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

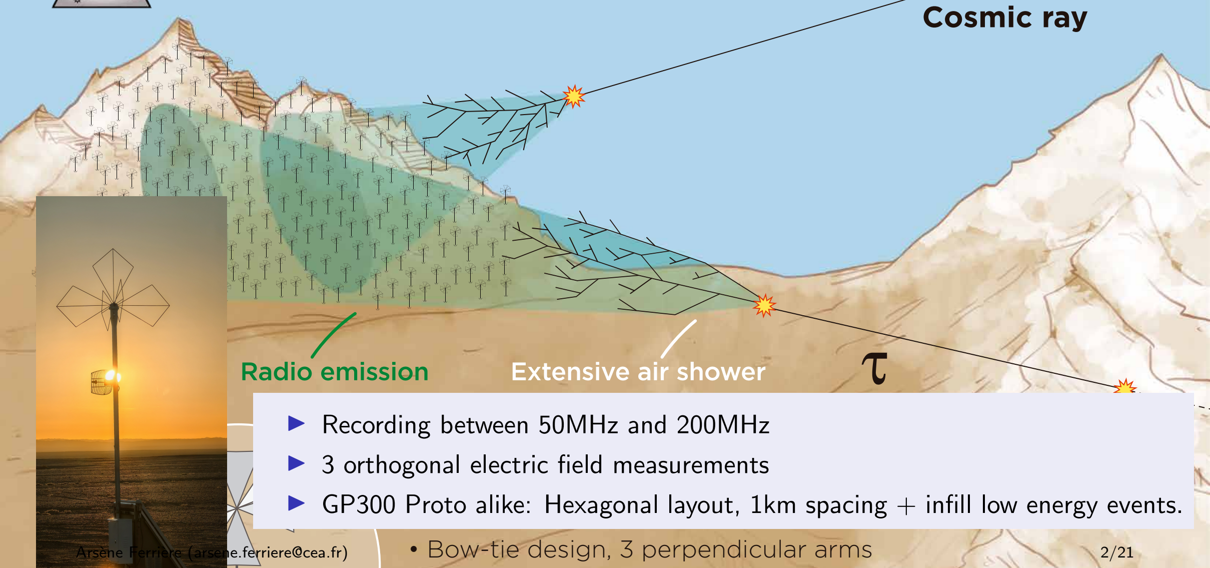
Workshop on machine learning for analysis of high-energy cosmic particles

Arsène Ferrière [1],[2] (arsene.ferriere@cea.fr)
on behalf of the GRAND collaboration
with contribution from Aurélien Benoit-Lévy[1]

[1] CEA-LIST, [2] LPNHE (Paris, France)

January 29, 2025

Giant Radio Array for Neutrino Detection



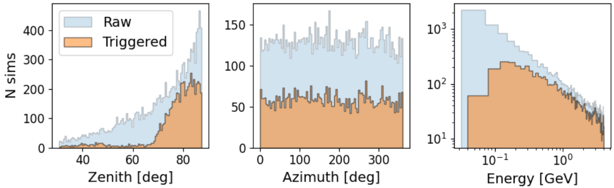
- Recording between 50MHz and 200MHz
- 3 orthogonal electric field measurements
- GP300 Proto alike: Hexagonal layout, 1km spacing + infill low energy events.

• Bow-tie design, 3 perpendicular arms

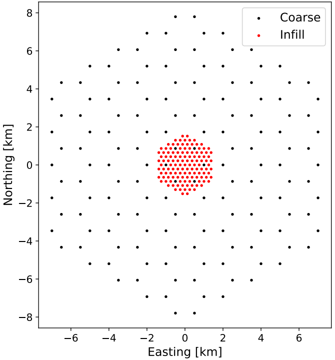
Simulation set

From each event:

1. ZHAireS simulations of proton and iron EAS.
2. 22 μV/m white noise mimicking galactic noise + 5ns jitter + RF chain
3. Compute Hilbert transform of the 2 horizontal components
4. Extract time and amplitude of maximum
5. Trigger condition : Signal $\geq$ 30ADC for more than 5 antennas.



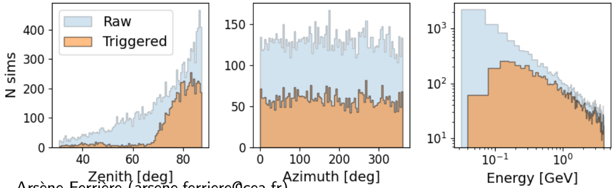
Trace example:



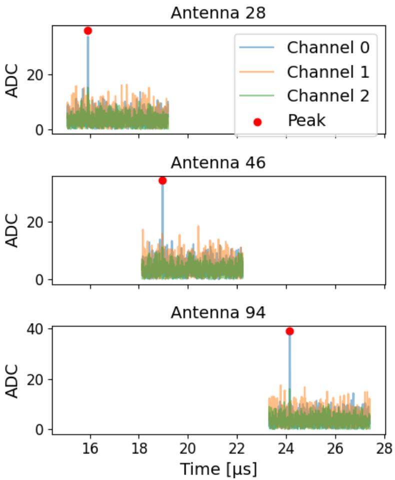
Simulation set

From each event:

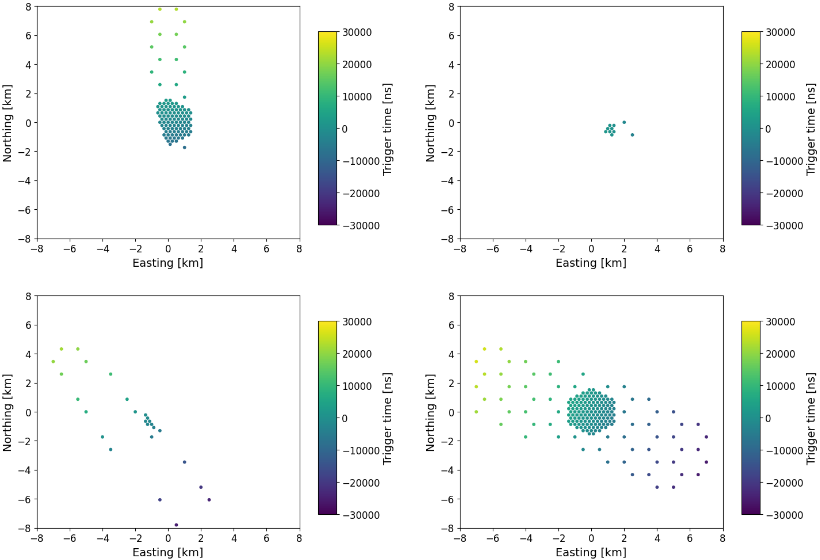
1. ZHAireS simulations of proton and iron EAS.
2. 22 μV/m white noise mimicking galactic noise + 5ns jitter + RF chain
3. Compute Hilbert transform of the 2 horizontal components
4. Extract time and amplitude of maximum
5. Trigger condition : Signal $\geq$ 30ADC for more than 5 antennas.



Antenna layout



Event examples



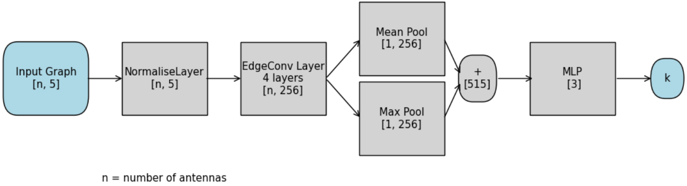
Wide variety of events:
Varying footprint size, shape, antenna multiplicity
Constraint on the possible NN architecture

Objectives

Aim of this work

- Reconstruct arrival direction using deep learning
- Reconstruct energy of primary using deep learning
- Estimate uncertainties for those reconstructions

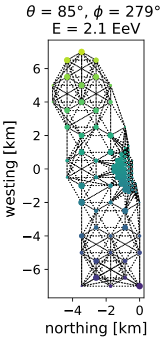
Architecture of Graph Neural Networks



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

EdgeConv Layer *[Abbasi et al.]*:

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



8 nearest neighbours

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).

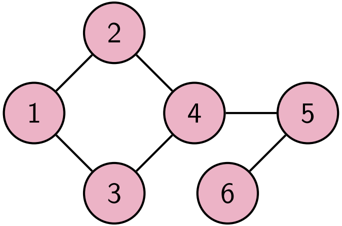
Mathematical Formulation:

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

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:

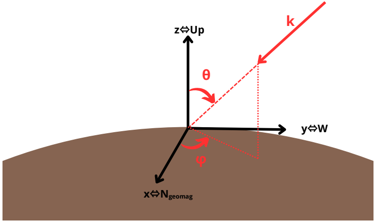


Training procedure

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



Parameters convention

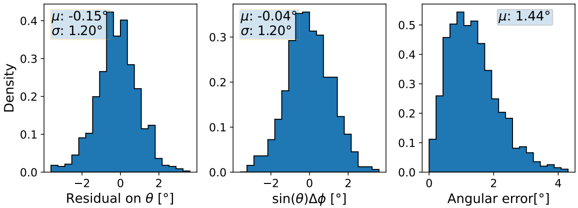
Ensemble methods:
With the N models, 1 "meta model".

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

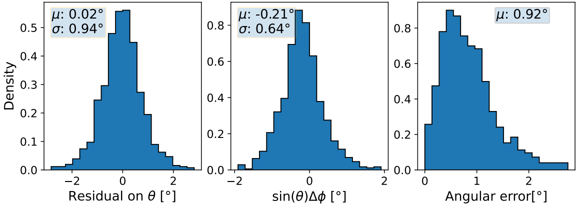
Performance

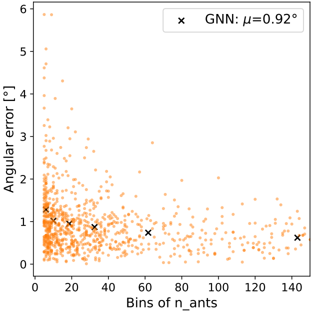
Direction reconstruction

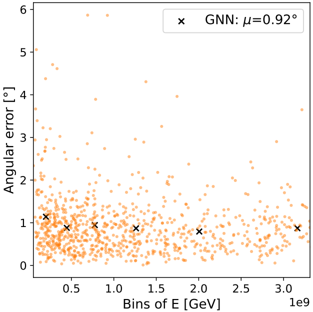
Single model predictions



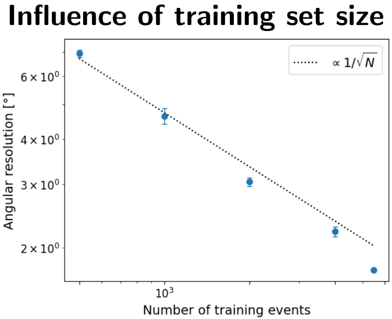
Ensemble method







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.

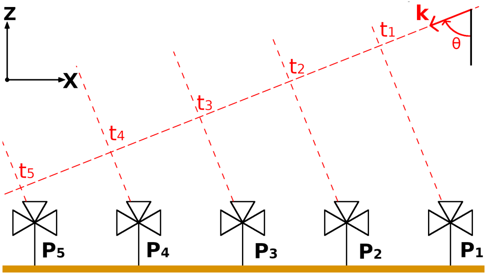
To avoid using more simulations : Feed physical knowledge to network.

Physical knowledge

What is PWF ?

Assume the wavefront is planar:

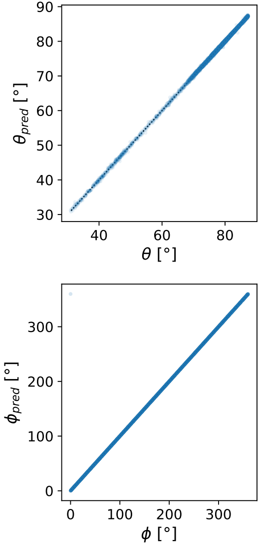
Linear relation between timings **T**, position **P** and propagation vector **k**:



Planar wavefront

Solution :

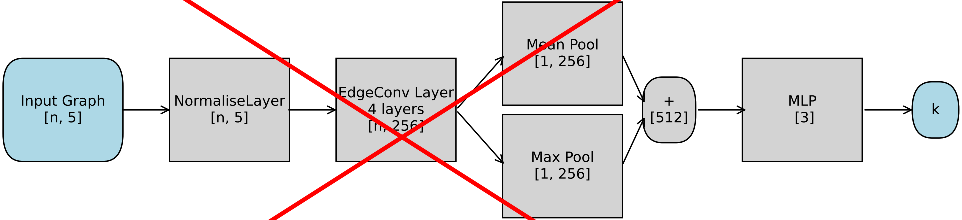
$$\mathbf{k} = \operatorname{argmin}(\mathbf{T} - \mathbf{P}\mathbf{k})^{\mathrm{T}}(\mathbf{T} - \mathbf{P}\mathbf{k}) \text{ s.t. } \|\mathbf{k}\| = c$$



Github repository

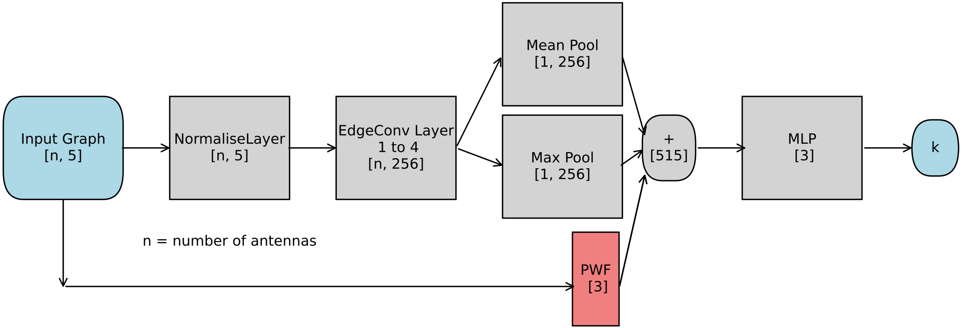
[Ferriere et al. 2025]

New Architecture of Graph Neural Networks : pGNN



n = number of antennas

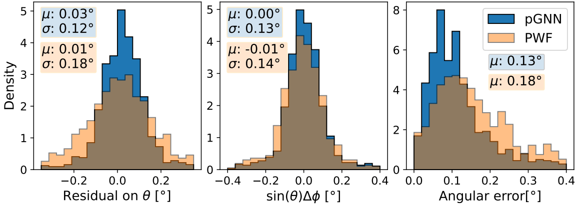
New Architecture of Graph Neural Networks : pGNN



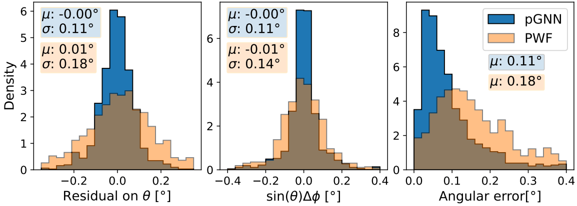
Performance

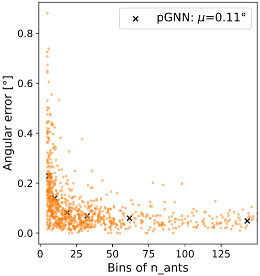
Direction reconstruction

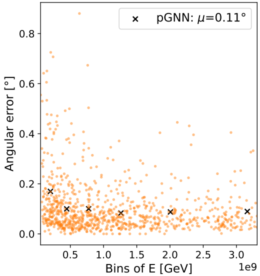
Single model predictions



Ensemble method





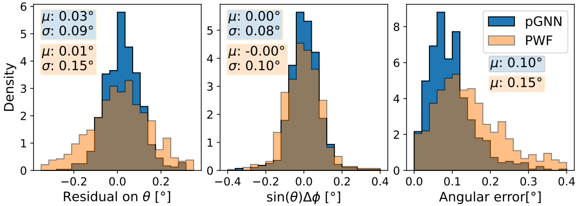


$\times 10$ improvement compared to raw GNN

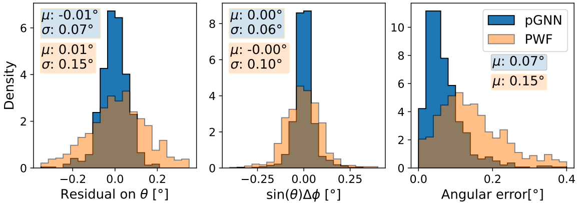
Performance > 10 antennas

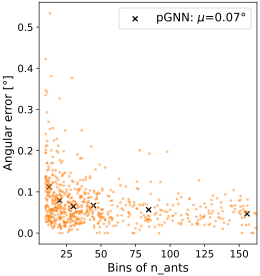
Direction reconstruction

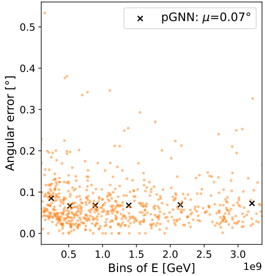
Single model predictions



Ensemble method



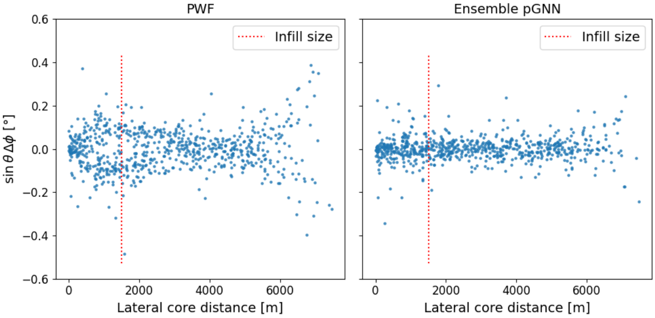


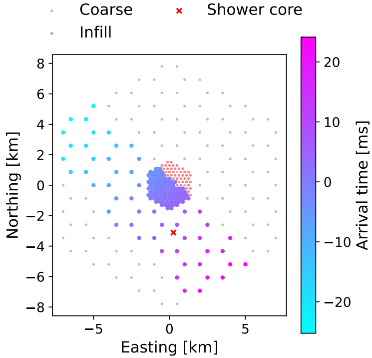


× 10 improvement compared to raw GNN

Correction of the PWF bias

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



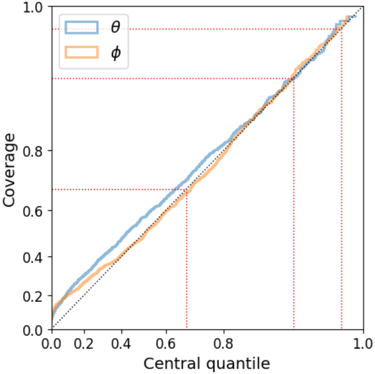


[Ferrière et al. in prep]

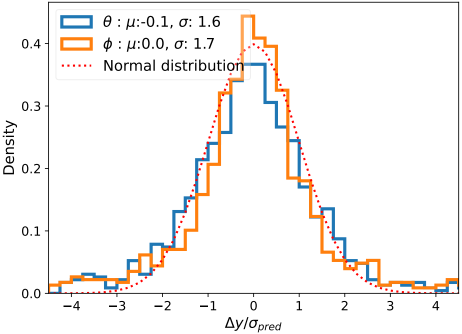
Uncertainty estimation

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 32 predictions of θ and σ_θ their std's.



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

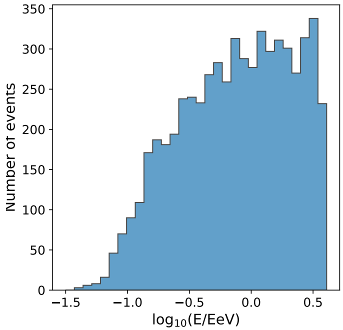


Training procedure and performances

Energy reconstruction - *Work in progress*

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

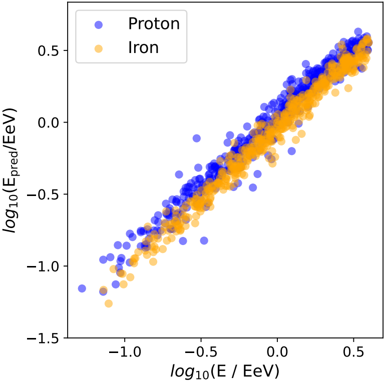
- Loss function $L(\omega) = \mathbb{E}[\log\left(\frac{E_{pred}}{E_{target}}\right)^2]$
- 5 input features: 3 antenna coordinates, arrival time, signal amplitude (normalised)
- **Training set**: 4937 events, **Validation set**: 928 events
- 10 models or more trained with different initialization/Train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

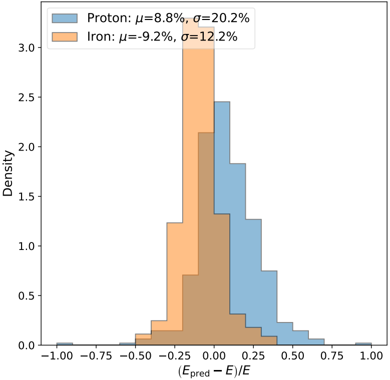


$\log E$ distribution

Energy resolution

Energy reconstruction - *Work in progress*





Total energy resolution : 19.5%

For primary energy!

Conclusion

Work on realistic simulations exclusively

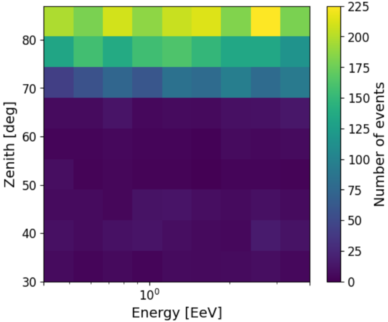
Results - Direction

- With GNN and ensemble methods, high direction reconstruction precision : 0.11° precision (0.07° if $n_{ants} >= 10$).
- Well calibrated uncertainty estimation

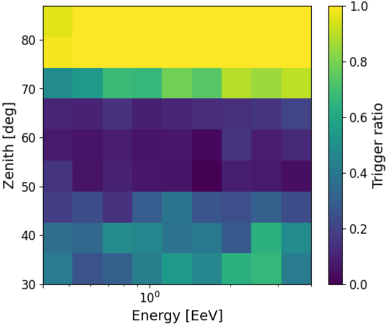
Results - Energy

- Energy resolution : 19.5%
- Need to be applied to EM energy
- Uncertainty estimation not yet calibrated.

Backup slides : joint trigger distribution



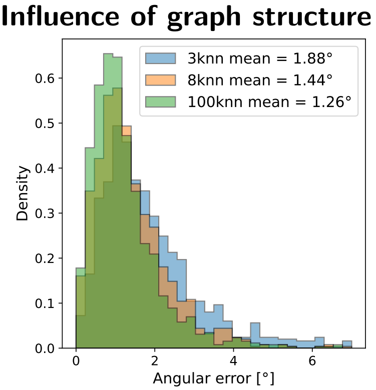
Joint trigger distribution



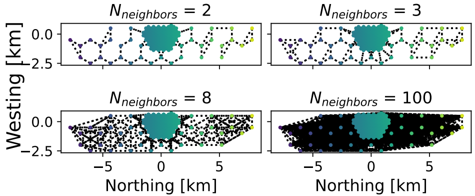
joint trigger ratio

Training set size and graph structure

Direction reconstruction



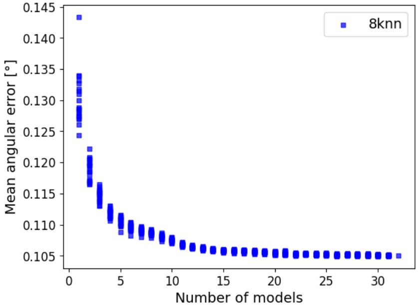
Degradation of performance when lowering number of neighbours



- 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