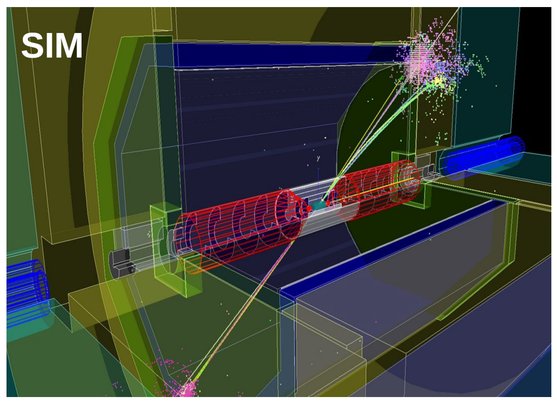
Towards Detector Agnostic Fast Simulation

Piyush Raikwar[1], Renato Cardoso[1], Vijay Ekambaram[3], Kristina Jaruskova[1], Jayant Kalagnanam[2], Peter McKeown[1], Nam Nguyen[2], Dalila Salamani[1], Mudhakar Srivatsa[2], Sofia Vallecorsa[1], Kyongmin Yeo[2], Anna Zaborowska[1]

[1]CERN [2]IBM Research, NY USA [3]IBM Research, India

Simulation

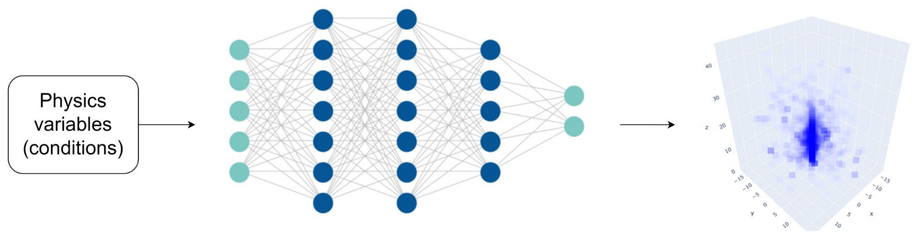


GEANT4 A SIMULATION TOOLKIT

Accurate but computationally expensive, especially for **calorimeter showers**.

Fast Simulation

Generating these showers with a Machine Learning (ML) model while retaining sufficient accuracy.



What's different?

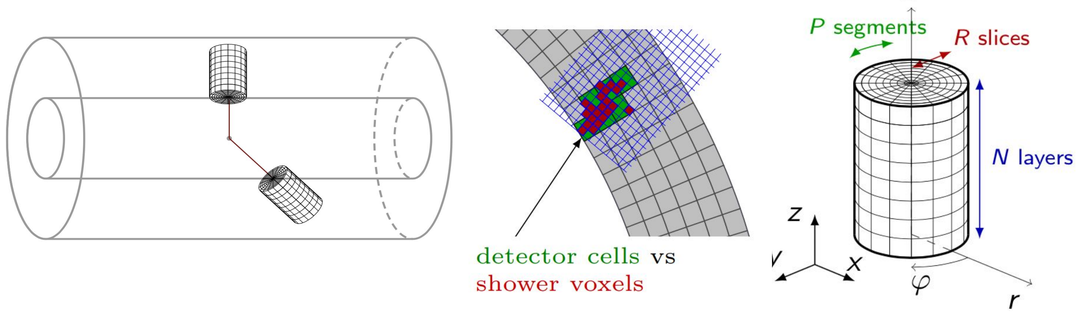
Most prior ML-based fast simulation methods are detector dependent - such a tailored design requires considerable amounts of **compute** and **manpower**.

Our approach: adopt a **detector agnostic** fast simulation scheme.

Make FastSim readily adaptable without access to ML expertise.

(a) Generic scoring method

Collect energy independent of the readout geometry by constructing a *physics-inspired* detector agnostic mesh to contain showers.

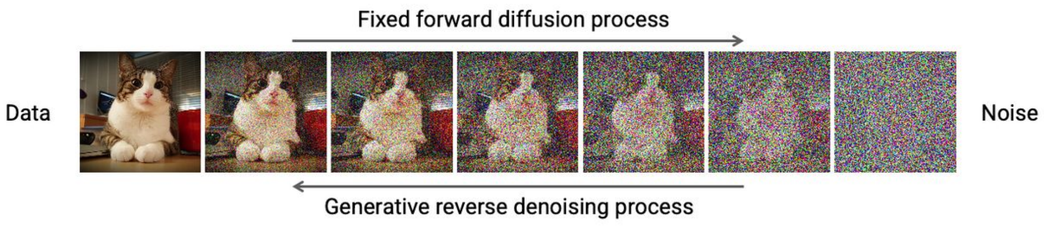


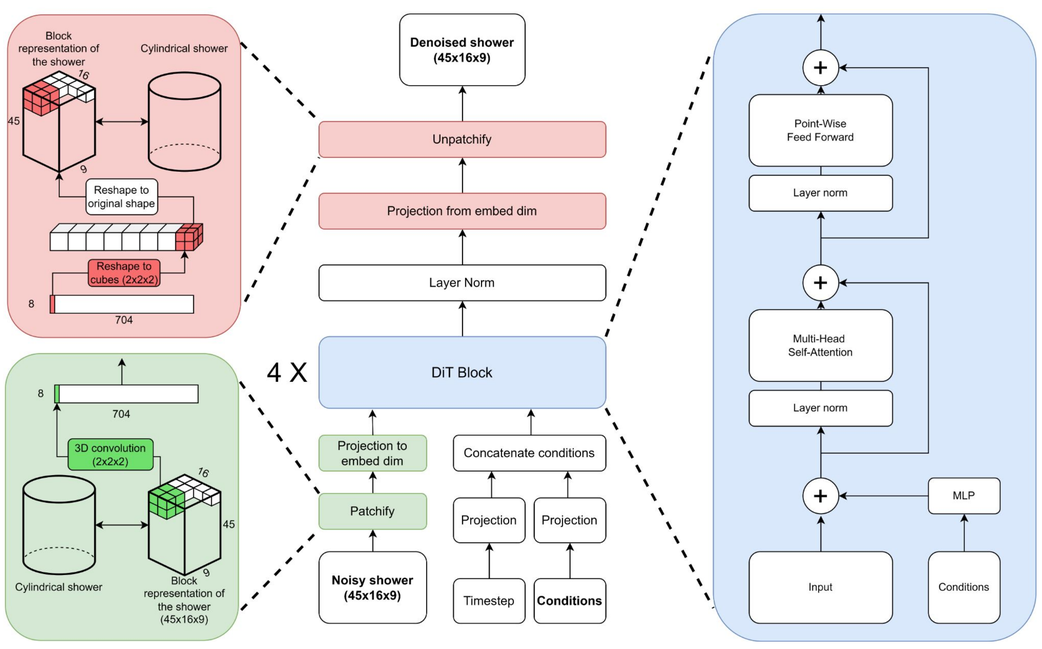
To achieve an appropriate detector response, the incident angles φ & θ, and the energy of the primary particle are also recorded for each shower.

(b) Generalizable model

Utilize knowledge from multiple detectors to learn generalized shower characteristics. This is accomplished by **training once** on multiple geometries, **then adapting** to the desired geometry.

We use a transformer-based diffusion model for higher accuracy and higher diversity.

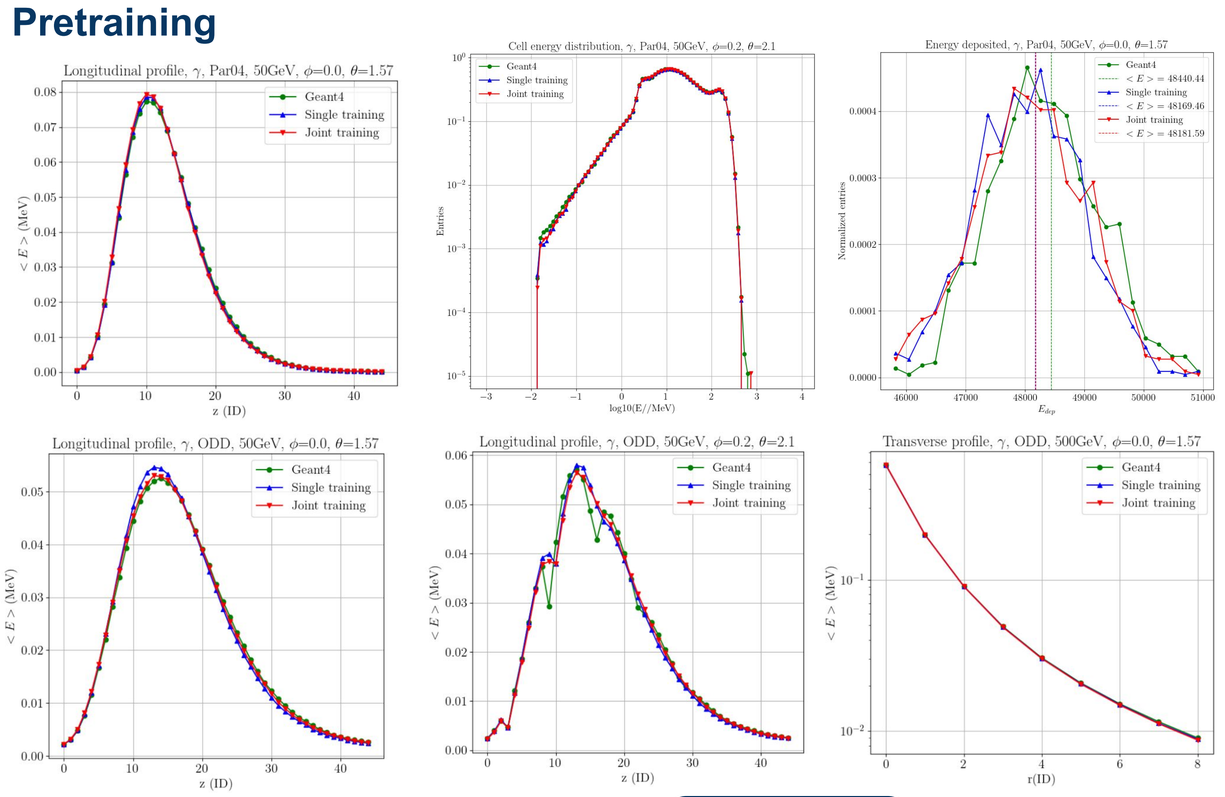




Experiments

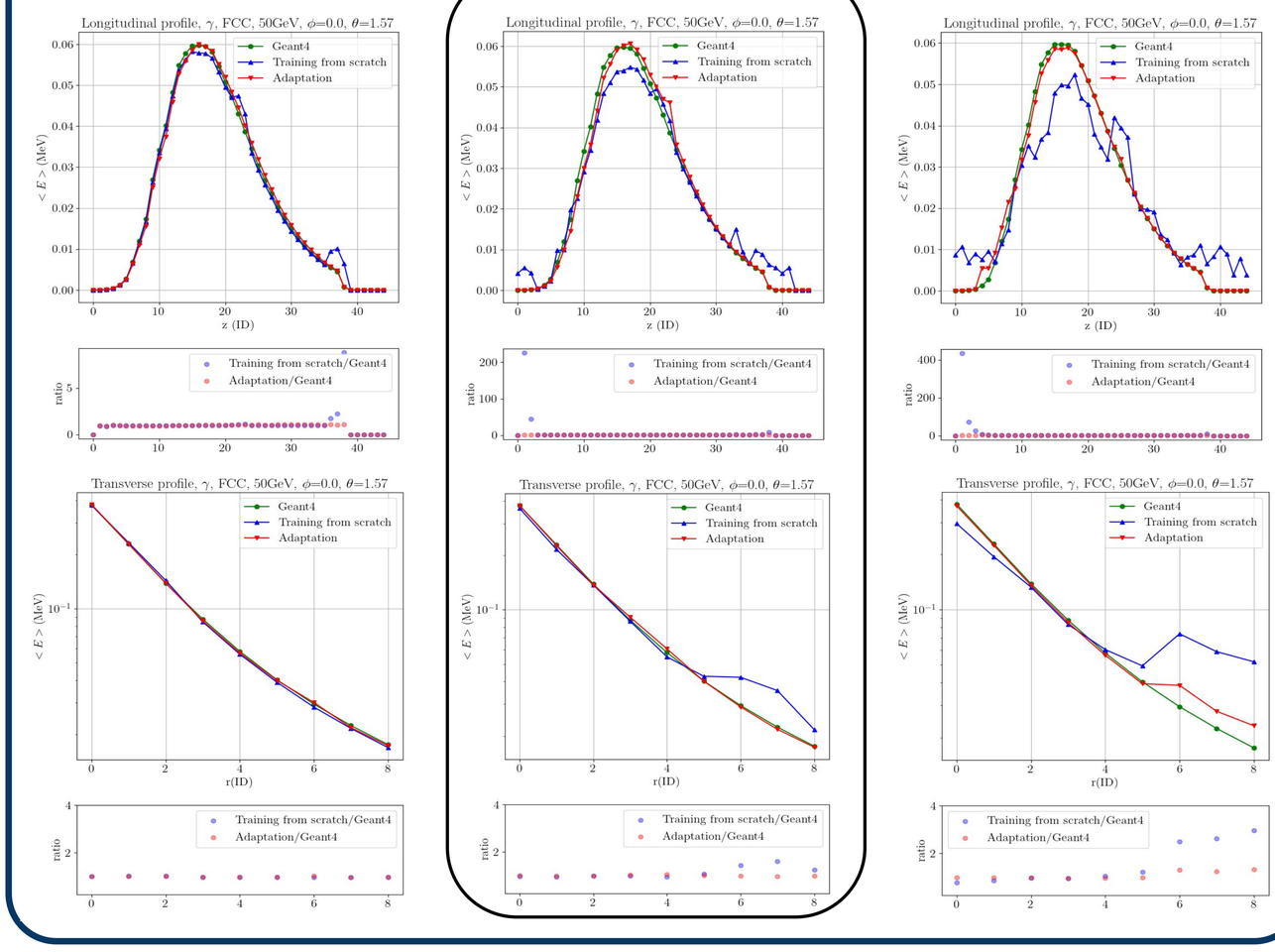
Pretraining on Par04 & ODD (Open Data Detector) and adaptation on FCCeeALLEGRO.

Pretraining



Adaptation

At 200K samples
~25x less training time
<50% of the data



Takeaway

We present an easy to use detector agnostic ML model for fast simulation. It is based on a transformer architecture and a diffusion process. Preliminary results indicate up to ~25x faster training time while requiring ~50% less data to achieve the same performance.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under GA no 101004761.