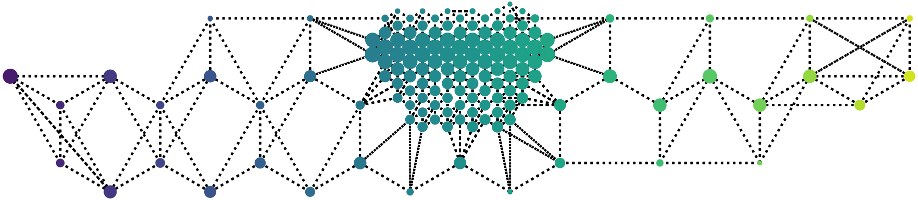
Direction and energy reconstruction with uncertainty quantification for GRAND using graph neural network

Warsaw collaboration meeting

Arsène Ferrière



Advisors : Aurélien Benoit-Lévy, Olivier Martineau

Why Machine Learning ?

- Work directly from data-like inputs
- Fast and reliable data analysis : Direction and Energy reconstruction
- Background event rejection

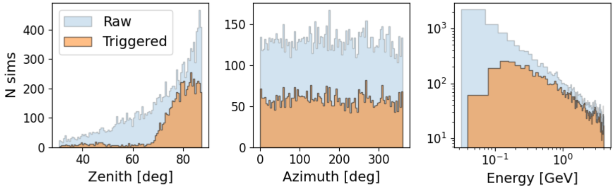
But, need for labelized data $\Longrightarrow$ Training on simulations = complications :

- How to make sure the reconstruction will work on real data ?
- How to properly simulate the measurement noise ?
- What level of confidence to give to the predictions ?

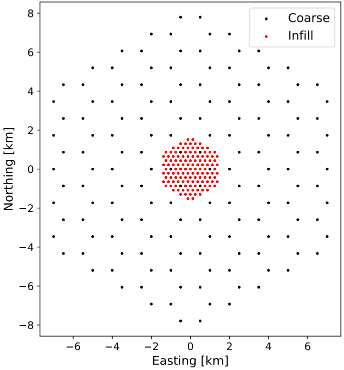
Developing reliable AI with uncertainty quantification.

Simulation set

1. DC2 ADC simulations
2. Low galactic noise + 5ns jitter
3. Extract time and amplitude of maximum
4. Trigger condition : Signal $\geq$ 30ADC (5σ) for more than 5 antennas.

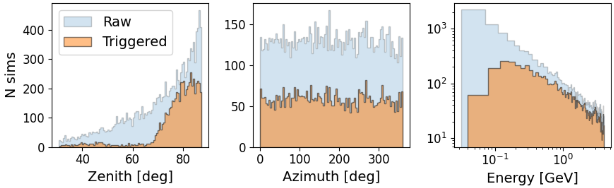


Antenna layout

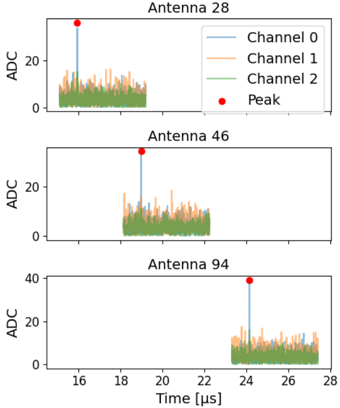


Simulation set

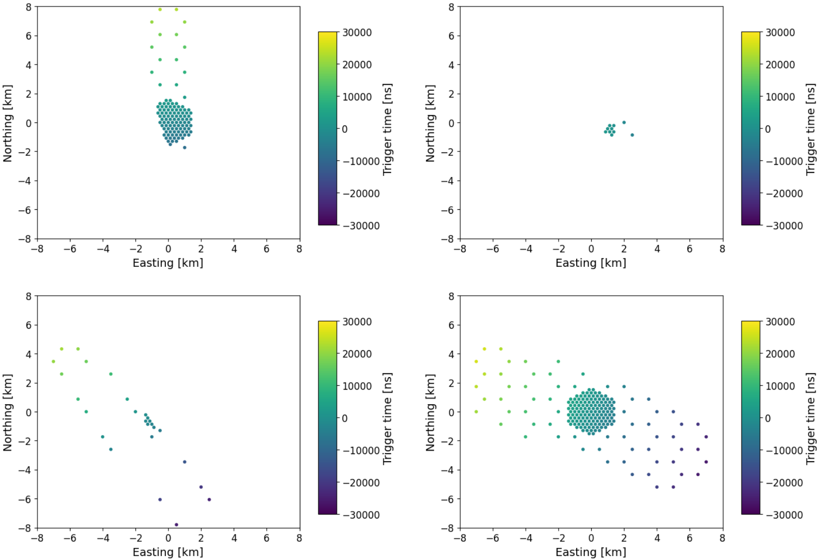
1. DC2 ADC simulations
2. Low galactic noise + 5ns jitter
3. Extract time and amplitude of maximum
4. Trigger condition : Signal $\geq$ 30ADC (5σ) for more than 5 antennas.



Trace example



Event examples



Wide variety of events:
Varying footprint size, shape, antenna multiplicity
Constraint on the possible NN architecture

GNN explanation

Key Steps in a GNN Layer: *[Scarselli et al.]*

- **Message Passing:** Each node aggregates information from its neighbors.
- **Aggregation:** Information is combined using a function (e.g., sum, mean, max).
- **Update:** Node embeddings are updated using a neural network (e.g., MLP).

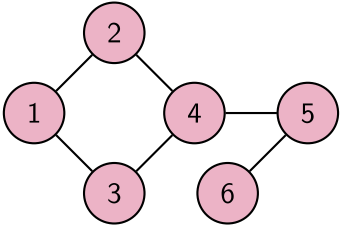
Mathematical Formulation:

$$h_i^{(l+1)} = \sigma\big(\mathrm{AGG}_{j\in\mathcal{N}(i)}\big(f_{\omega_l}[h_i^{(l)}, h_j^{(l)}]\big)\big),$$

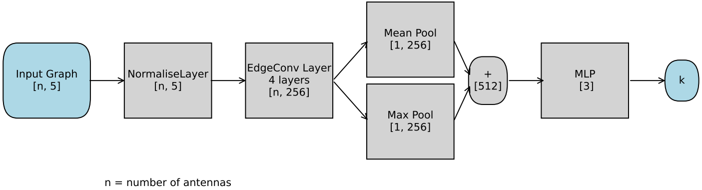
where:

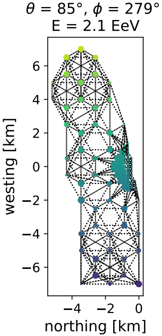
- $h_i^{(l)}$: Node embedding at layer l.
- $\mathcal{N}(i)$: Neighbors of node i.
- f_{ω_l}: Is a function with trainable parameters (MLP).
- σ: Non-linear activation (e.g., ReLU).

Visualization:



Architecture of Graph Neural Networks





Architecture of graph neural networks:
780k parameters (can be reduced to 60k)

8 nearest neighbours

EdgeConv Layer *[Abbasi et al.]*:

$$x_\mu^{k+1} = \sigma\left(\frac{1}{|N_\mu|}\sum_{\nu\in N_\mu} f_\theta(x_\mu^k\,,\, x_\nu^k)\right)$$

Training procedure

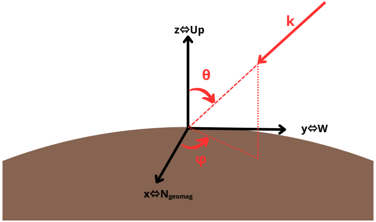
Direction reconstruction

Reconstructing θ, ϕ : No direct reconstruction.
Better in cartesian coordinates: **k**

- Loss function $L(\omega) = \mathbb{E}[||\mathbf{k}_{\mathrm{pred}} - \mathbf{k}_{\mathrm{sim}}||^2]$
- 5 input features: 3 antenna coordinates, arrival time, signal amplitude (normalised)
- **Training set**: 4937 events, **Validation set**: 928 events
- 10 models or more trained with different initialization/Train dataset order

Ensemble methods:
With the N models, 1 "meta model".

$$\mathbf{k}_{pred} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{k}_{pred,i}$$

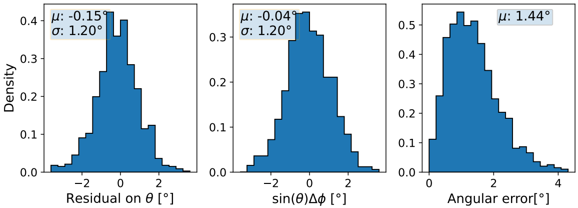


Parameters convention

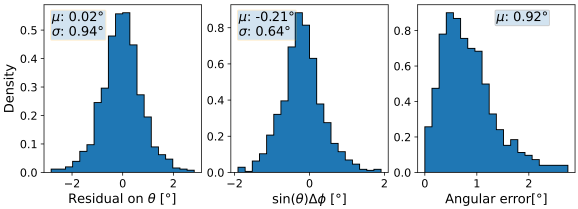
Performance

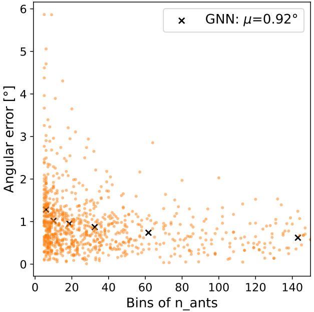
Direction reconstruction

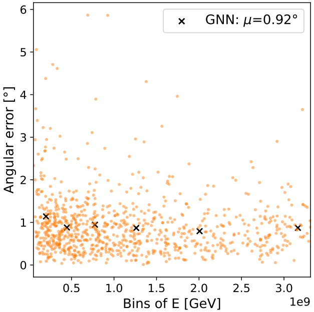
Single model predictions



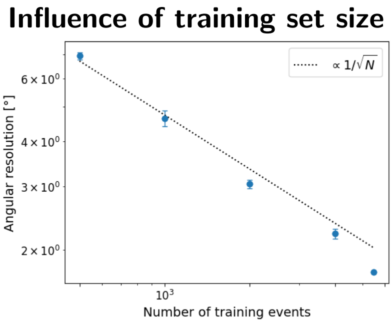
Ensemble method







Training set size



Evolution of the error when increasing the training set size.

Extrapolation is no reason yet :
Suggests that larger training set => Better performances.

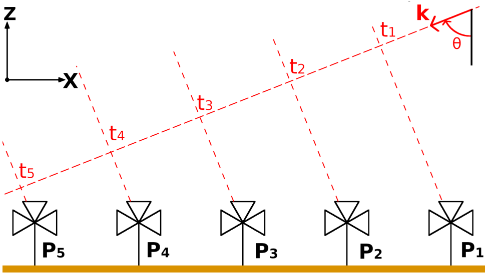
To avoid using more simulations : Feed physical knowledge to network.

Physical knowledge

What is PWF ?

Assume the wavefront is planar:

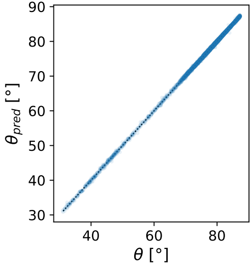
Linear relation between timings **T**, position **P** and propagation vector **k**:

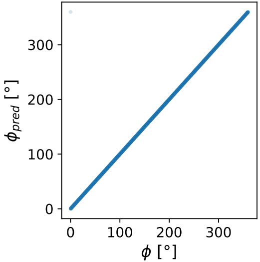


Planar wavefront

Solution :

$\mathbf{k} = \operatorname{argmin}(\mathbf{T} - \mathbf{P}\mathbf{k})^{\mathrm{T}}(\mathbf{T} - \mathbf{P}\mathbf{k})$ s.t. $\|\mathbf{k}\| = c$

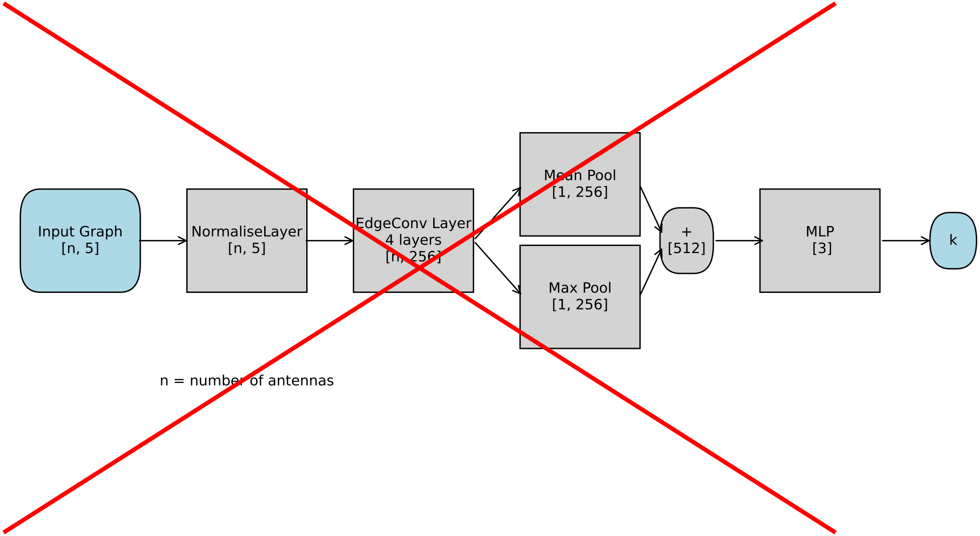




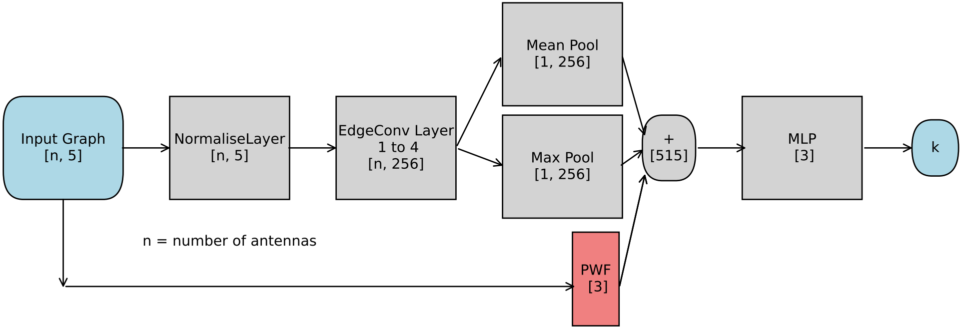
Github repository

[Ferriere et al. 2025]

New Architecture of Graph Neural Networks : pGNN



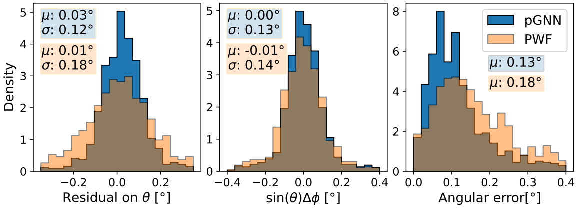
New Architecture of Graph Neural Networks : pGNN



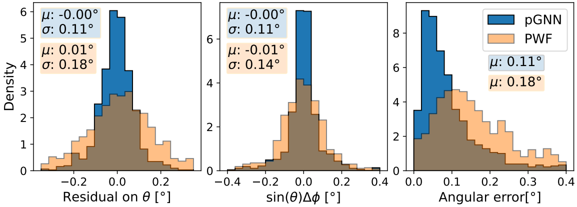
Performance

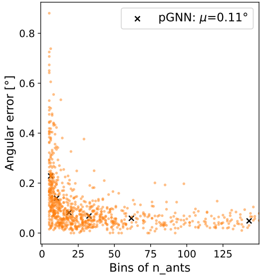
Direction reconstruction

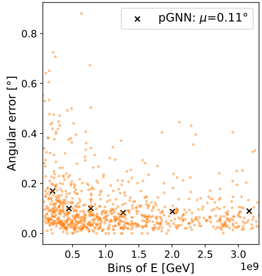
Single model predictions



Ensemble method



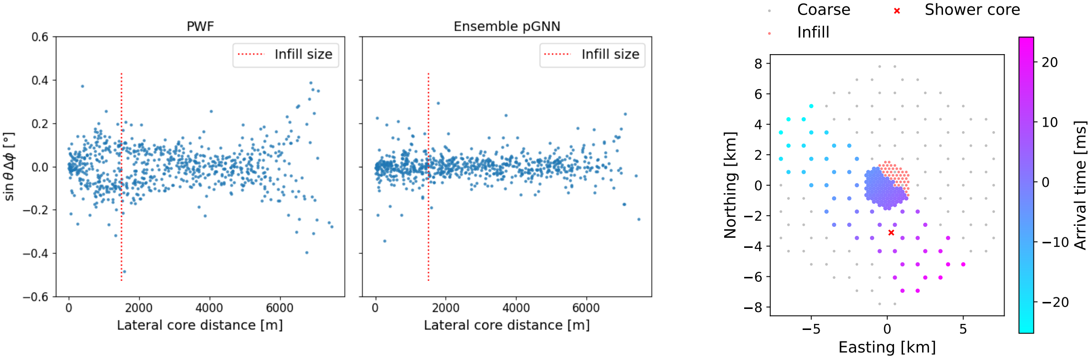




$\times$ 10 improvement compared to raw GNN

Correction of the PWF bias

PWF is known to be biased when asymmetries in the antenna footprint.



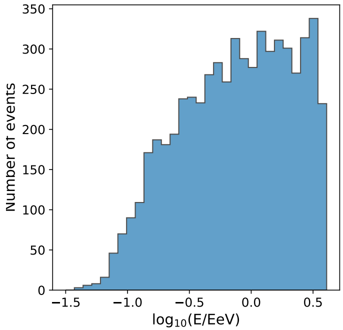
[Ferrière et al. in prep]

Training procedure and performances

Energy reconstruction - *Work in progress*

Reconstructing E ? No : Uneven distribution. Better $\log E$.

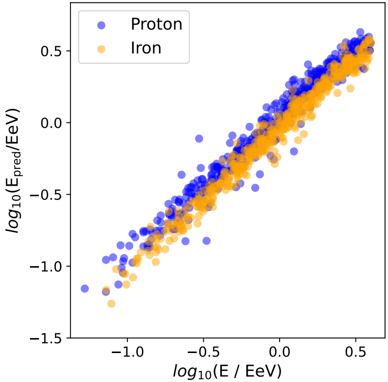
- Loss function $L(\omega) = \mathbb{E}[\log\left(\frac{E_{pred}}{E_{target}}\right)^2]$
- 5 input features: 3 antenna coordinates, arrival time, signal amplitude (normalised)
- **Training set**: 4937 events, **Validation set**: 928 events
- 10 models or more trained with different initialization/Train dataset order
- Secondary input : **PWF** + polynomial fit of energy from (average amplitude, maximum amplitude, number of antennas, PWF zenith angle)

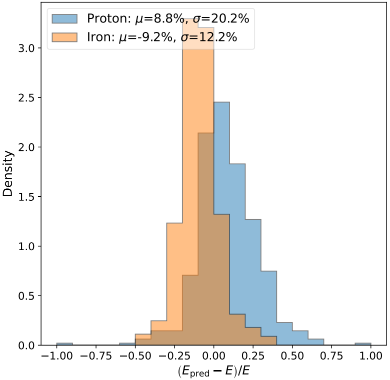


$\log E$ distribution

Energy resolution

Energy reconstruction - *Work in progress*



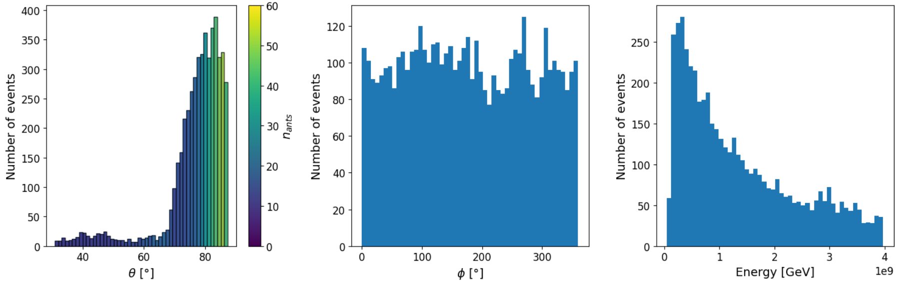


Total energy resolution : 19.5%
For primary energy!

On more realistic simulation

Direction reconstruction

On sims with the new RF chain and effective length: mult $\geq$ 5, trigger at 85 ADC:

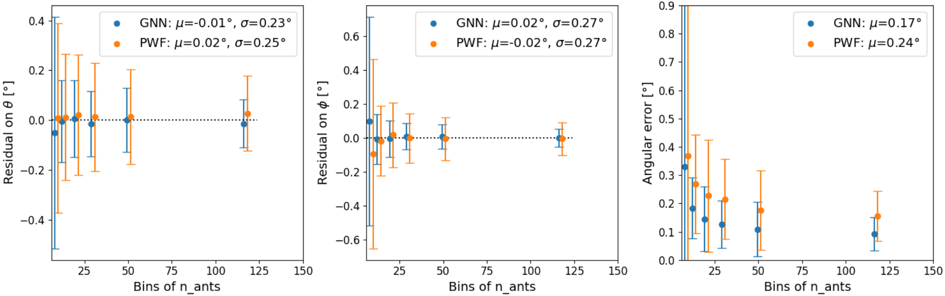


Distribution of triggered events

On more realistic simulation

Direction reconstruction

Difficult training as much lower SNR $\Longrightarrow$ *less dus per event*

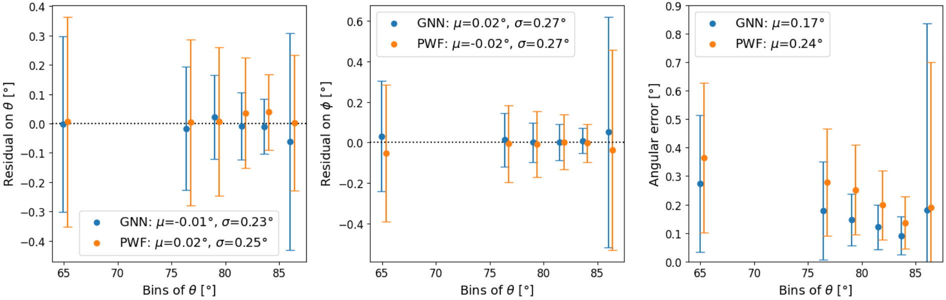


Performance vs antenna multiplicity (mult ≥ 6 : 91% of events)

On more realistic simulation

Direction reconstruction

Difficult training as much lower SNR $\Longrightarrow$ *less dus per event*

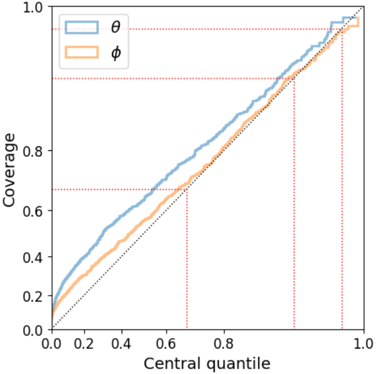


Performance vs incoming angles (mult ≥ 6 : 91% of events)

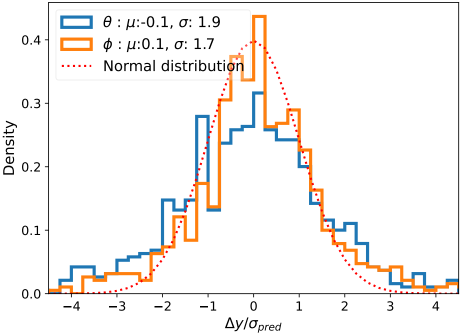
Uncertainty estimation

Direction reconstruction

Under Gaussian assumption : $\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$ and $\phi \sim \mathcal{N}(\mu_\phi, \sigma_\phi^2)$. With μ_θ the mean of our 30 predictions of θ and σ_θ their std's.



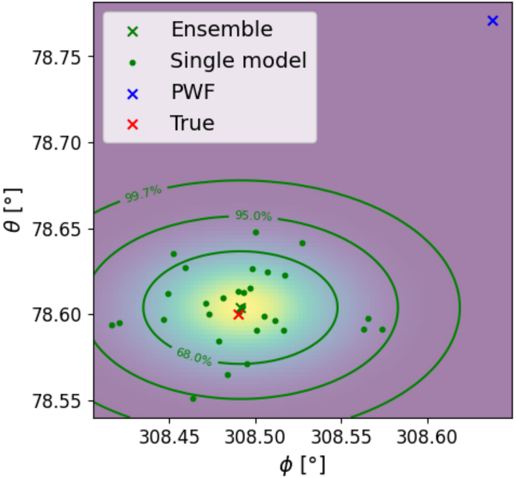
PP plot for our uncertainty estimator. We slightly overestimate our uncertainties for θ



Distribution of normalized residuals

Uncertainty estimation

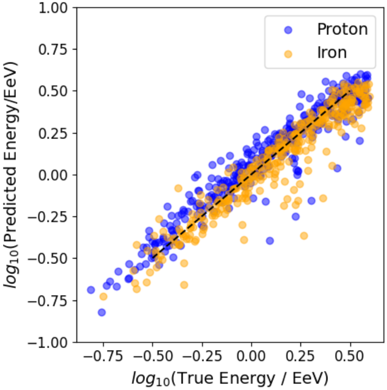
Direction reconstruction

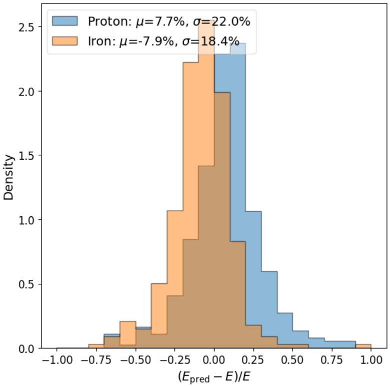


Example of reconstruction

Energy resolution

Energy reconstruction - *Work in progress*





Total energy resolution : 21.7%
For primary energy!

Left to do

Direction reconstruction

1. Improve energy reconstruction on the new set of simulation
2. More robust testing on triggering effects/smearing effects
3. Add features (spectral slope, polarization)

On real Data

Conclusion

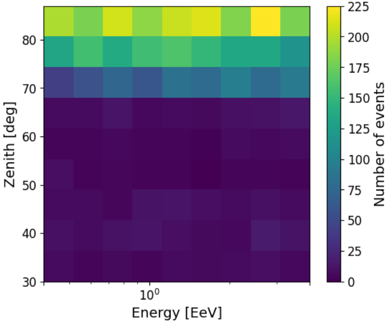
Results - Direction on data-like simulation

- With GNN and ensemble methods, high direction reconstruction precision : $0.17°$ precision for $n_{ants} >= 6$.
- Slightly overconfident uncertainty estimation.

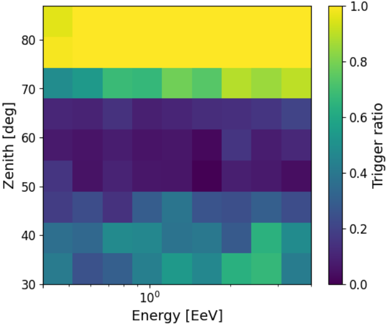
Results - Energy on simulation

- Energy resolution : 21.6%
- Uncertainty estimation not yet calibrated.

Backup slides : joint trigger distribution



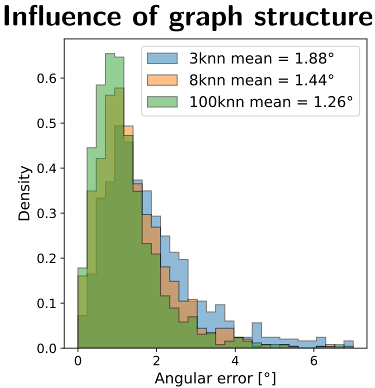
Joint trigger distribution

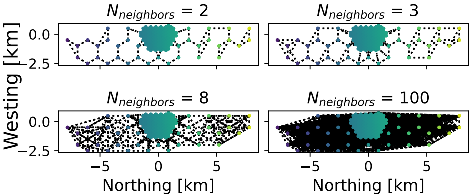


joint trigger ratio

Training set size and graph structure

Direction reconstruction



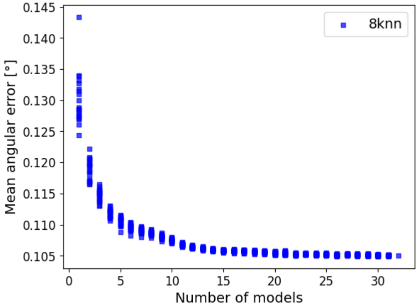


Degradation of performance when lowering number of neighbours

- if $N_{neighbors} = 3$: graph depth $= 8.18$
- if $N_{neighbors} = 8$: graph depth $= 3.9$
- if $N_{neighbors} = 100$: graph depth $= 1.14$

Ensemble size

We have a net improvement of the performances of the direction reconstruction with the same training size.



Evolution of the precision with the number of models