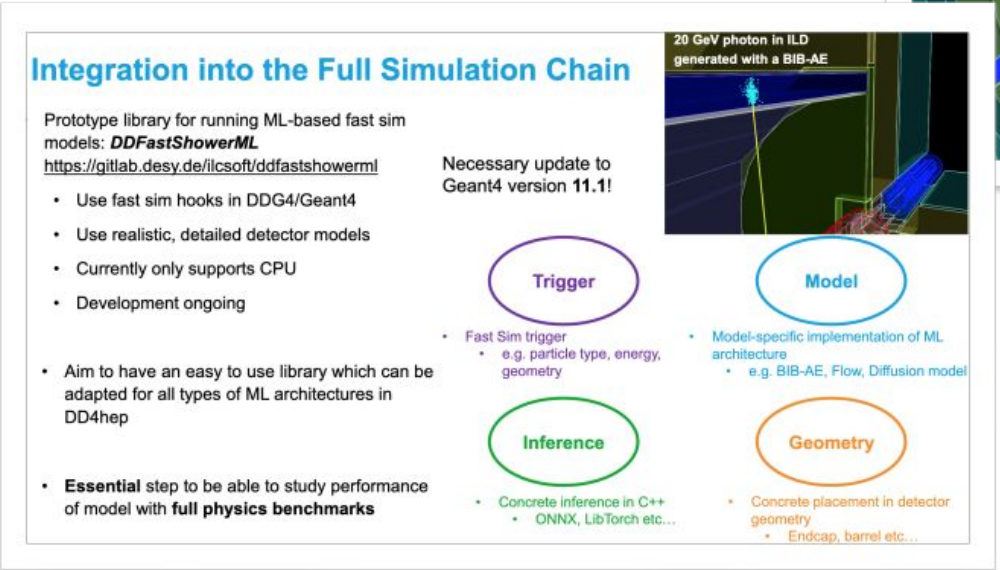
Progress on 3D Regular Grid

Two-Angle Conditioned BIB-AE

P.McKeown



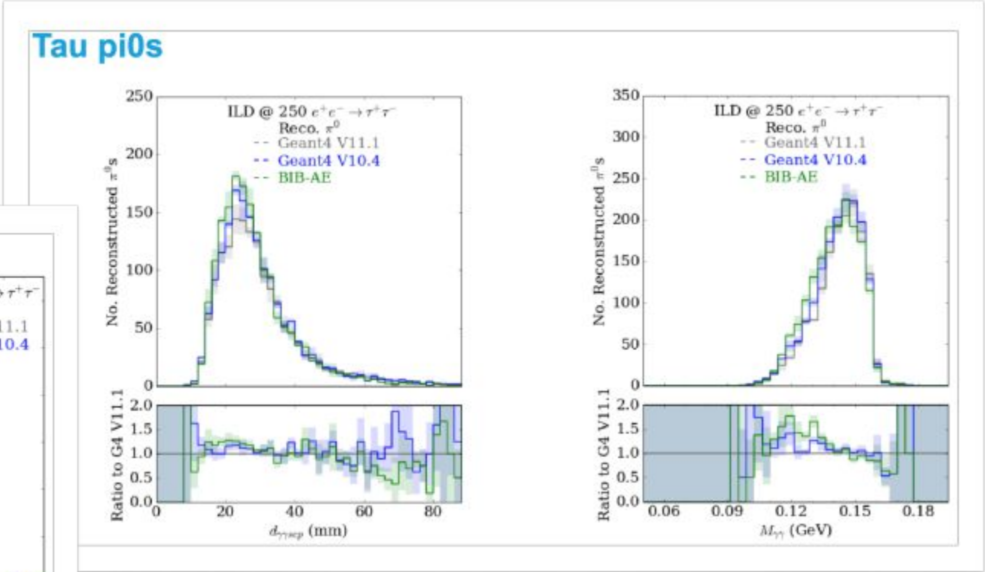
- used BIB-AE for fast simulation of photons from tau-pairs in ILD w/ DDML
- run full reconstruction (w/ PandoraPFA) of these events and compare G4 and BIB-AE

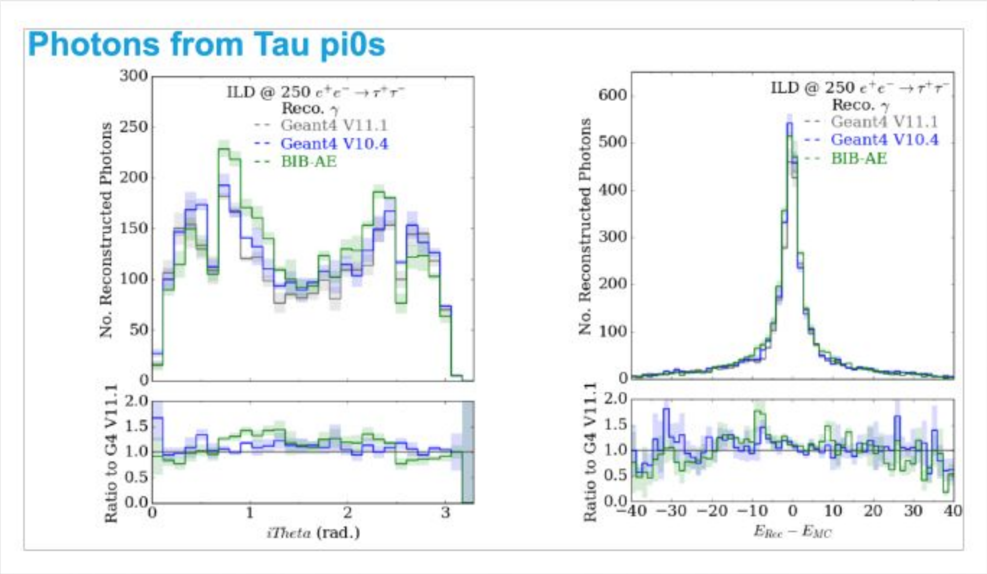


Progress on 3D Regular Grid

Two-Angle Conditioned BIB-AE

P.McKeown





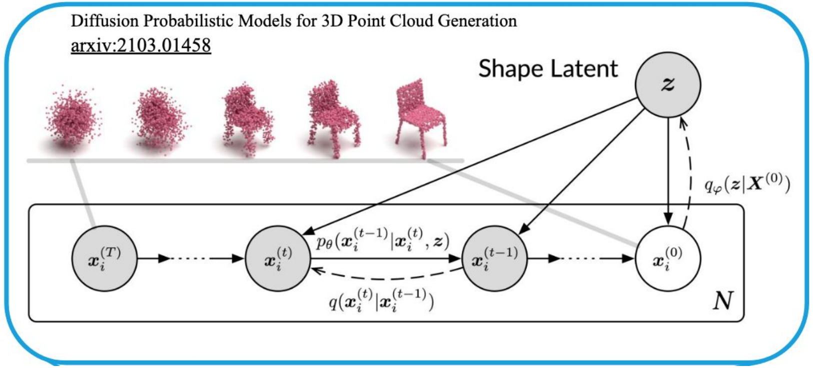
- reasonably good description of pi0s from tau decays
- differences between BIB-AE of similar order as G4-10.4 (used for training) and G4-11.1 (used for reco)

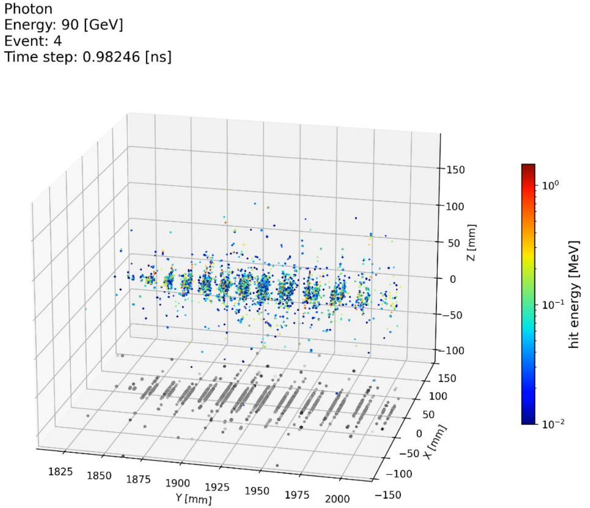
Point Cloud Representation of the EM Showers

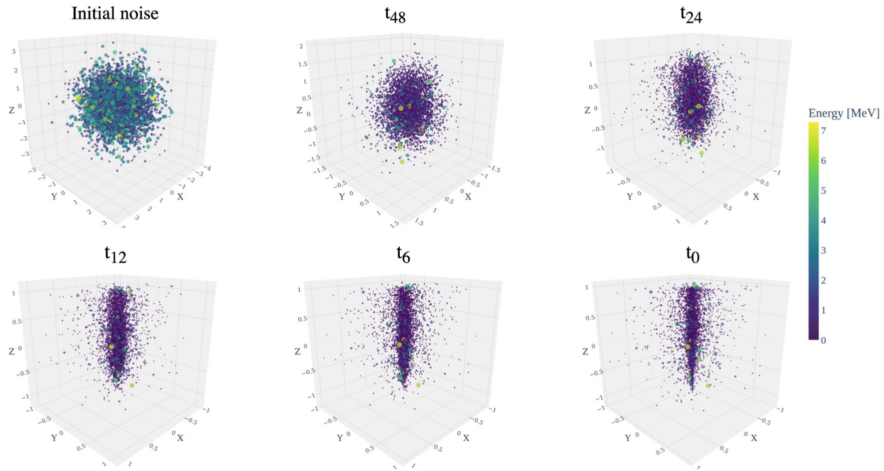
GEANT4 Steps

A way to overcome potential issues from irregular (realistic) cell geometry is use of much higher granularity/resolution

- All G4 interactions, ultimate resolution
- Detached from detector layer geometry



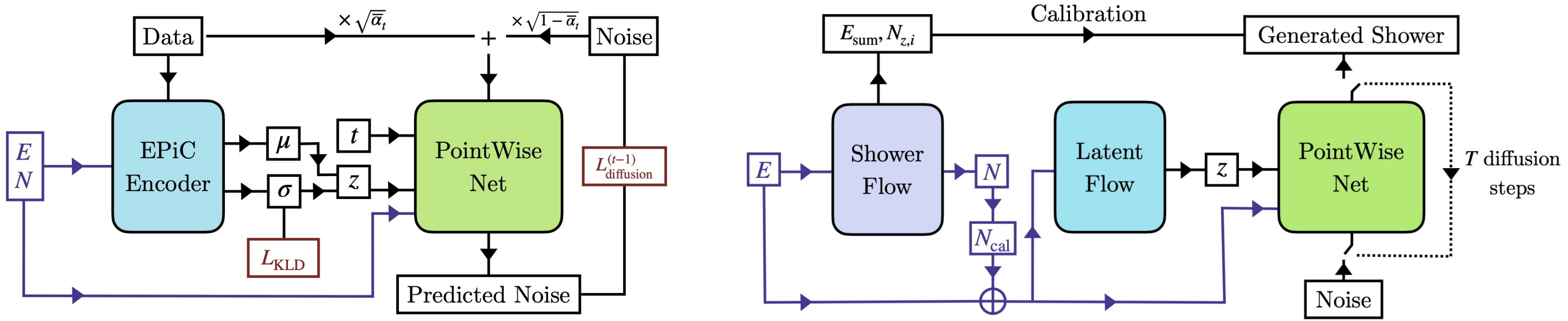




Point Cloud + Diffusion Model

CaloClouds, Model Overview

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2305.04847**



(a) Training at random time step t

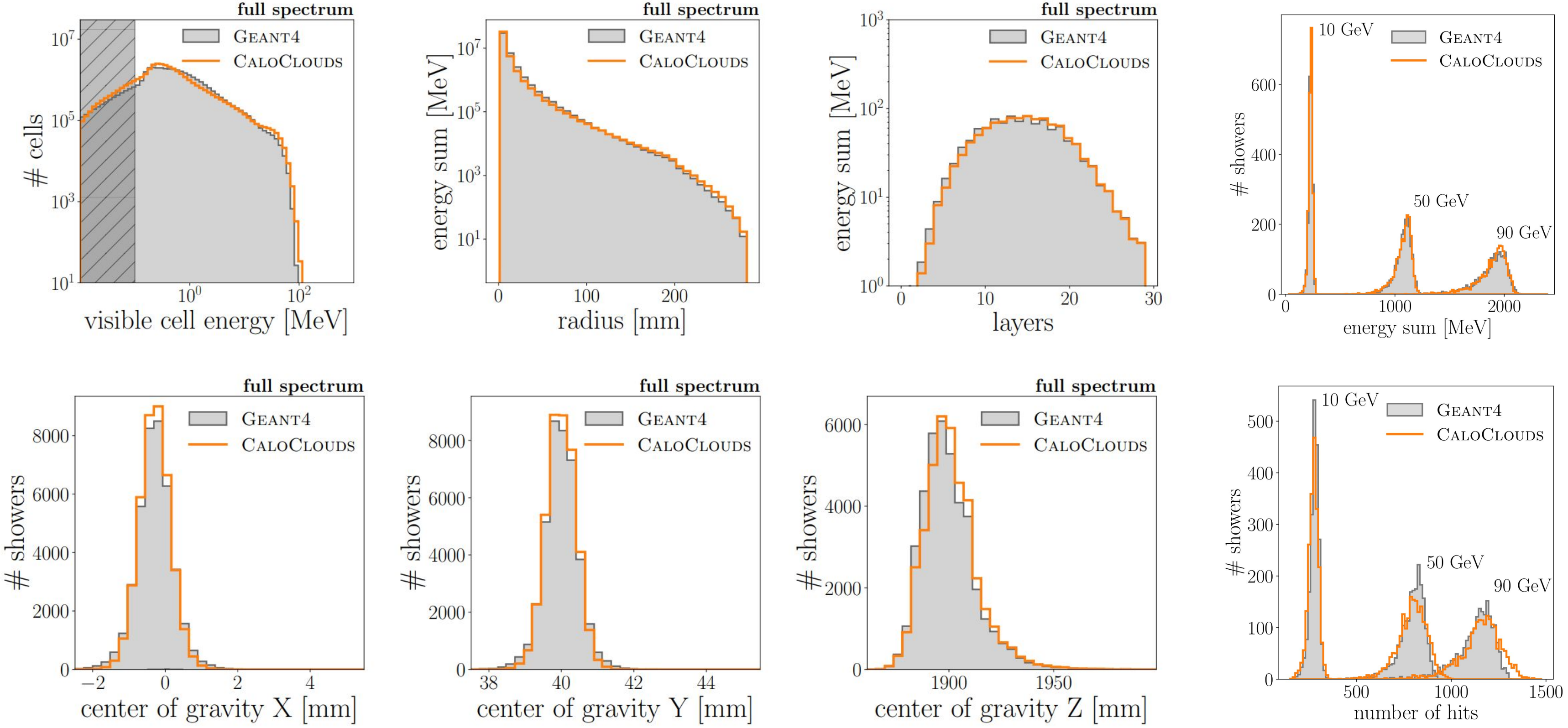
(b) Sampling with reverse diffusion through all time steps T

- GANs and VAEs convert noise from some simple distribution to a data sample
- **DMs learn to gradually denoise data starting from noise**

Point Cloud + Diffusion Model

Results

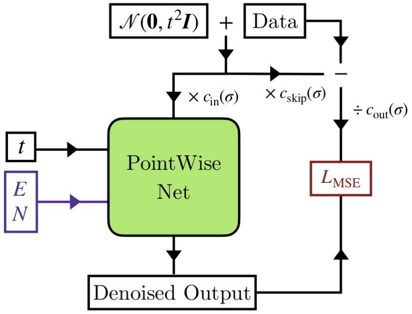
CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2305.04847**



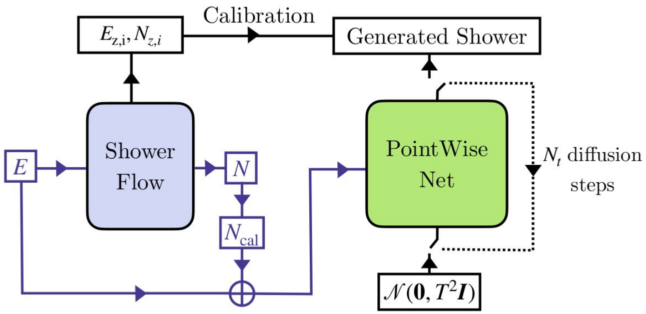
Point Cloud + Diffusion Model

CalоClouds II, Model Overview

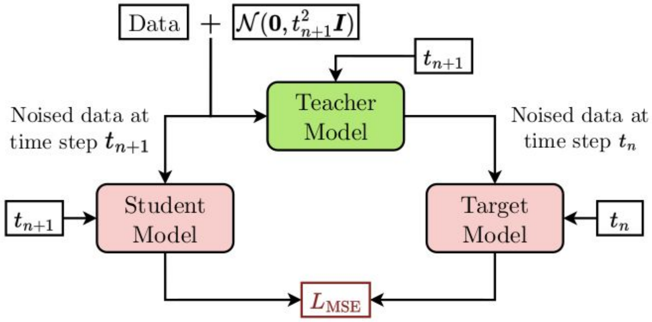
CaloClouds II: Ultra-Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2309.05704**



(a) Training



(b) Sampling



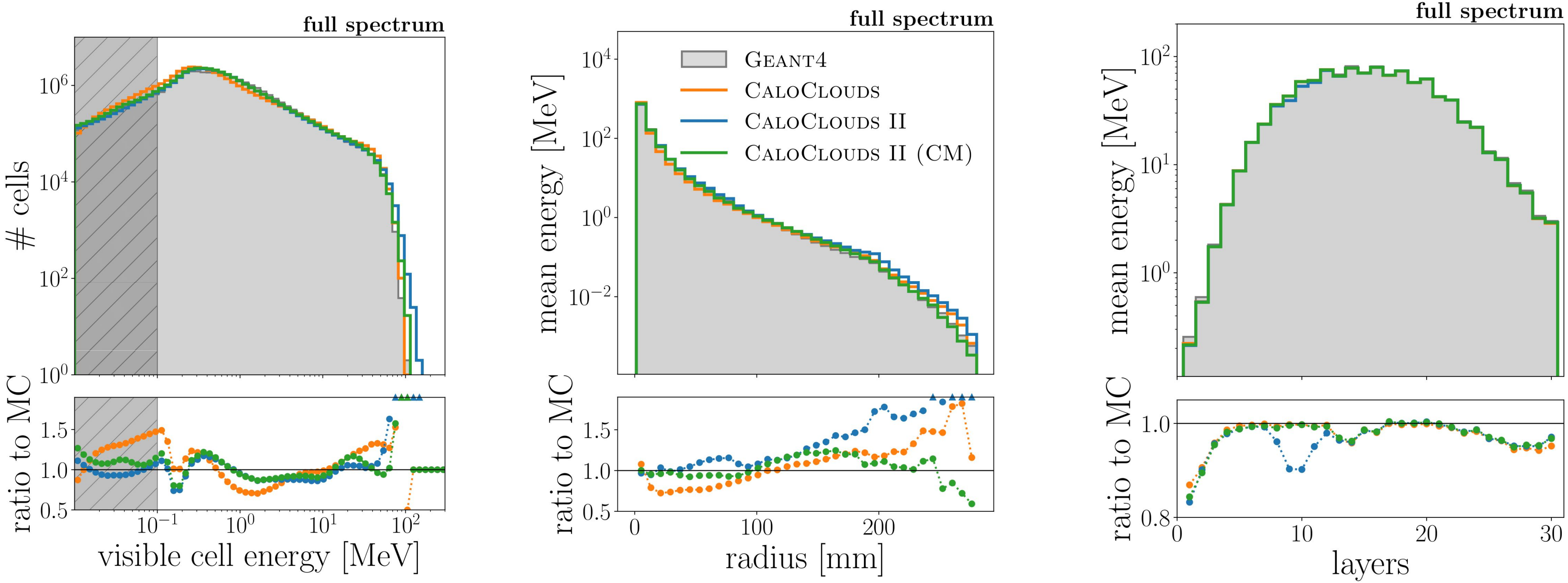
Consistency Distillation

Modified version of CaloClouds + Consistency Distillation → significantly reduced inference time

Point Cloud + Diffusion Model

Results

CaloClouds II: Ultra-Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2309.05704**



Point Cloud + Diffusion Model

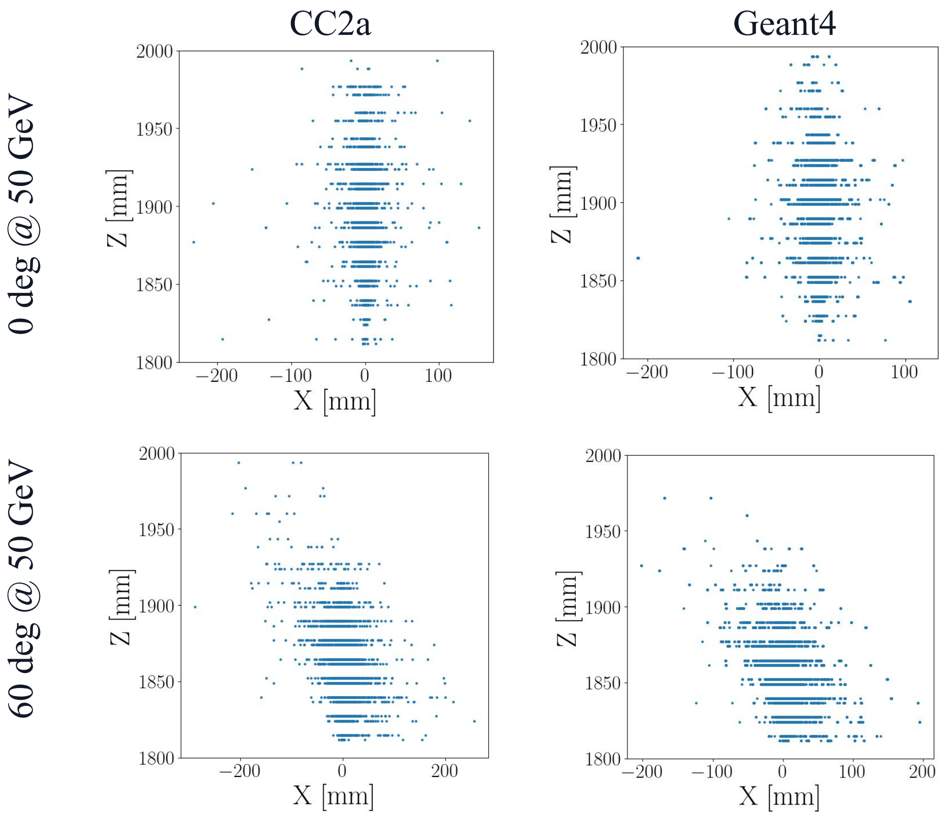
Results

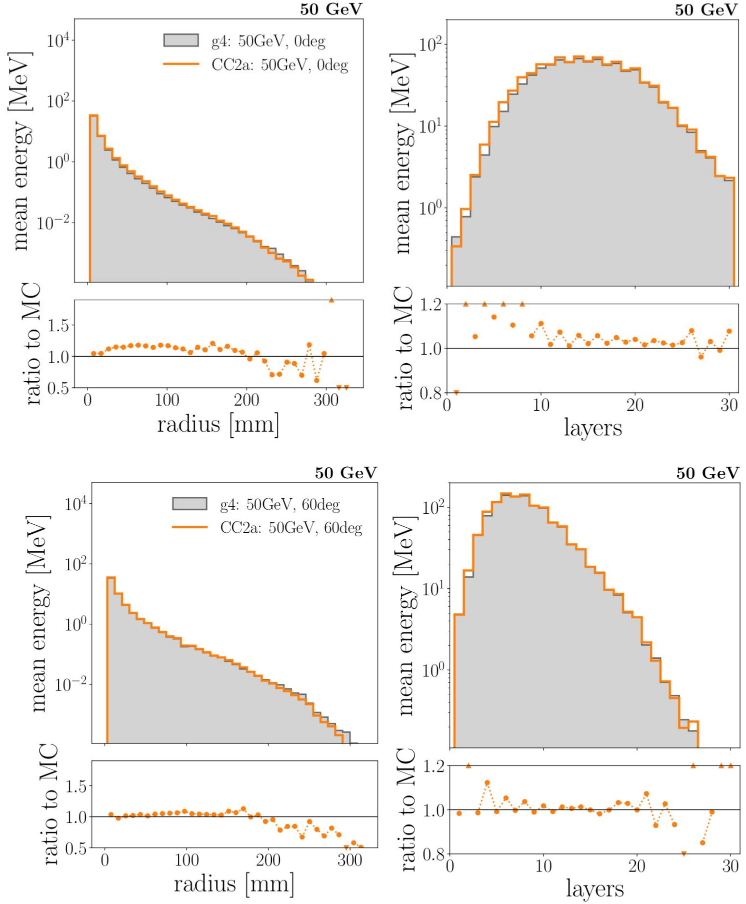
CaloClouds II: Ultra-Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2309.05704**

Hardware	Simulator	NFE	Batch Size	Time / Shower [ms]	Speed-up
CPU	GEANT4			3914.80 ± 74.09	$\times 1$
	CALOCLOUDS	100	1	3146.71 ± 31.66	$\times 1.2$
	CALOCLOUDS II	25	1	651.68 ± 4.21	$\times 6.0$
	CALOCLOUDS II (CM)	1	1	84.35 ± 0.22	$\times 46$
GPU	CALOCLOUDS	100	64	24.91 ± 0.72	$\times 157$
	CALOCLOUDS II	25	64	6.12 ± 0.13	$\times 640$
	CALOCLOUDS II (CM)	1	64	2.09 ± 0.13	$\times 1873$

Point Cloud + Diffusion Model

Work in Progress, Adding Two-Angle Conditioning





Summary

BiB-AE Family

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)

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, et al., 2023 MLST in press **DOI 10.1088/2632-2153/acefa9**, **arXiv: 2303.18150**

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

High Fidelity / **High speed** / **Challenging to Scale** / **Challenging to Integrate**

CaloClouds Family

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2305.04847**

CaloClouds II: Ultra-Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2309.05704**

Fair Fidelity / **High speed** / **Easy to Scale** / **Layer Geometry Independent** / **Straightforward to Integrate**

Flows Family

Convolutional L2LFlows: Generating Accurate Showers in Highly Granular Calorimeters Using Convolutional Normalizing Flows, Buss et al., **coming soon on arXiv…**

L2LFlows: Generating High-Fidelity 3D Calorimeter Images, Diefenbacher et al., **arXiv:2302.11594**

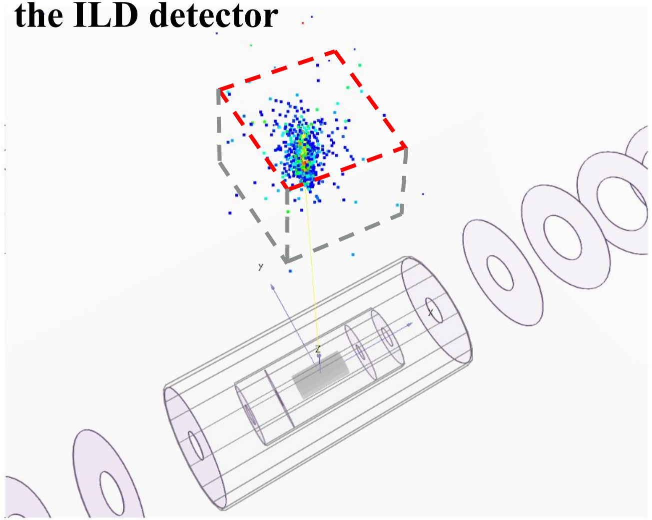
Ultimate Fidelity / **Fair speed** / **Challenging to Scale** / **Layer Geometry Independent** / **Straightforward to Integrate**

BACKUP SLIDES

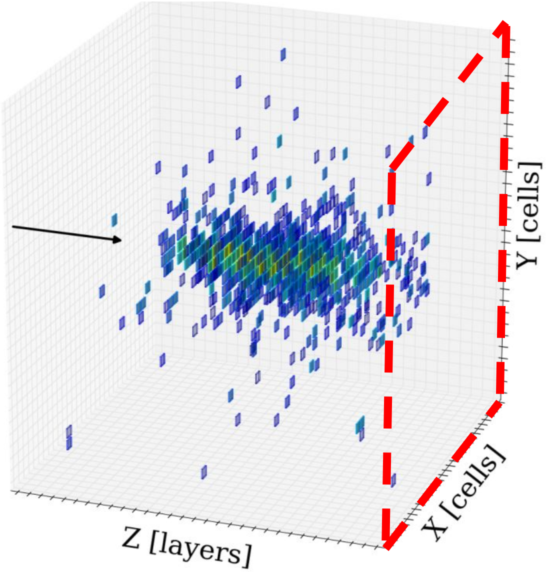
Image Representation of the EM Showers

ILD Detector

A simulated 60 GeV photon shower in the ILD detector



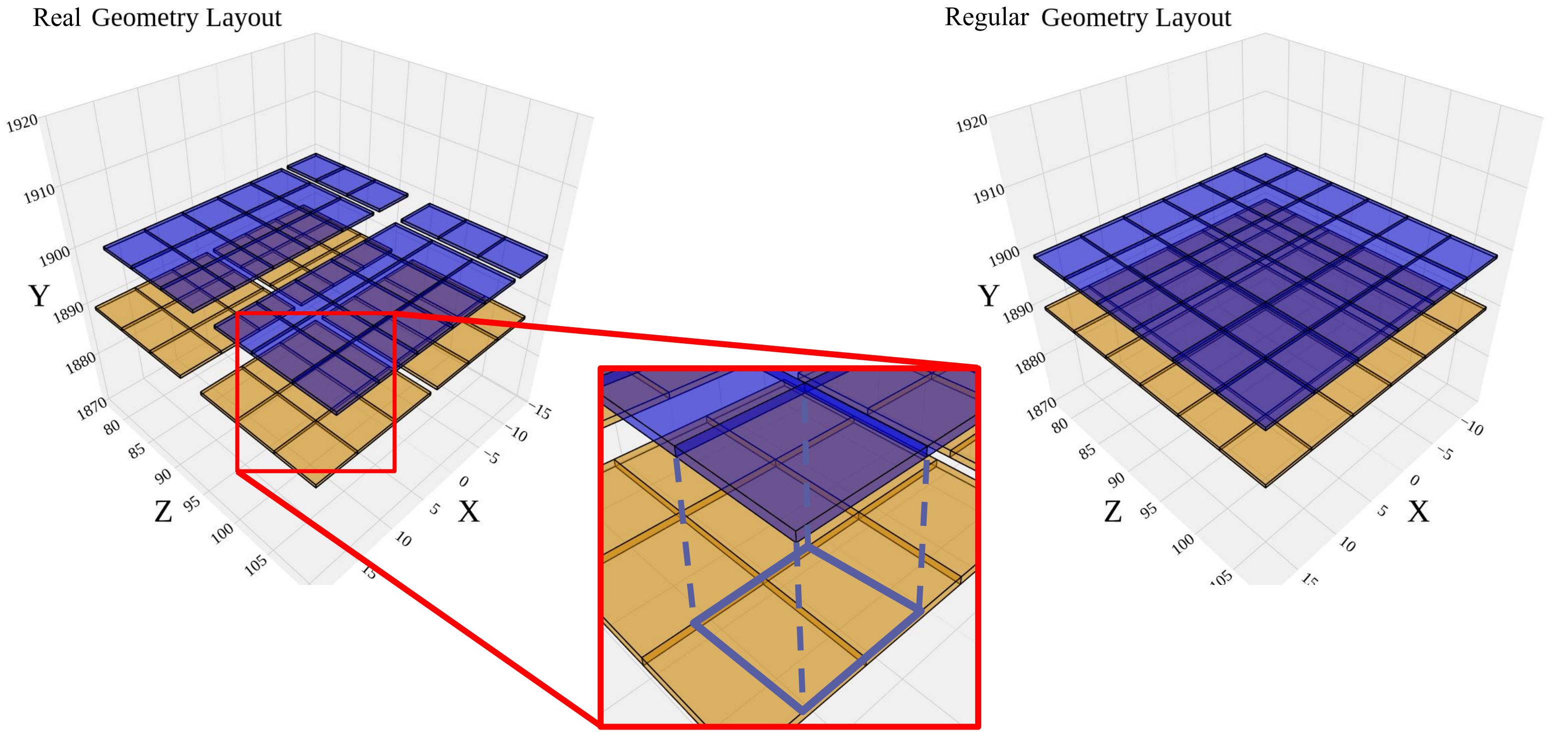
Regular grid 30x30x30



One to one mapping from detector geometry to a regular grid

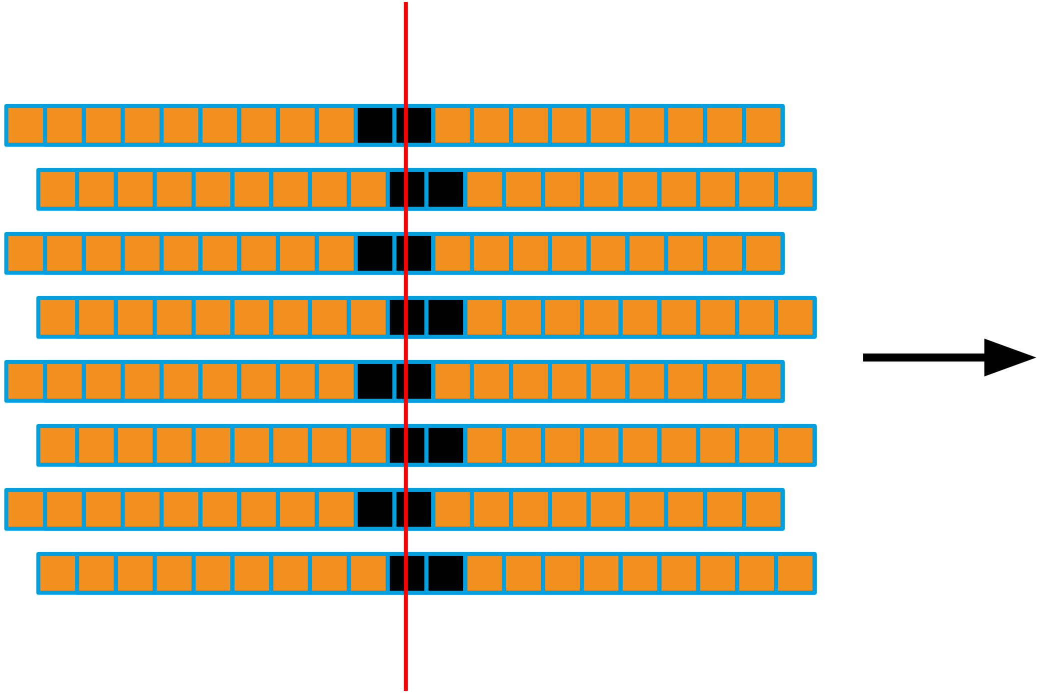
Problems with Image Representation of the EM Showers

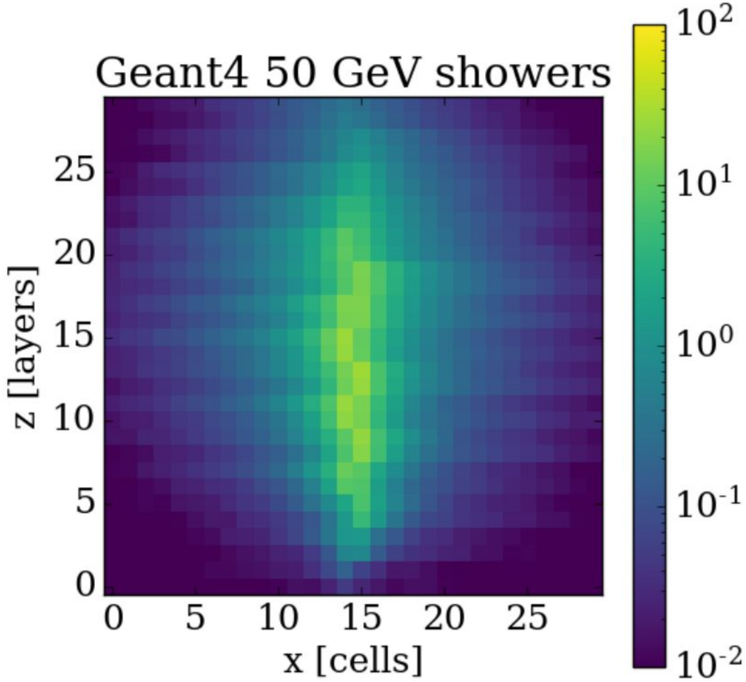
ILD Detector, ECAL Layers Structure



Problems with Image Representation of the EM Showers

ILD Detector, ECAL Layers Structure, Staggering Effect

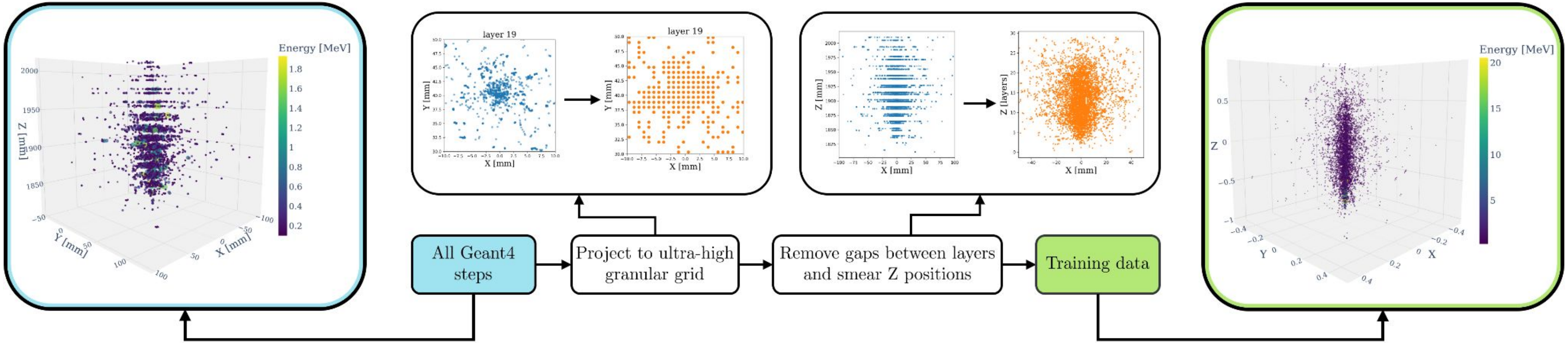




Models have to learn not only EM shower properties, but also geometry “artifacts”, like **staggering effect**

Point Cloud Representation of the EM Showers

Data Preprocessing



Number of points reduced to ~6k per shower, high enough resolution to move the shower in different place without harming physical properties of the shower

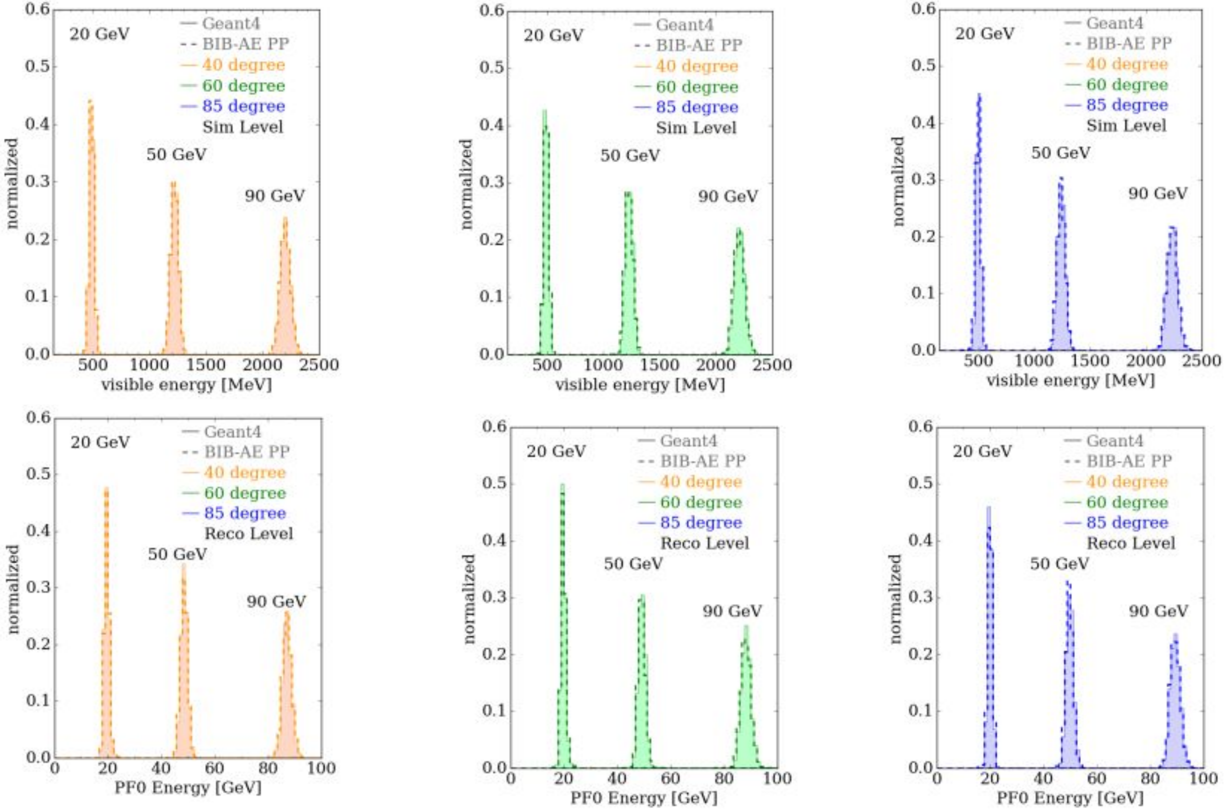
New Angles

Energy Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, et al. 2023 MLST in press
DOI 10.1088/2632-2153/acefa9, **arXiv: 2303.18150**

- **Sim** level visible energy

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



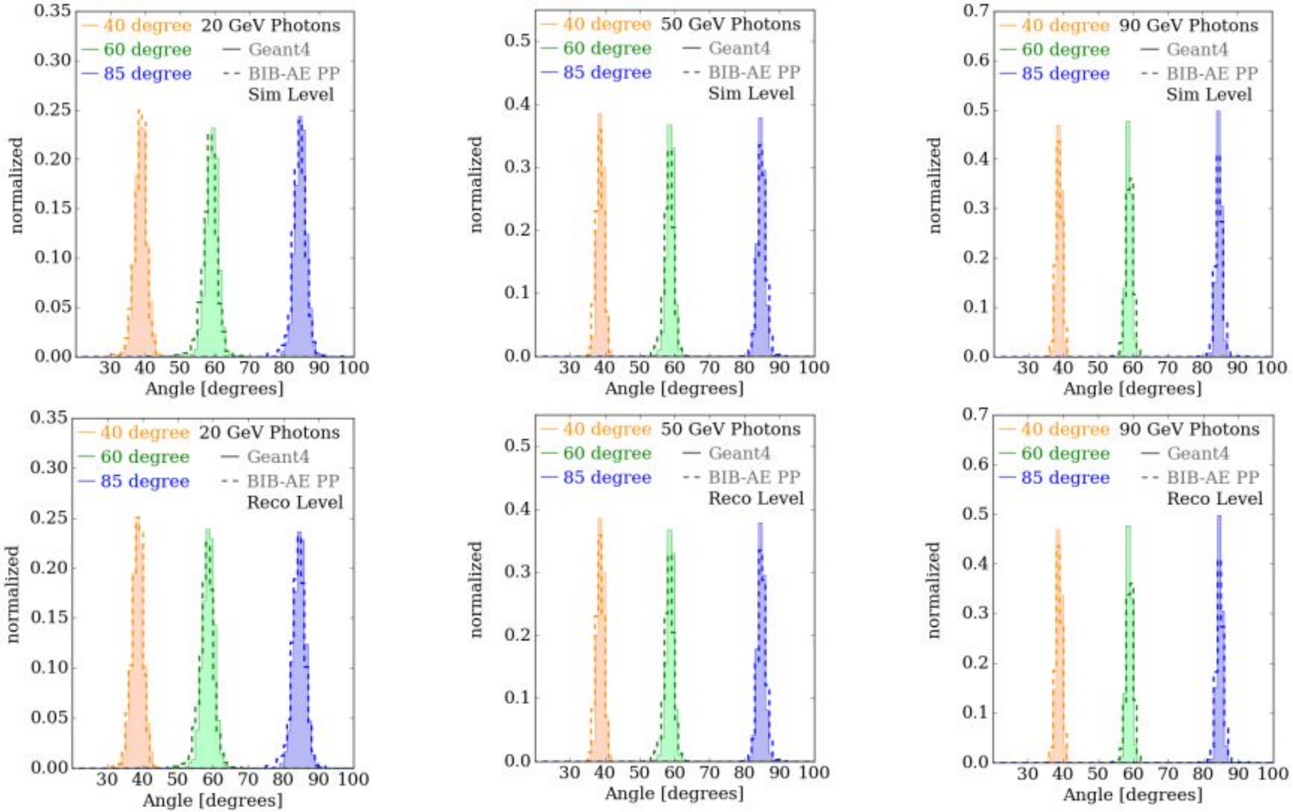
New Angles

Angular Conditioning Performance

New Angles on Fast Calorimeter Shower Simulation, Diefenbacher, et al. 2023 MLST in press
DOI 10.1088/2632-2153/acefa9, **arXiv: 2303.18150**

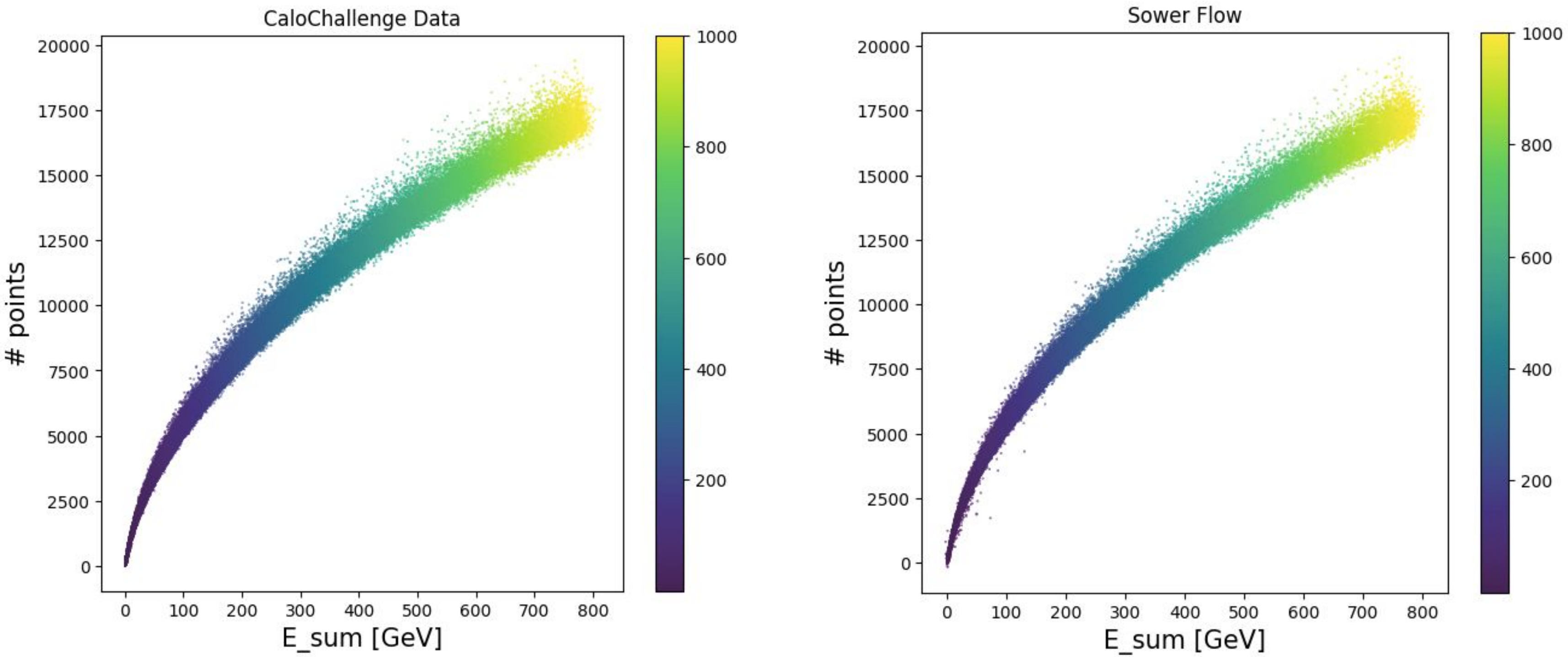
- **Sim** level angle reconstruction

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

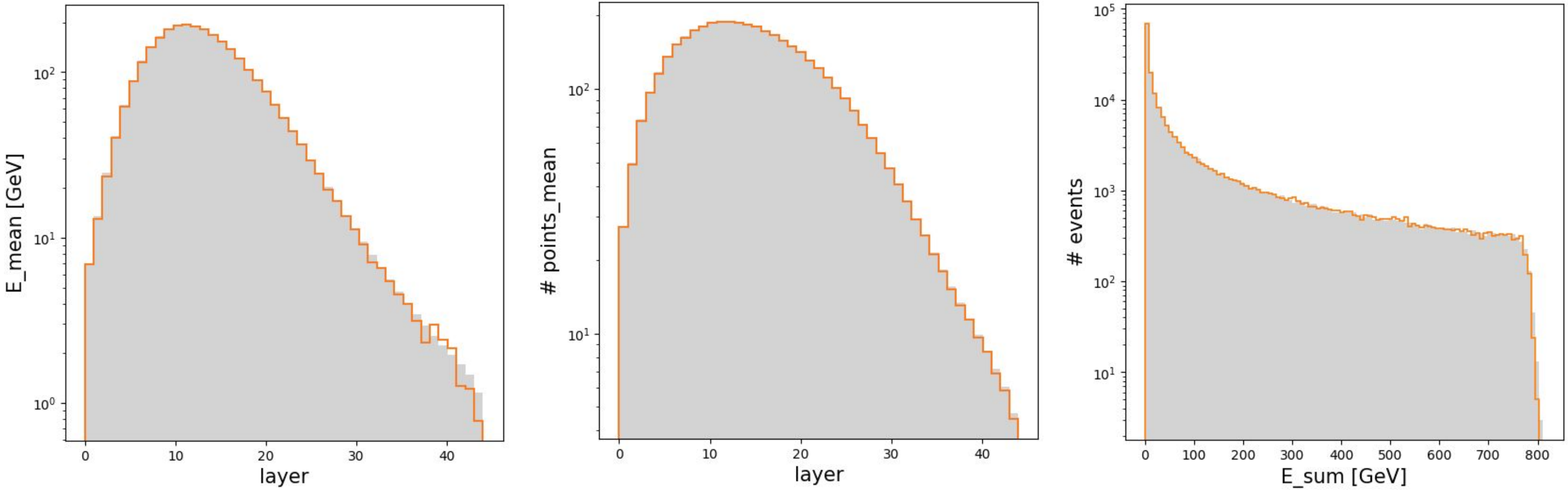


Shower Flow

Results

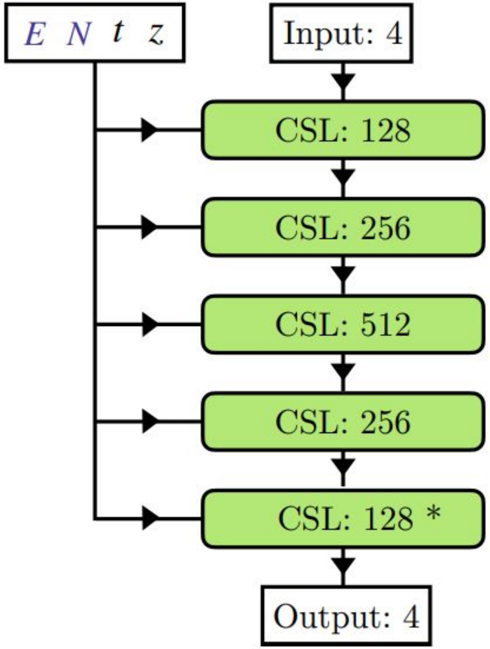


Shower Flow
Results

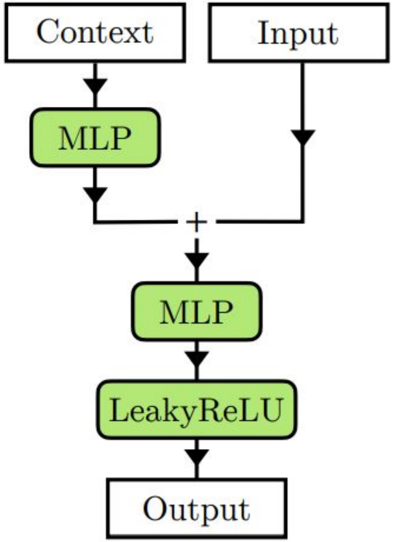


Point Cloud + Diffusion Model

PointWise Net



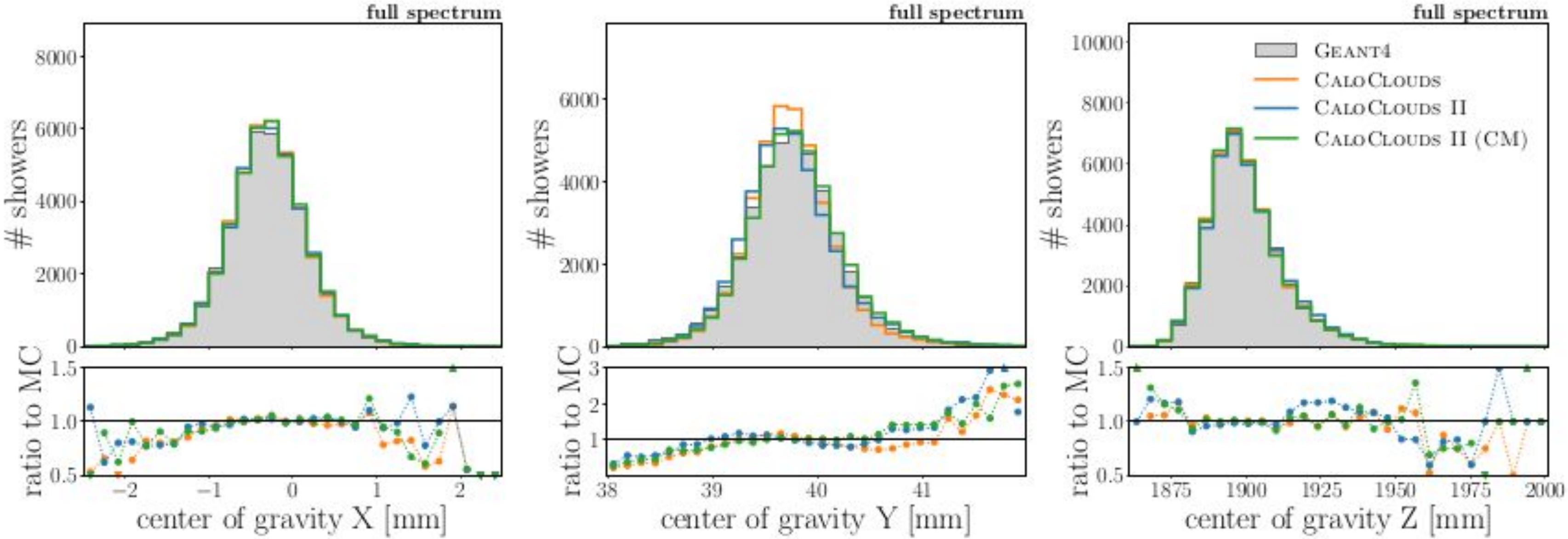
(a) PointWise Net



(b) ConcatSquash Layer (CSL)

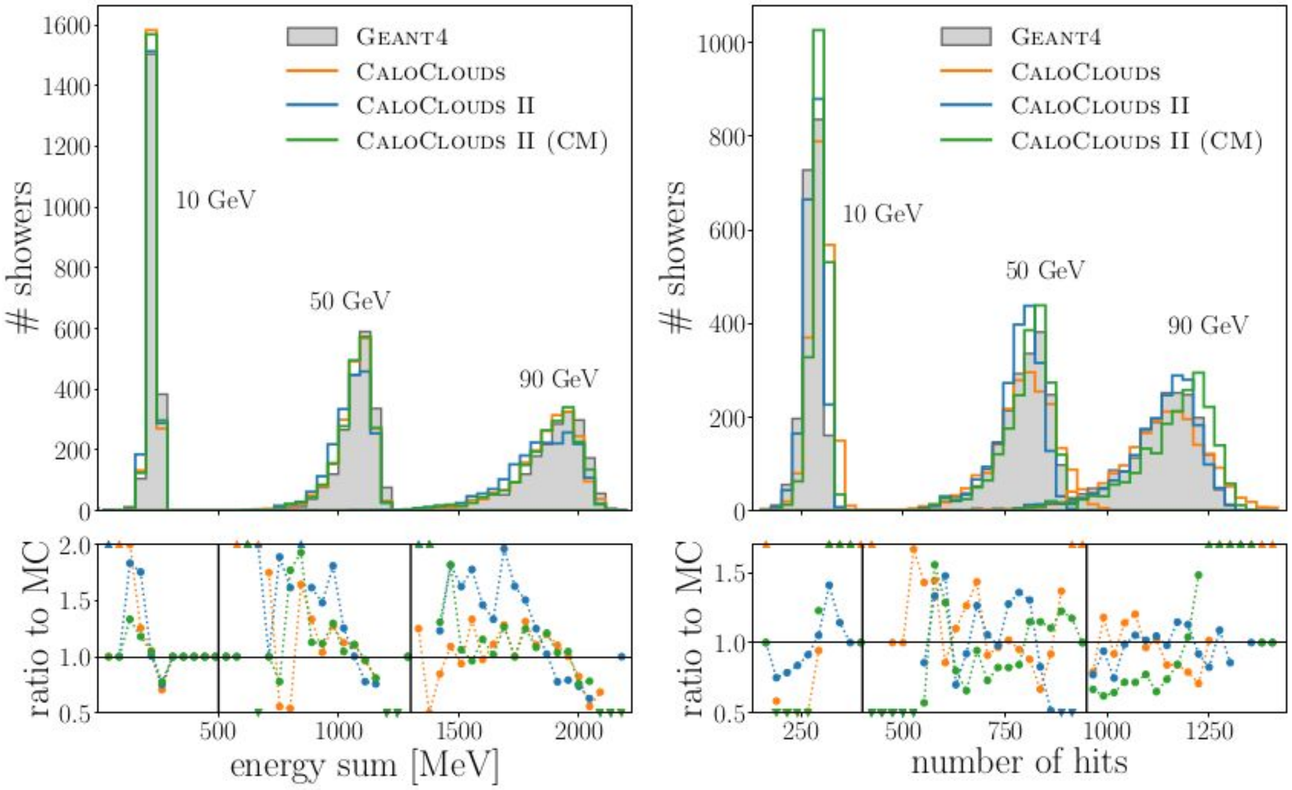
Point Cloud + Diffusion Model

Results, Position of the Center of Gravity



Point Cloud + Diffusion Model

Results, Visible Energy and the Number of Hits

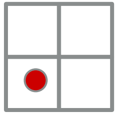


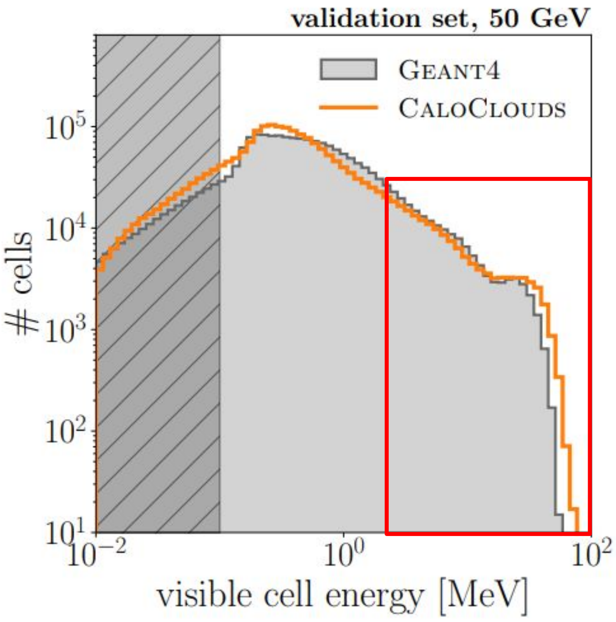
Point Cloud + Diffusion Model

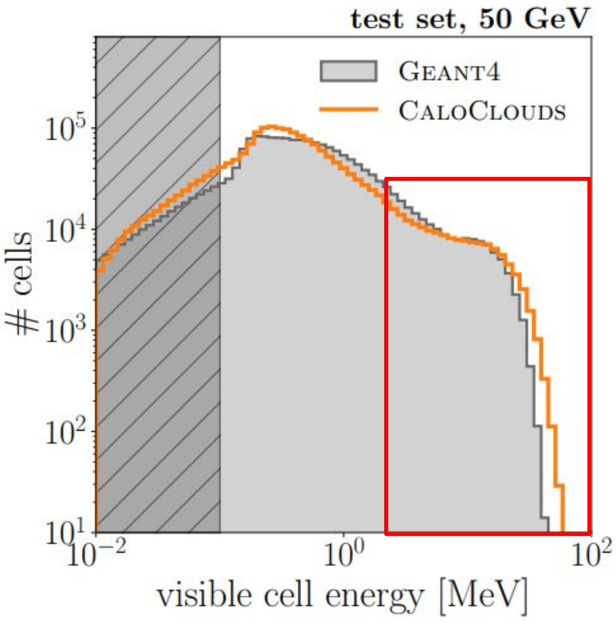
Results

CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation, Buhmann, et al. 2023, **arXiv:2305.04847**

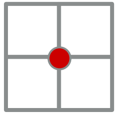
validation incident point pos.







test incident point pos.

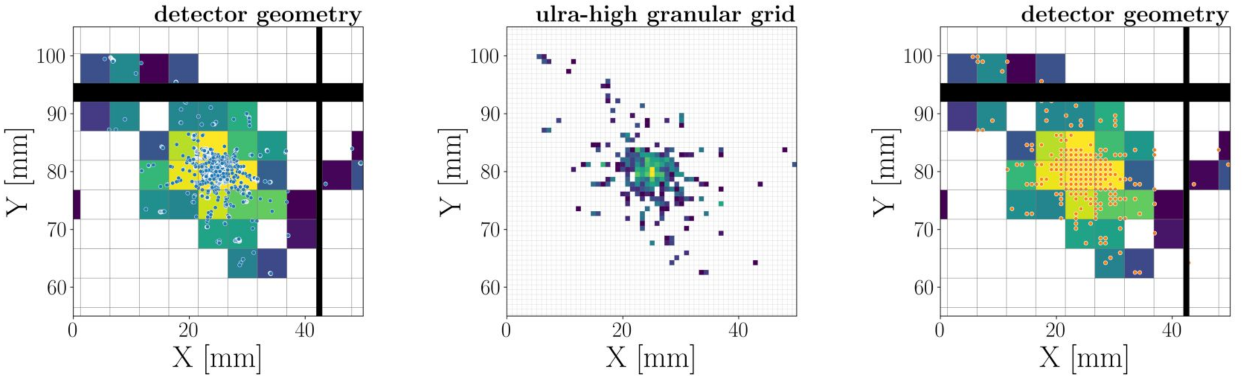


Per-cell energy distribution for the 50 GeV validation (left) data set, created at the same position as the training data set and for a 50 GeV test (right) data set simulated at a different position with the generated point cloud translated to this position

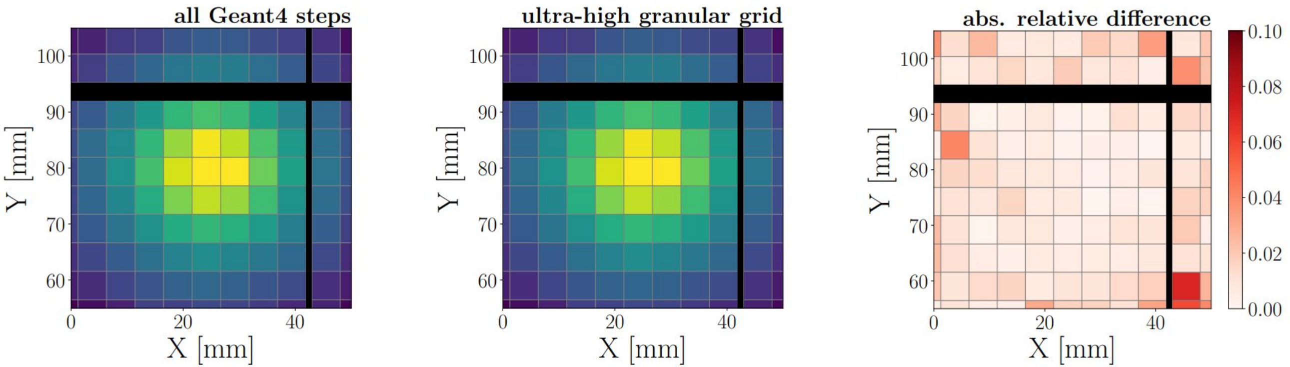
Point Cloud Representation of the EM Showers

Effects of the Pre-Clustering

Single event of 90 GeV shower in 21th layer



2k events of 90 GeV showers in 21th layer, overlay



Point Cloud Representation of the EM Showers

Effects of the Pre-Clustering