Université Paris Cité

Foundation-like models for Astroparticle Physics Experiments

Signal Identification via Machine Learning Techniques

–

Maximilian Eff

Laboratoire Astroparticule et Cosmologie
Data Intelligence Institute of Paris

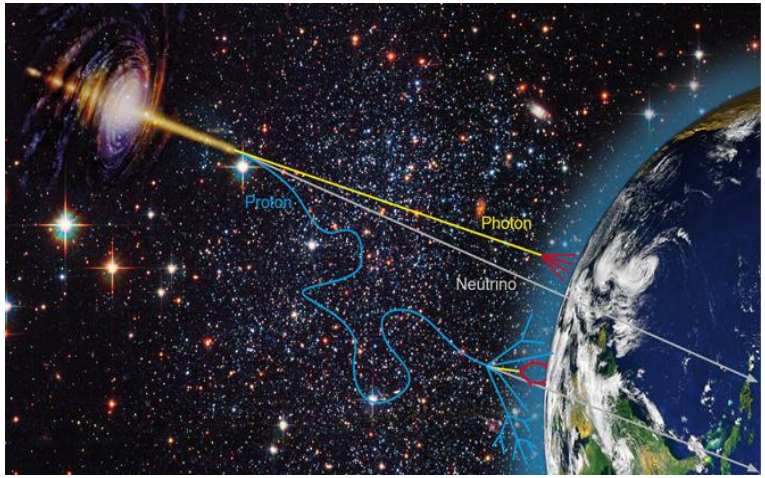
Current Challenge in Astroparticle Physics Analyses:

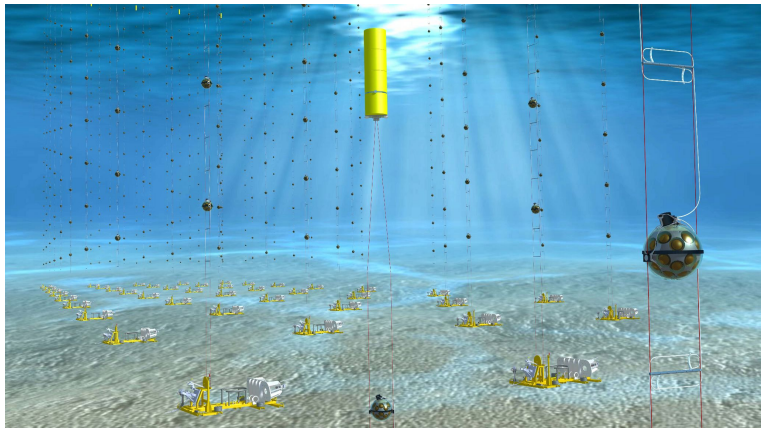
- Huge amount of Data is recorded
- Data is dominated by various types of noise
- Blows up processing time

Opportunity to accelerate analyses with Machine Learning

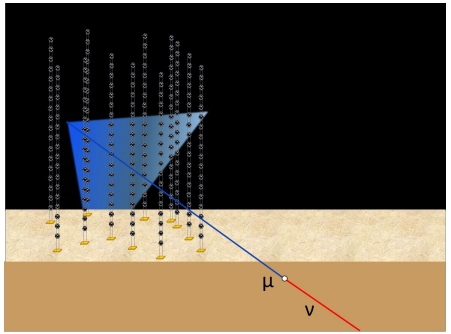
Project:

- Develop a Machine Learning pipeline for astroparticle physics experiments.
- Case study: KM3NeT

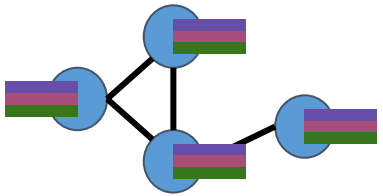




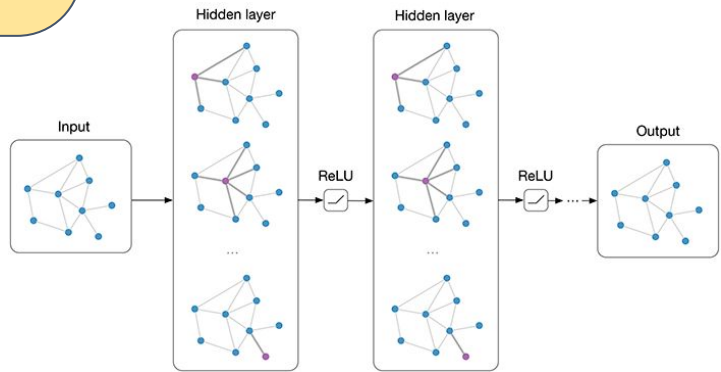
Data can be best described by a Graph



- Nodes with features (e.g. x, y, z, t, A, ...)
- Edges defined via
  k nearest neighbors
  in Minkowski Spacetime
  (Captures spatial and temporal patterns)

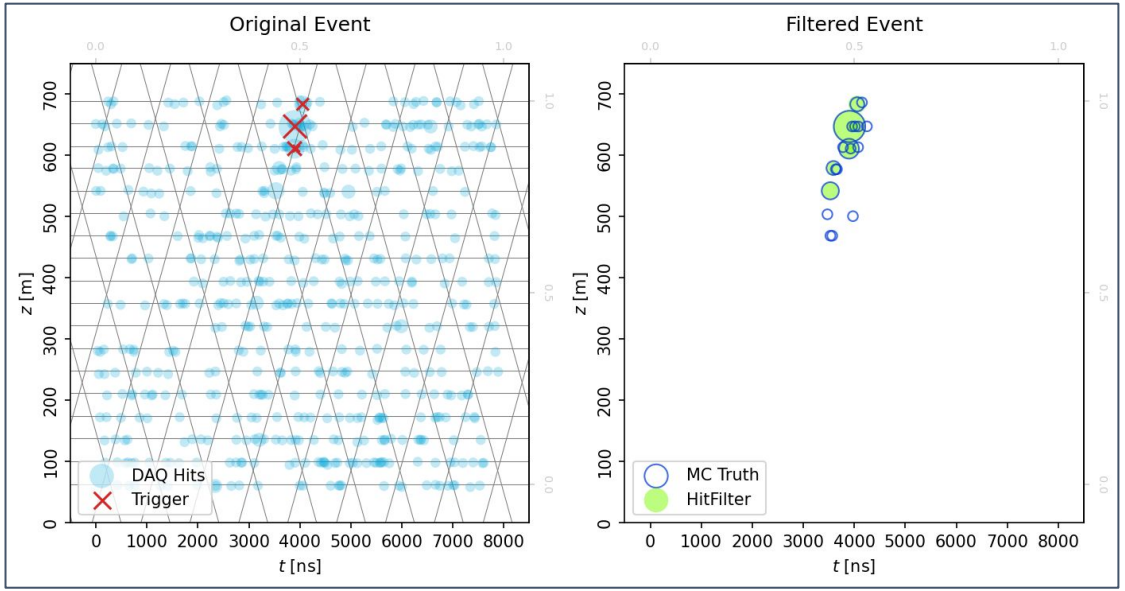


⇒ Use **Graph Neural Networks**



Challenge in KM3NeT:
Snapshot events are dominated by noise (K40)

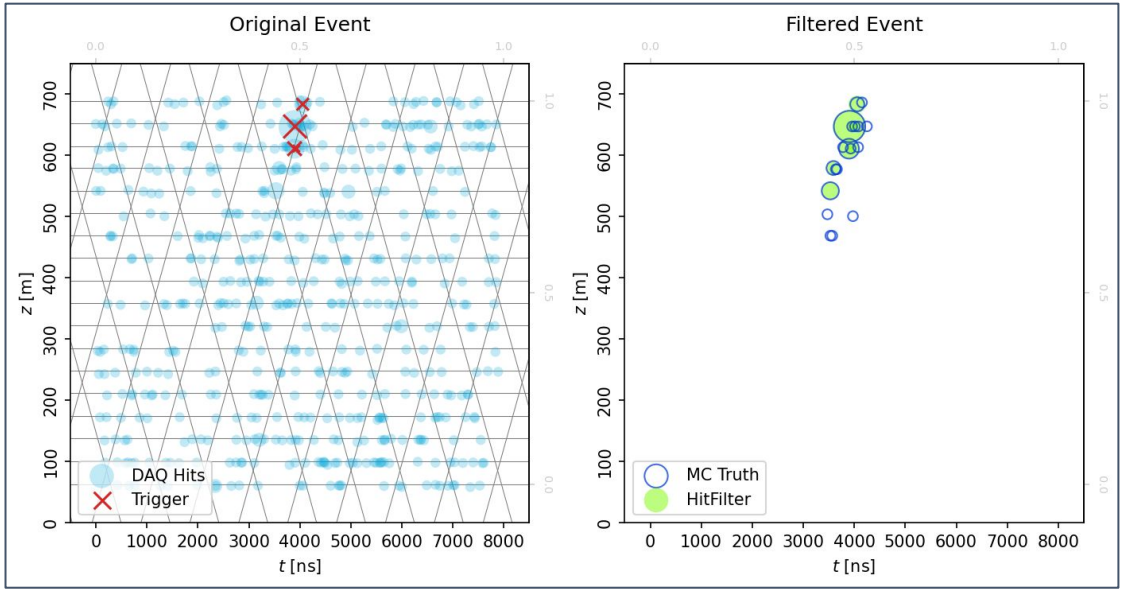
Hit Filter:
Train a model to discriminate signal from noise hits



Continue analysis with cleaned event

Challenge in KM3NeT:
Snapshot events are dominated by noise (K40)

Hit Filter:
Two Sub-Models
1. **Feature Extractor**
2. **Hit Classifier**



Continue analysis with cleaned event

Hit Filter: Feature Extractor

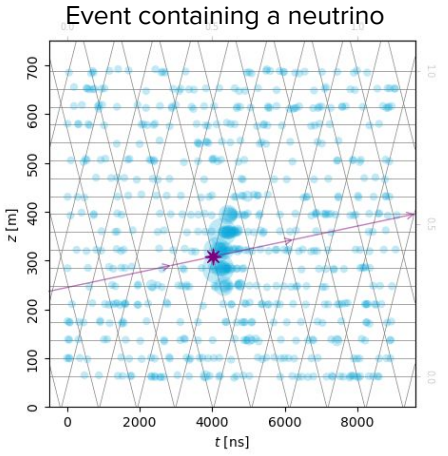
1. Feature Extractor

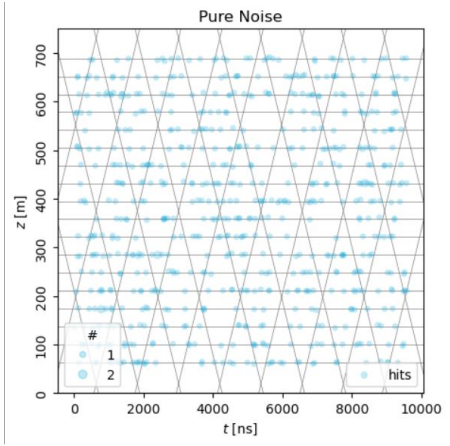
Train Binary Classifier to identify Events as “Signal” or “Noise”

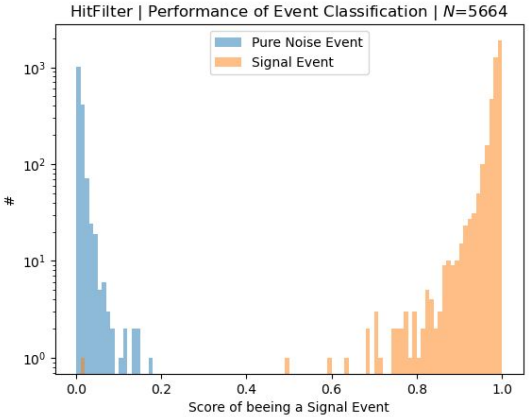
Obtain internally, high-dimensional & information rich representation of

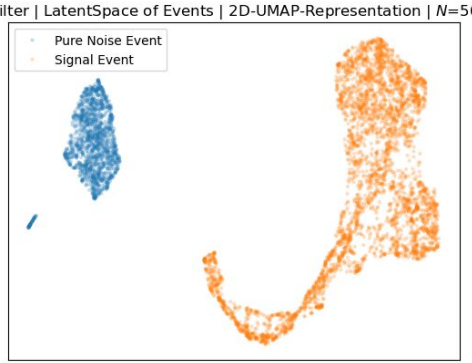
- the full Event
- each individual Hit

Building on this result,
explore how this applies to the pattern recognition of individual hits.

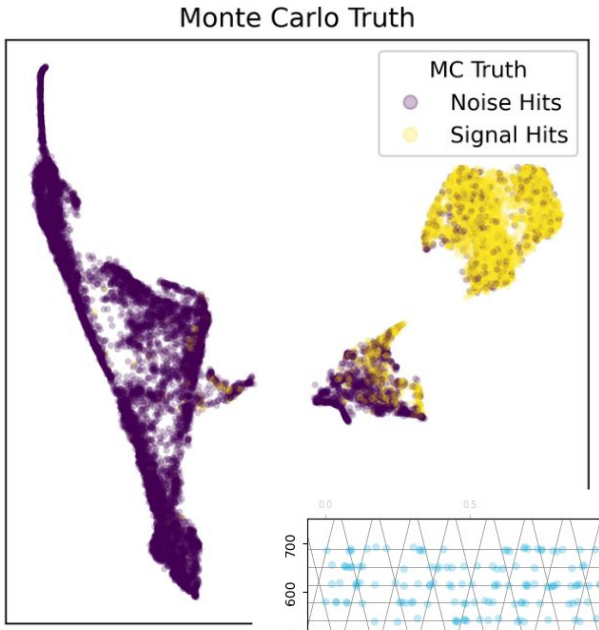


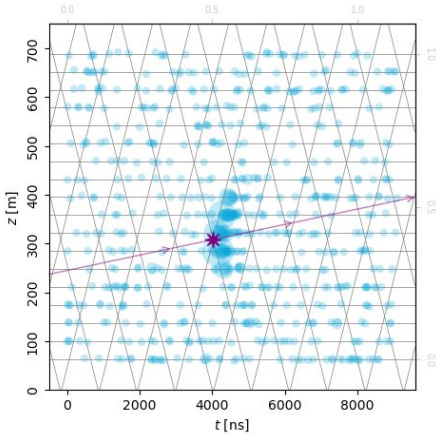






Hit Filter: Latent Space of Hits





The latent space encodes both the representation of individual signals from the devices and the aggregated information within an event.

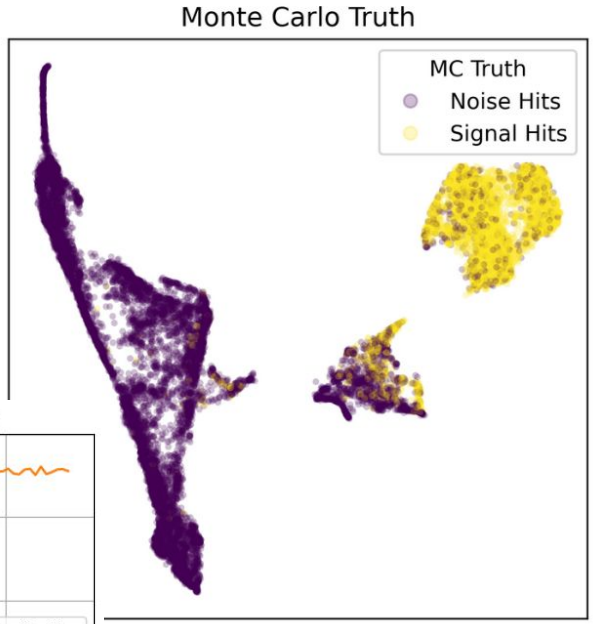
The remarkable classification accuracy of events containing neutrinos compared to pure noise events aligns closely with the strong separation achieved for individual signals in the devices.

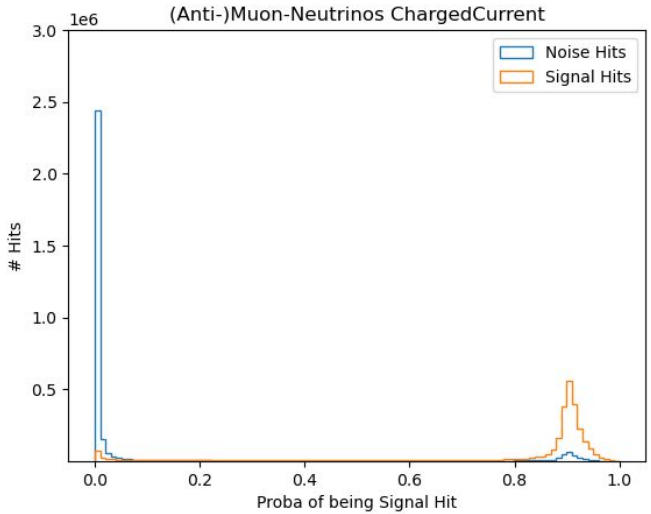
The next step involves the identification the neutrino signals, which can be accomplished using a clustering approach or a straightforward supervised classifier.

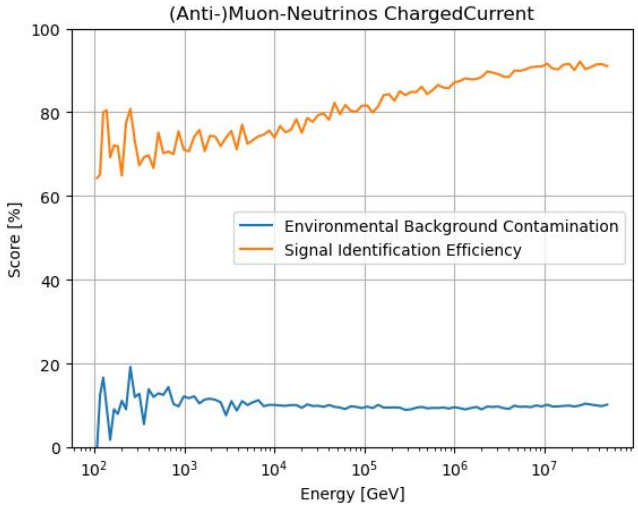
2. Hit Classifier

Train Binary Classifier to identify Hits as “Signal” or “Noise”

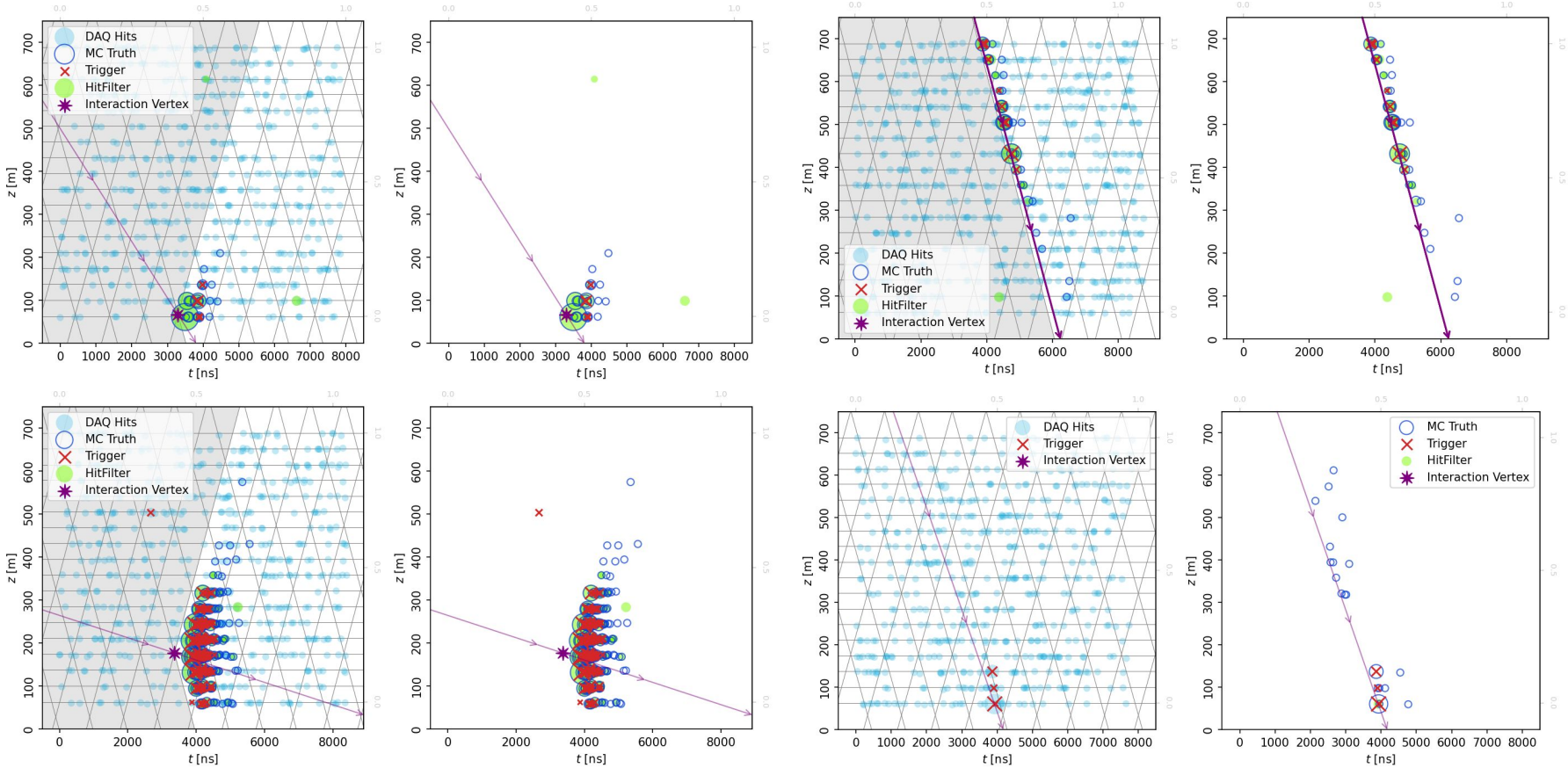
Starting directly from the latent space of the individual Hits





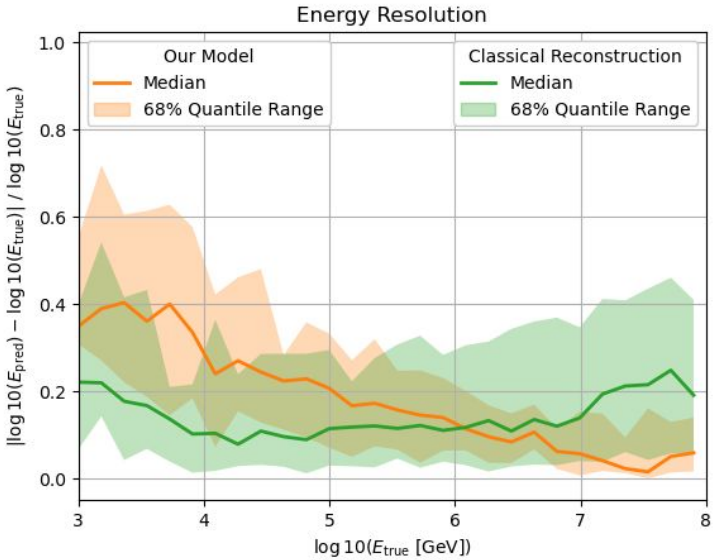


Noise Filter: Examples



Pipeline Components

1. **Environmental Noise Filter**
   Signal vs. noise discrimination
2. **Neutrino Event Selection**
   Atmospheric muon rejection
3. **Parameter Reconstruction**
   Energy (E), direction (θ), and azimuthal angle (φ)



Thank you for your attention!