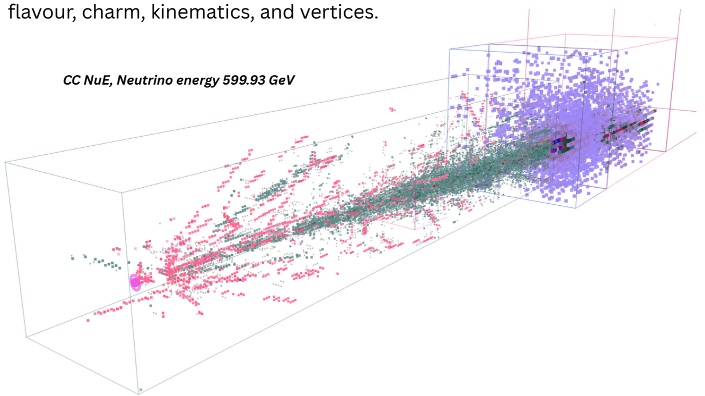
Towards foundation-style models for energy-frontier heterogeneous neutrino detectors via self-supervised pre-training

Fabio Cufino, *Saúl Alonso-Monsalve, Umut Kose, Anna Mascellani, André Rubbia*

1. Physics Challenge - Energy-frontier neutrino events are dense, sparse, and heterogeneous

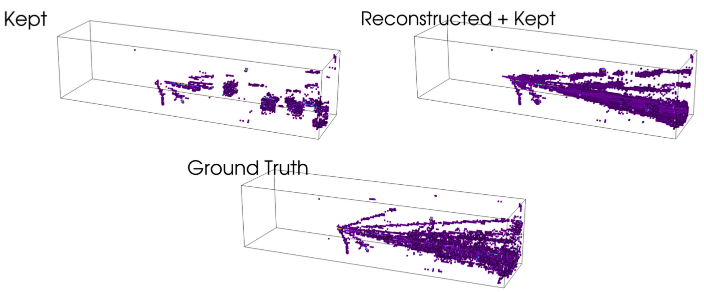
TeV-scale forward neutrino interactions produce sparse, highly collimated, and overlapping detector signatures, making event reconstruction a representation-learning problem. FASERCal is a challenging case: the model must extract information from a very sparse 3DCal volume and combine it with ECAL, AHCAL, and muon-spectrometer data to reconstruct flavour, charm, kinematics, and vertices.



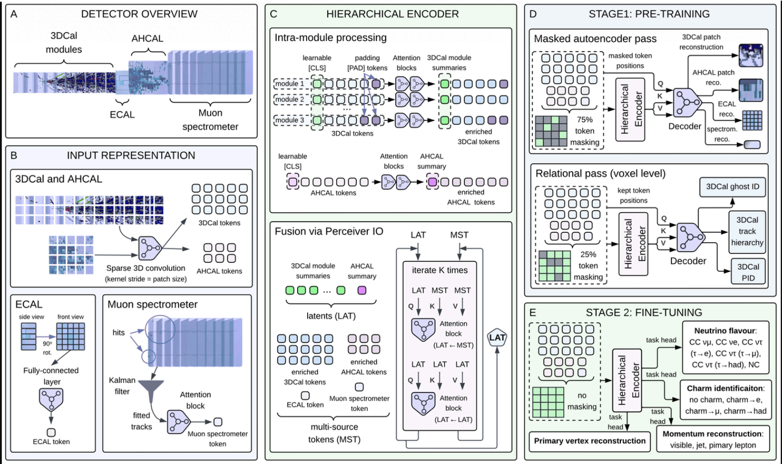
2. Model & Pre-training - A sparse multimodal encoder learns detector representations

We use a sparse ViT-like encoder designed for heterogeneous detector data. Sparse 3D convolutions convert active 3DCal and AHCAL voxels into patch tokens. Perceiver-IO fusion combines calorimetric tokens with ECAL and muon-spectrometer streams. Pre-training combines:

Masked reconstruction: masking 75% of occupied calorimeter patches, reconstructs missing occupancy and charge.



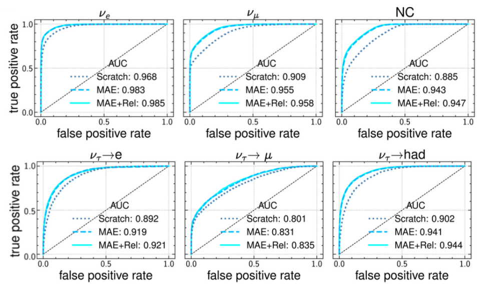
Relational voxel-level learning: The encoder also predicts ghost labels, interaction hierarchy, and particle category. These targets force the representation to learn local detector semantics, not only global geometry.



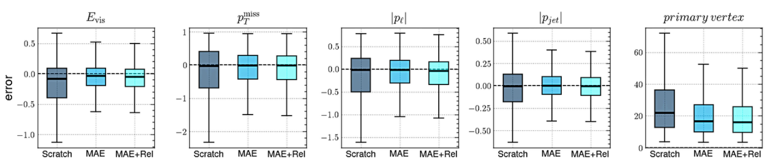
3. Downstream Results - Pre-training improves the hardest physics channels

All the fine-tuned models use the same architecture and differ only in initialization of the encoder: training from Scratch, MAE pre-training, or MAE+Rel pre-training:

- MAE already improves dominant channels with respect to Scratch, while MAE+Rel gives the largest gains in topologically complex and low-yield signatures.

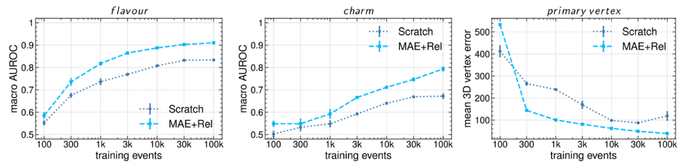


MAE pretraining improves vertex and kinematic reconstruction, with MAE+Rel giving the strongest overall gains: smaller primary-vertex displacement, tighter and better-centered E_vis and p_jet residuals, especially in CC and NuTau->had channels, while missing p_T and p_lep improve more moderately and remain broad for intrinsically difficult tau samples.



4. Data Efficiency & Transfer - Towards foundation-style detector models

The strongest evidence for a foundation-style behaviour is not only better performance, but reusable structure. With roughly 10^3 labelled events, the MAE+Rel encoder already exceeds the flavour-classification performance of a scratch model trained on about 10^4 events. Vertex reconstruction also improves strongly, with the mean 3D vertex error reduced from about 240 mm to 100 mm at 10^3 labelled events.



Data-efficiency study: *MAE+Rel is compared with Scratch across training budgets from 100 to 100k events. Error bars show the spread over three random seeds. Macro area under the receiver-operating-characteristic curve (macro-AUROC) for flavour and charmed-quark classification, and mean three-dimensional vertex error.*

In-domain gains are encouraging — but the foundation-style claim holds if the representation transfers to other detectors, tasks and energy scales. We test two target domains at increasing distance from FASERCal:

- SuperFGD-like: Plastic scintillator — close technology, GeV single particles
- PILArNet LArTPC — different technology, energy regime and a different particle-ID benchmark.

What we reuse: the transferable core of the source encoder — attention blocks, latent cross-/self-attention, normalisation layers, global query token (inherited when shapes & semantics match).

What we re-initialise: detector-specific patch embeddings, positional encodings, detector-branch embeddings, and the task heads.

SuperFGD class	Best previous	MAE+Rel
p	0.907	**0.943**
$\pi^{\pm}$	**0.643**	0.609
$\mu^{\pm}$	0.595	**0.748**
$e^{\pm}$	0.772	**0.787**

SuperFGD transfer: *Per class accuracy from the confusion-matrix diagonal*

PILArNet transfer: *The transferredMAE+Rel encoder matches or exceeds the strongest published ensemble baseline on a different detector technology and particle-ID task.*

PILArNet benchmark	Best previous	MAE+Rel
Single-particle Acc.	0.9014	**0.9154**
Single-particle AUROC	0.842	**0.891**
Multi-particle Acc.	0.9644	**0.9662**
Multi-particle AUROC	0.944	**0.951**

Beats the strongest published baseline!