When does Diffusion Models learn Physics?

A case study of diffusion in the Calo Challenge

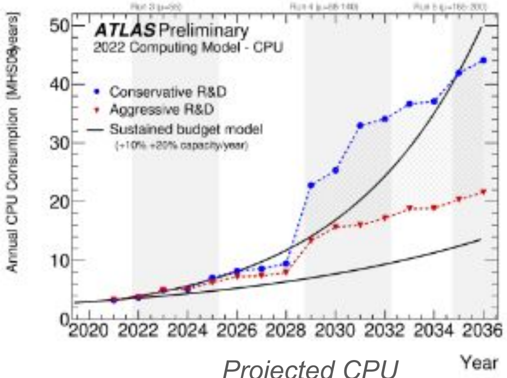
Group 1

Mentor : - Liv Helen Vage

Computational Bottleneck in Calorimeter Simulation

Why Fast Detector Simulation ?

- ❖ Geant4 is the gold standard.
  - ➢ Accurately simulates the particle interactions with the detector material.
  - ➢ Tracks both the primary and the secondary particles produced in that shower.
- ❖ Why is it computationally expensive?
  - ➢ Showers contain thousands of secondary particles.
  - ➢ Each interaction is simulated at every step.
  - ➢ This simulation contains a major fraction of computing resources at the collider experiments.
- ❖ Challenge at the HL - LHC.
  - ➢ Higher event rates, increased pile - up.
  - ➢ More complex, high granular detectors.
  - ➢ Require more computational resources to simulate.



Projected CPU requirements

https://cds.cern.ch/record/2802918

How to overcome these challenges ?

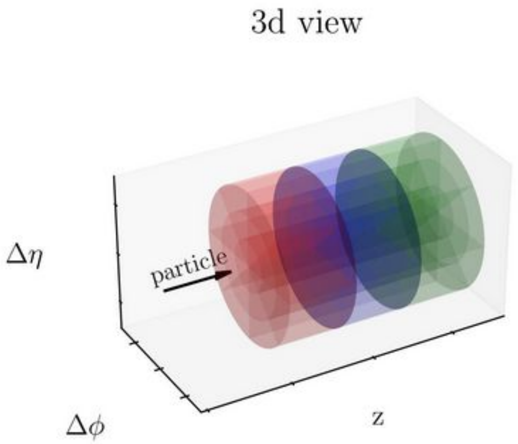
- ➢ Replace computationally expensive Geant4 simulations with **ML-based surrogate models**
- ➢ In this work
  - ○ using denoising diffusion models which will learn from Geant4-generated showers and check how fast realistic calorimeter responses are derived while preserving the essential physics.

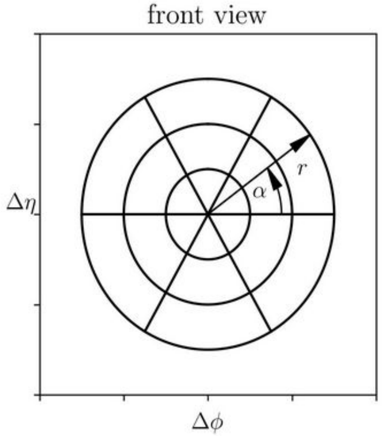
Dataset & Detector geometry

- **Dataset format**
  - Using the voxelization format of ATLAS fast simulation
  - The spatial energy deposits in each layer of the Geant4 datasets are represented as **voxels**, i.e., discrete three-dimensional volume elements that store the deposited energy

- **Detector geometry**
  - Concentric cylinder with particles propagating along the z-axis segmented along the z-axis into N_z discrete layers
  - Each layer has N_r bins along the radial direction and N_α bins in the angle α
  - ***Voxel*** $= N_z \times N_r \times N_\alpha$

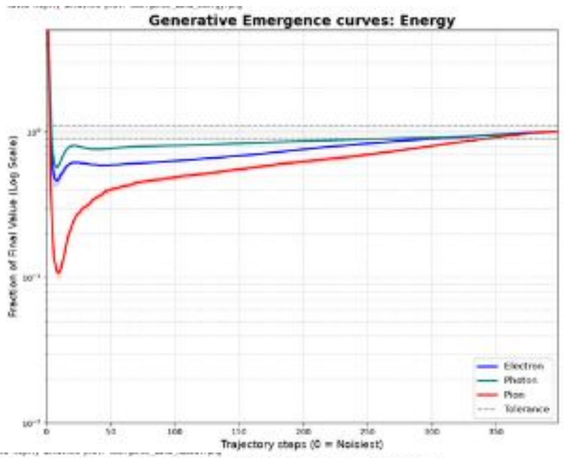
	Number of Voxels	N_z	N_α	N_r
γ	360	5	[1,10,10,1,1]	[8,16,19,5,5]
π^+	533	7	[1,10,10,1,10,10,1]	[8,10,10,5,15,16,10]
e^-	6480	45	[9]	[16]

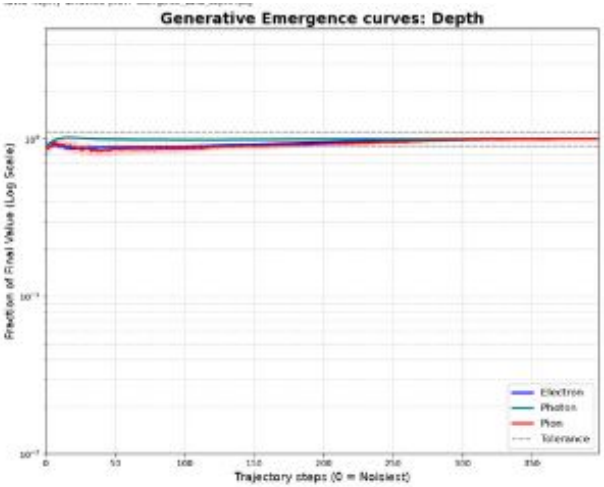


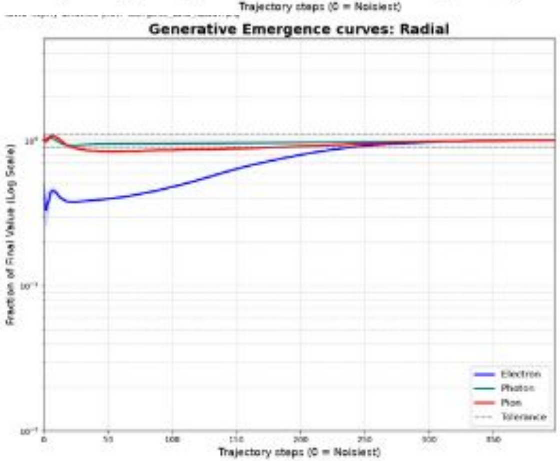


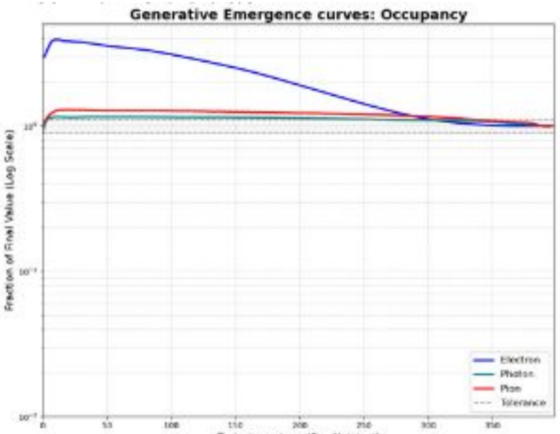
Schematic view of the voxelization in all datasets

Emergence Plots









FFT Analysis

- ❖ Fast Fourier Transforms: A mathematical operation that takes a 1D energy array (specifically, the total transverse energy distributed in a continuous circle around the beamline, or the azimuthal axis) and decomposes it into a sum of sine and cosine waves.
- ❖ By applying this 1D Real FFT, you translate the physical energy data into Fourier Modes (m). This allows you to break down the complex shape of a particle shower into distinct frequencies:

- **Low frequencies (m=0, 1):** Represent the smooth, macroscopic physics [Depth].
- **High frequencies (m=4):** Represent the jagged, microscopic physics [like the granular clusters in a hadronic pion shower(Occupancy)].

Azimuthal Fourier Modes Visualization

Visualizing the energy distribution equation: $E(\varphi) = \Sigma \cdot \cos(m\varphi)$

The dashed circle represents the baseline detector ring. The colored shape represents the energy distribution.

m = 0 (Monopole)

Total isotropic energy.
0 waves around the ring.

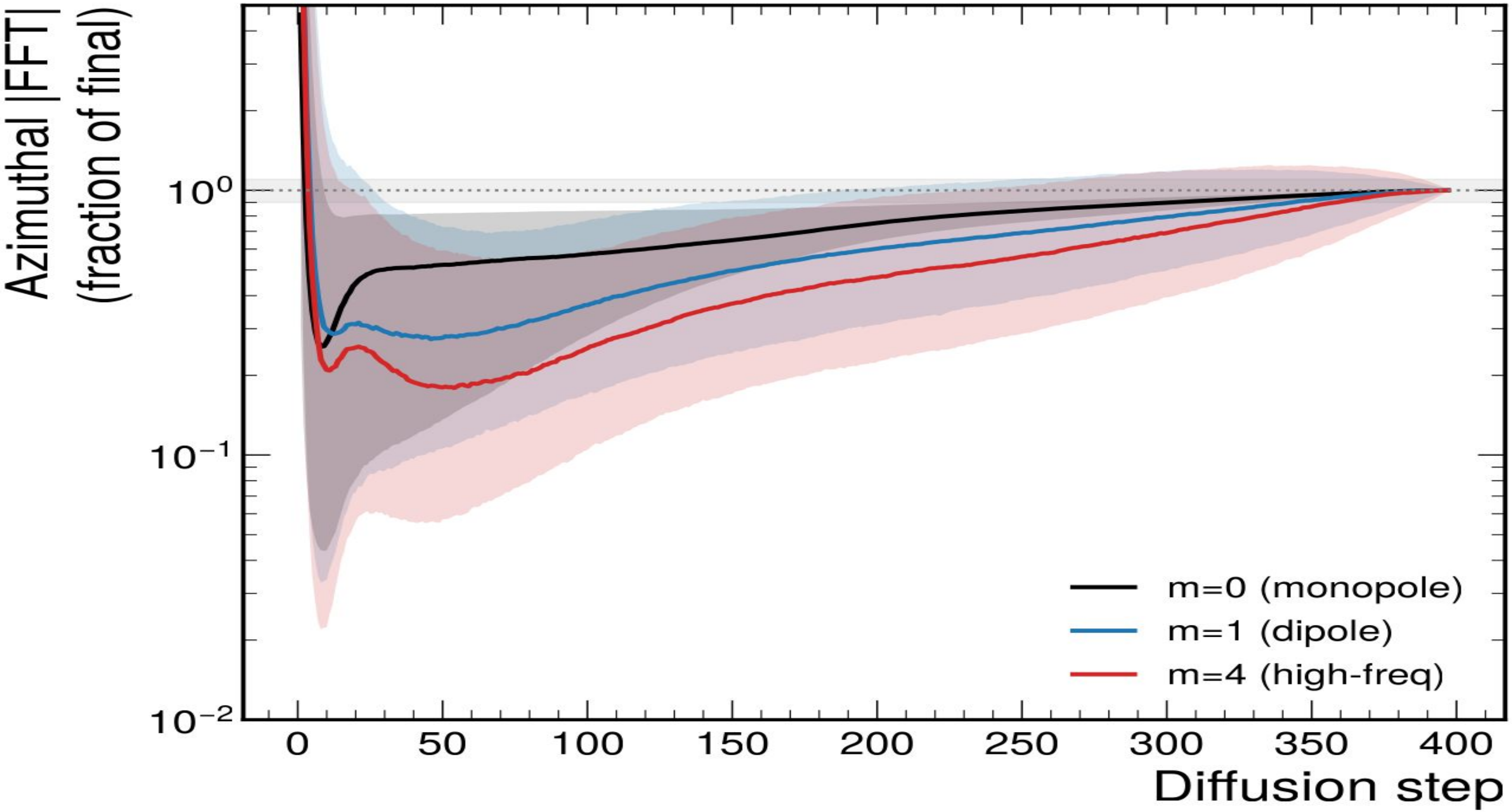
m = 1 (Dipole)

Macroscopic lean / Asymmetry.
1 full wave around the ring.

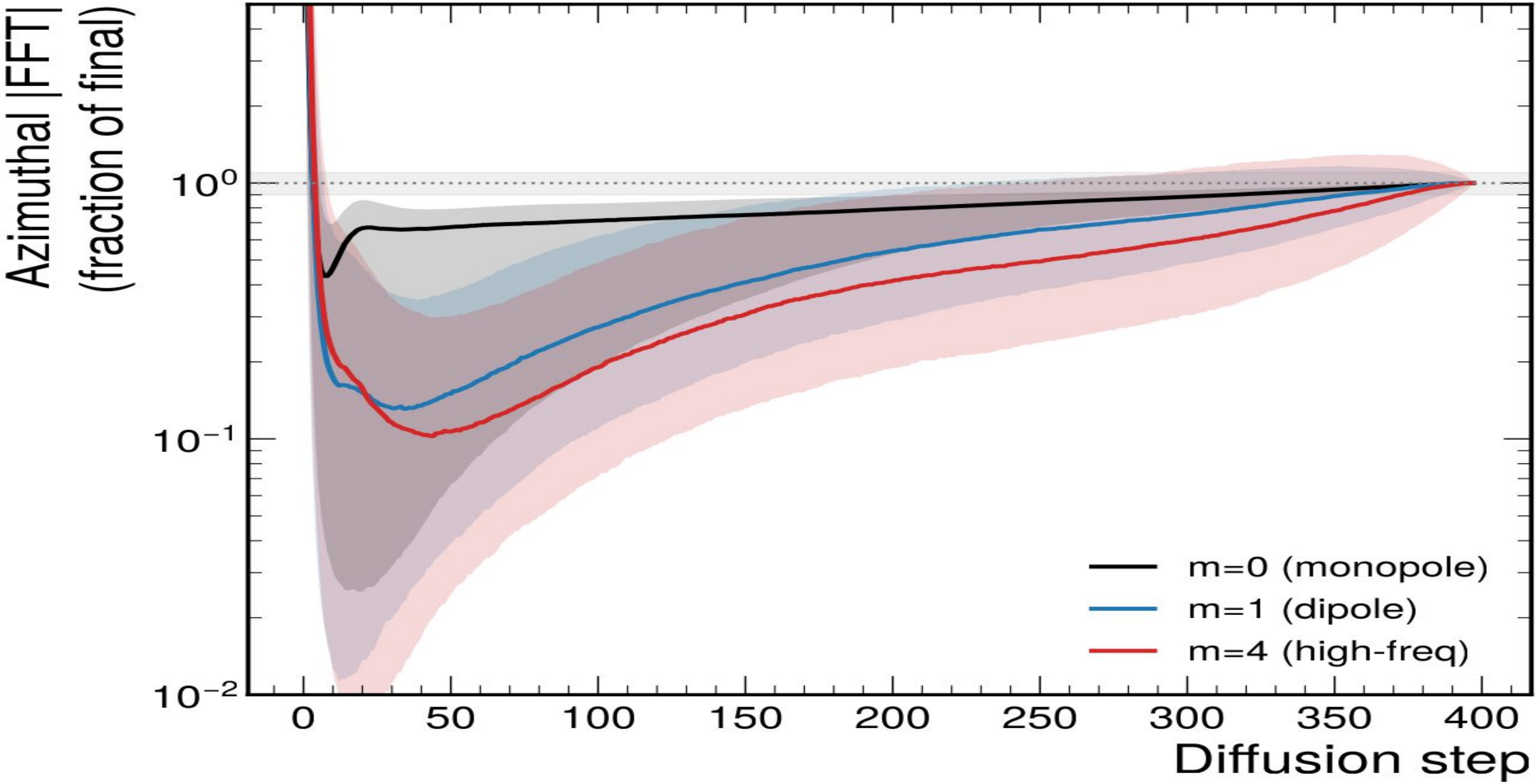
m = 4 (High-Frequency)

Microscopic granularity / Chaos.
4 full waves around the ring.

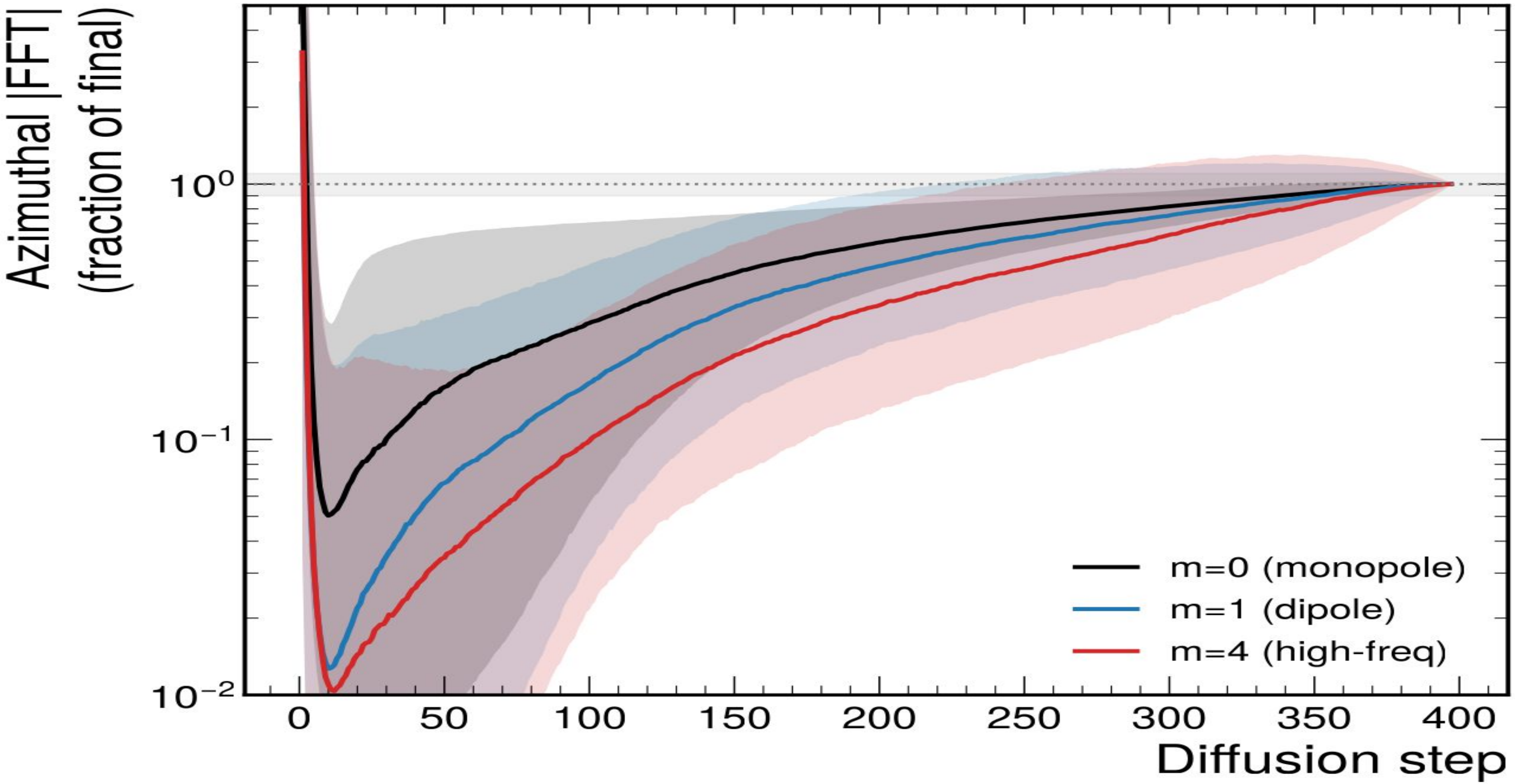
FFT Plot for Electron



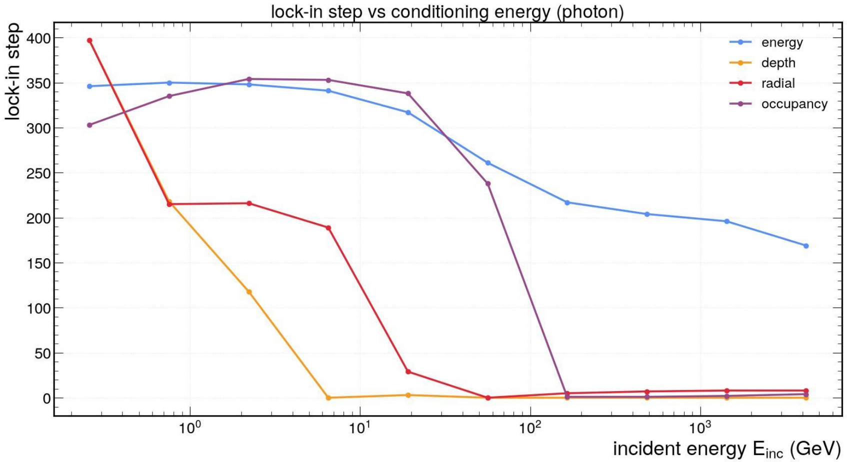
FFT Plot for Photon

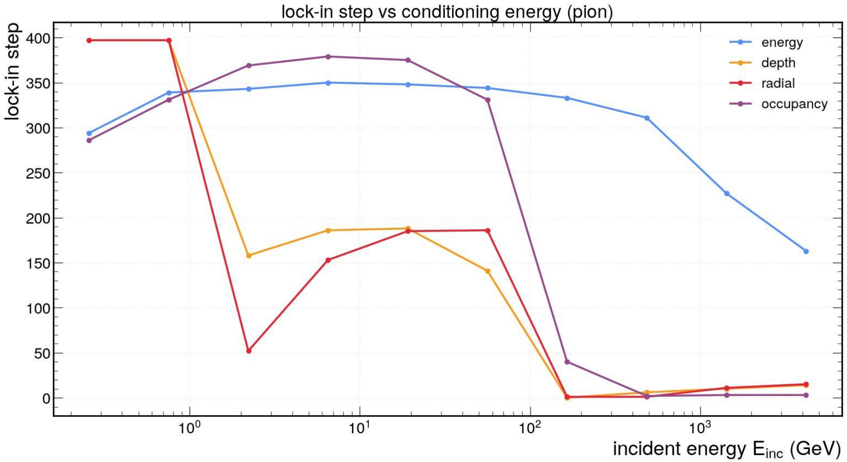


FFT Plot for Pion

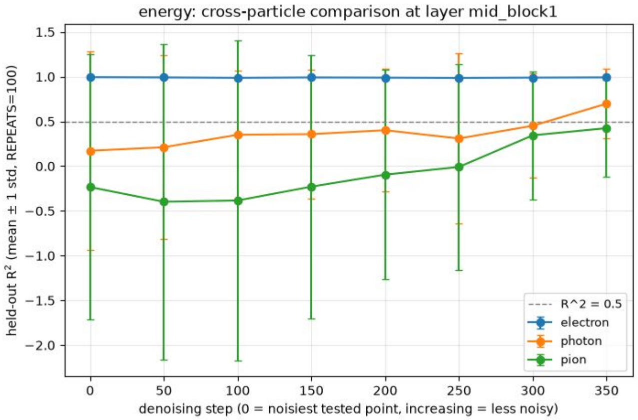


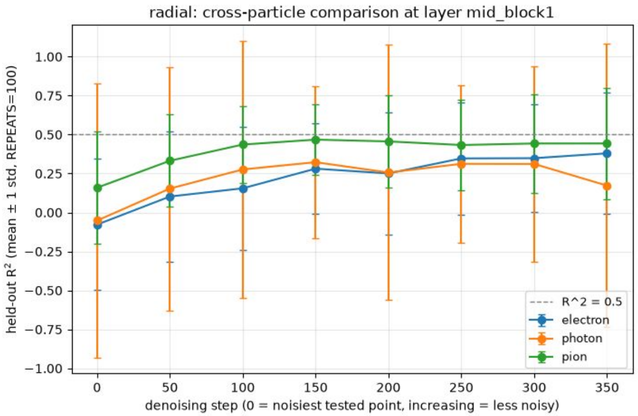
Locking point vs Energy: How does incident energy affect NN "Learning" ?



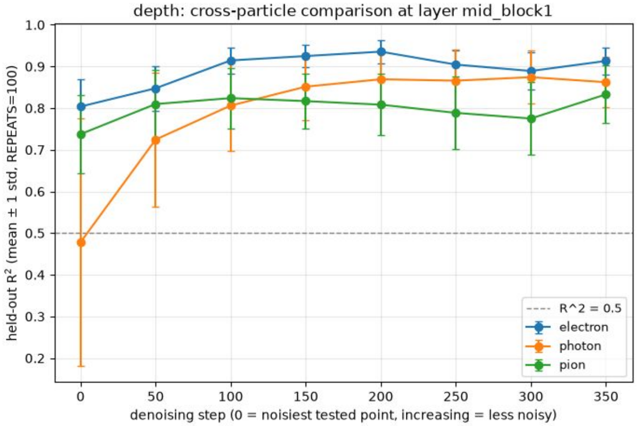


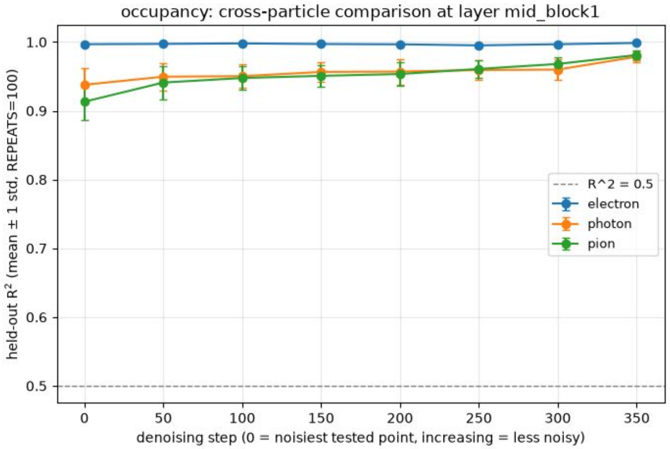
Linear probing the intermediate layers: When does it learn what?





Occupancy: strongest and most stable for all particles;
Depth: strong and stable for all particles;
Energy: excellent for electron, decent for photon, poor for pion;
Radial spread: weakest overall.

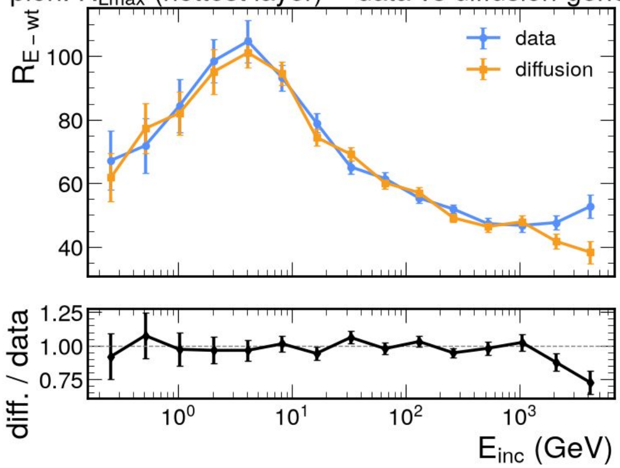




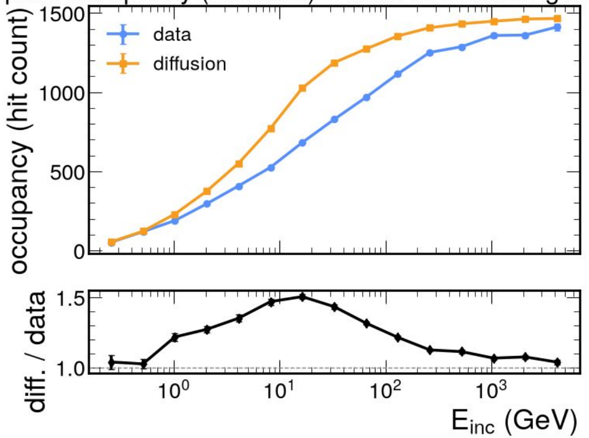
Easiest particle overall: electron
Intermediate: photon
Hardest: pion

Diffusion vs Data: How reliable are the diffusion networks for calorimetry ?

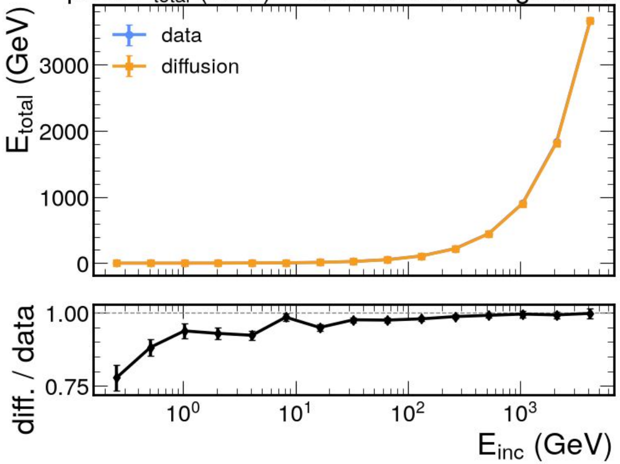
pion: R_{Lmax} (hottest layer) -- data vs diffusion-generated



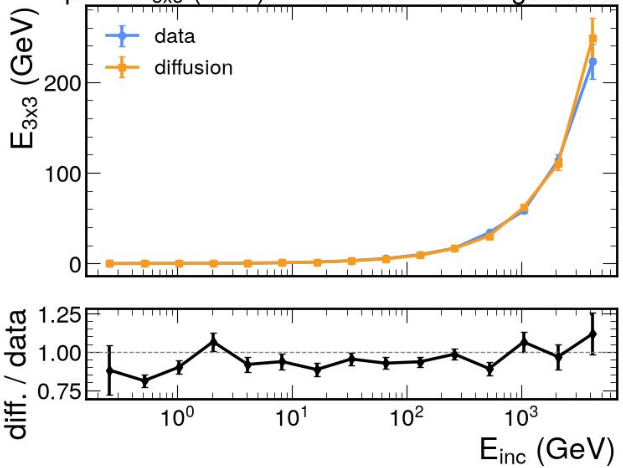
pion: occupancy (hit count) -- data vs diffusion-generated



pion: E_{total} (GeV) -- data vs diffusion-generated



pion: E_{3x3} (GeV) -- data vs diffusion-generated

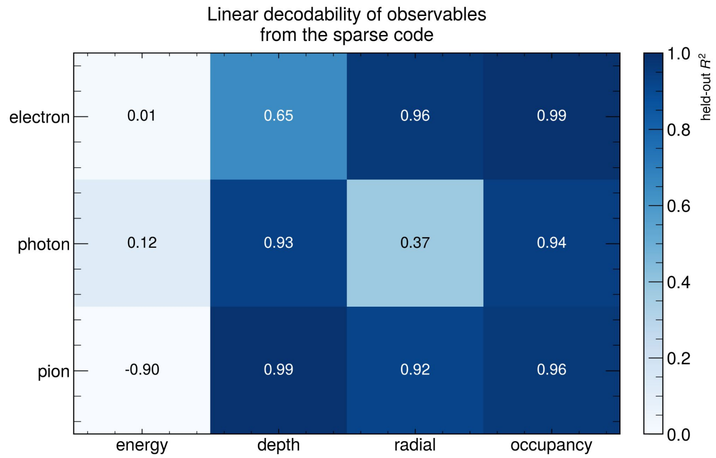


A Mechanistic Detour

What does the model encode internally?

We use sparse autoencoder (SAE) to find out which observables are linearly decodable. SAEs are trained only to reconstruct the bottleneck activation.

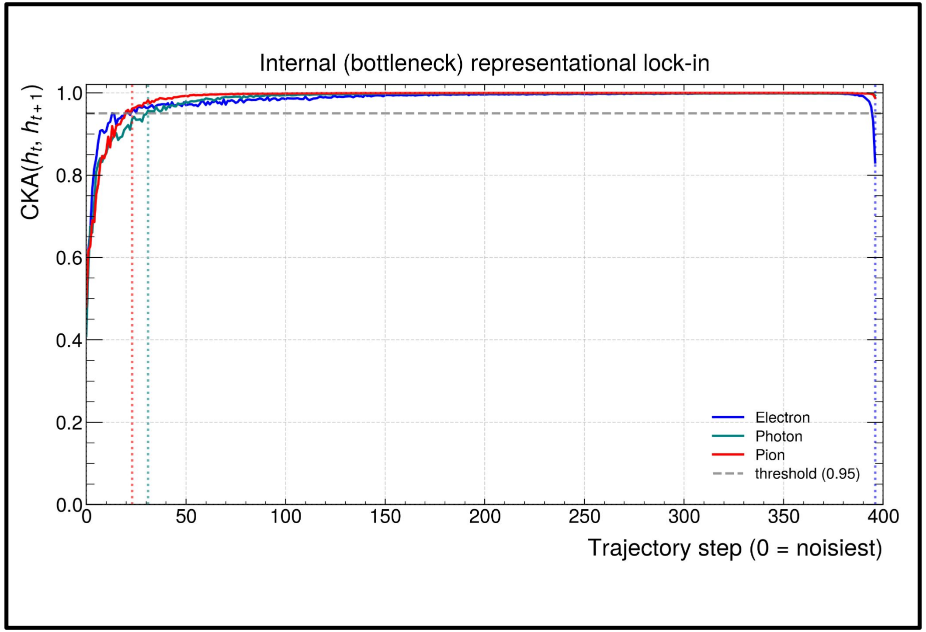
Energy is not linearly decodable!



Centered Kernel Alignment (CKA)

CKA is a tool to estimate the overlap between two latent representation. Being a kind of cosine similarity, a value close to unity would imply large overlap.

Here we find the overlap between internal representations of the same model across steps. The question we try addressing is “when does the representation lock-in happen?”



Summary

- Inspected how the diffusion model predict the physics at each step.
- Plotted Emergence Curves at Matching Energy to study the lock-in steps for different observables for each particle type.
- Performed FFT Analysis to prove Spectral Emergence phenomenon in Diffusion models
- Used SAE to find linearly decodable physics.
- Used CKA to estimate lock-in of internal representation.
- The model learns “coarser” things better, the true PDFs of nHits and energy distribution does not agree well with the data.

Interpreting CaloDiffusion: when physics emerges and how it is represented

Bhagya S.Babu
Department of Physics
Indian Institute of Technology, Hyderabad

Avik Das
Department of High Energy Physics
Tata Institute of Fundamental Research
avik.das@tifr.res.in

Dikshita Hazarika
Department of High Energy Physics
Tata Institute of Fundamental Research, Mumbai

Saurabh Rai
Department of Physics
Indian Institute of Technology, Kanpur

Anju Sharma
Department of Physics
Indian Institute of Technology, Hyderabad

Vishal S
Department of Physics
St. Joseph's University, Bengaluru

Khushbu Zala
Department of Basic Science and Humanities
MPSTME, NMIMS University

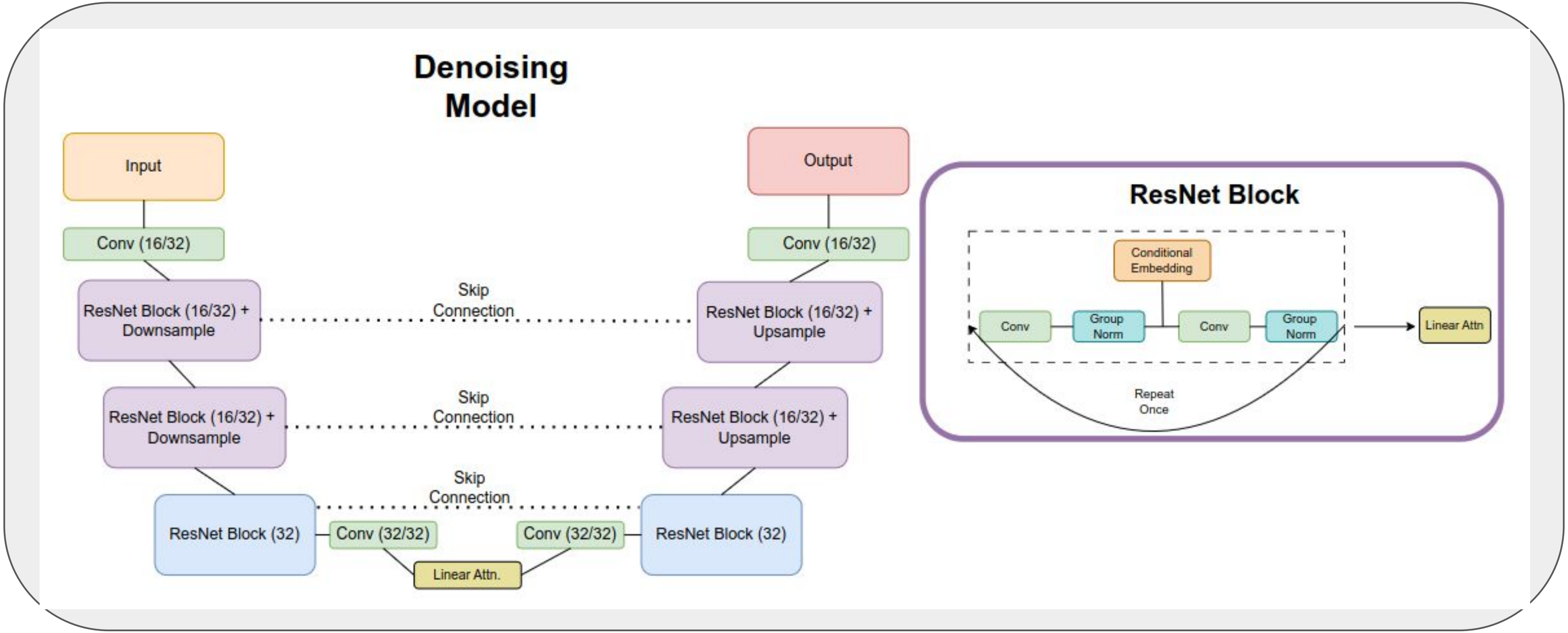
Liv Våge
Department of High Energy Physics
Princeton University

Abstract

Diffusion models have become one of the most powerful classes of generative models, and their use across the sciences is growing rapidly. Unlike most generative models, they expose their generation process: the sample can be inspected at every denoising step, making the generation directly observable. This has enabled a growing body of interpretability work in machine learning and other domains, but it has not yet been applied in high energy physics. We study CaloDiffusion, a diffusion model for calorimeter shower simulation, using both behavioural techniques and mechanistic metrics. This is a demanding test case: calorimeter showers are sparse and structurally unlike the natural images for which these behaviours were characterised. Some findings confirm general diffusion behaviour — coarse, global structure such as energy resolves early. Others are less expected: the model's internal representation of the final shower forms early in the trajectory, well before that structure is visible in the output. Beyond the specific results, these techniques offer a way to see when and where a diffusion surrogate captures the physics it is meant to reproduce — information relevant to trusting and accelerating such models in simulation pipelines.

BACKUP

Network Architecture



A schematic of network architecture. Here the the number in parentheses are the number of filters used in that module for the network.

https://arxiv.org/pdf/2308.03876

Longitudinal Shower Profiles

Visualize the average energy deposited across successive calorimeter layers **$< E_{layer} >$.**