Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

DESY.

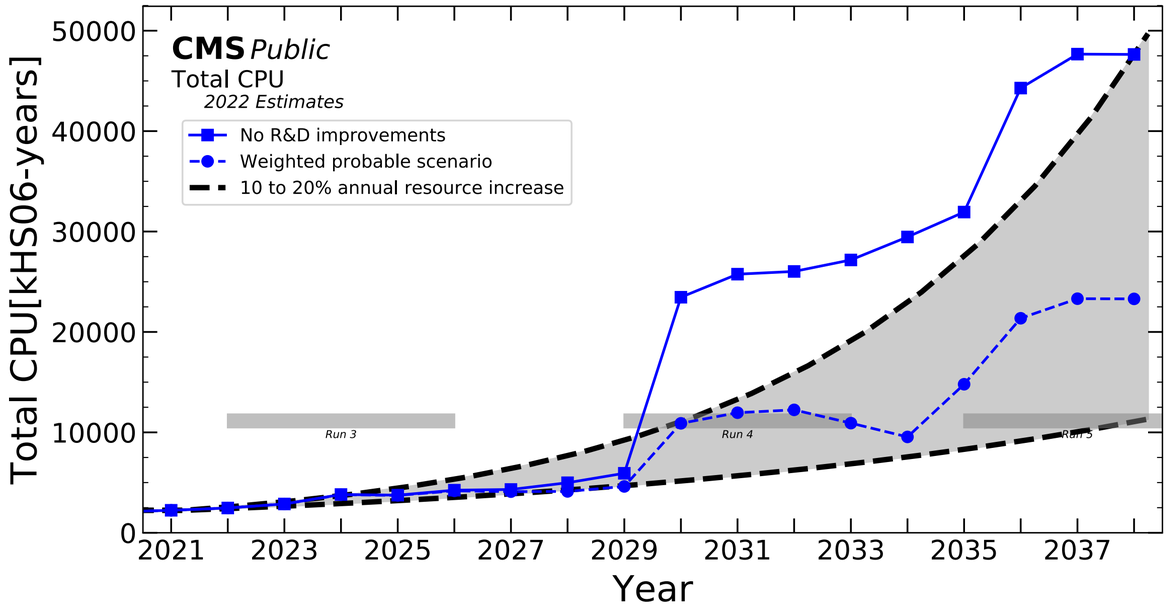
CaloHadronicVAE : Fast Point Cloud Generation via Latent Diffusion

Henry Day-Hall, Frank Gaede, Gregor Kasieczka, Katja Krüger,
and Martina Mozzanica

martina.mozzanica@uni-hamburg.de

The Need for Fast Simulation

- **Simulation is fundamental** to compare theory to measurements
- Simulations will have to match the statistics of collected data x10
- Computational requirements expected to **exceed available resources** soon
- **Detector simulation** is the most expensive block of the chain
- Without R&D, increase in computing time exceeds the one in resources

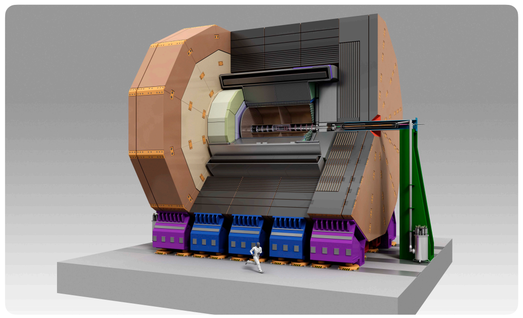


CMS Collaboration, Offline and Computing Public Results (2021)

The ILD Calorimeter and Data Representation

- International Large Detector (ILD) proposed for FCC-ee

- High-Granularity sampling calorimeters:
  - ECal: Si-W - 5mm x 5mm - 30 layers
  - HCal: Sci-Fe - 30mm x 30mm - 48 layers

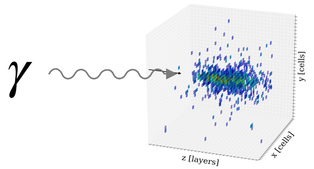


The International Linear Collider: A Global Project

High granularity calorimeter → High fidelity simulation

Fixed Grid

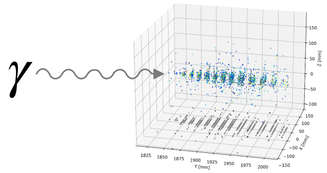
- 3D array filled with energy values
- entries correspond to calorimeter cells



Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed

Point Clouds

- variable-length, permutation-invariant sets
- only generation of non-zero points

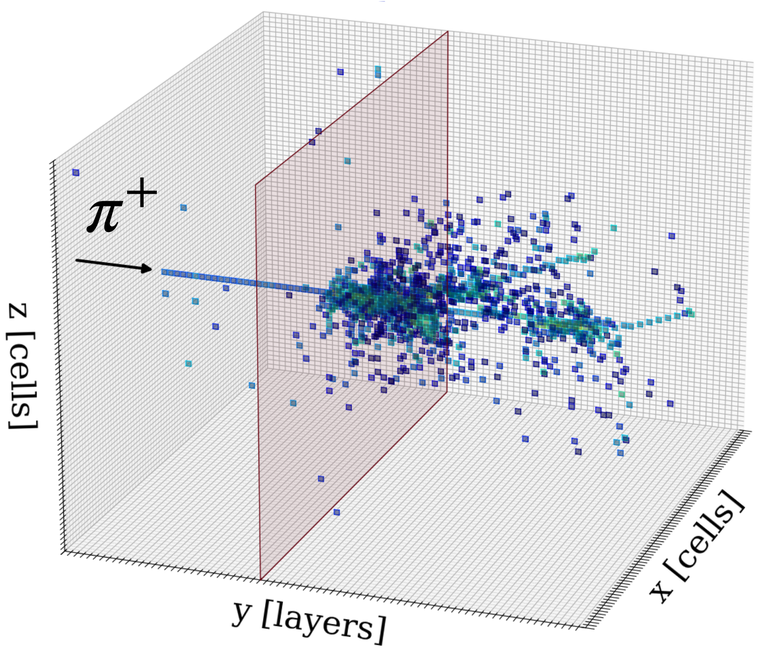


CaloClouds: Fast Geometry-Independent Highly-Granular Calorimeter Simulation

Point Clouds allow multiple points per cell & cell-geometry independence

Dataset

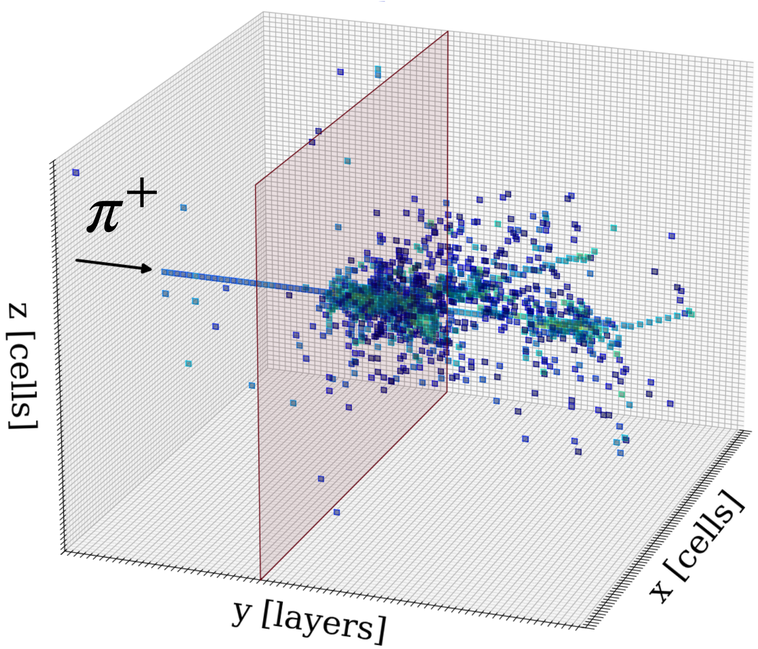
- Only **hadronic showers** are considered (the harder ones to model)
- **10-90 GeV π^+ showers** with direction perpendicular to the face of the ECal
- **Time feature included** - first time for high granular calorimeters
- Cut on the transversal plane: -450mm to 450 mm
- **x9** higher granularity



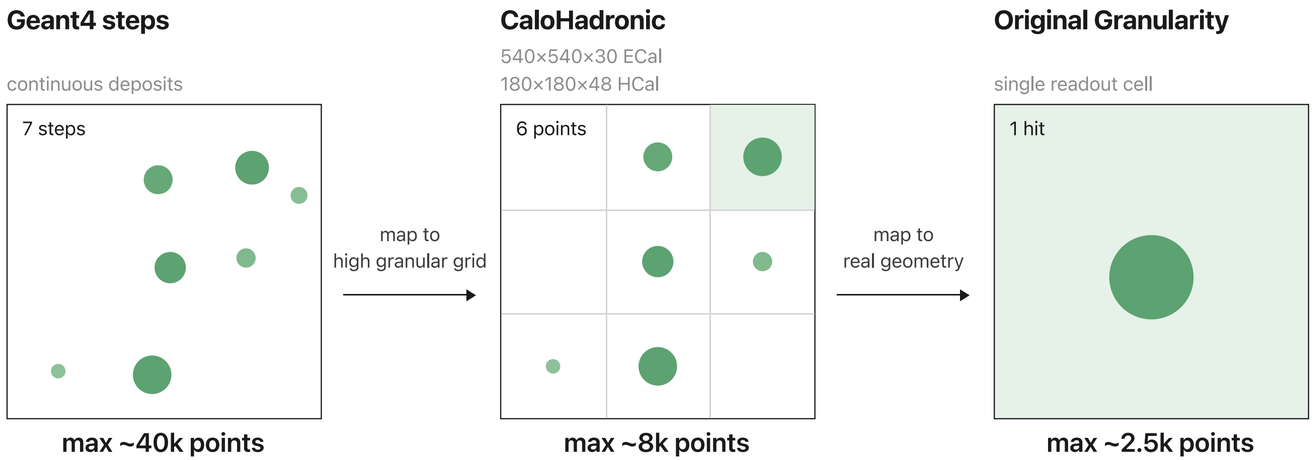
CaloHadronic: a diffusion model for the generation of hadronic showers

Dataset

- Only **hadronic showers** are considered (the harder ones to model)
- **10-90 GeV π^+ showers** with direction perpendicular to the face of the ECal
- **Time feature included** - first time for high granular calorimeters
- Cut on the transversal plane: -450mm to 450 mm
- **x9** higher granularity

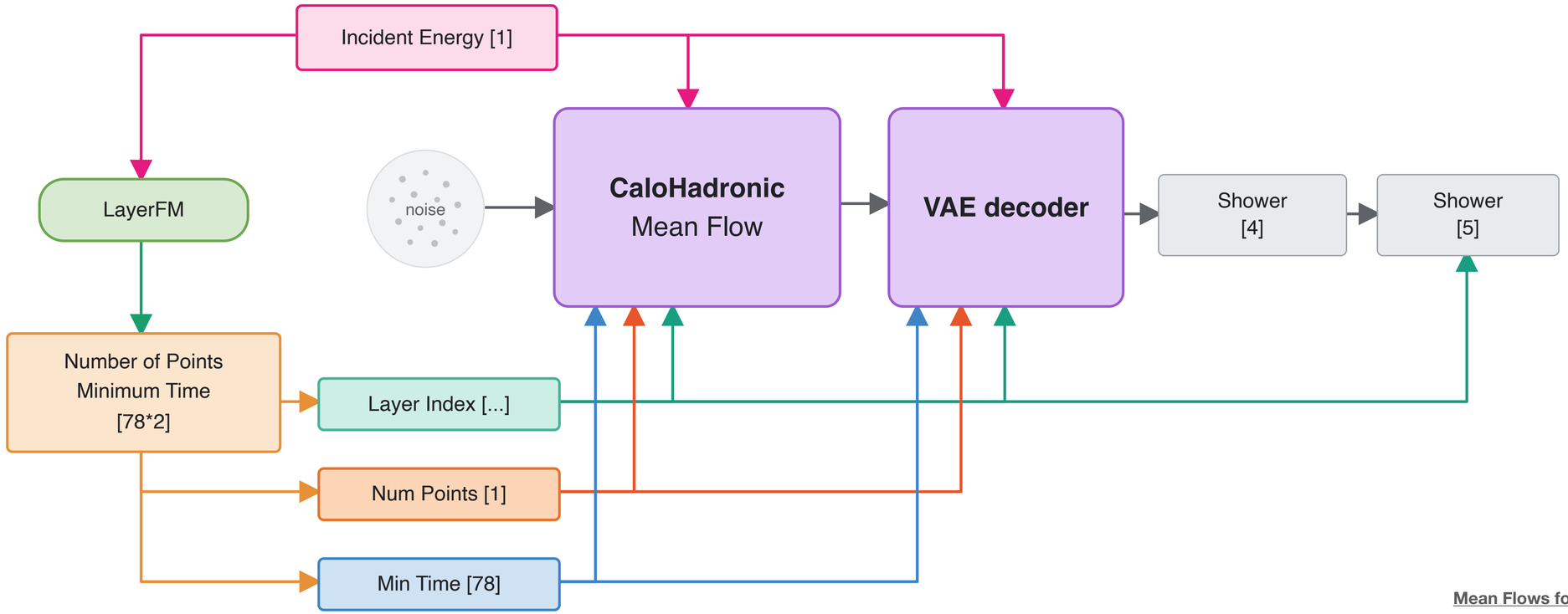


CaloHadronic: a diffusion model for the generation of hadronic showers



CaloHadronicVAE

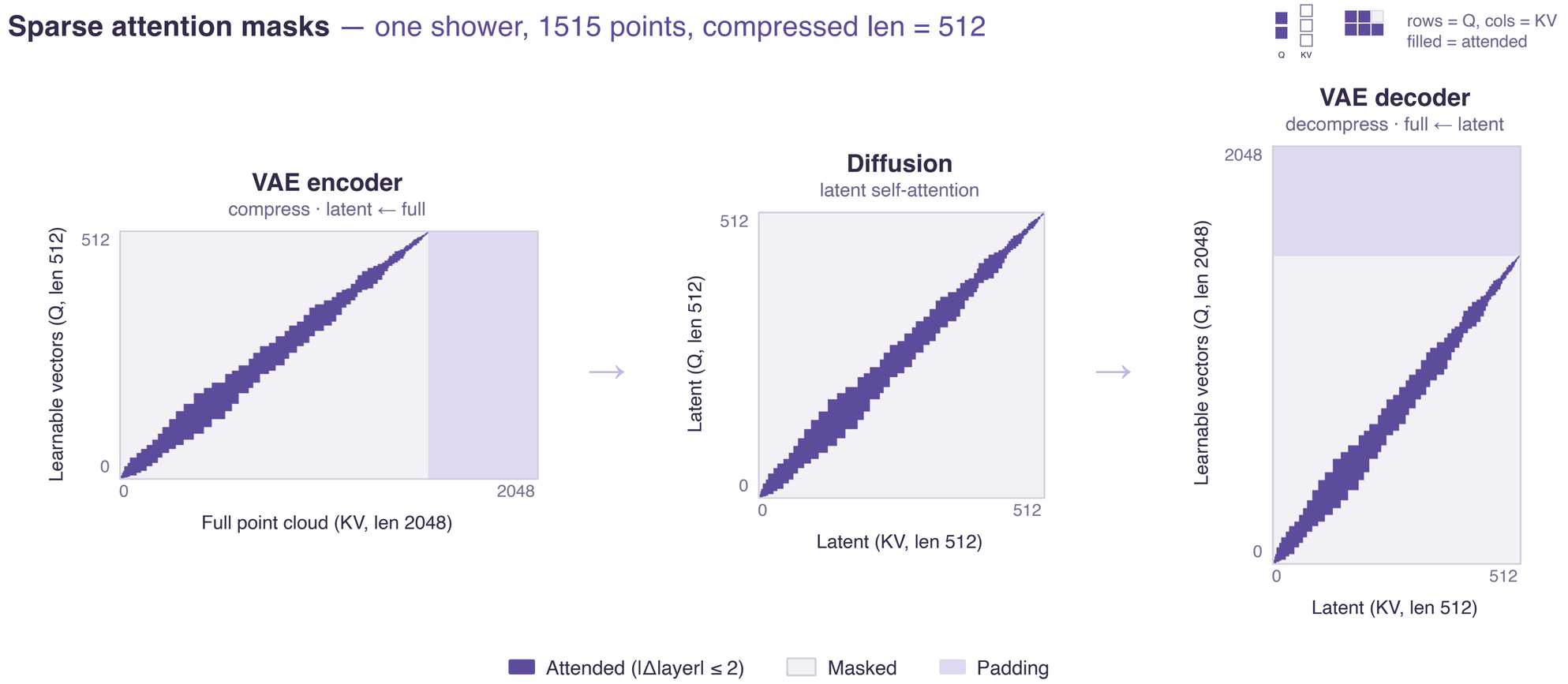
- The model comprises three components: a continuous normalising flow + a VAE + a diffusion model.
- It generates 5D point clouds [x, y, z, energy, time], conditioned on the incident energy.
- The flow generated layer-wise quantities: the number of points per layer (split into total points and layer index) and the minimum time per layer.



Mean Flows for One-step Generative Modeling

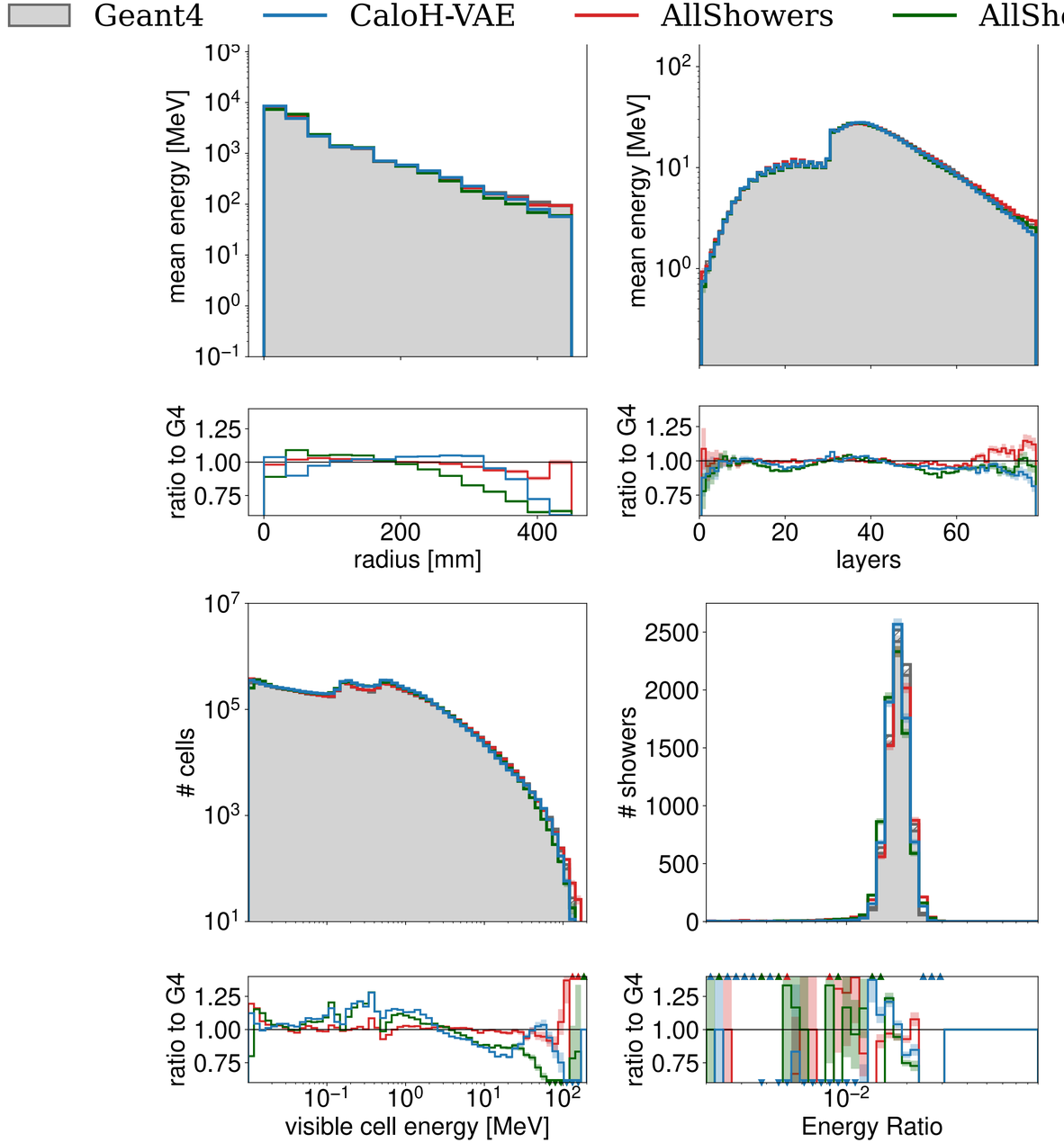
Sparse Attention Mask for speed up

- The model avoids the cost of direct point-cloud generation by encoding every shower into a **latent vector**.
- This approach is **significantly cheaper** than direct point-cloud generation



Results

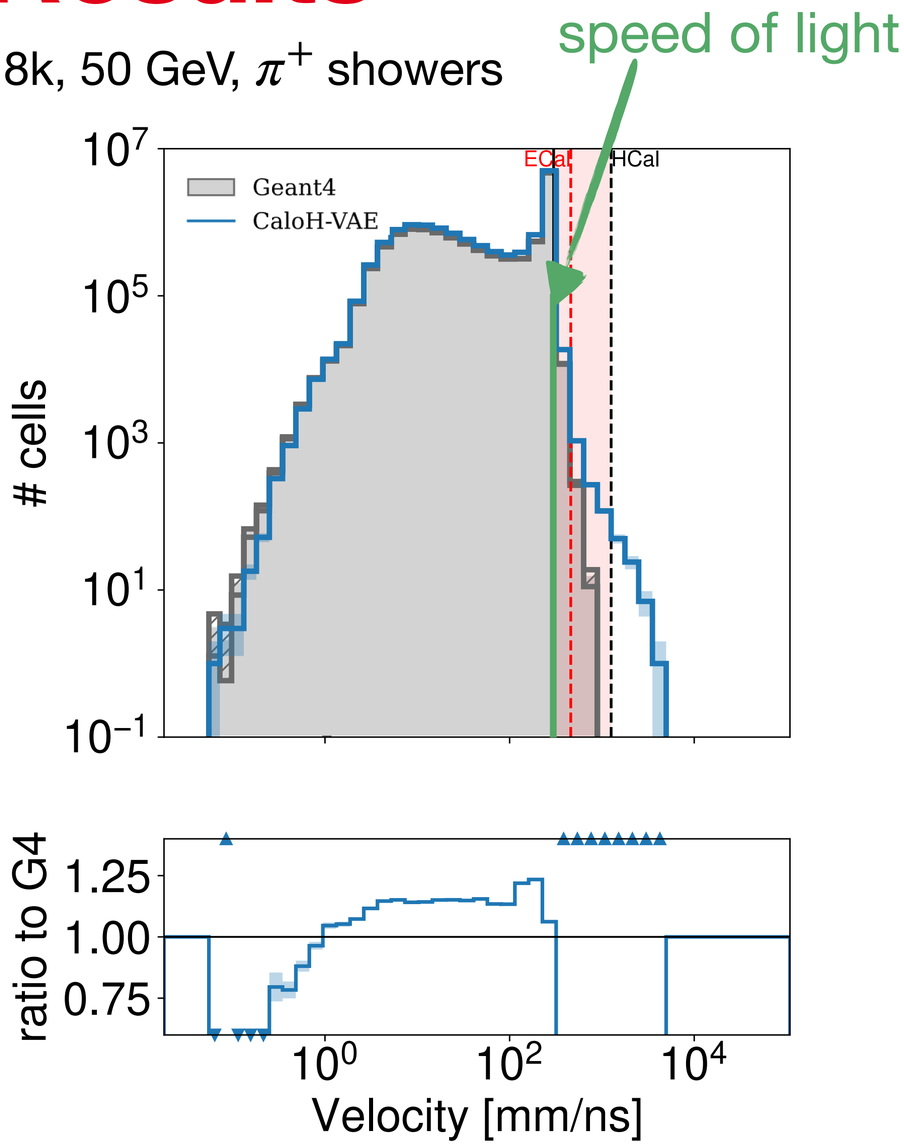
8k, 50 GeV, π^+ showers

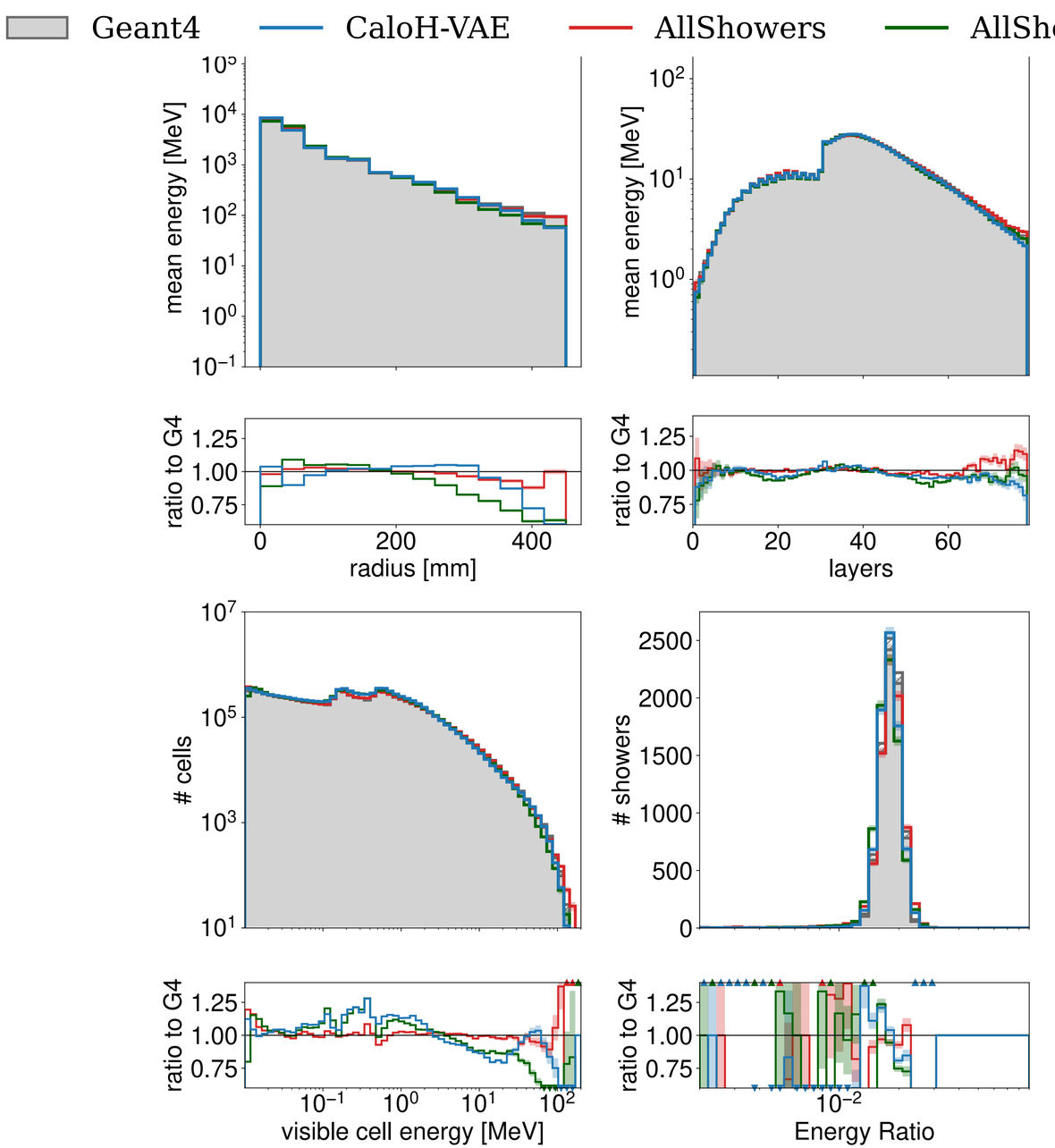


AllShowers: One model for all calorimeter showers

Results

8k, 50 GeV, π^+ showers

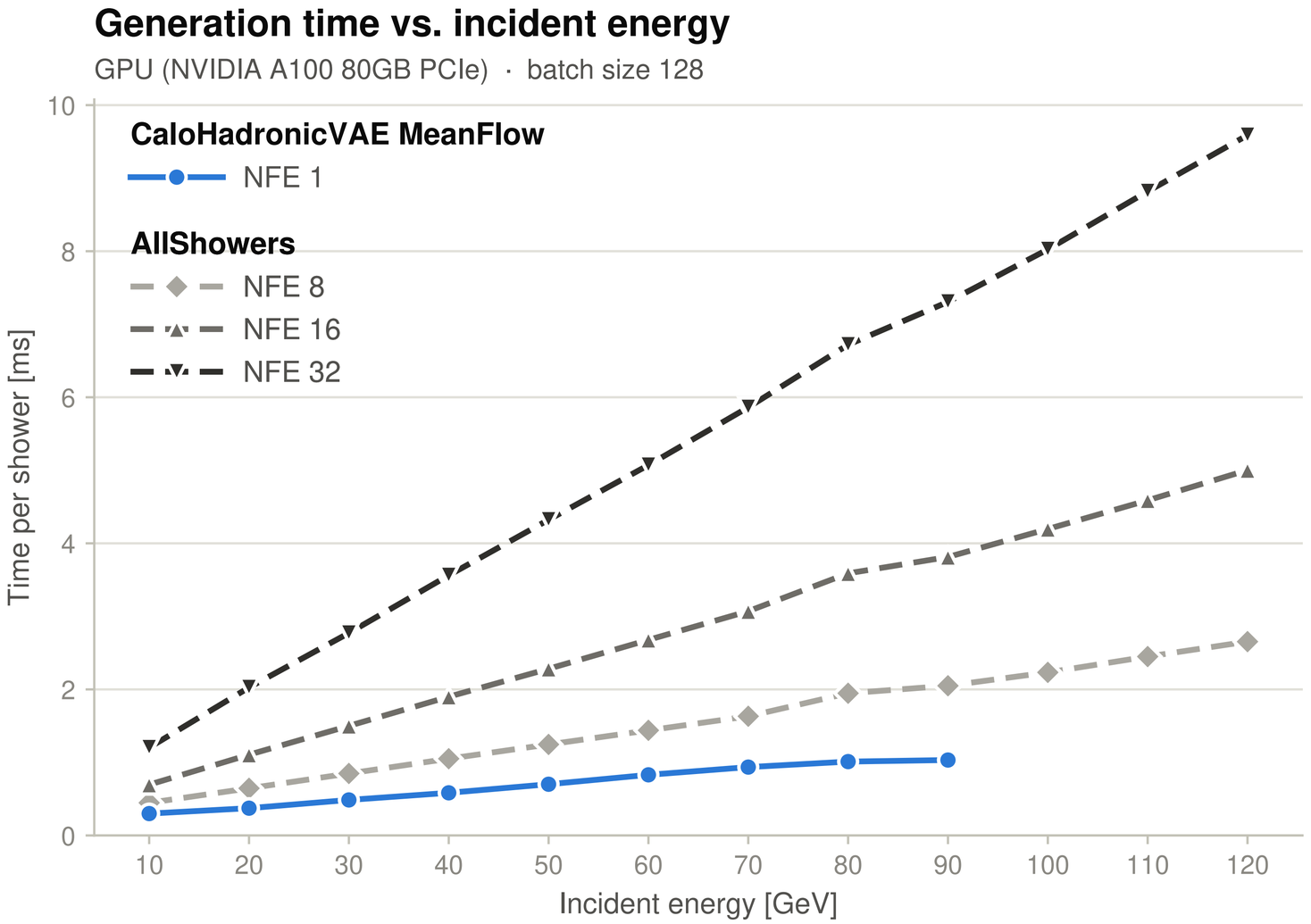




AllShowers: One model for all calorimeter showers

Results: generation time

NFE = number of function evaluations



Summary

- **Point cloud** based model are the best solution for speed purposes
- Time feature included!
- **Latent diffusion** enables to have a compact representation of a shower to further improve speed up
- Scales very well with the number of points/ incident energy
- Outlook: em showers, different hadronic showers, angular conditioning, …
- **Paper out soon!**