

# Fast Simulation of Calorimetry showers (FastSim)

- Piyush Raikwar, Peter Mckeown, Anna Zaborowska  
EP-SFT, CERN

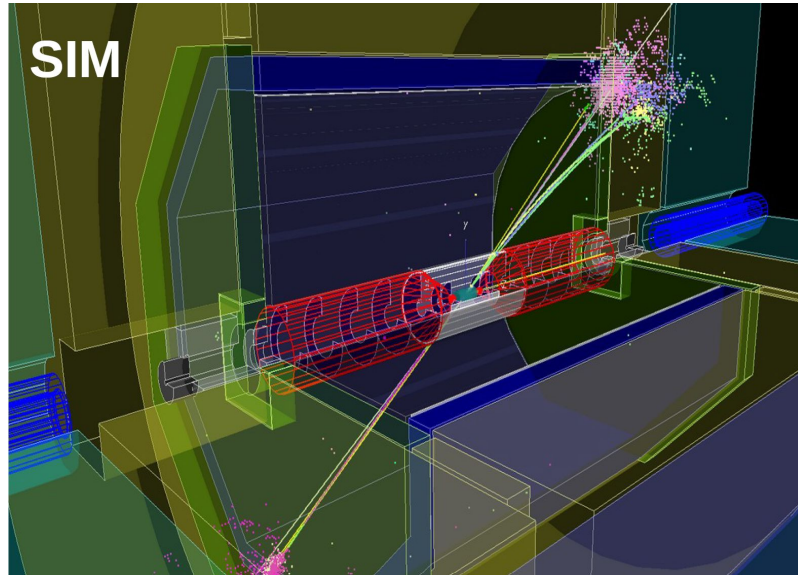
ML4EP meeting

16.05.2024

---

# Particle showers

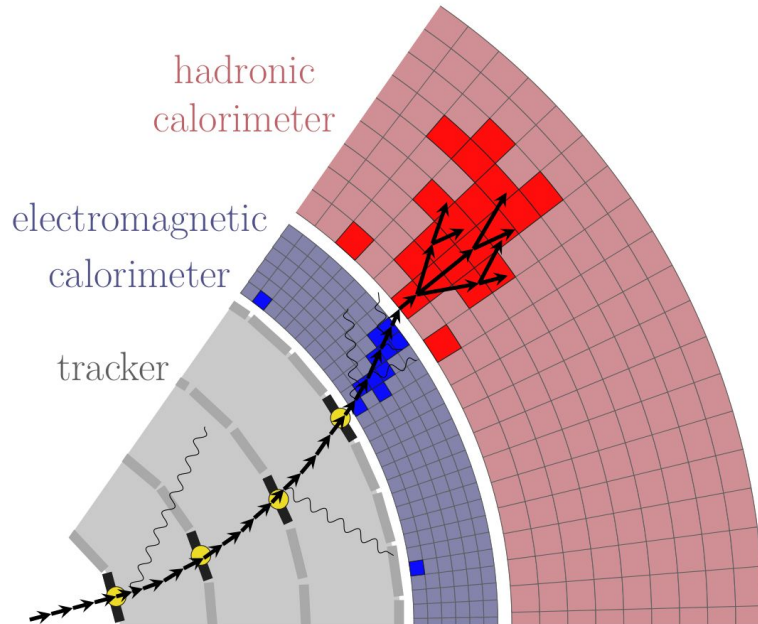
Inside different detectors at LHC



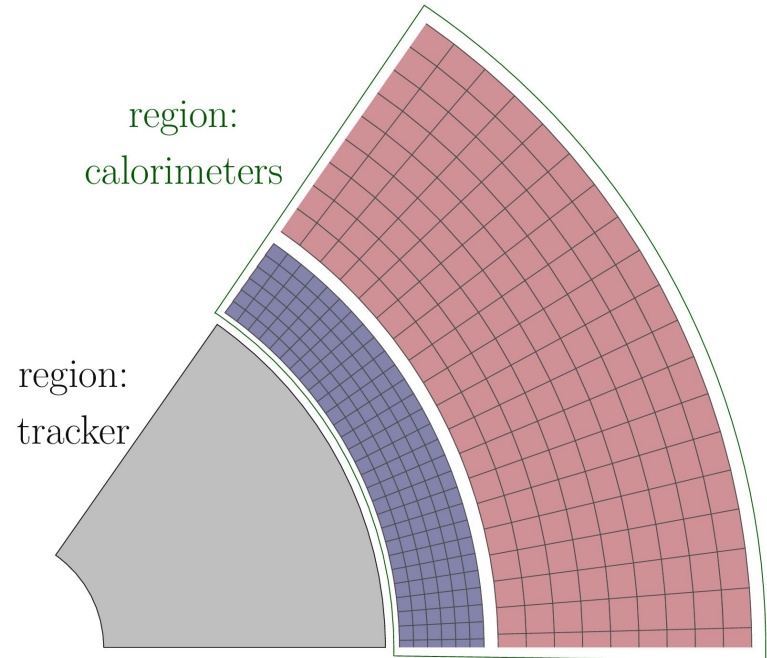
# Fast shower simulation



FullSim

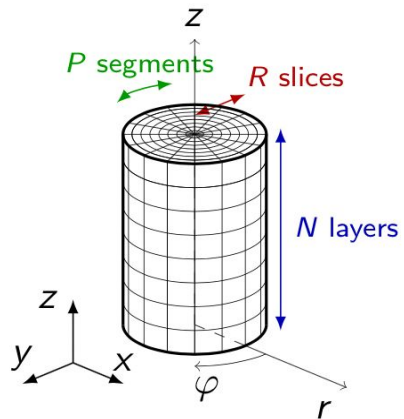
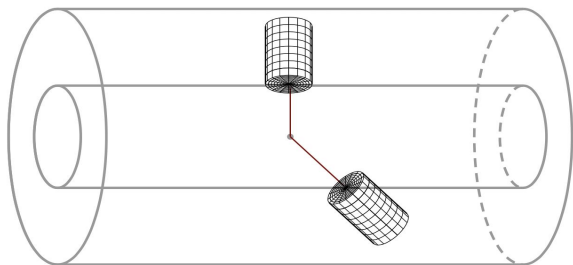


FastSim



# Energy scoring

A detector agnostic mesh is constructed to contain the largest shower. ([Par04 example](#))

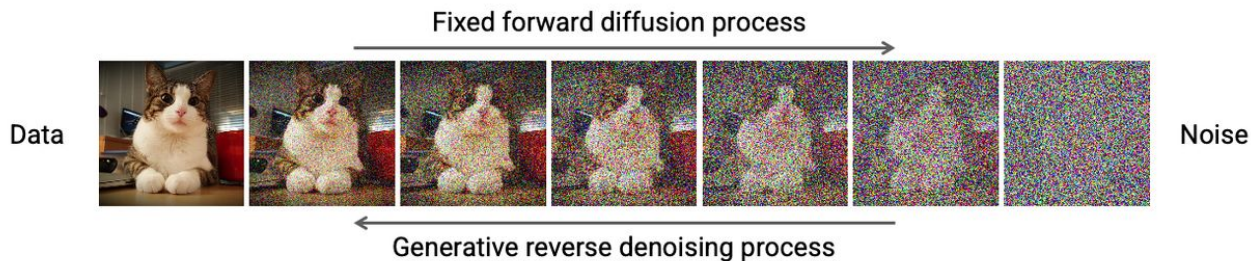
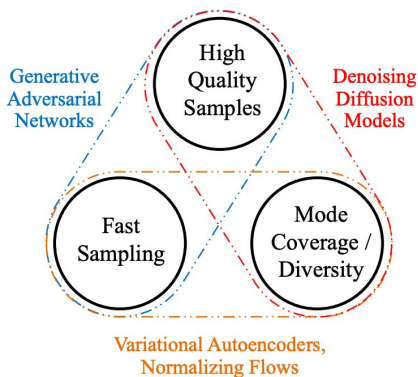


- The mesh aligns with the direction of incident particle.
  - The direction, i.e., the angles are recorded.

---

# Generative model

We use a **diffusion model** for higher accuracy and higher diversity.

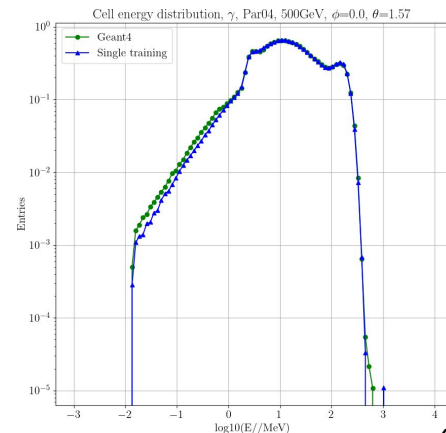
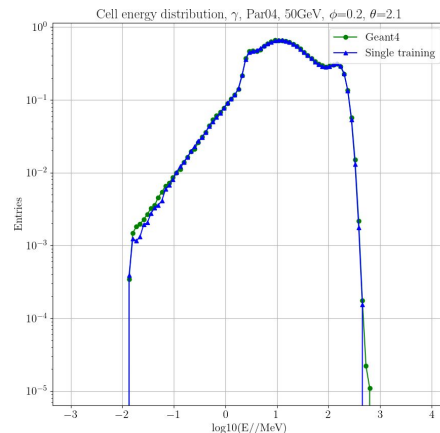
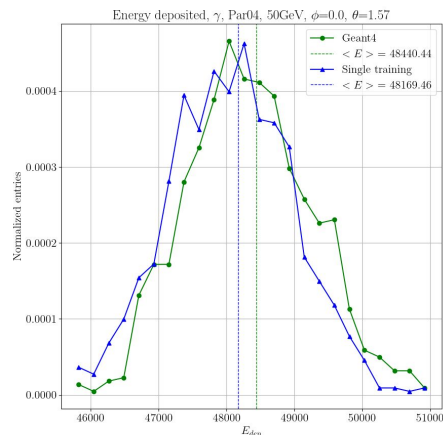
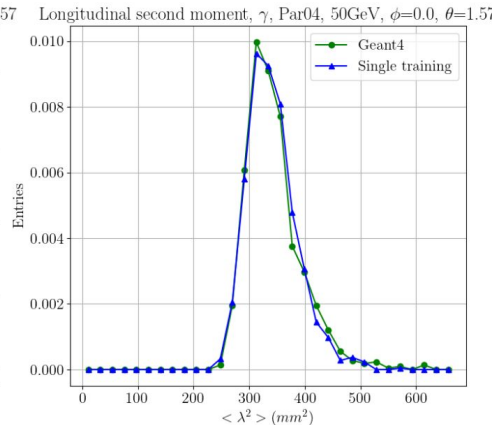
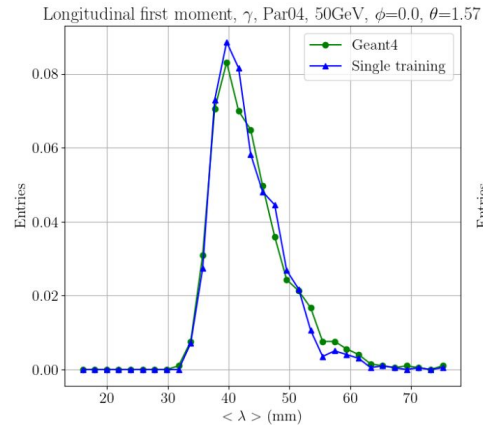
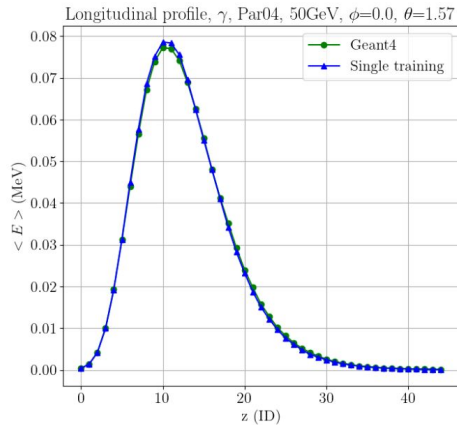


As for the architecture, we apply **transformer** blocks.

- A **generalized architecture** that works with any type of data, e.g., text, images, audio, etc.
- Models long-range dependencies (Attention mechanism).

# Results

- Cell energy distribution harder to get
- Marginal profiles are relatively easy



---

# What's different?

- Development of machine learning models for fast shower simulation is computationally expensive.
- Moreover, designing model for *each experiment* requires dedicated expertise.

**Make FastSim easily available without access to ML expertise.**

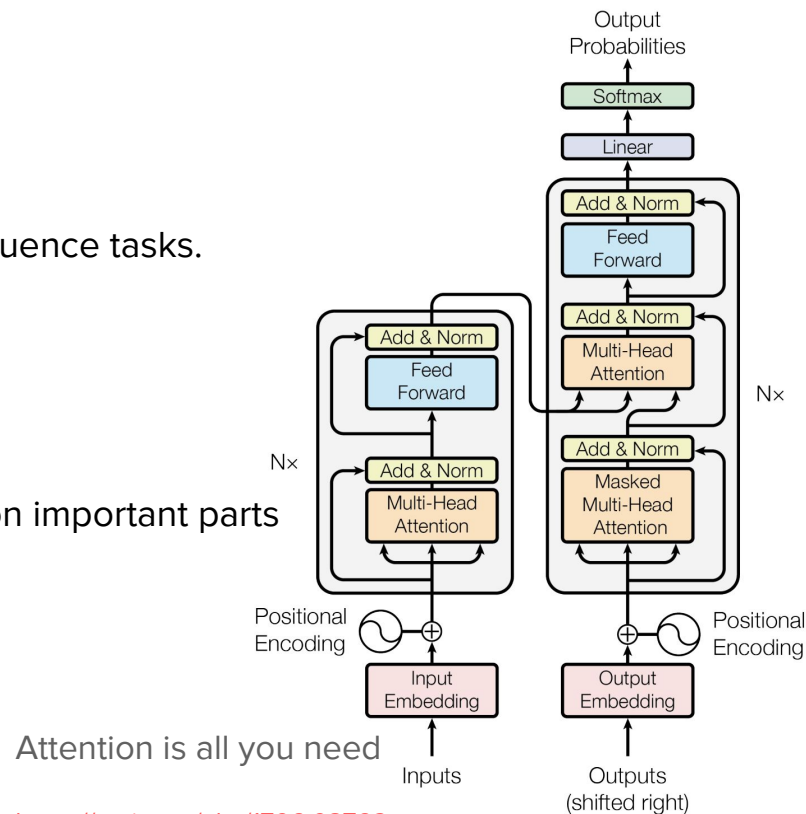
1. *Generic energy scoring mesh*
  - Collect energy irrespective of the detector geometry.
2. *Generalizable ML model*
  - **Train once** on very large & diverse datasets to learn rich representations.
  - **Then adapt** to new detectors, quickly.

**Backup**



# Transformer

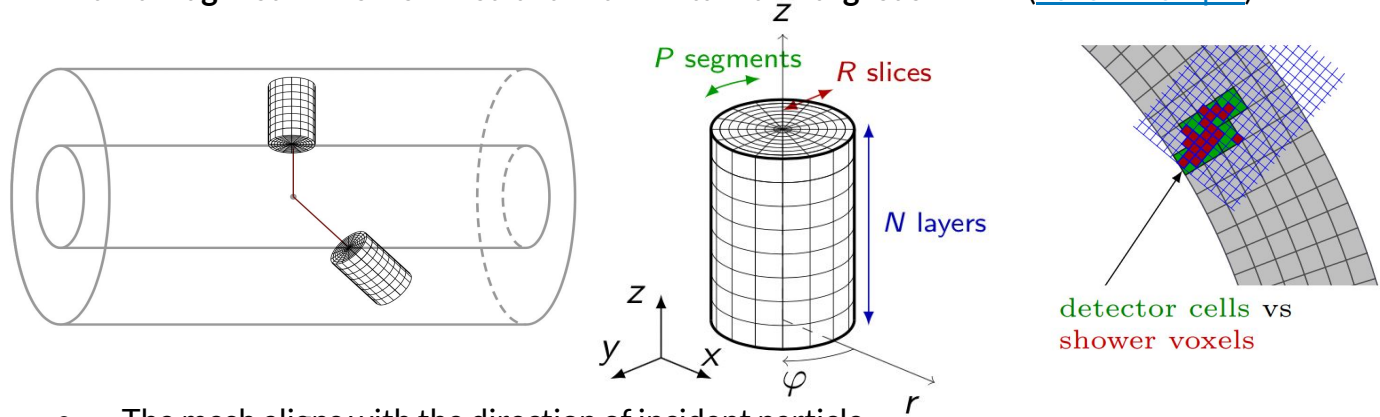
- Proposed for sequence-to-sequence tasks.
- I/O is any type of sequences.
- Encoder-Decoder blocks.
- Positional embeddings.
- **Attention:** Dynamically focus on important parts in the input.
- Multi-headed attention.



<https://arxiv.org/abs/1706.03762>

# Energy scoring

A detector agnostic mesh is constructed to contain the largest shower. ([Par04 example](#))



- The mesh aligns with the direction of incident particle.
  - The direction, i.e., the angles are recorded.
- The size of the cells can vary across detectors according to its  $X_0$  &  $R_M$ , but the number of cells remains constant<sup>1</sup>.

<sup>1</sup> i.e., for a particular model

---

# Experiments

1. Training on single geometry
  - Par04 (CaloChallenge<sup>1</sup>), simplistic cylindrical geometry
  - 1M samples
2. Joint training on multiple geometries
  - Par04 and [Open Data Detector](#) (ODD, realistic geometry)
  - 1M samples each
  - Geometry condition - one hot encoding
3. Adaptation on [FCCeeALLEGRO](#)
  - Checkpoint from 2
  - Baseline - training from scratch
  - 100K, 200K, 400K samples in each case

*Note: The results are preliminary*

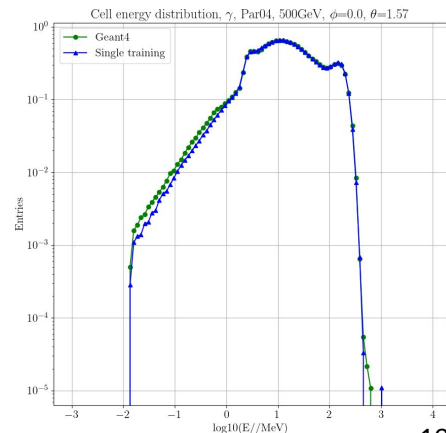
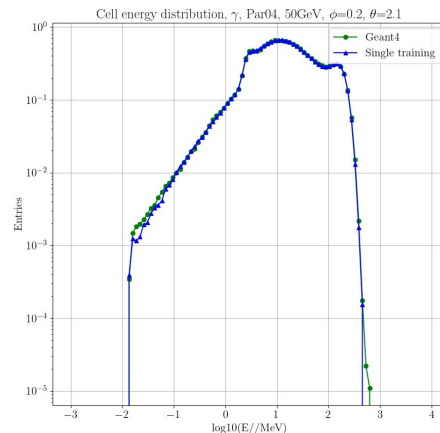
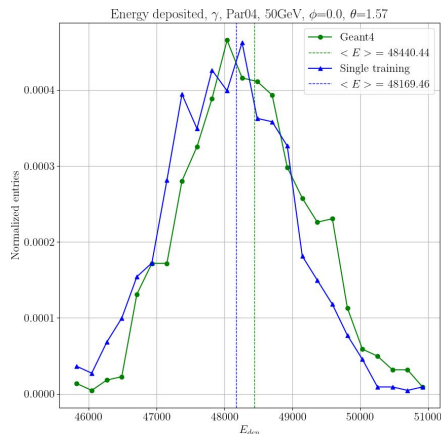
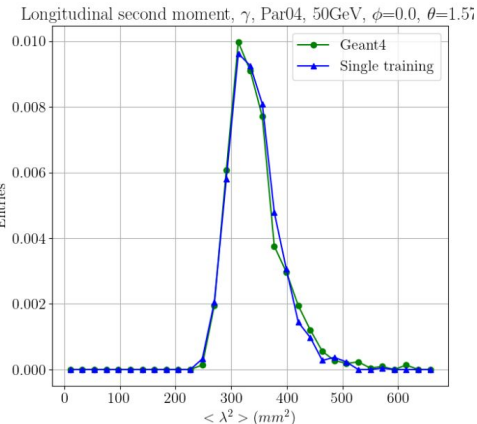
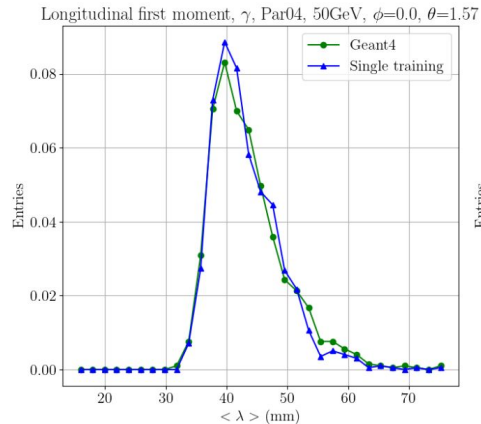
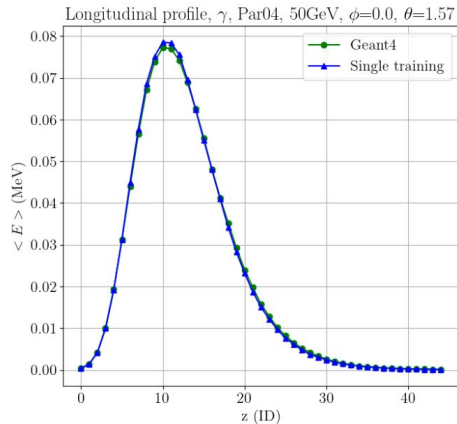
---

<sup>1</sup> More samples and more conditions (phi, theta) compared to CaloChallenge Dataset 2 11

# 1. Training on single geometry

Par04

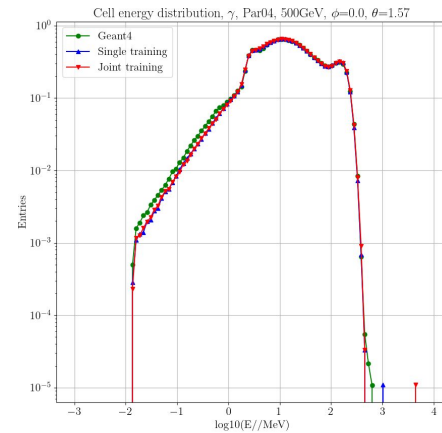
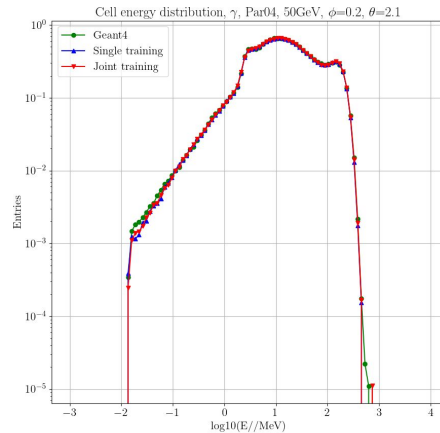
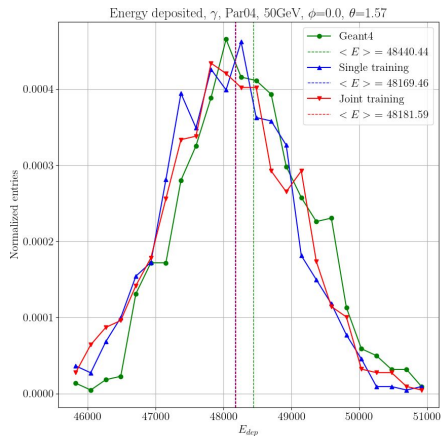
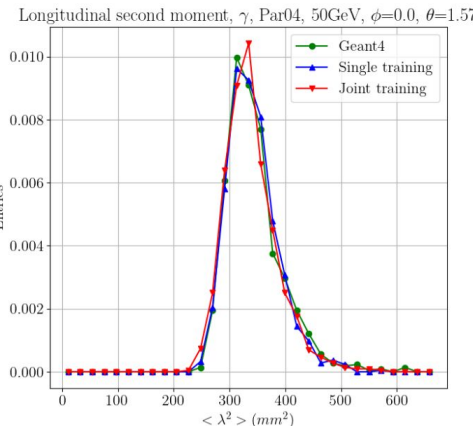
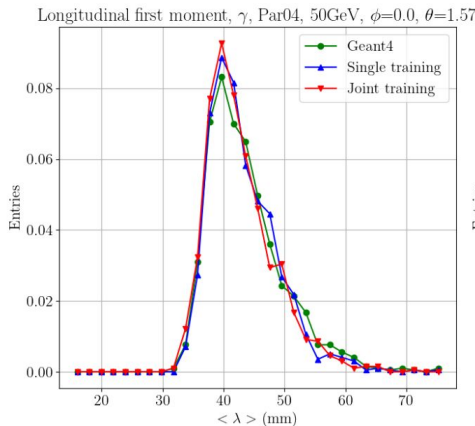
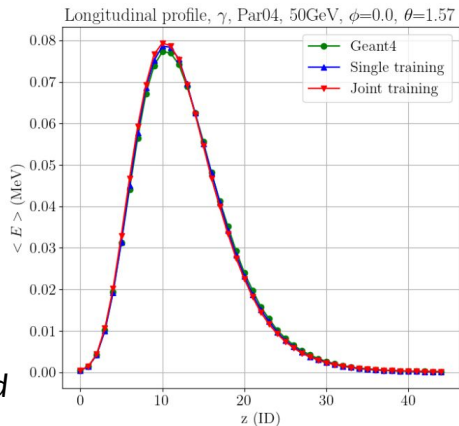
- Good accuracy with diffusion models
- Even cell energy



# 2. Joint training

Par04

- Almost no degradation after adding geometry condition
- Even if there was, not intended to use directly



# 3. Adaptation

FCCeeALLEGRO

250 epochs for training from scratch  
20 epochs for adaptation

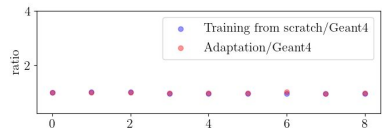
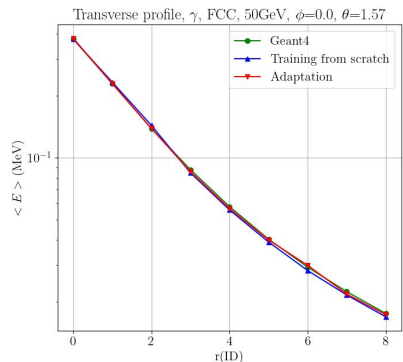
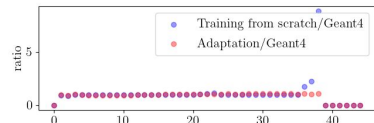
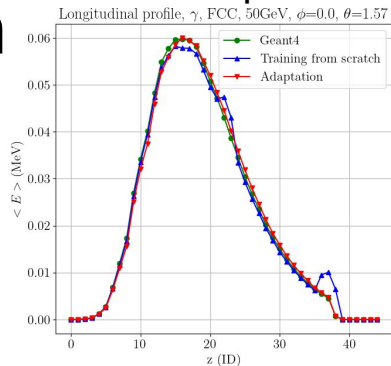
At 200K samples

~25x less training time

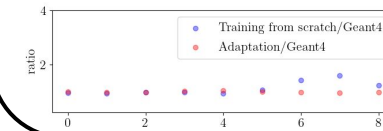
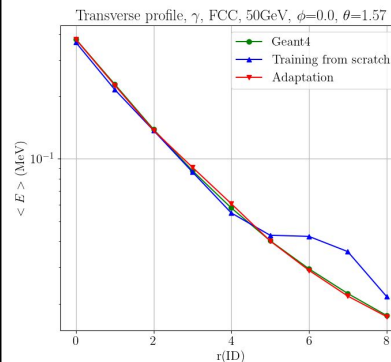
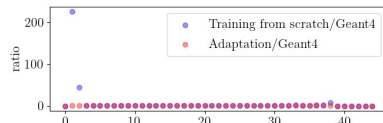
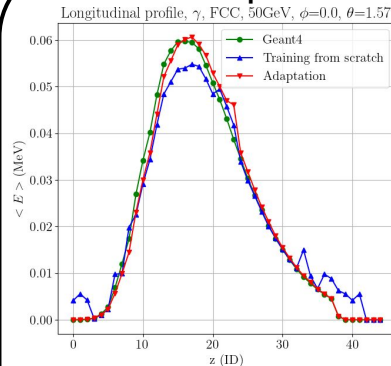
<50% of the data

Preliminary results

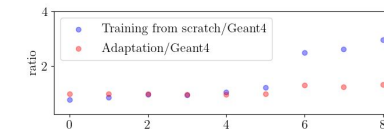
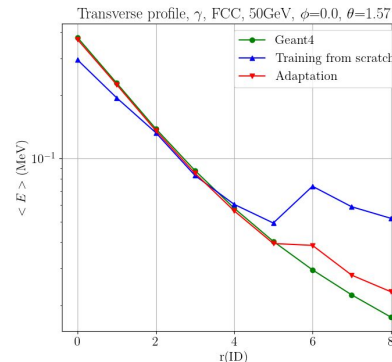
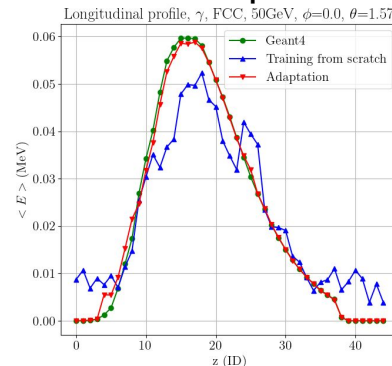
## 400K samples



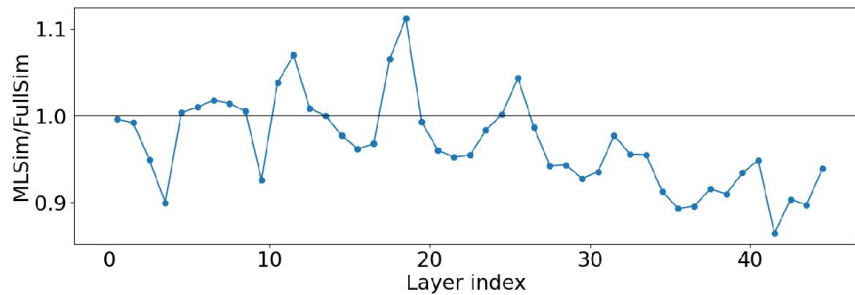
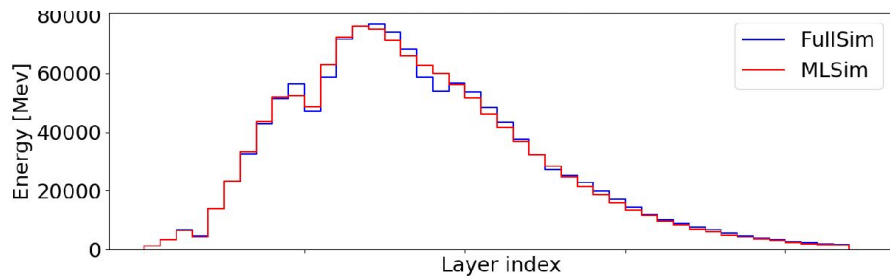
## 200K samples



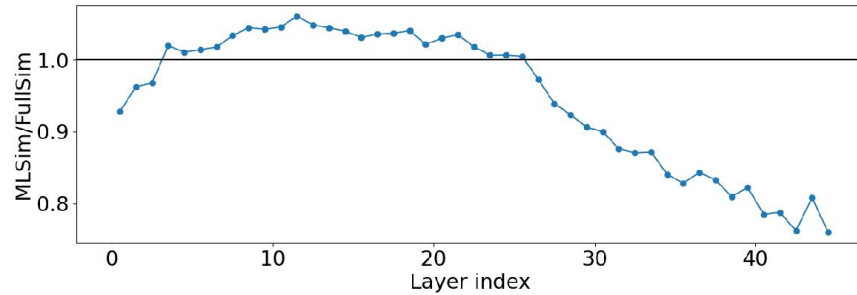
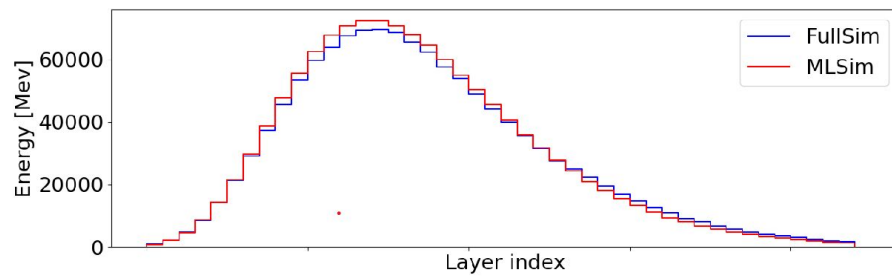
## 100K samples



# Angle conditioning



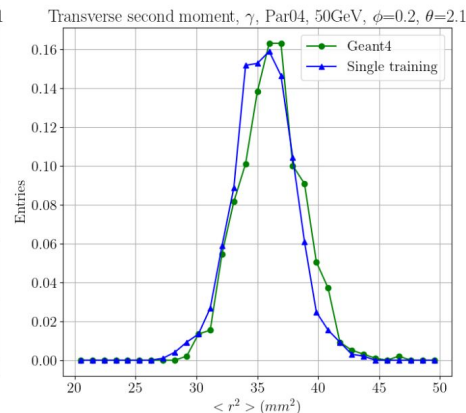
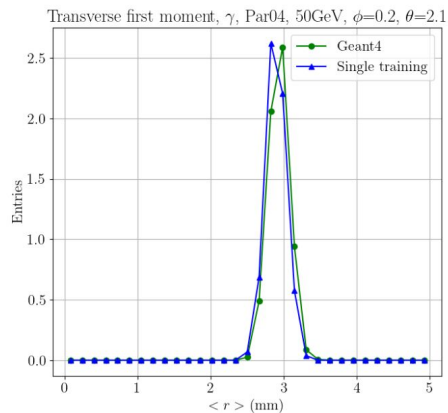
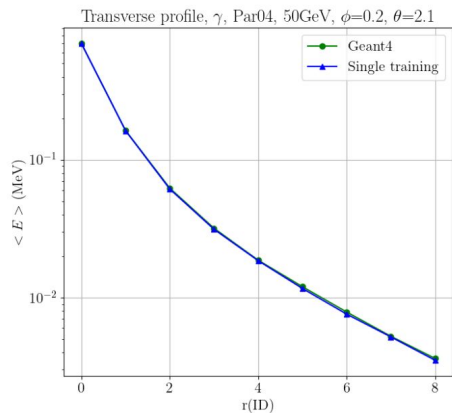
$\varphi = 0.2, \theta = 2.1$



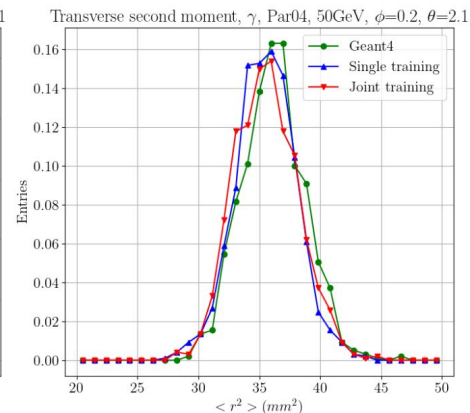
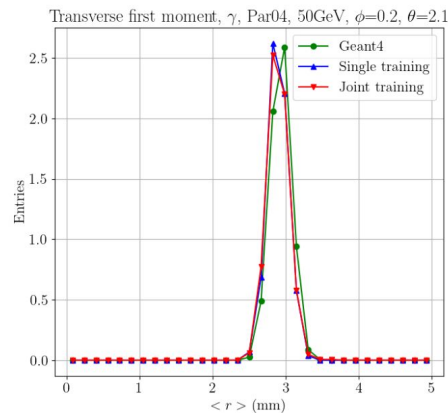
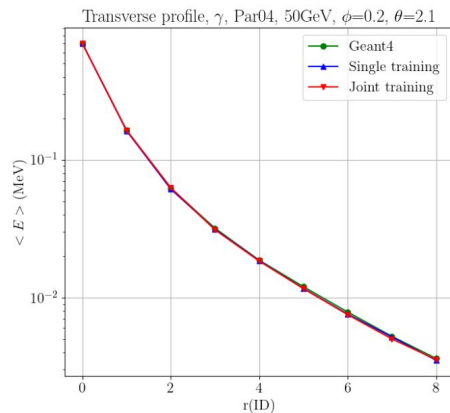
$\varphi = 0.0, \theta = 1.57$

# Transverse profiles - Par04

Single geometry training

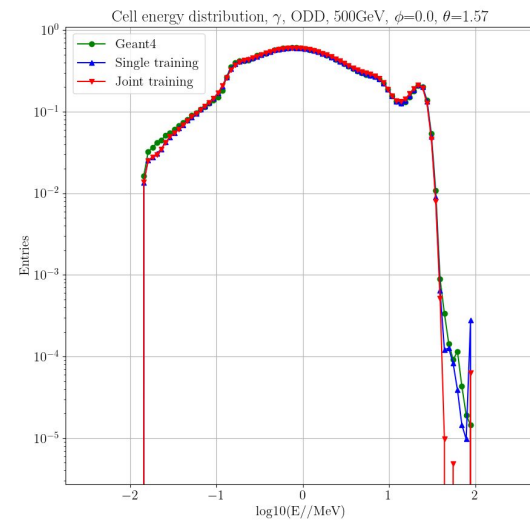
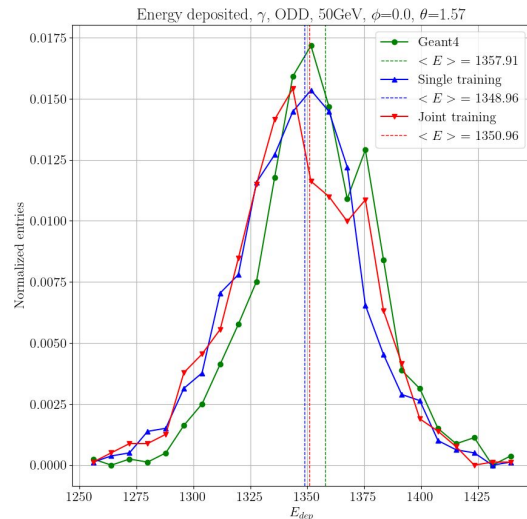
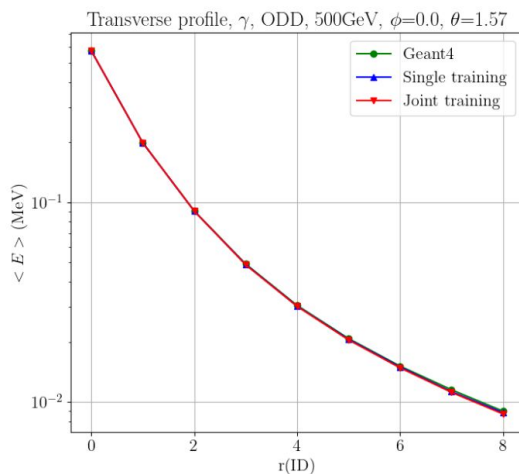
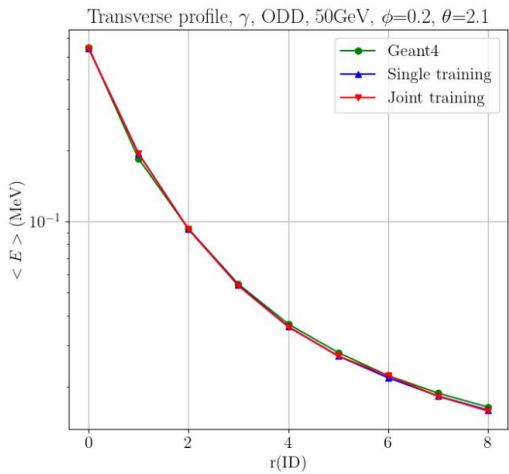
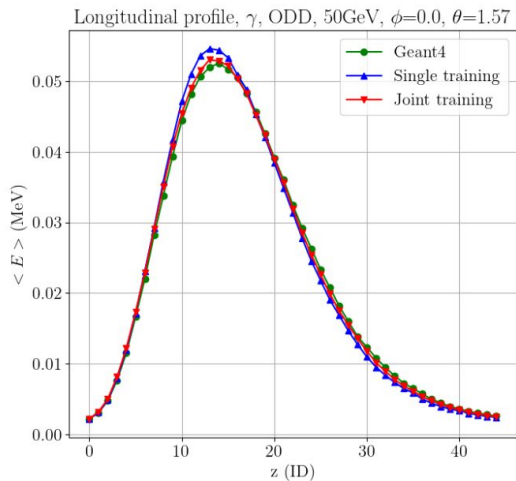
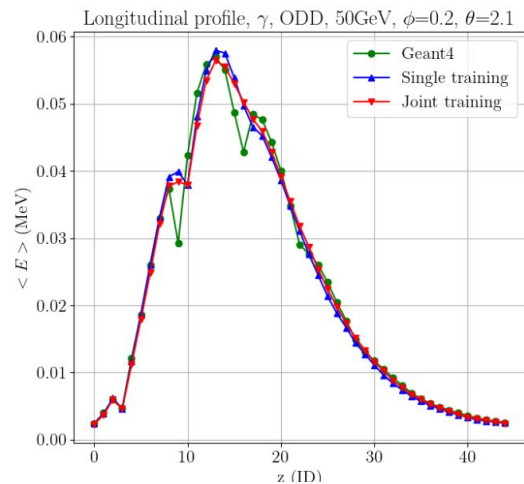


Joint training





# Joint training - ODD

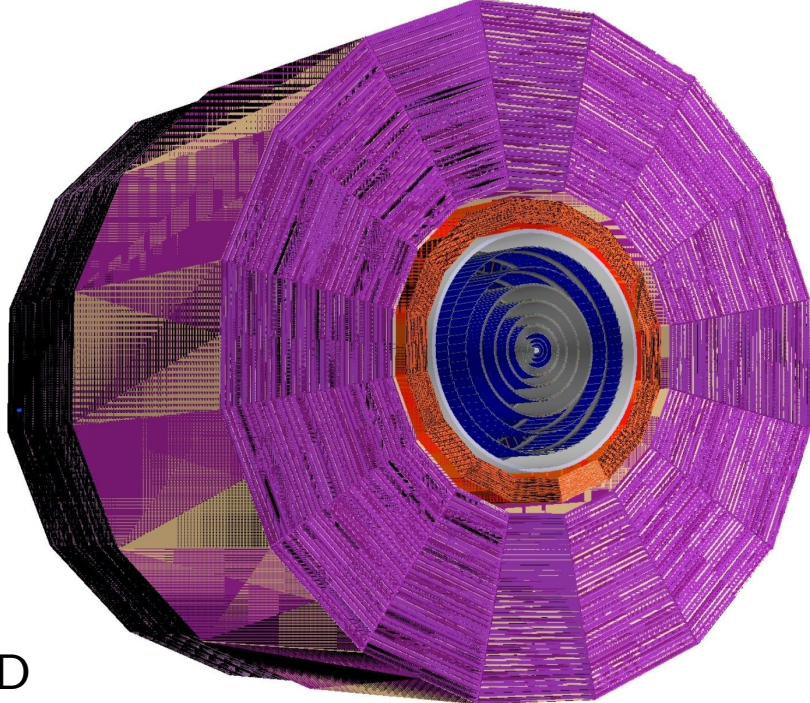


---

# Future work

- [WIP] Initial architecture tuning
- [WIP] ONNX/LibTorch conversion and integrating with Par04
- [GSoC project] Optimization of the diffusion process (Distillation, DDIM, etc.)
- Pretraining on more geometries (hence generating around 4M samples)
  - Par04 SiW
  - Par04 SciPb
  - Par04 PbWO<sub>4</sub>
  - Open data detector (ODD)
  - FCCee CLD
  - FCCee ALLEGRO

ODD



FCCeeALLEGRO

