Friedrich-Alexander-Universität
Erlangen-Nürnberg

Application of graph networks to a next generation wide-field gamma-ray observatory in the southern sky

Franziska Leitl
Jonas Glombitza, Martin Schneider, Stefan Funk and Christopher van Eldik
Erlangen Centre for Astroparticle Physics (ECAP)
30.01.25

Friedrich-Alexander-Universität
Erlangen-Nürnberg

Application of Graph Networks to a wide-field Water-Cherenkov-based Gamma-Ray Observatory

J. Glombitza,[a] M. Schneider,[a] F. Leitl,[a] S. Funk,[a] and C. van Eldik[a]

[a]Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics, Nikolaus-Fiebiger-Str. 2, 91058 Erlangen, Germany

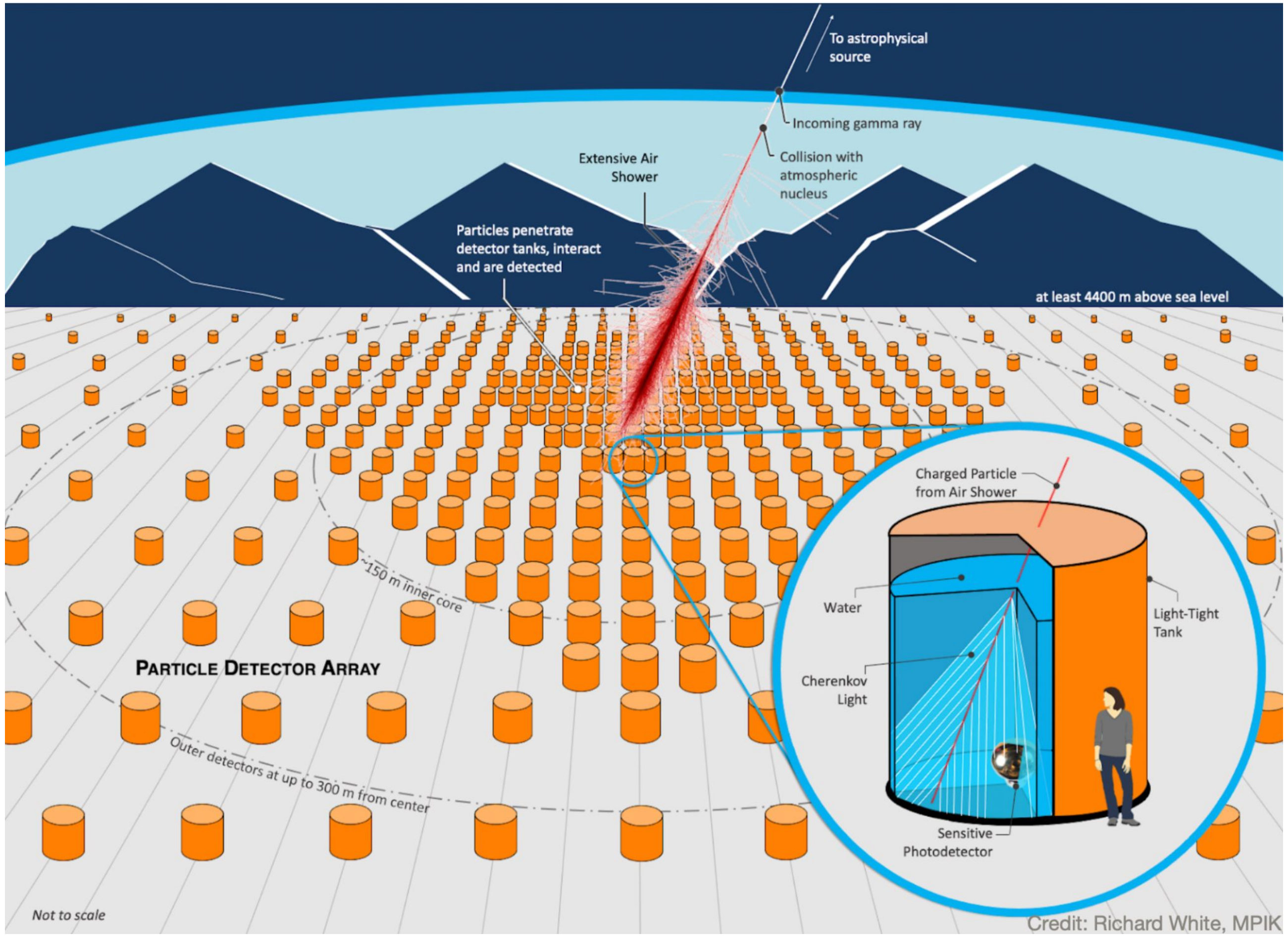
E-mail: jonas.glombitza@fau.de, martin.friedrich.schneider@fau.de, franziska.leitl@fau.de

https://doi.org/10.48550/arXiv.2411.16565

The Southern Wide-field Gamma-Ray Observatory (SWGO)

The Southern Wide-field Gamma-Ray Observatory (SWGO)

- Future **gamma-ray detector** located in Atacama Astronomical Park, **Chile**.
- Ground-level detector array primarily based on **water-Cherenkov detector** units
- **Altitude**: 4770m
- Energy range from **hundreds of GeV up to the PeV scale**
- Close to **100% duty cycle** and **order steradian field of view**.



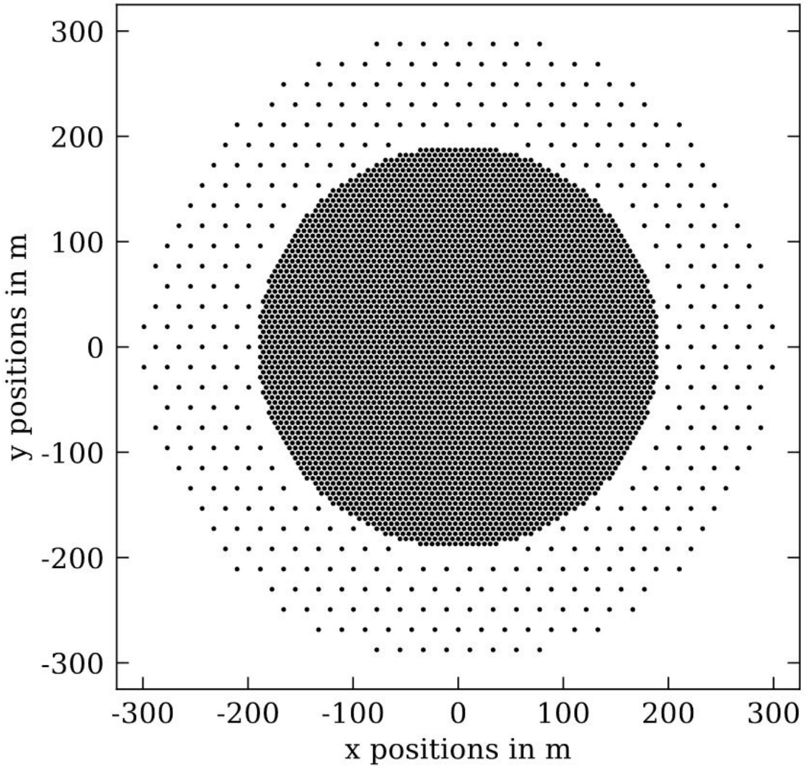
SWGO (2024)

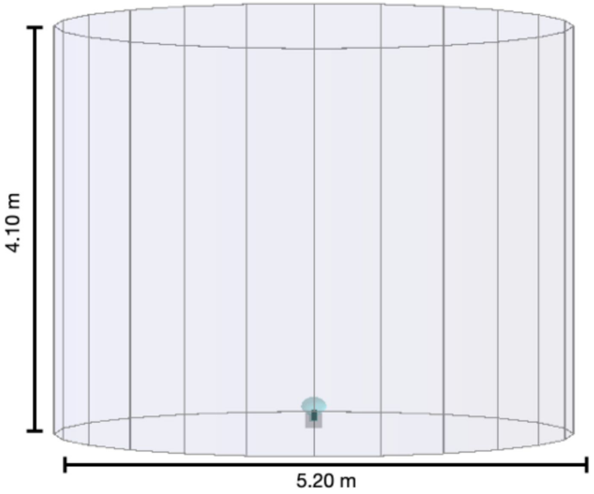
The Southern Wide-field Gamma-Ray Observatory (SWGO)

SWGO still in the **R&D phase**
→ Testing different detector and array designs

One of the possible **candidate designs**:

- Roughly 280,000m^2 (~ 4600 units)
- Two zones: Fill factor 80% and 5%
- Similar style to HAWC tanks





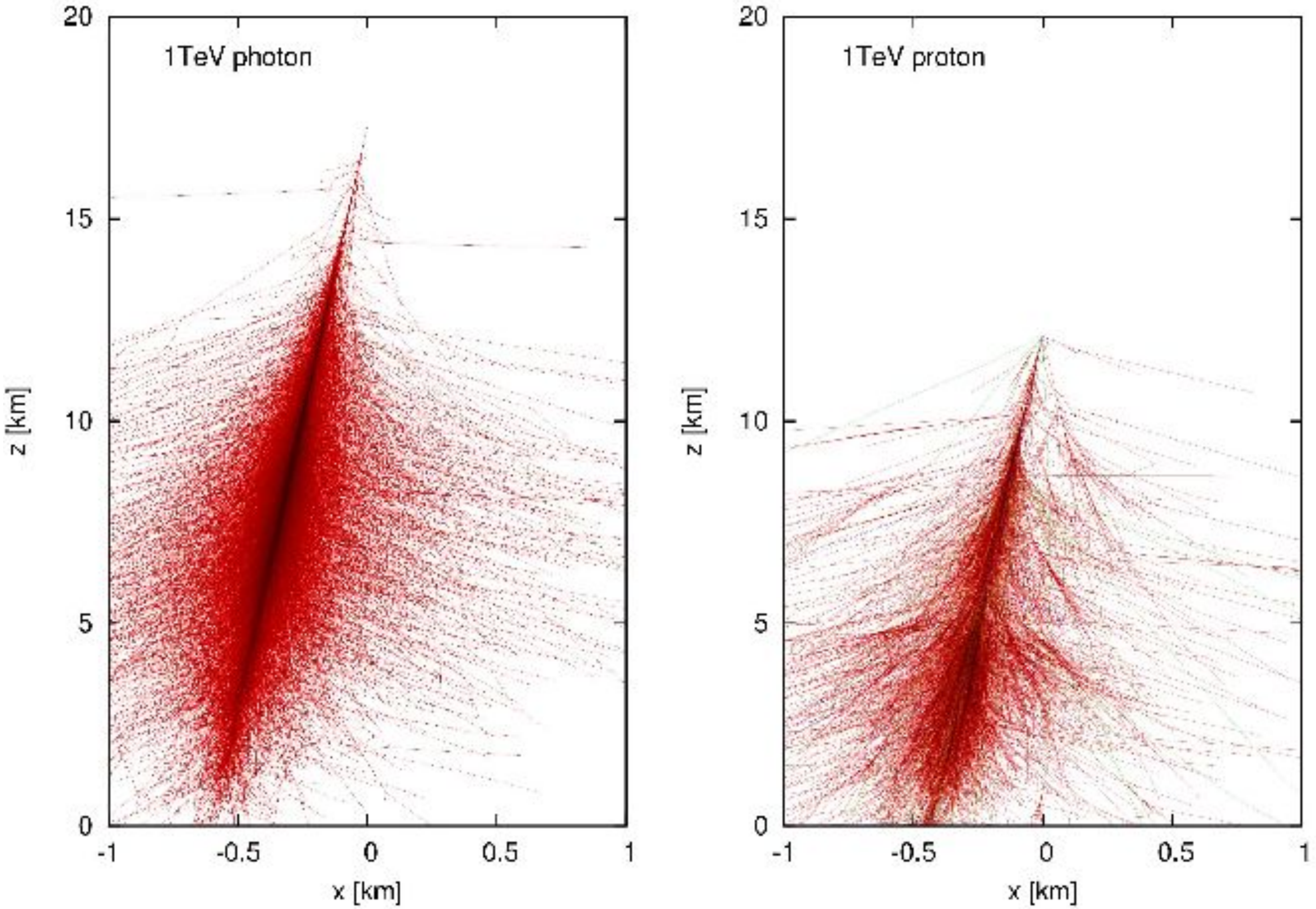
The Southern Wide-field Gamma-Ray Observatory (SWGO)

- Want to **characterize gamma rays** as good as possible!

- Need to **avoid cosmic rays**!

Our tasks:

1. γ / hadron separation
2. Energy reconstruction of gamma rays



Barnacka et al. (2012)

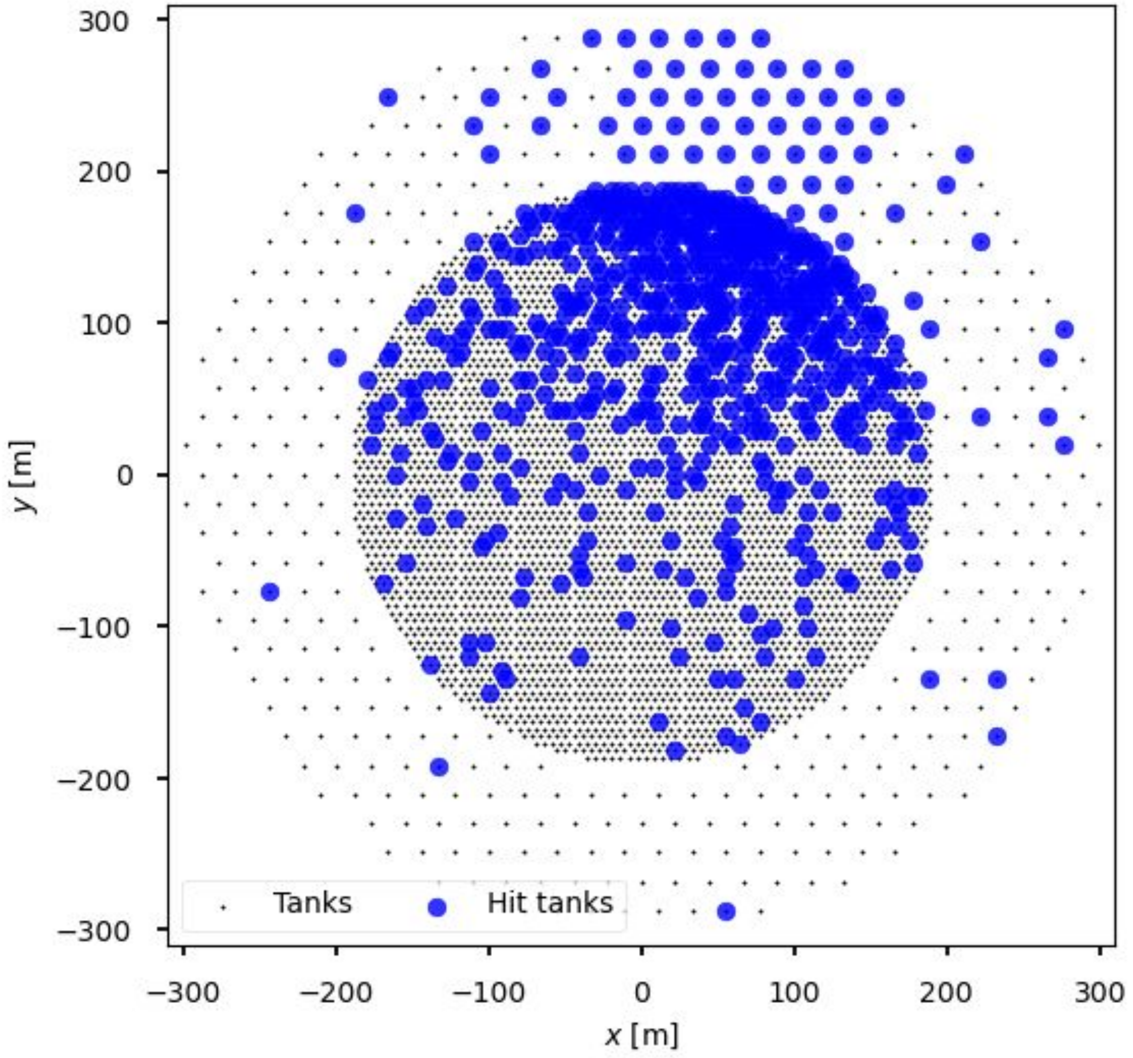
Why Graph Neural Networks?

Example shower footprint

Measured shower profile on the ground depends on:

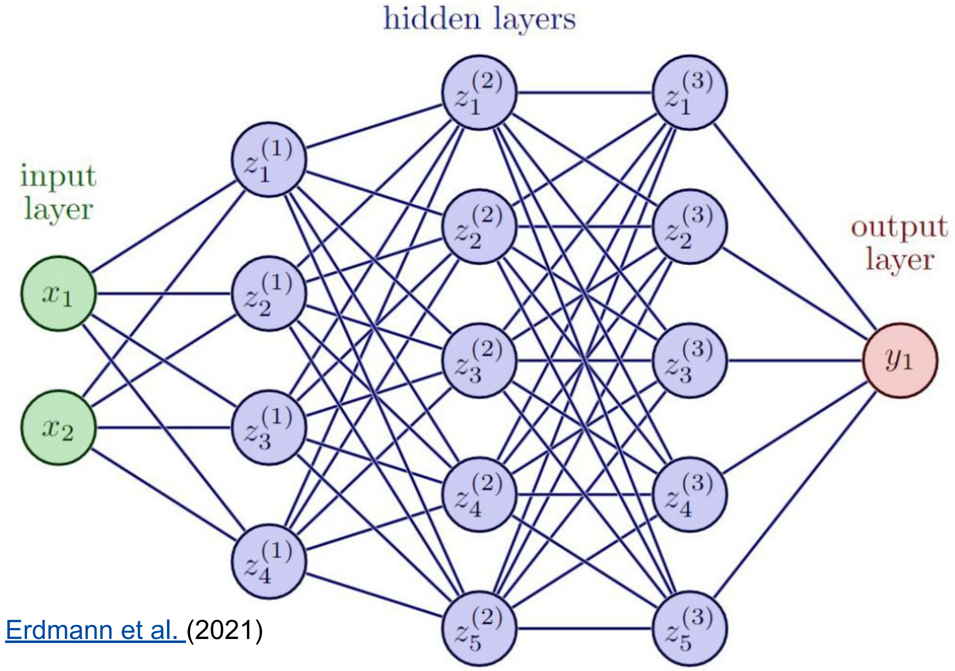
- Energy of shower
- Incoming zenith angle
- Position

→ Can trigger between tens and thousands of tanks



Deep neural networks

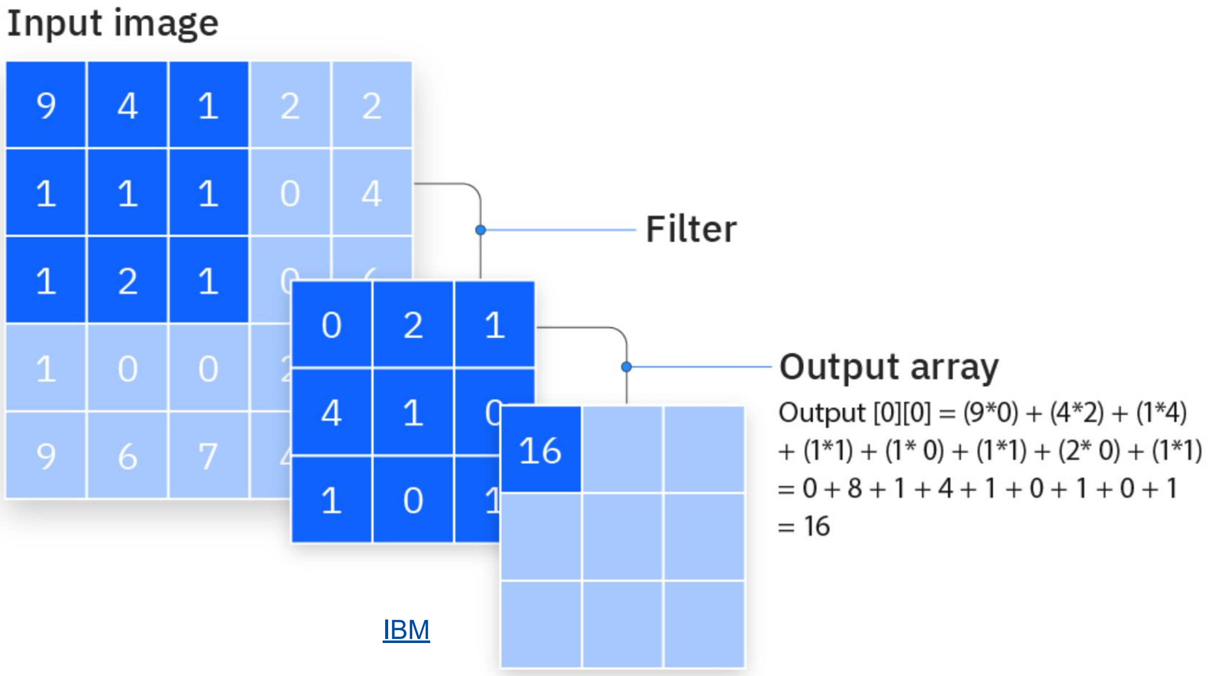
Fully connected network



Erdmann et al. (2021)

→ **Problematic for large datasets**

Convolutional Neural Networks (CNNs)

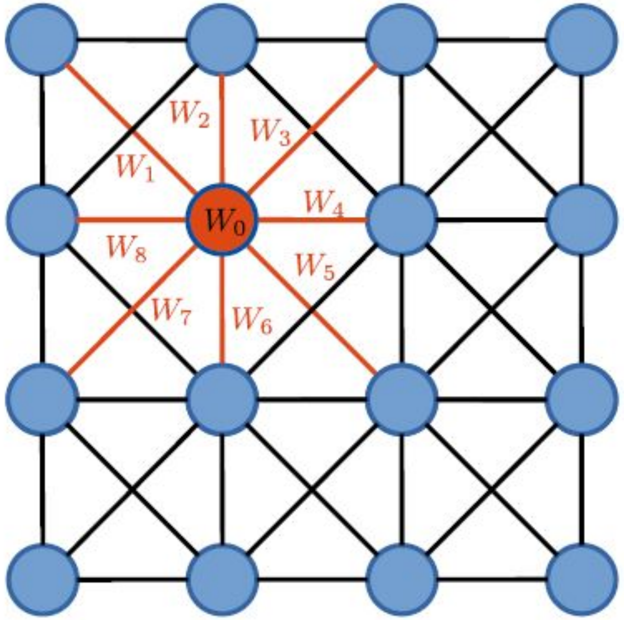


IBM

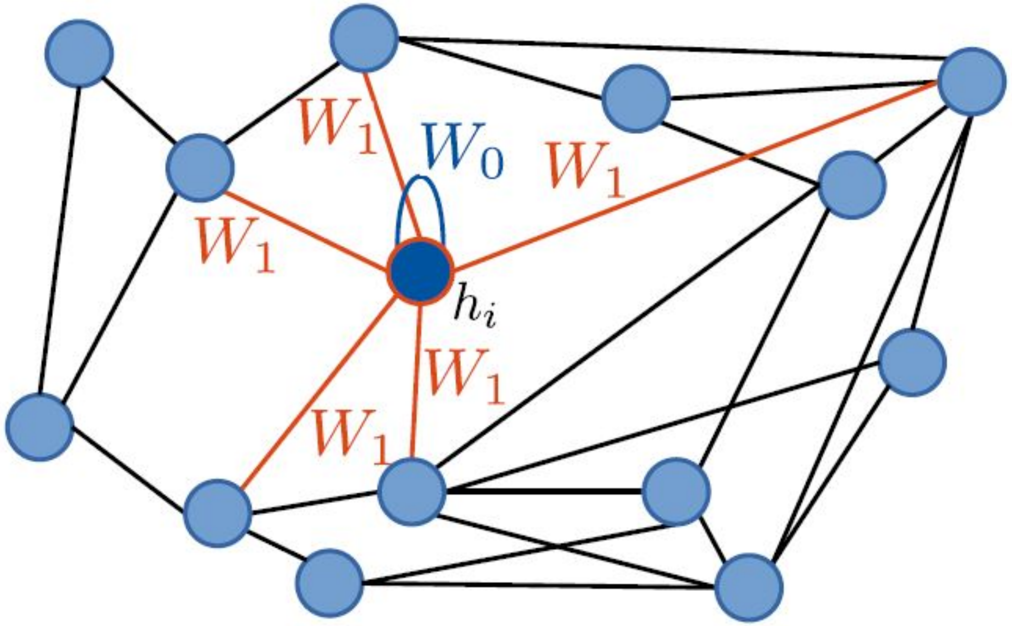
→ **Bound to regular grid structure**

Graph Neural Networks (GNNs)

CNNs



GNNs



Propagation: $h_i^{(l+1)} = \sigma(h_i^{(l)} W_0^{(l)} + \sum_{j \in \mathcal{N}_i} \frac{1}{c_{ij}} h_j^{(l)} W_1^{(l)})$

→ **GNNs can be applied to non-regular grids**

Erdmann et al. (2021)

Graph Neural Networks (GNNs)

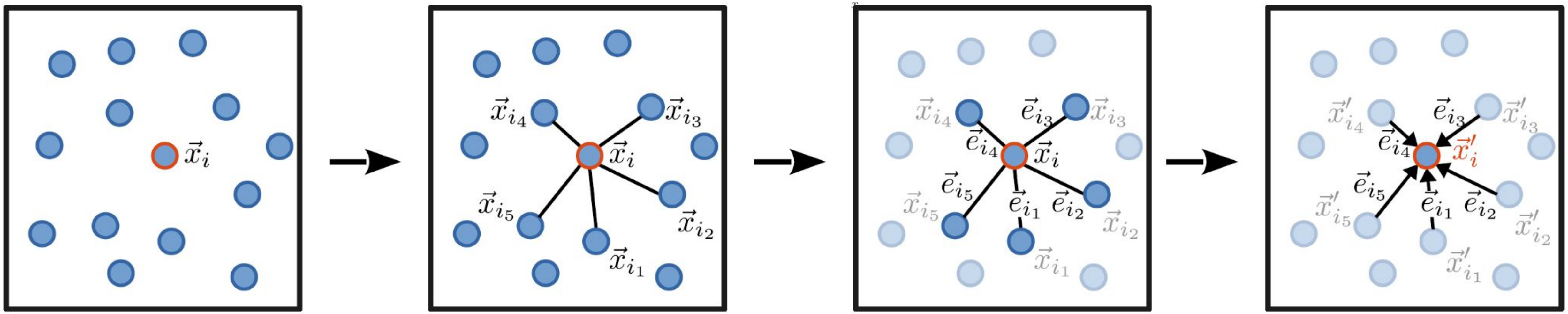
EdgeConvolutions

Create point cloud

Construct directed graph for each point

Estimate edge features with kernel function

Aggregate over neighborhood



Erdmann et al. (2021)

$$\mathbf{x}'_i = \sigma\left(\mathop{\square}_{j \in \mathcal{N}(i)} h_{\mathbf{\Theta}}(\mathbf{x}_i \,\|\, \mathbf{x}_j - \mathbf{x}_i) \right)$$

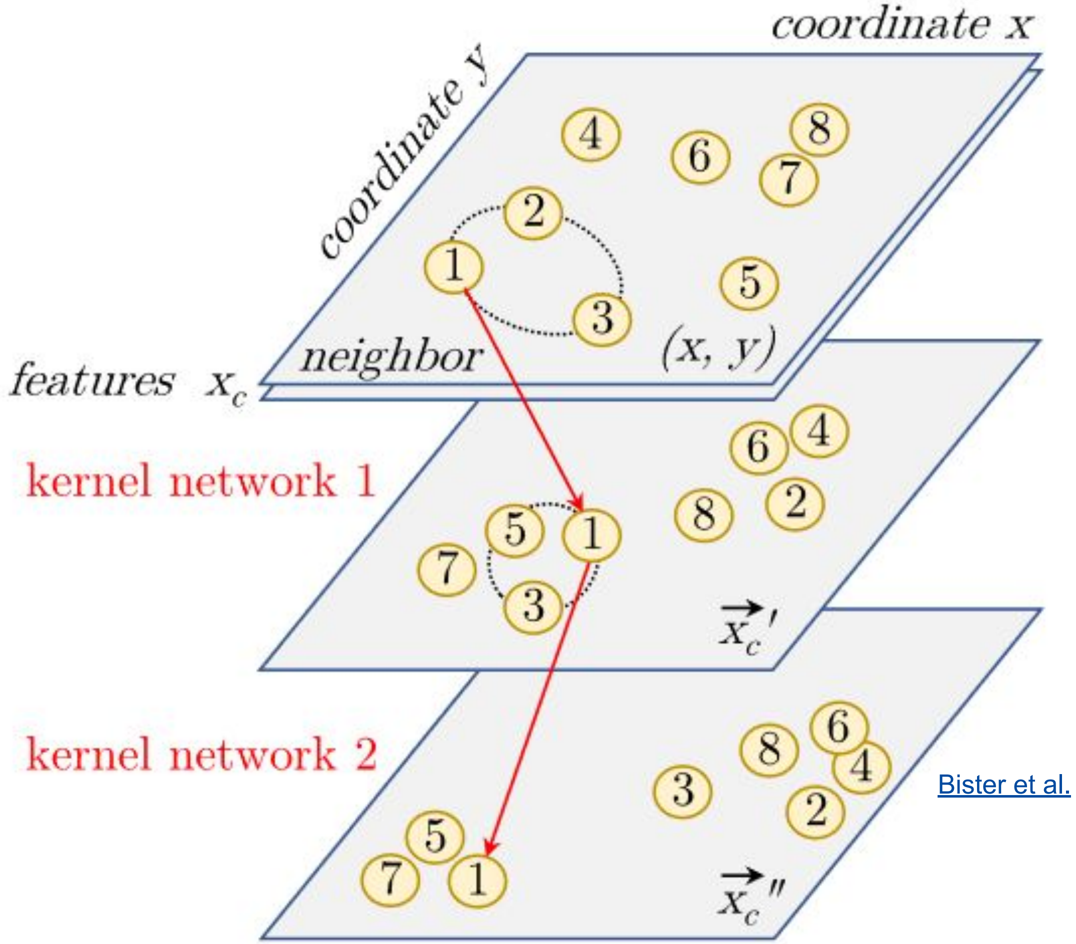
Graph Neural Networks (GNNs)

DynamicEdgeConvolutions

- Update neighbours in feature space for each layer

- Nodes with similar features become neighbours

→ Extract global features



Bister et al. (2021)

Friedrich-Alexander-Universität
Erlangen-Nürnberg

SWGO
The Southern Wide-field
Gamma-ray Observatory

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

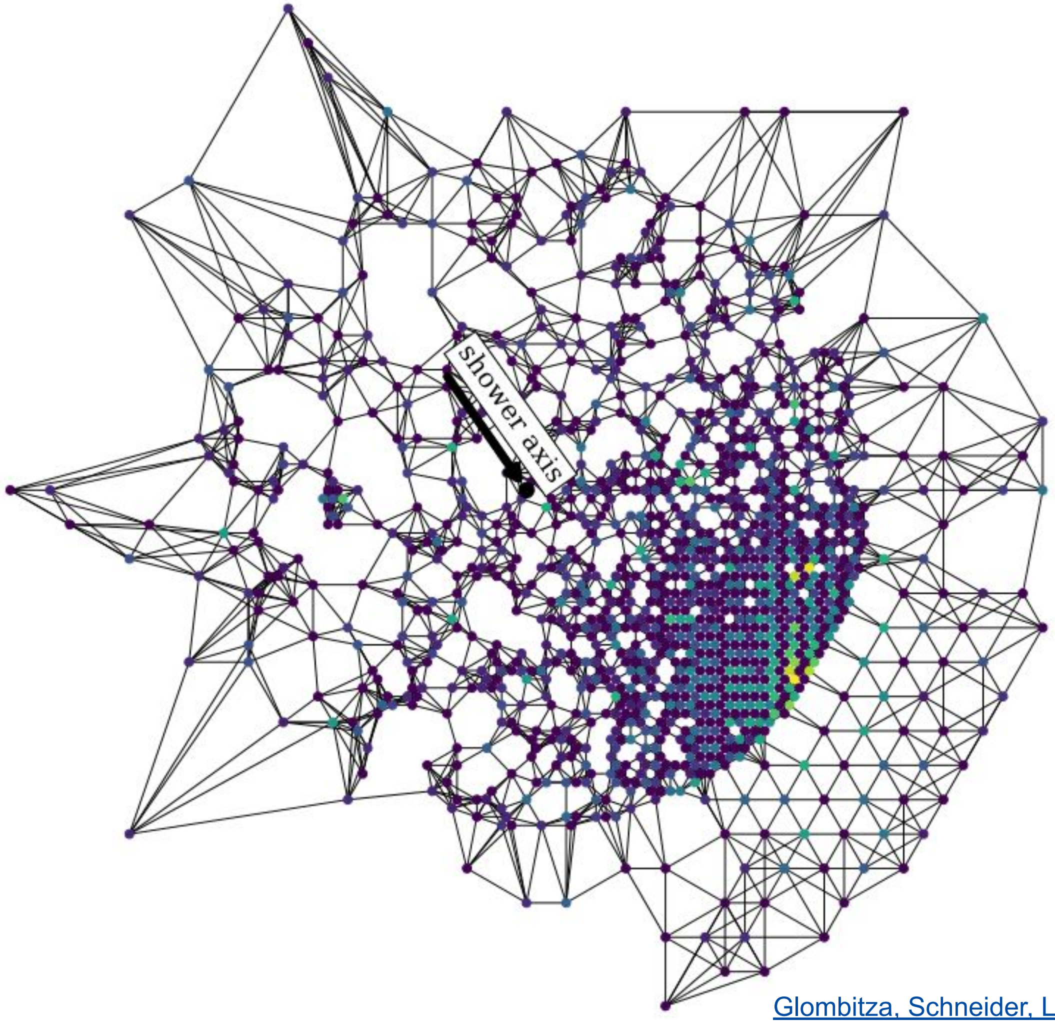
GNNs for SWGO

Application of GNNs to SWGO
Inputs

- **Nodes as triggered tanks**
  → Different for each event

- **4 inputs** per tank:
  - *x and y position* of each triggered tank
  - *Arrival times* of first Cherenkov photon measured in PMT
  - *Charge* measured in each tank

- ***k*NN** clustering:
  6 neighbours and self loop



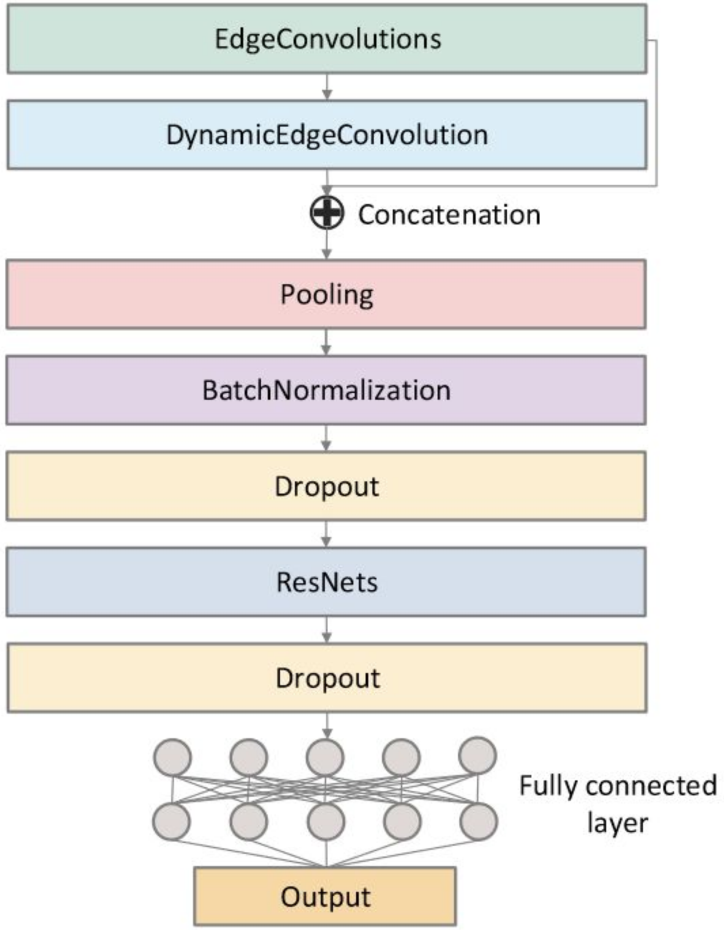
Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Application of GNNs to SWGO

Network structure

- NVIDIA A40/A100 GPUs
- Implementation using PyTorch_Geometric
- ~ 500k trainable parameters

→ Performed a small hyperparameter search for the γ / **hadron separation** and **energy reconstruction**



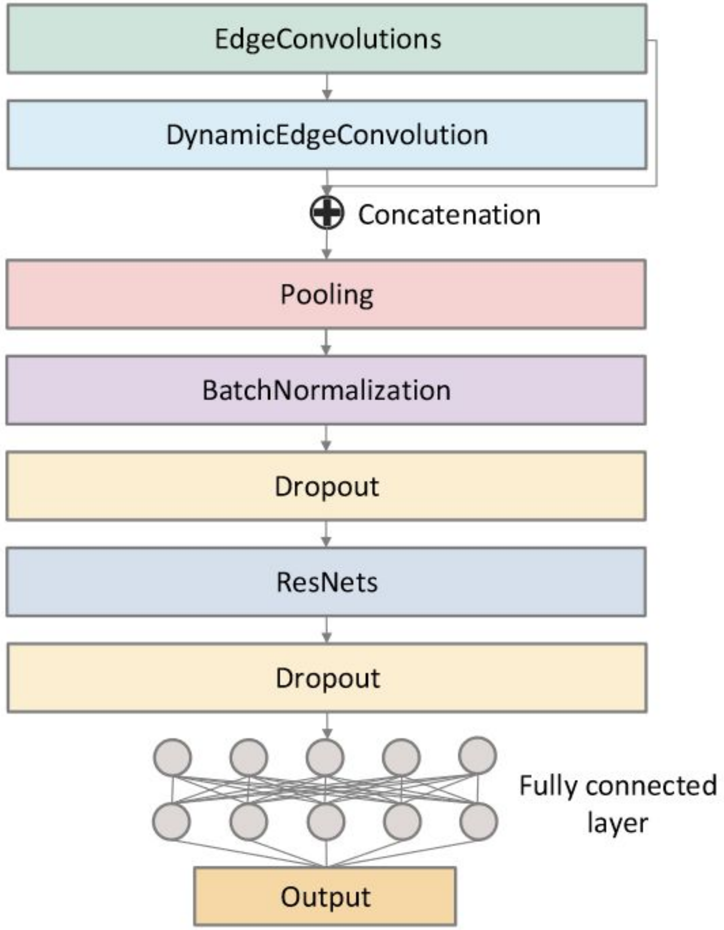
Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Application of GNNs to SWGO

Network structure

- **Hyperparameter search** (70 trainings each):
  - Learning rate
  - Decay factor
  - Batchsize
  - Weight decay
  - n_{EdgeConv}
  - $n_{\text{DynEdgeConv}}$
  - n_{feat}
  - n_{ResNet}
  - Batchnorm
  - Dropout
  - $n_{k\text{NN,DynEdgeConv}}$

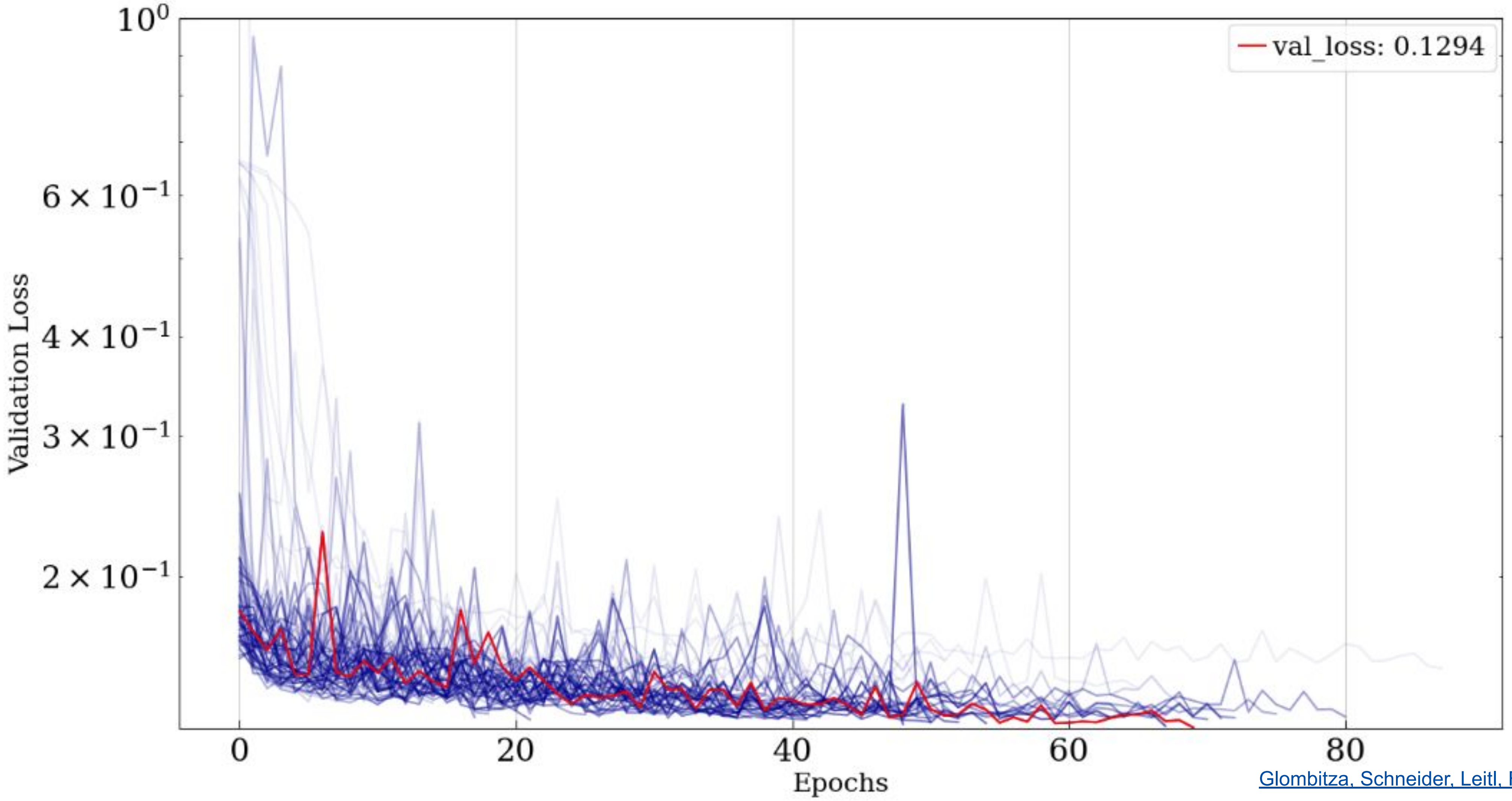
- Trainings running up to 24h



Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Hyperparameter search

Validation losses



Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

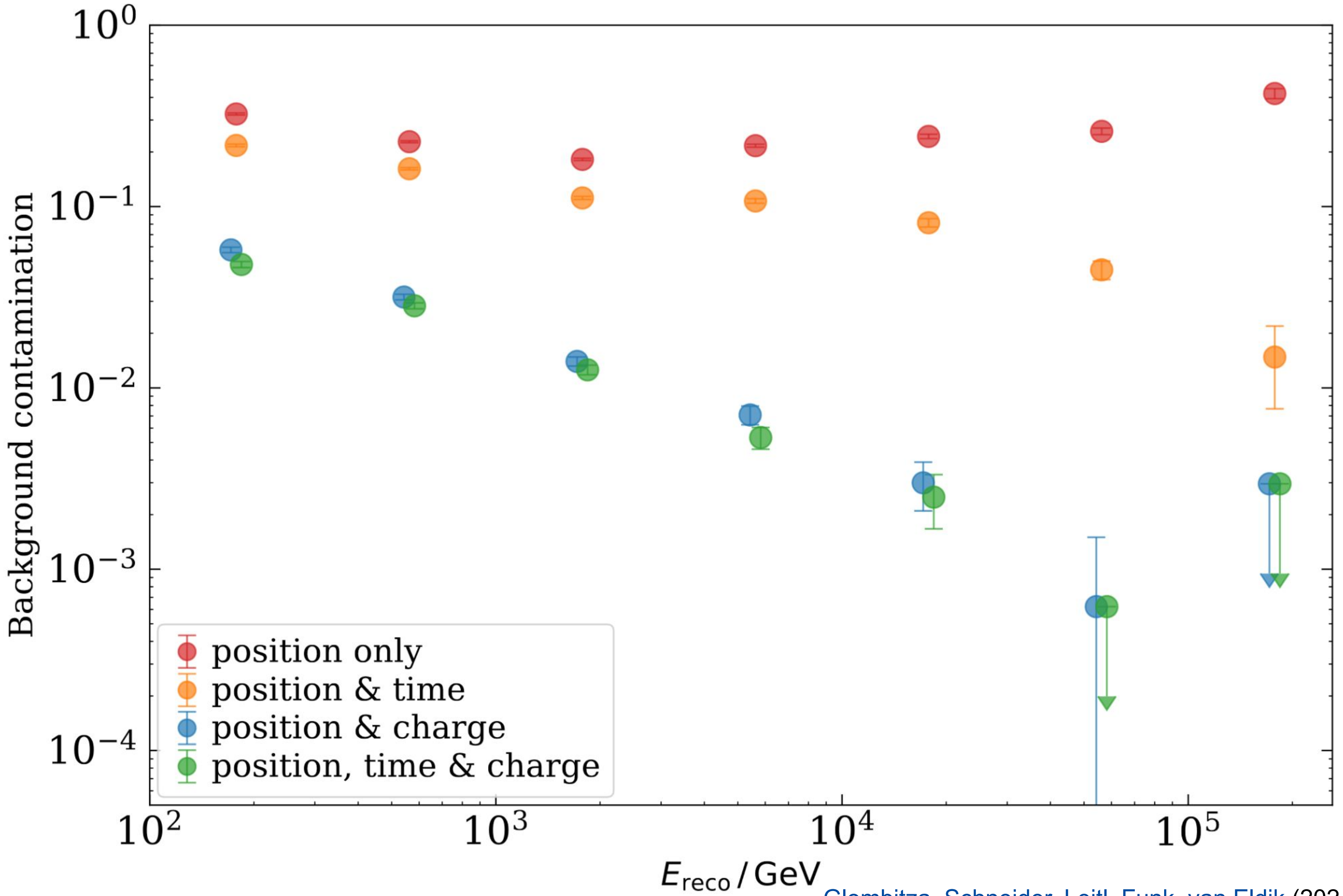
γ / hadron separation

Performance for different input features

- **Training output:**
  Score that indicates likelihood of particle type

- **Background contamination:**
  Fraction of protons relative to the simulated number of protons misclassified as gamma rays

→ **Position, time & charge** information yields the **best results**



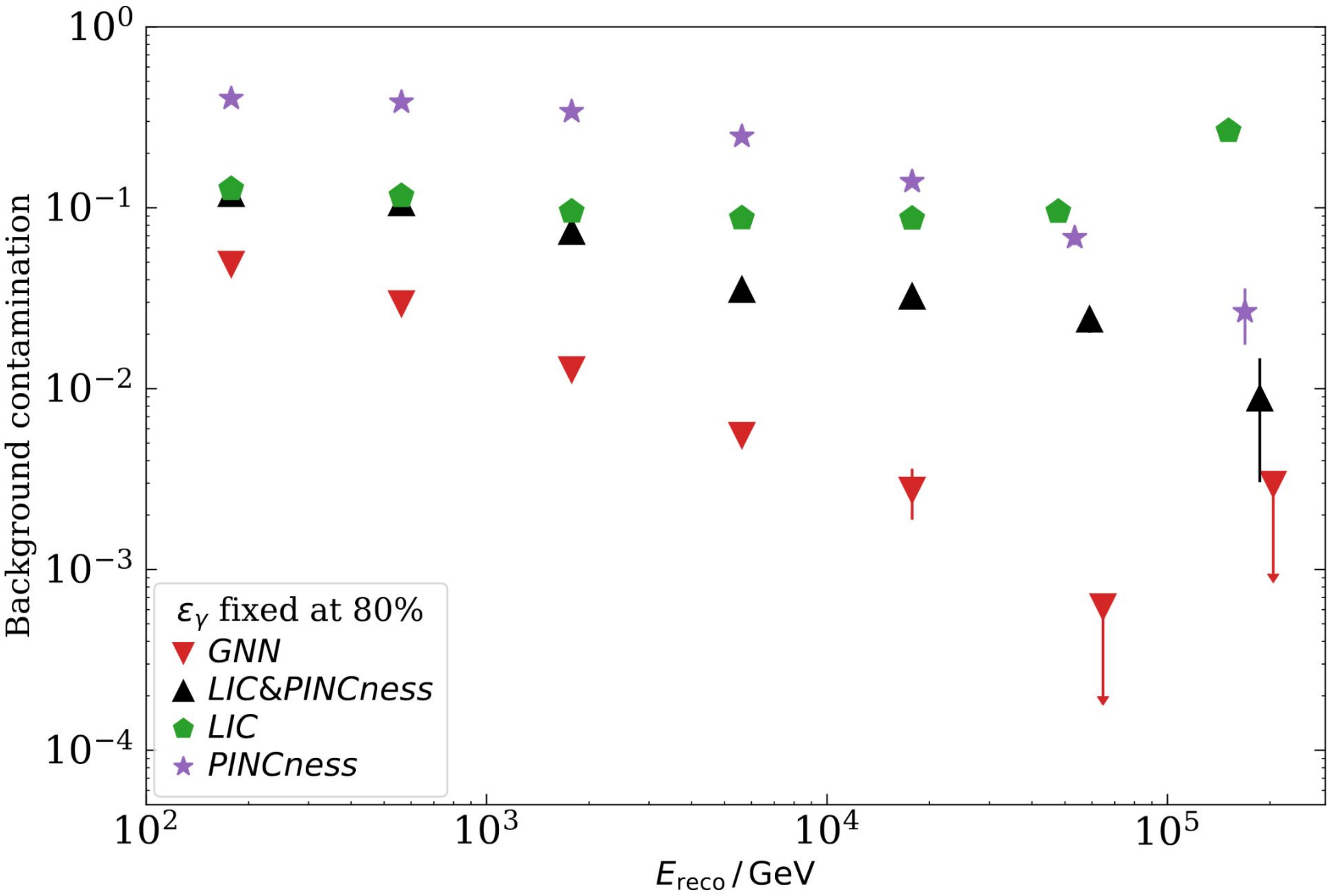
Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

γ / hadron separation

Performance estimation

- LIC and PINCness as **standard parameters** originating from HAWC
- Background rejection of **GNN surpasses standard methods** in all energy ranges

→ **Improvement by a factor of two** at low energies and **one order of magnitude** at high energies



Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

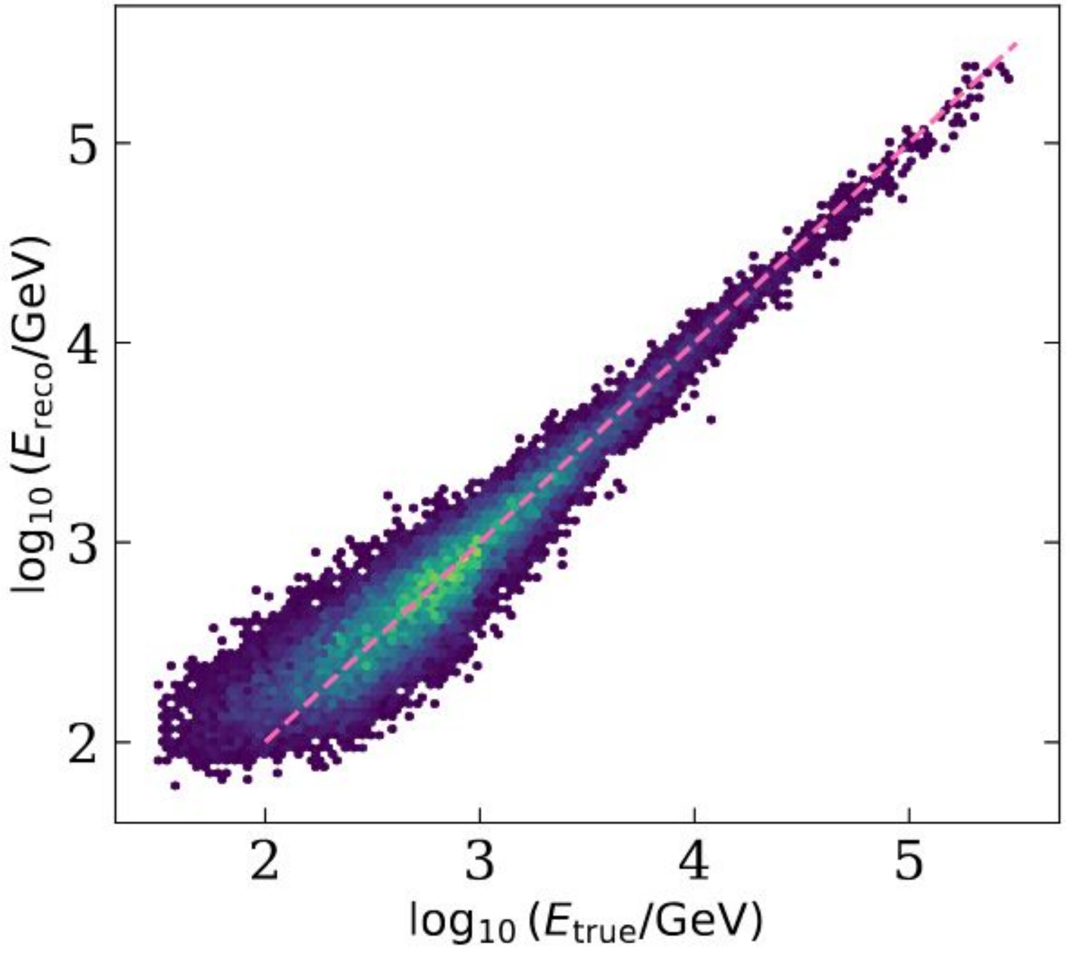
Energy reconstruction

Dispersion matrix

- **Training output**: $\log_{10}(E_{reco}/\mathrm{GeV})$

- Ideal curve: $\log_{10}(E_{true}/\mathrm{GeV}) = \log_{10}(E_{reco}/\mathrm{GeV})$

→ **Good agreement** between **reconstructed** and **true energy** in dispersion matrix



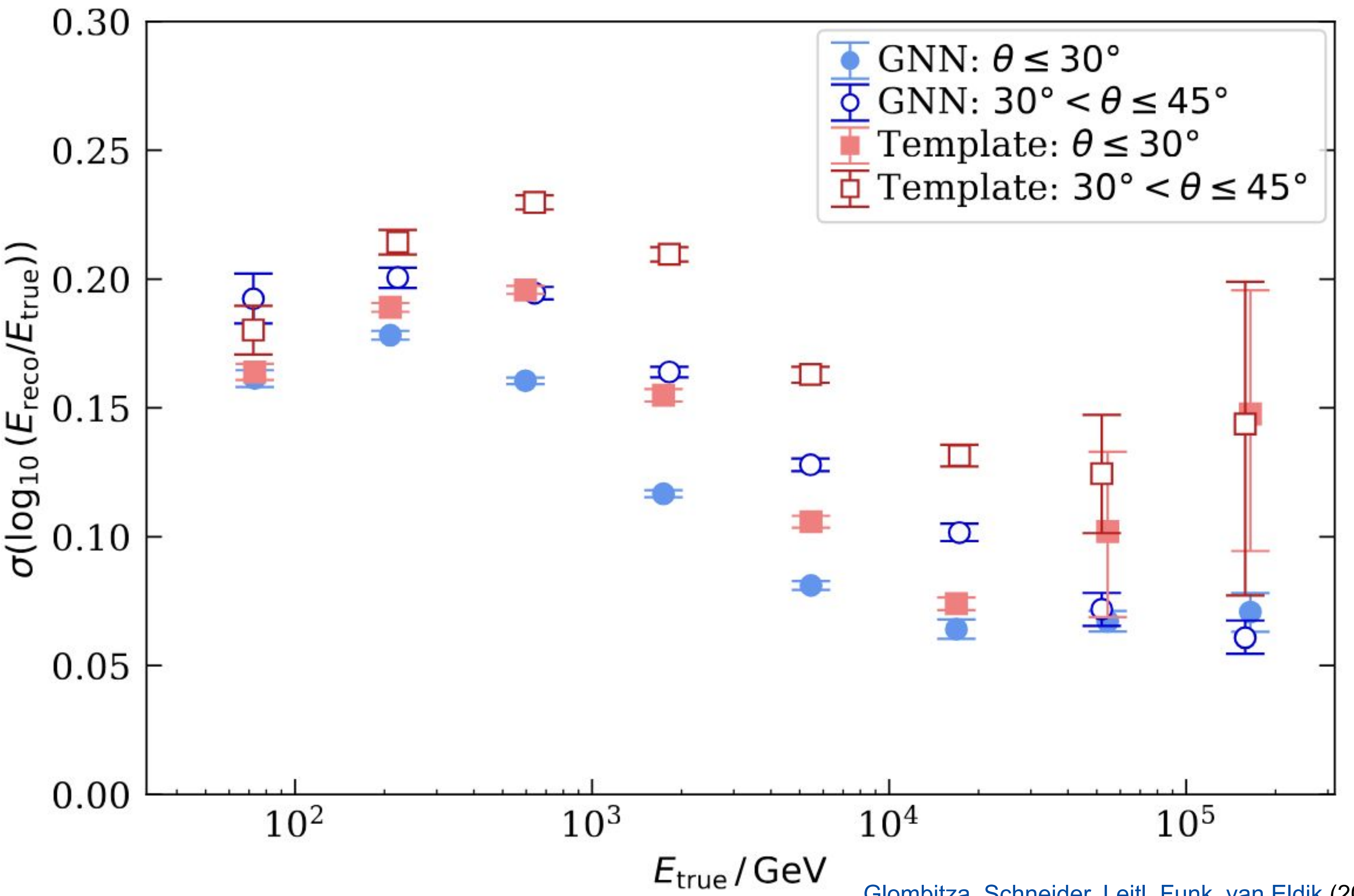
Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Energy reconstruction
Performance estimation

- **Energy resolution** as RMS of $\log_{10}(E_{reco}/E_{true})$

- Compare both to the current **template-based standard method**

→ **GNN outperforming** current **standard method** over the whole energy range



Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

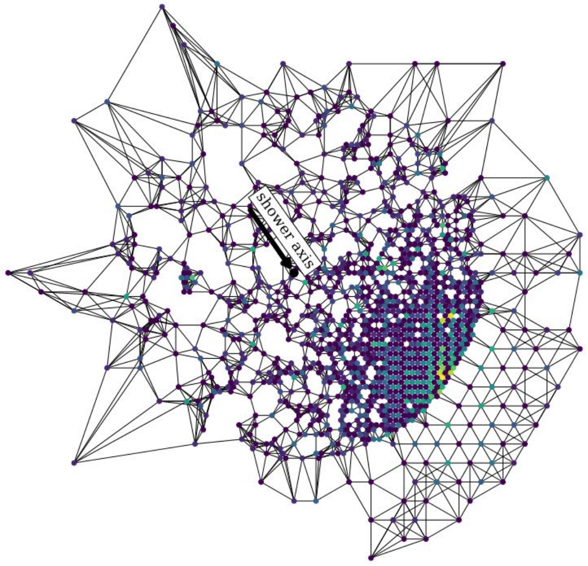
Summary

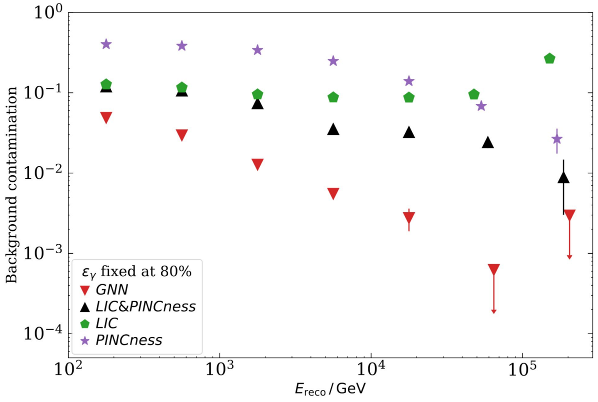
Successfully applied Graph Neural Networks (GNNs) to an SWGO candidate configuration:

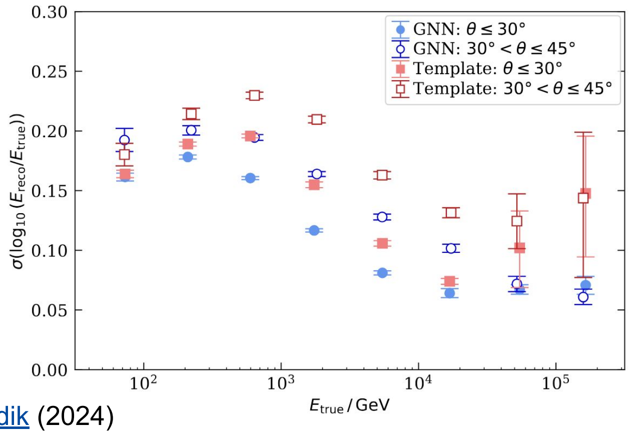
- γ / hadron separation
- Energy reconstruction

→ **Great improvements** compared to **current standard methods** over the whole energy range

→ Further Details in our paper: Glombitza, Schneider, Leitl, Funk, van Eldik (2024)







Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

<u>What we are working on now:</u>

- Applied GNNs to other candidate configurations
  → Networks robust when changing configurations
- GNNs for direction reconstruction, Vision Transformers, …

Thank you for your attention!

Backup

Architecture for γ / hadron separation

	Layer	features	setting	input shape	output shape
	EdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times 4$	$n_{\text{nodes}} \times n_{\text{feat}}$
	ReLU	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
$3 \times \{$	EdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
	ReLU	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
	DynEdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
	Concat	–	layer outputs	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	MaxPooling	–	–	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	Dropout	–	$p = 0.13$	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	ResNet	n_{feat}	–	$4 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	n_{feat}
$4 \times \{$	ResNet	n_{feat}	–	n_{feat}	n_{feat}
	Dropout	–	$p = 0.13$	n_{feat}	n_{feat}
	Linear	2	–	n_{feat}	2
	SoftMax	–	–	n_{feat}	2

Table 1: Network architectures for the separation task. n_{nodes} describes the number of nodes in the input graph, which depend on the number of triggered tanks, and n_{feat} the number of features. We further use a bottleneck layer in the first ResNet layer to adjust the dimensionality.

Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Backup

Architecture for energy reconstruction

	Layer	features	setting	input shape	output shape
	EdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times 4$	$n_{\text{nodes}} \times n_{\text{feat}}$
	Linear	n_{feat}	–	$n_{\text{nodes}} \times 4$	$n_{\text{nodes}} \times n_{\text{feat}}$
	Addition	–	–	$n_{\text{nodes}} \times 4$	$n_{\text{nodes}} \times n_{\text{feat}}$
	ReLU	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
$3 \times$	EdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
$3 \times$	Linear	n_{feat}	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
$3 \times$	Addition	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
$3 \times$	ReLU	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}}$
	DynEdgeConv	n_{feat}	$\square_j$: mean	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}} \cdot 2$
	Linear	n_{feat}	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}} \cdot 2$
	Addition	–	–	$n_{\text{nodes}} \times n_{\text{feat}}$	$n_{\text{nodes}} \times n_{\text{feat}} \cdot 2$
	ReLU	–	–	$n_{\text{nodes}} \times n_{\text{feat}} \cdot 2$	$n_{\text{nodes}} \times n_{\text{feat}} \cdot 2$
	Concat	–	layer outputs	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	MaxPooling	–	–	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	Batchnorm	–	momentum $= 0.1$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	Dropout	–	$p = 0.24$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$
	ResNet	n_{feat}	–	$5 \cdot (n_{\text{nodes}} \times n_{\text{feat}})$	n_{feat}
$2 \times$	ResNet	n_{feat}	–	n_{feat}	n_{feat}
	Dropout	–	$p = 0.24$	n_{feat}	n_{feat}
	Linear	1	–	n_{feat}	1

Table 2: Network architectures for the energy reconstruction task. n_{nodes} describes the number of nodes in the input graph, which depend on the number of triggered tanks, and n_{feat} the number of features. The "Addition" denotes a residual connection that connects consecutive convolutional and linear layers. We further use a bottleneck layer in the first ResNet layer to adjust the dimensionality.

Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Backup

Kernel network h_Θ

layer	features	settings
Linear	n_{feat}	no bias
Batchnorm	–	momentum=0.9
Activation	–	–
Linear	n_{feat}	no bias
Batchnorm	–	momentum=0.9
Activation	–	–
Linear	n_{feat}	no bias
Batchnorm	–	momentum=0.9
Activation	–	–

(a) Architecture of the kernel function h_Θ, where n_{feat} is to be chosen by the user. For the γ/hadron separation, we used SiLU as an activation function and for the energy reconstruction ReLU.

Residual Module "ResNet"

layer	features	settings
Linear	n_{feat}	no bias
SiLU	–	–
Batchnorm	–	momentum=0.1
Linear	n_{feat}	no bias
SiLU	–	–
Batchnorm	–	momentum=0.1
Add	–	–
SiLU	–	–

(b) Details of our used ResNet modules.

Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

γ / hadron separation

Current standard method: LIC and PINCness

Compare performance of GNN wit parameters originating from HAWC:

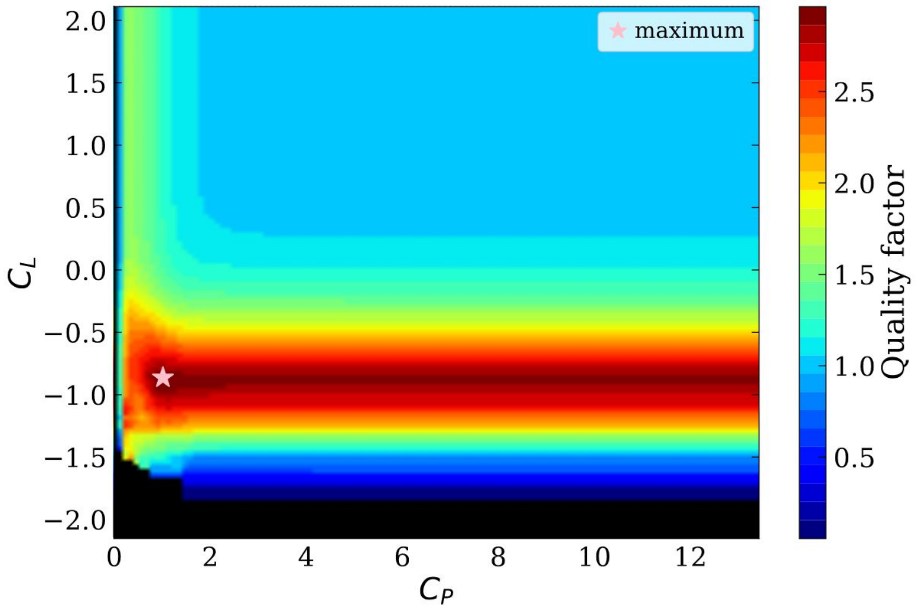
$$\mathrm{LIC} = \log_{10}\left(\frac{CxPE_{40}}{n_{\mathrm{hit}}}\right)$$

with $CxPE_{40}$ as largest charge measured in PMT at least 40m from shower core and n_{hit} number or tanks hit in event

and
$$\mathrm{PINCness} = \frac{1}{N}\sum_{i=0}^{N}\frac{\left(\log_{10}(q_i) - \langle\log_{10}(q_i)\rangle\right)^2}{\sigma^2}$$

with σ as charge uncertainty and q_i as the charge of the i-th PMT triggered in an event.

→ Find the cuts for LIC (C_L) and PINCness (C_P) for our dataset by optimizing the quality factor in each energy bin



(a) $2.5 \leq \log_{10}\left(E_{\mathrm{reco}}/\mathrm{GeV}\right) < 3.0$

Glombitza, Schneider, Leitl, Funk, van Eldik (2024)

Energy reconstruction

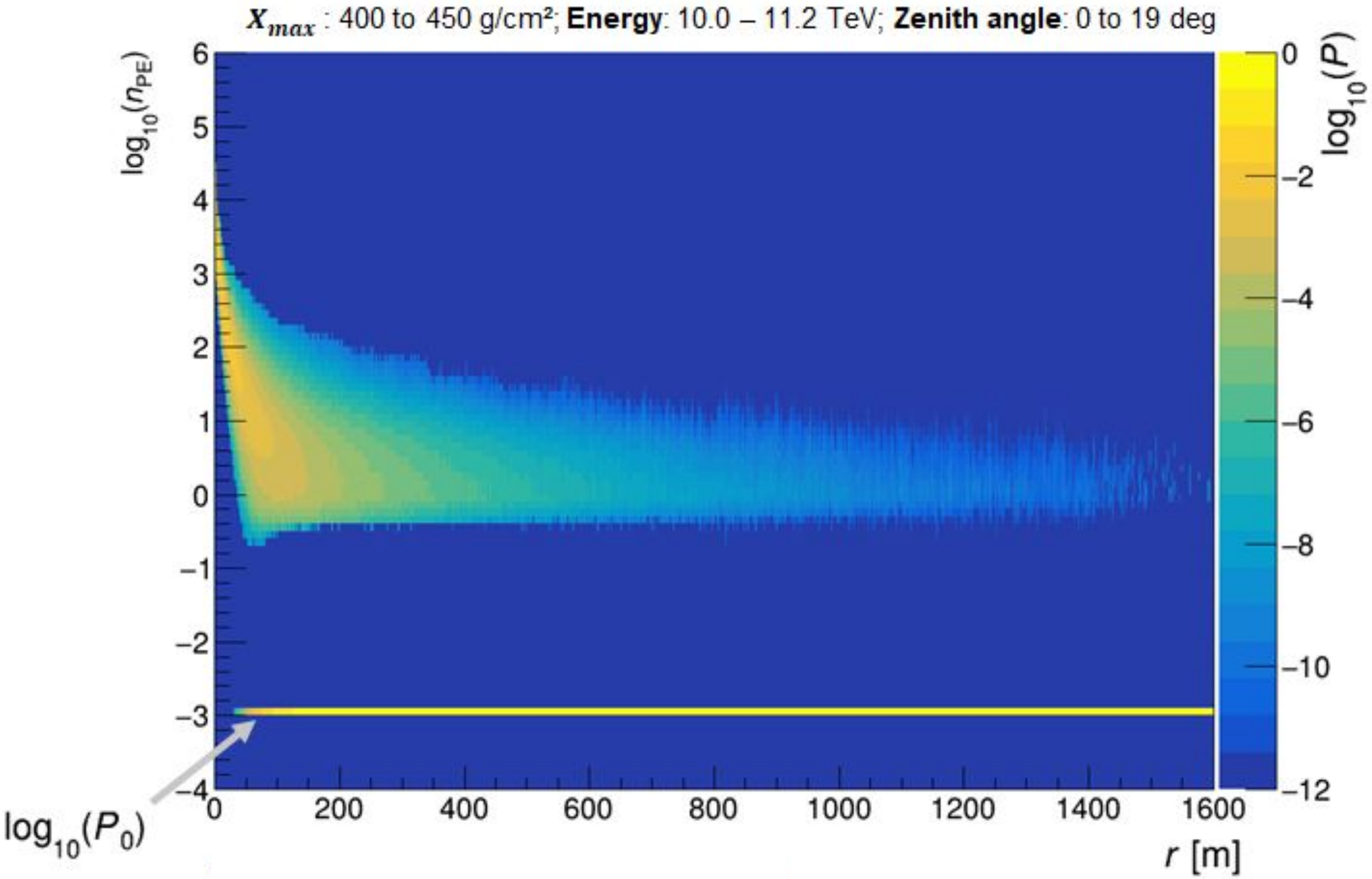
Current standard method: Template-based reconstruction

- Templates:
  MC simulations of gamma-induced EAS binned in X_{max}, E and θ

- Save the information if a tank did not measure any charge $\log_{10}(P_0)$

- Minimise log-likelihood to get best fit parameters

$$\log L = -2 \sum_i \log(F(\log_{10}(N_{\mathrm{PE}})_i, r_i, X_{\mathrm{max}}, E|\theta, \phi))$$



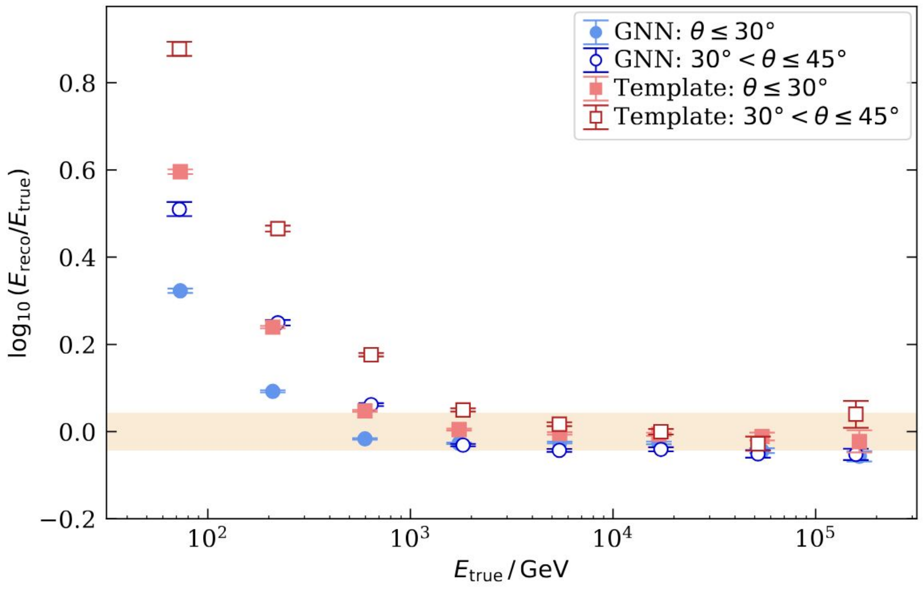
Energy reconstruction

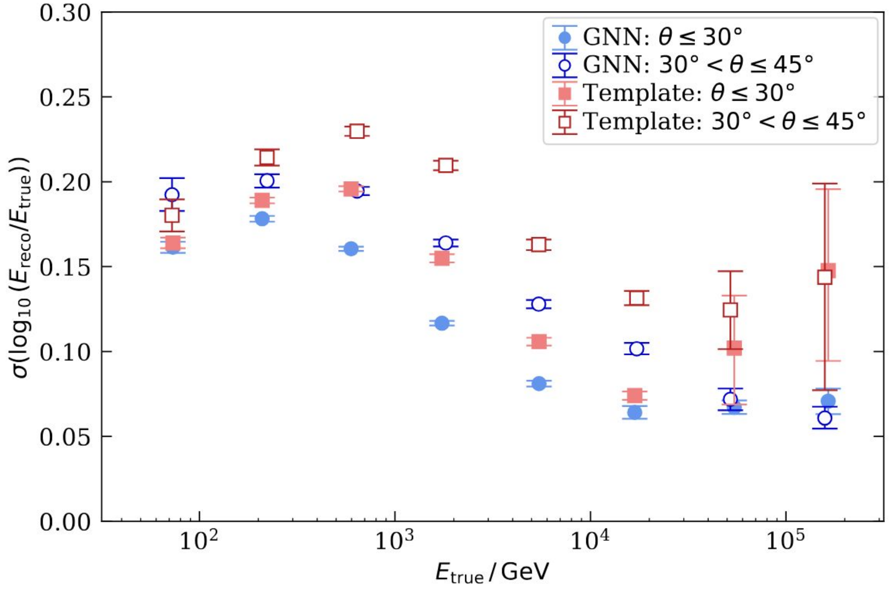
Performance estimation

Estimate performance via

- **Energy bias** as mean of $\log_{10}(E_{reco}/E_{true})$
- **Energy resolution** as RMS of $\log_{10}(E_{reco}/E_{true})$

and compare both to the **current template-based standard method**





→ **GNN outperforming current standard method over the whole energy range**

Glombitza, Schneider, Leitl, Funk, van Eldik (2024)