AIDAinnova - 12.3 FastSim Activities at DESY

Annual Meeting 2023

Valencia

F. Gaede, DESY
24.04.2023

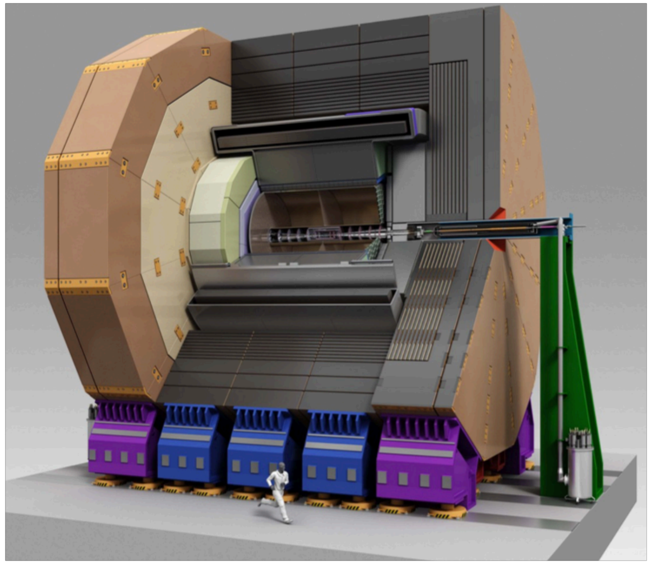
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

AIDA innova

FTX DESY.

Calo-ML Working Group at DESY

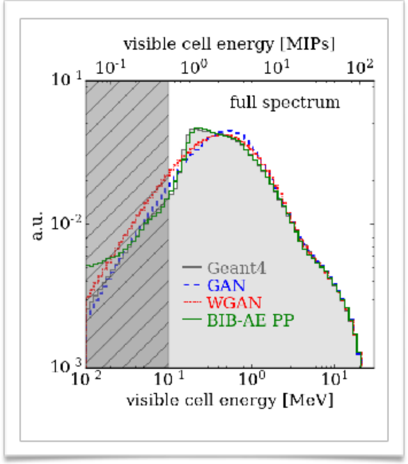
- goal: study the suitability of using **generative networks** for the simulation of **calorimeter showers** in **highly granular** calorimeters with **high fidelity**
  - ideally such that they **can replace G4 in full simulation for physics analyses** for Future Higgs Factories
  - work carried out 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, 3 PhD students
  - joined with **UHH** (G.Kasieczka) in **Quantum Universe** cluster of excellence
  - **AIDAinnova**, ACCLAIM (Helmholtz Innovation Pool), CDCS,…

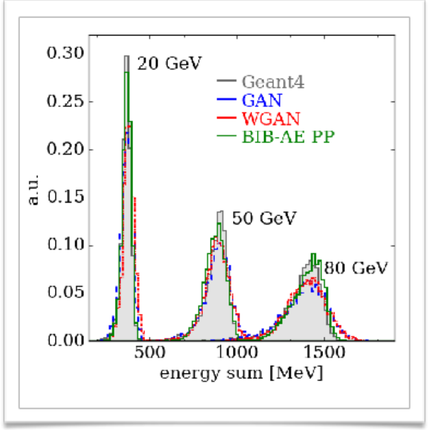


Previous Work

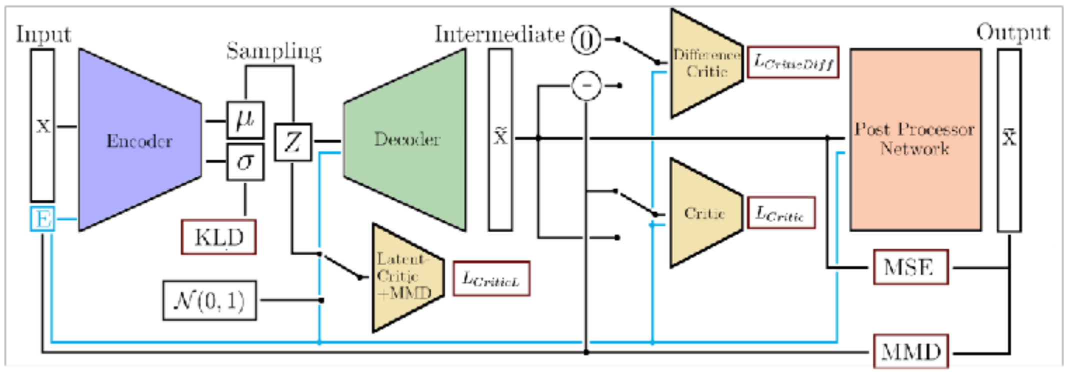
High fidelity simulation of Photons

- started with photons at **90 deg impact** angle in ILD Ecal (5x mm^2, SiW) w/ uniform energies 10-100 GeV
    - achieved **high fidelity** in distributions of relevant physical variables
    - using **Bounded-Information-Bottleneck Auto Encoder** (**BIB-AE**) w/ post-processing
- extend approach to hadrons, different angles, reco, …



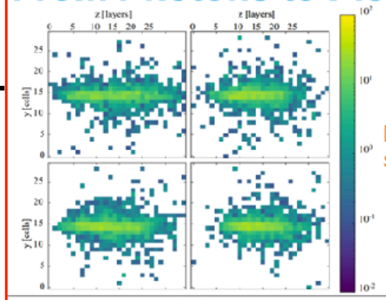


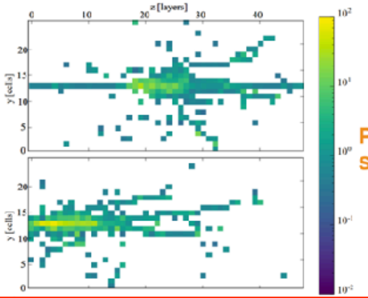
Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed, E. Buhmann et al. e-print: [2005.05334](2005.05334), published in Comput Softw Big Sci 5, 13 (2021)



Bounded-Information-Bottleneck Auto Encoder with Post Processing

From Photons to Pions



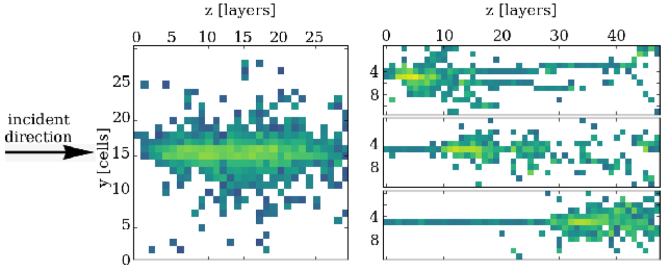


		Energy	Angles θ, ϕ	ECAL +HCAL	Reco
Photon showers	• Predominantly governed by EM interactions • Compact structure ↓ Relatively easy to generalise	✓	+	N/A	+
Pion showers	• Hadronic and EM interactions • Complex structure • Large event-to-event fluctuations ↓ Hard to learn	+	X	X	+

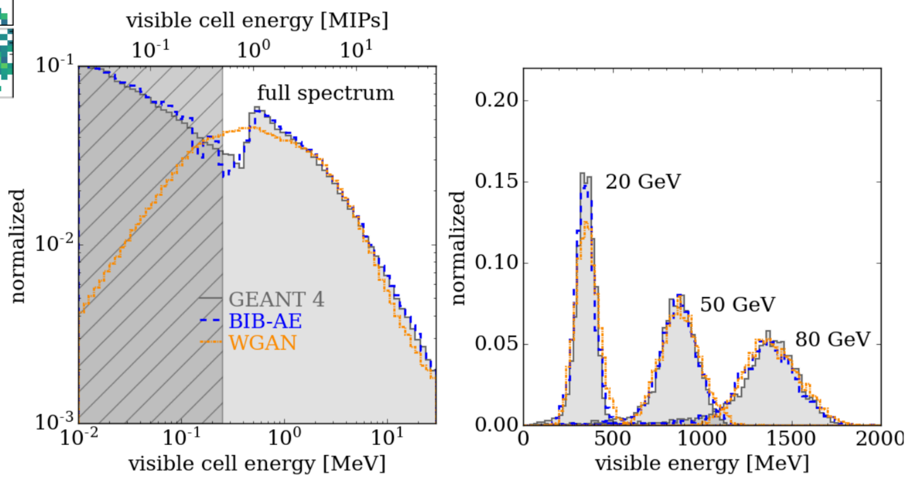
✓ = Achieved X = Yet to be done + = Addressed here

Previous Work

Hadronic Shower Simulation

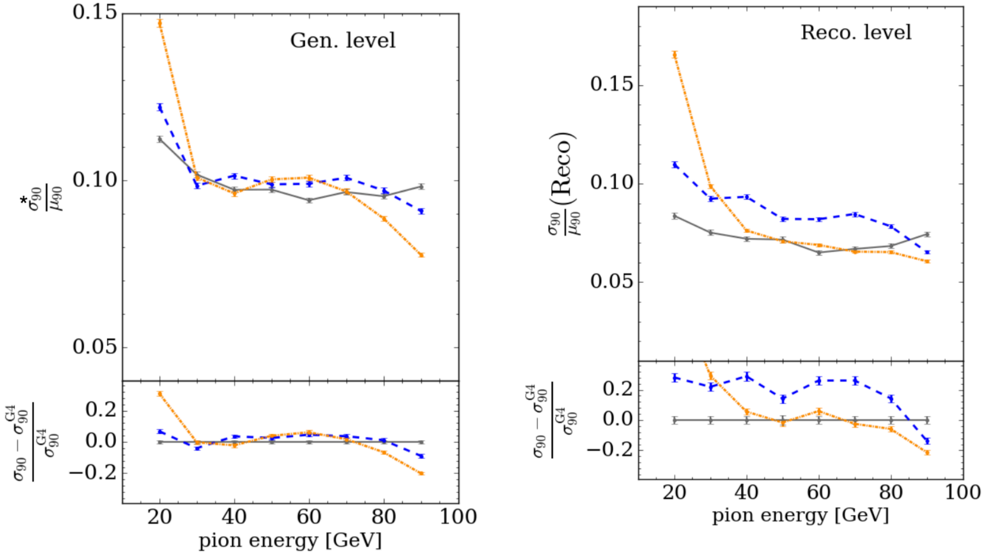


- use charged pion sample in ILD SciFe AHcal
- 500k showers [10-100GeV] @ 90 deg
  - **no Ecal in front**
  - projected into 25x25x48 grid
- investigate **WGAN** and **BIB-AE**
- achieve ~high fidelity for pion showers in highly granular calorimeter
- applied full PandoraPFA **reconstruction** to single showers
- eventually the important benchmark for ML based fast sim….



Very good agreement of MIP peak for **BIB-AE** with Post-Processing!

Great agreement with Geant4

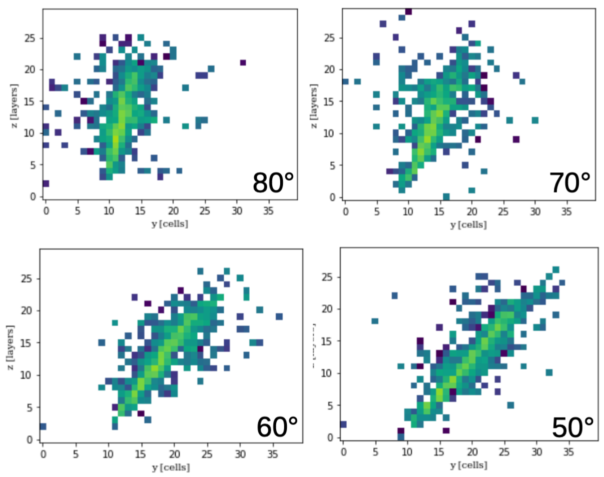


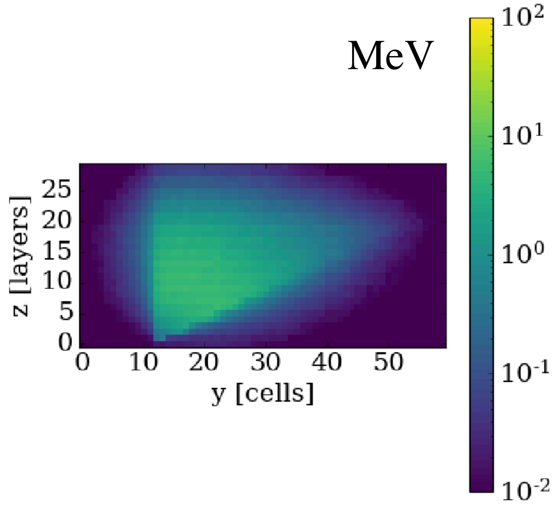
Hadrons, Better, Faster, Stronger
E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede et al,
Published in: Mach.Learn.Sci.Tech. 3 (2022) 2, 025014,
arXiv:2112.09709

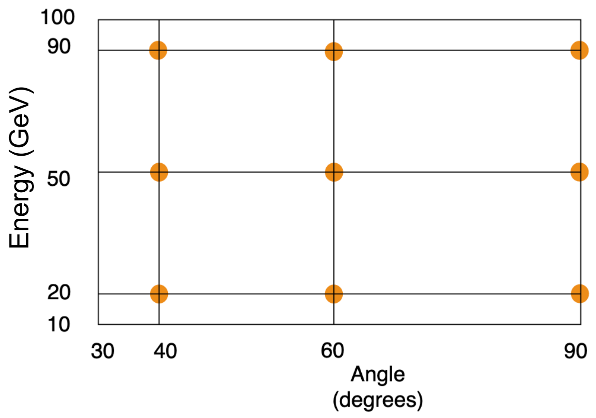
Energy and Angular Conditioning

P.McKeown

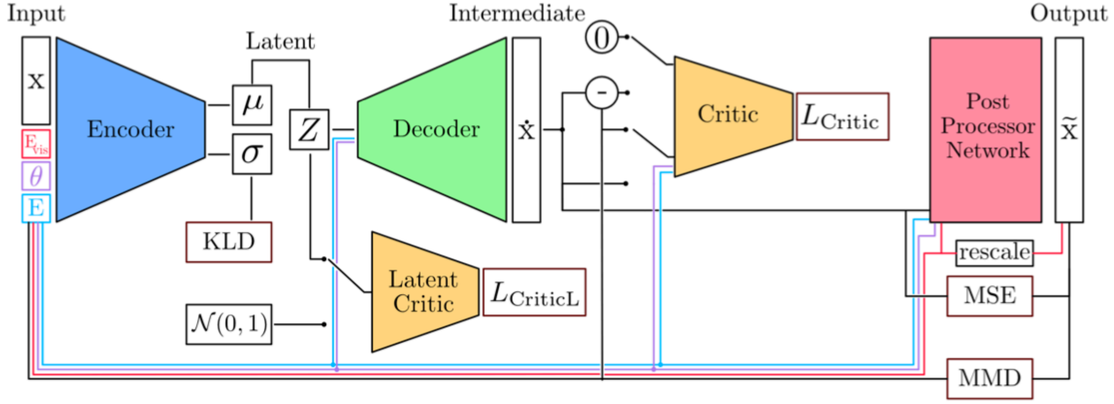
- using photon sample in ILD SiW Ecal
  - **angular and energy conditioning**
  - large training sample, uniform in [10-100GeV, 30-90deg]
- train a **BIB-AE** w/140 k showers (500 k used for PP)
  - significant GPU resources needed for training







training and validation samples - 30x60x30 grid

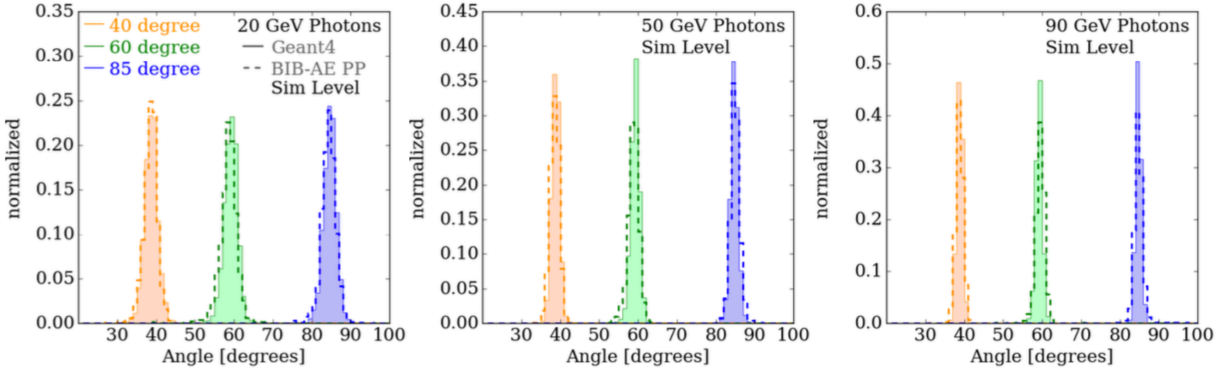


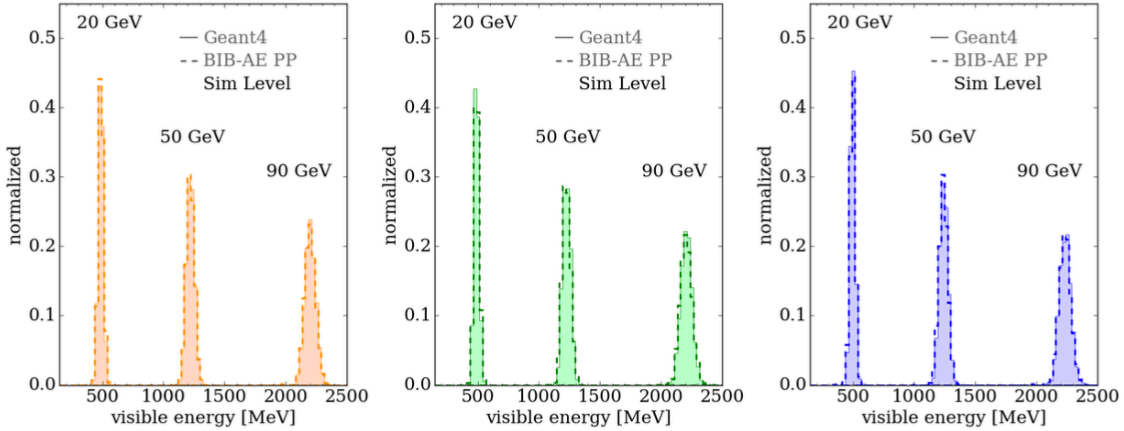
New Angles on Fast Calorimeter Shower Simulation, S.Diefenbacher, P.McKeown et al arXiv: 2303.18150, submitted to MLST

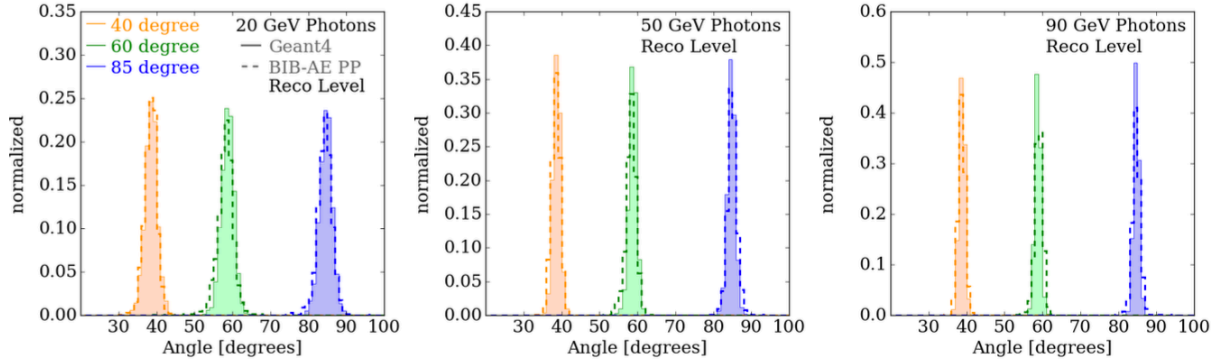
Energy and Angular Conditioning

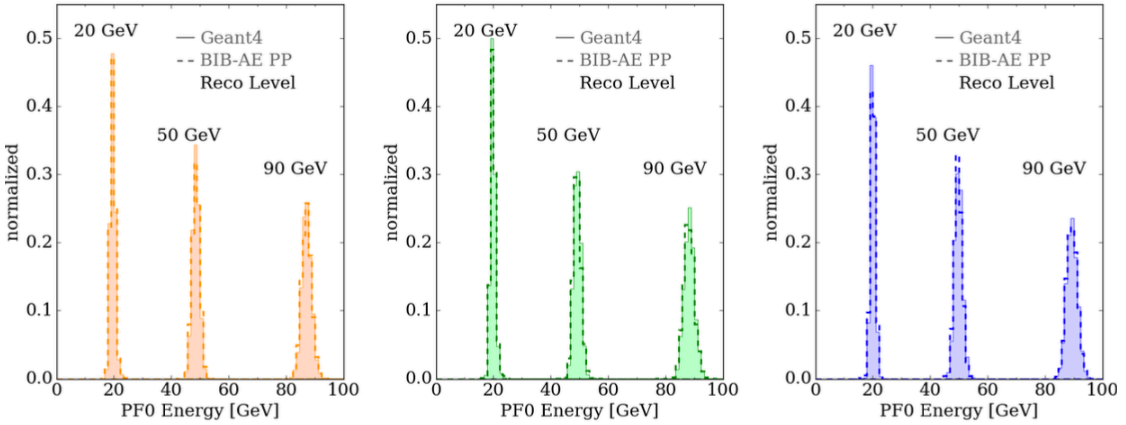
P.McKeown

after simulation/generation





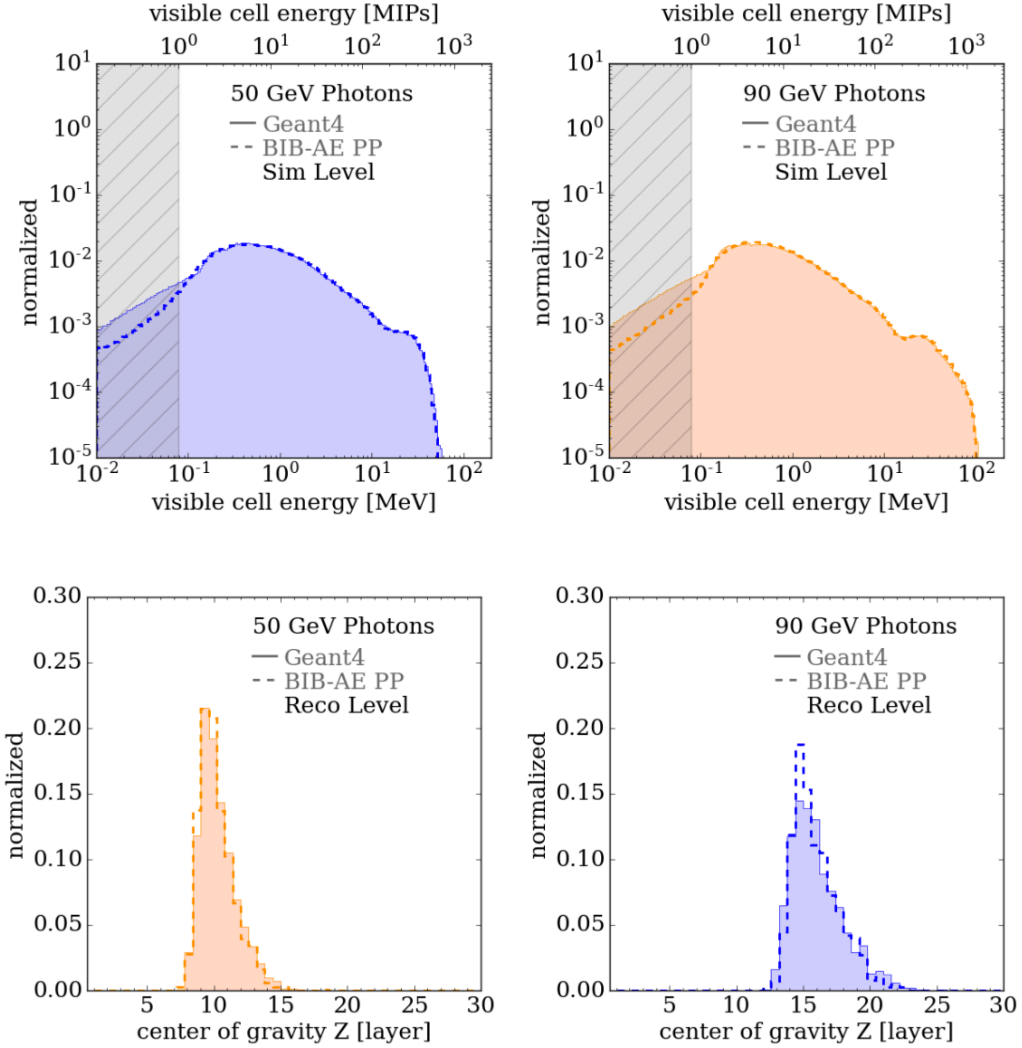


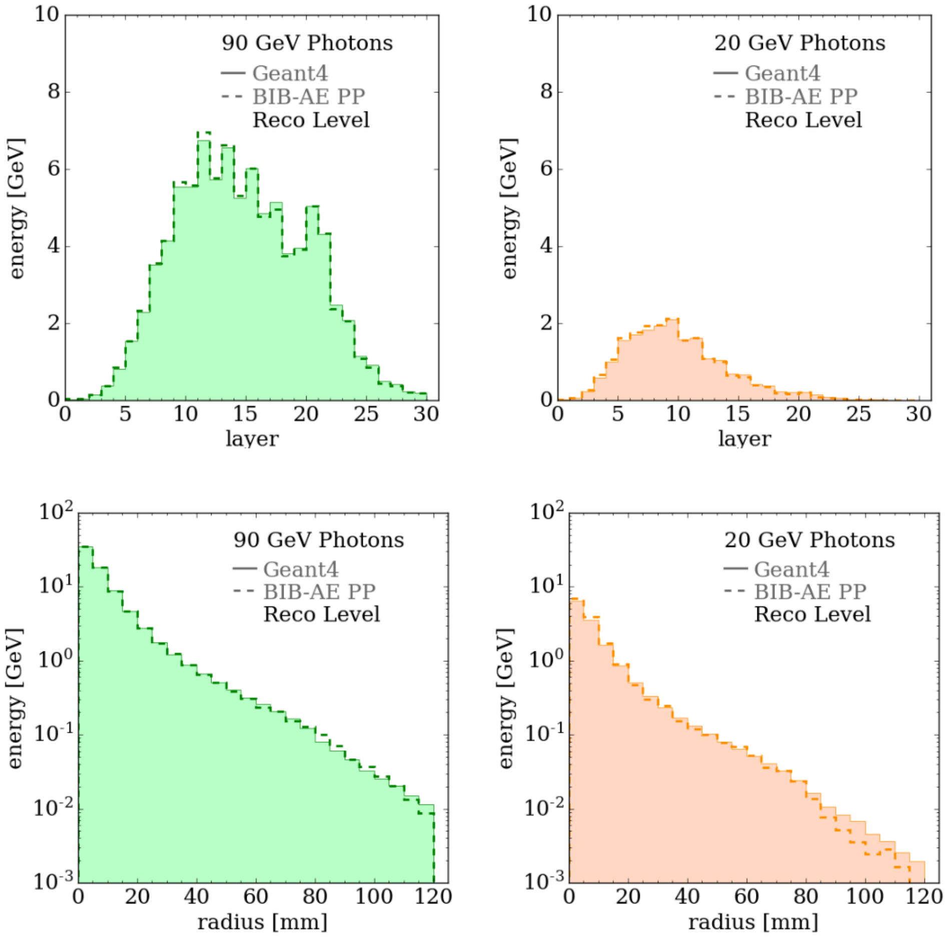


after full reconstruction w/ PandoraPFA

Energy and Angular Conditioning

P.McKeown



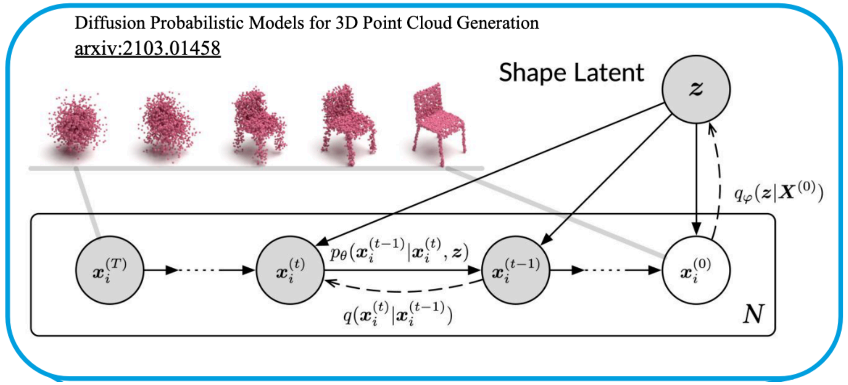


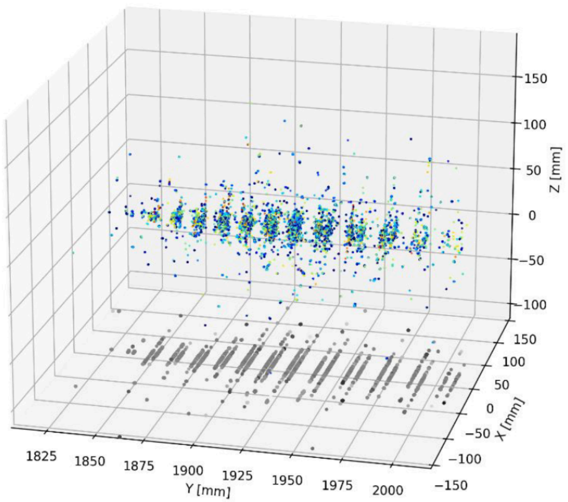
- some example distributions - best (left) and worst (right) in validation data set
- achieve excellent physics fidelity throughout

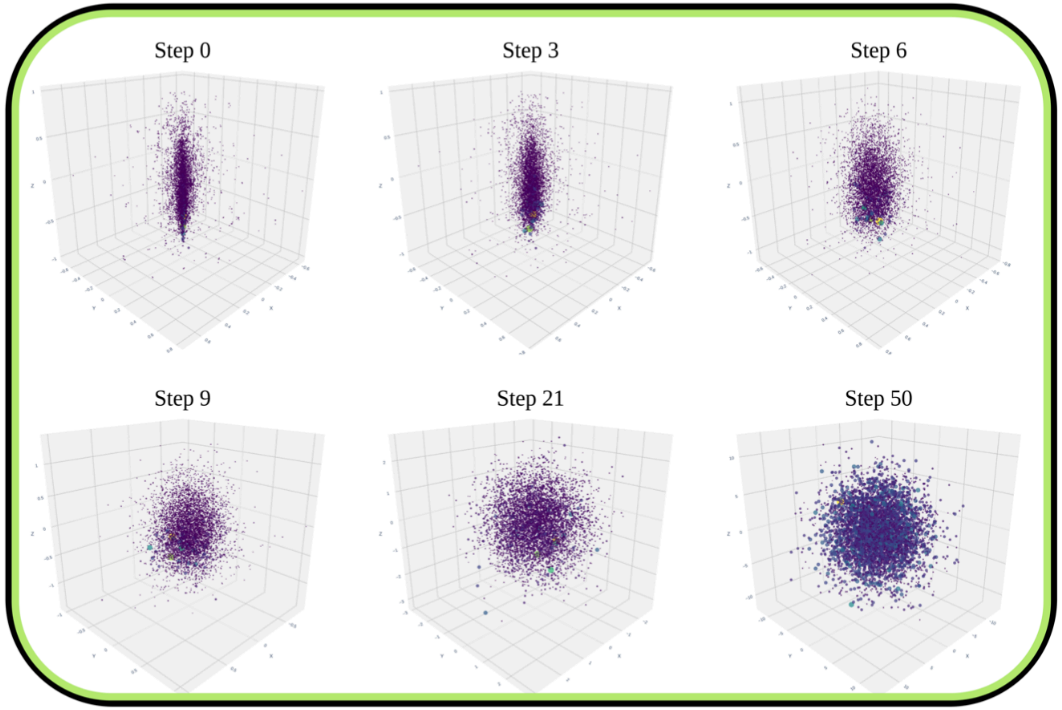
Point Cloud Diffusion

A.Korol

- recently good success with using **point clouds** and **diffusion models**
  - using much higher granularity from Geant4 simulation to address issue with irregular cell geometries in 'real' calorimeters



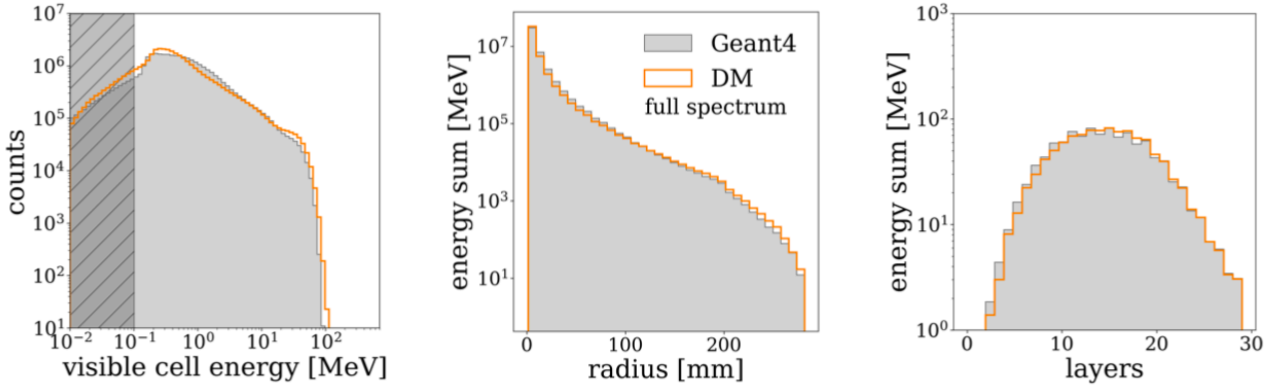


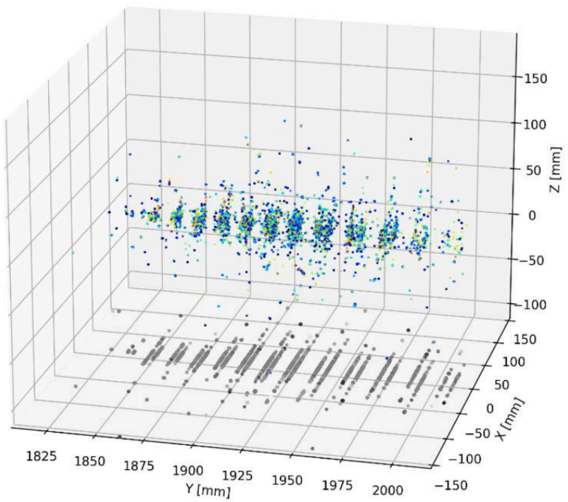


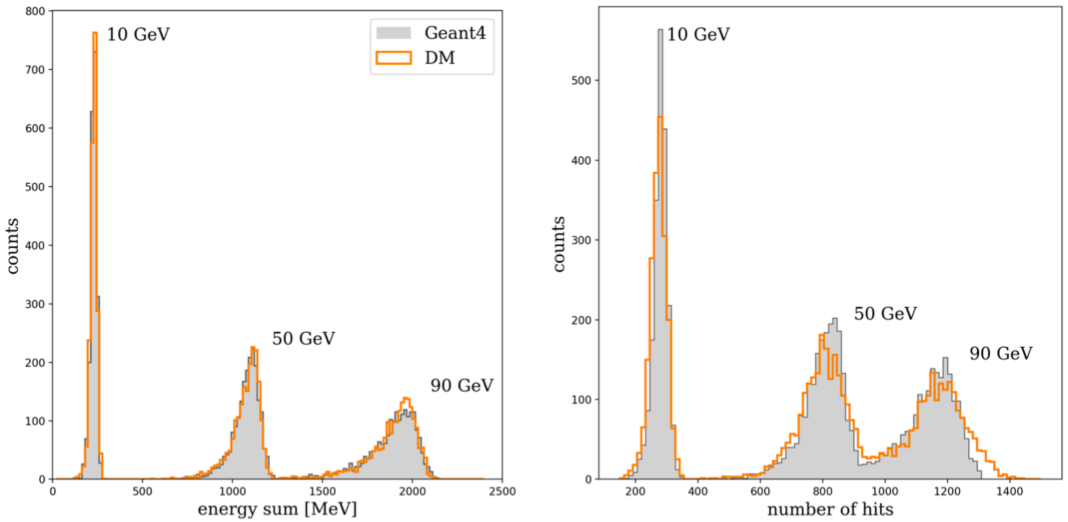
Point Cloud Diffusion

A.Korol

- recently good success with using **point clouds** and **diffusion models**
  - using much higher granularity from Geant4 simulation to address issue with irregular cell geometries in ‘real’ calorimeters
- achieve promising first results
  - publication in preparation …



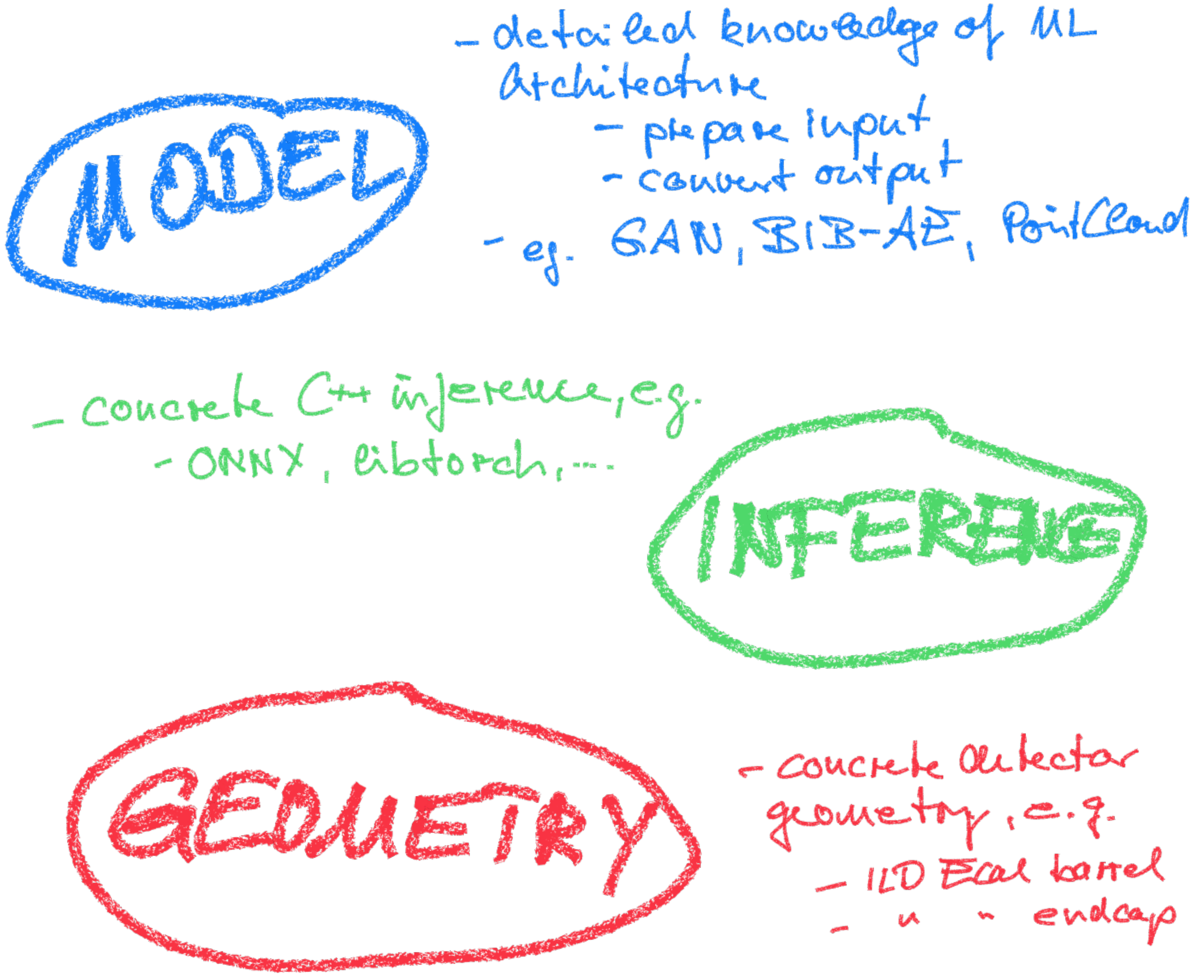




ML inference in DDG4

running ML fast sim in ‘real’ simulations

- prototype library for running fast ML inference in DDG4/Geant4 based full simulation: *DDFastShowerML*:
  - https://gitlab.desy.de/ilcsoft/ddfastshowerml
- goal to have an easy to use library that can easily be extended to all types of ML architectures for all detectors in DD4hep
- will use some time of the **hackathon discussion** to iterate this idea and see if this is the best way to integrate ML in DDG4/Geant4



ML inference in DDG4

running ML fast sim in ‘real’ simulations

- prototype library for running fast ML inference in DDG4/Geant4 based full simulation: *DDFastShowerML*:
  - https://gitlab.desy.de/ilcsoft/ddfastshowerml
- goal to have an easy to use library that can easily be extended to all types of ML architectures for all detectors in DD4hep
- will use some time of the **hackathon discussion** to iterate this idea and see if this is the best way to integrate ML in DDG4/Geant4

```
/// User callback to model the particle/energy shower - details defined in ML_MODEL
virtual void modelShower(const G4FastTrack& track, G4FastStep& step) override {

  // remove particle from further processing by G4
  step.KillPrimaryTrack();
  step.SetPrimaryTrackPathLength(0.0);
  G4double energy = track.GetPrimaryTrack()->GetKineticEnergy();
  step.SetTotalEnergyDeposited(energy);


  _input.clear() ;
  _output.clear() ;
  for( auto& layerSPs : _spacepoints )
    layerSPs.clear() ;

  fastsimML.model.prepareInput( track, _input , _output ) ;

  fastsimML.inference.runInference(_input, _output ) ;

  fastsimML.model.convertOutput( track, _output , _spacepoints) ;

  fastsimML.geometry.localToGlobal( track, _spacepoints ) ;

  // now deposit energies in the detector using calculated global positions

  for( auto& layerSPs : _spacepoints )
    for( auto& sp : layerSPs ) {
      fastsimML.hitMaker->make( G4FastHit( G4ThreeVector(sp.X,sp.Y,sp.Z) , sp.E ), track);
    }
}
```

Summary and Outlook

- calo ML group in Hamburg (DESY/UHH) focuses on high fidelity generation of showers in highly granular calorimeters
- investigating different ML architectures and topics:
  - first **generation of pion showers** in high granular calorimeter - incl. **reconstruction**
  - **energy and angular conditioning** for EM showers with high fidelity
  - **point cloud diffusion** model for EM showers
  - started to develop *DDFastShowerML* library to run inference in DDG4

- continue work towards final goal: full calorimeter simulation w. ML (in Geant4) for ILD
  - test ML model with physics observables after reconstruction, e.g. di-boson mass or JER
  - continue work on combining ECal&HCal calorimeters for hadron simulation
  - …