Friedrich-Alexander-Universität
Erlangen-Nürnberg

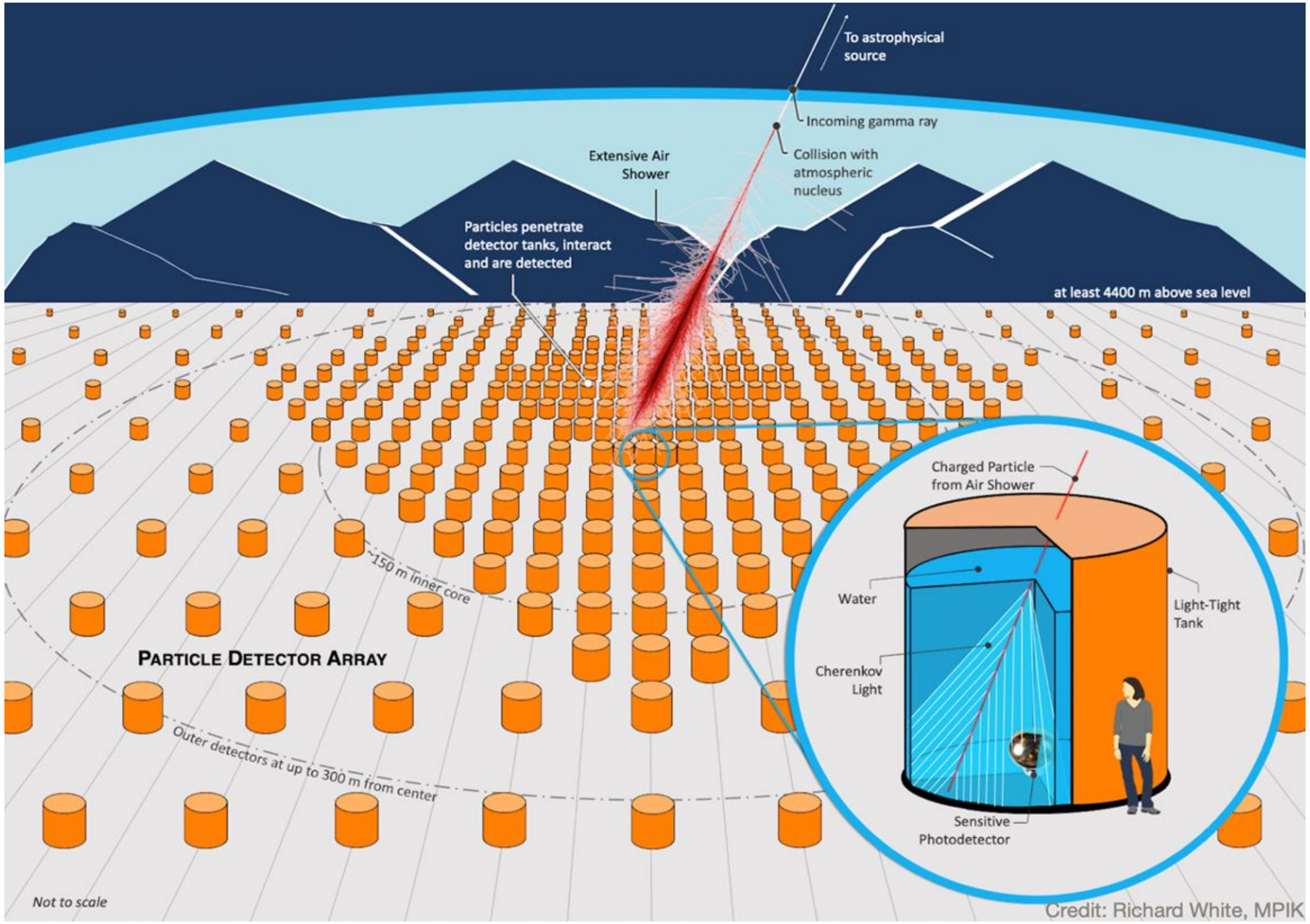
Deep-Learning-Based Gamma/Hadron Separation for the Southern Wide-field Gamma-ray Observatory

Martin Schneider, Christopher van Eldik,
Jonas Glombitza, Franziska Leitl, Markus Pirke

HEAMM workshop - 09.04.24

Introduction to SWGO

The Southern Wide-field Gamma-ray Observatory



What is SWGO?

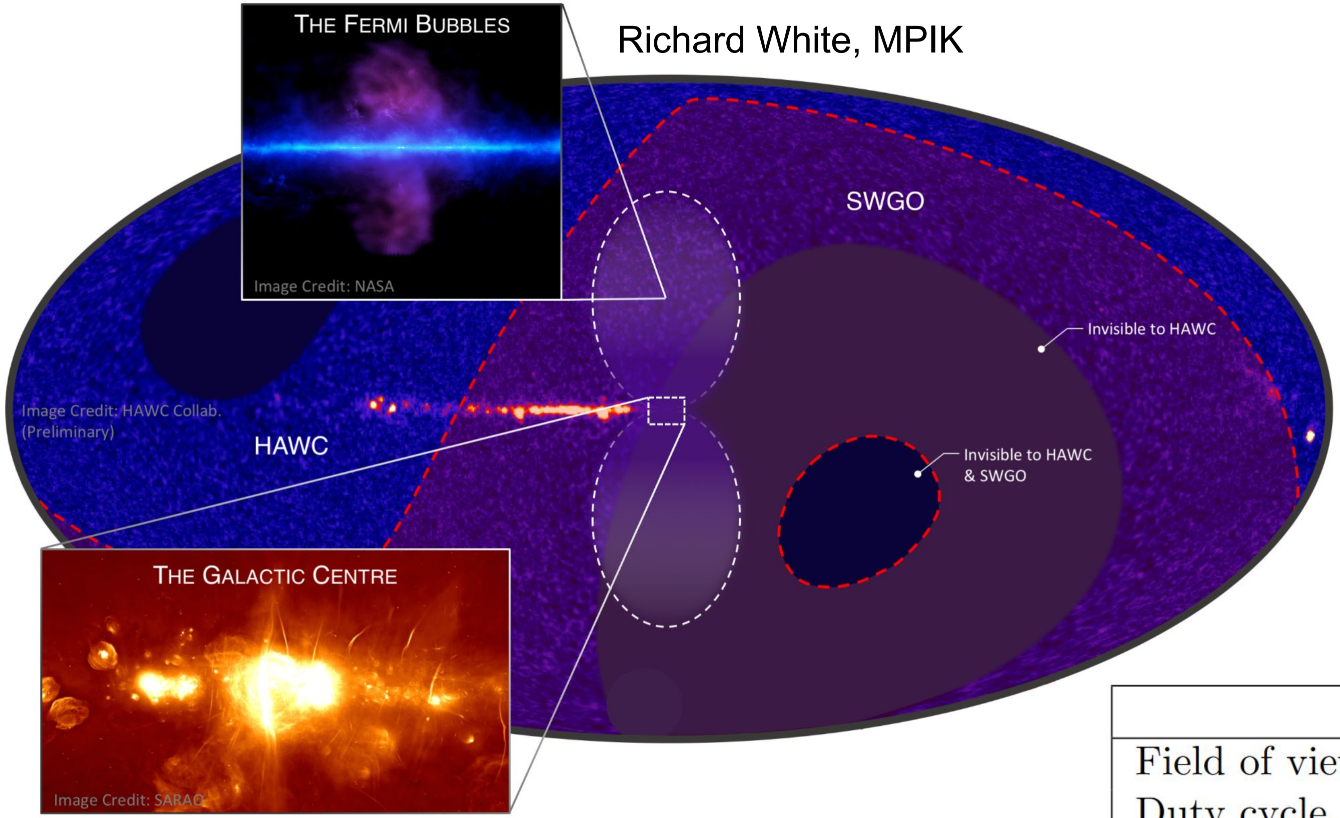
- Array of Water Cherenkov Detectors (**WCDs**) to measures extensive air showers at ground level

- Will complement Imaging Air Cherenkov Telescopes (**IACTs**) like H.E.S.S. and future CTA south

- Detection principle successfully demonstrated by the **HAWC** and **LHAASO** experiments

Motivation

Science case and sky coverage



Richard White, MPIK

IACTs vs WCDs

Ground-level particle detection with >95% duty cycle and inherent wide fov

(precision and instant sensitivity from IACTs will still be unmatched)

SWGO whitepaper

	IACT Arrays	Ground-particle Arrays
Field of view	3°–10°	90°
Duty cycle	10%–30%	>95%
Energy range	30 GeV – >100 TeV	∼500 GeV – >100 TeV
Angular resolution	0.05°–0.02°	0.4°–0.1°
Energy resolution	∼7%	60%–20%
Background rejection	>95%	90%–99.8%

Science Cases:

- PWNe, Pulsar Halos, PeVatron sources
- Fermi Bubbles, DM from GC halo

…

Gamma Hadron Separation

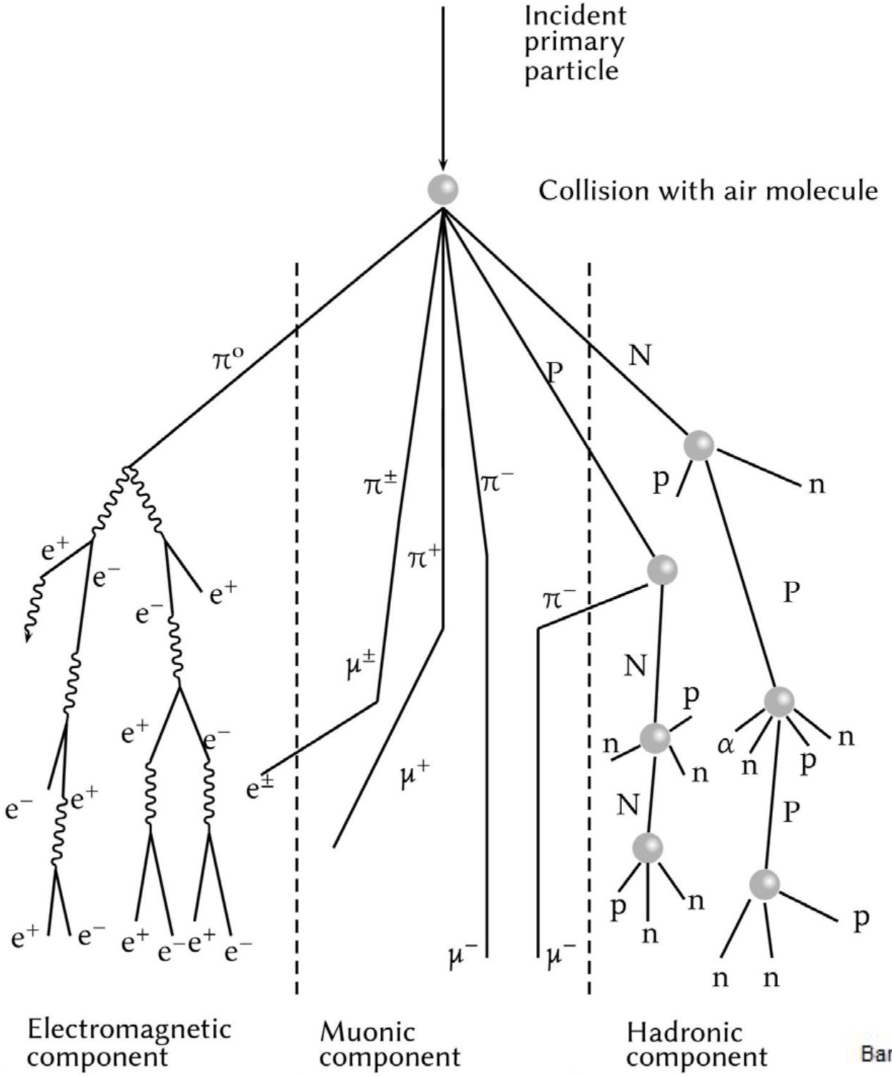
For HAWC like detectors

Common challenge with IACTs:

- Rejection of the huge background of air showers from charged, close to isotropic, cosmic rays.

SWGO still in design phase:

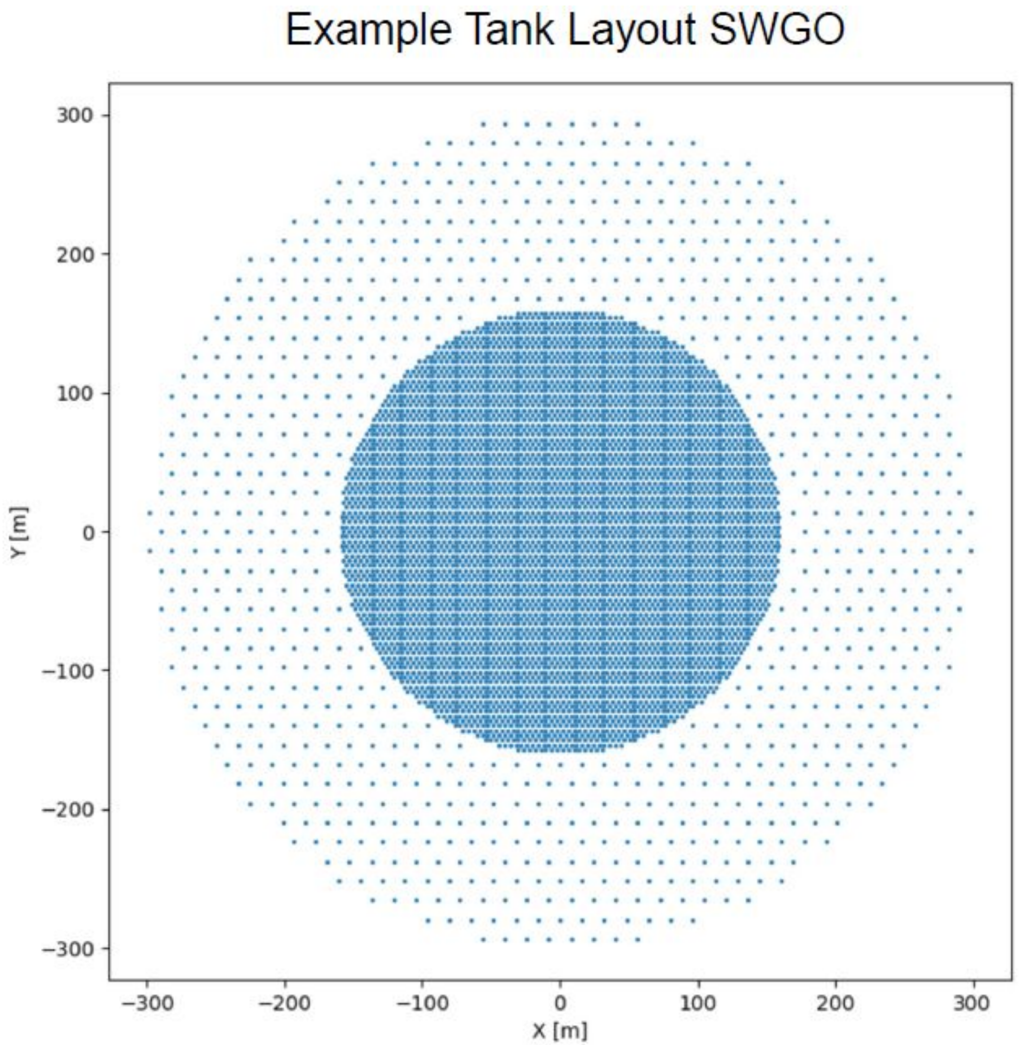
- Muon tagging power (and thus G/H separation) varies by detector design
- Different configurations evaluated at fixed cost



Barrantes et al. 2018

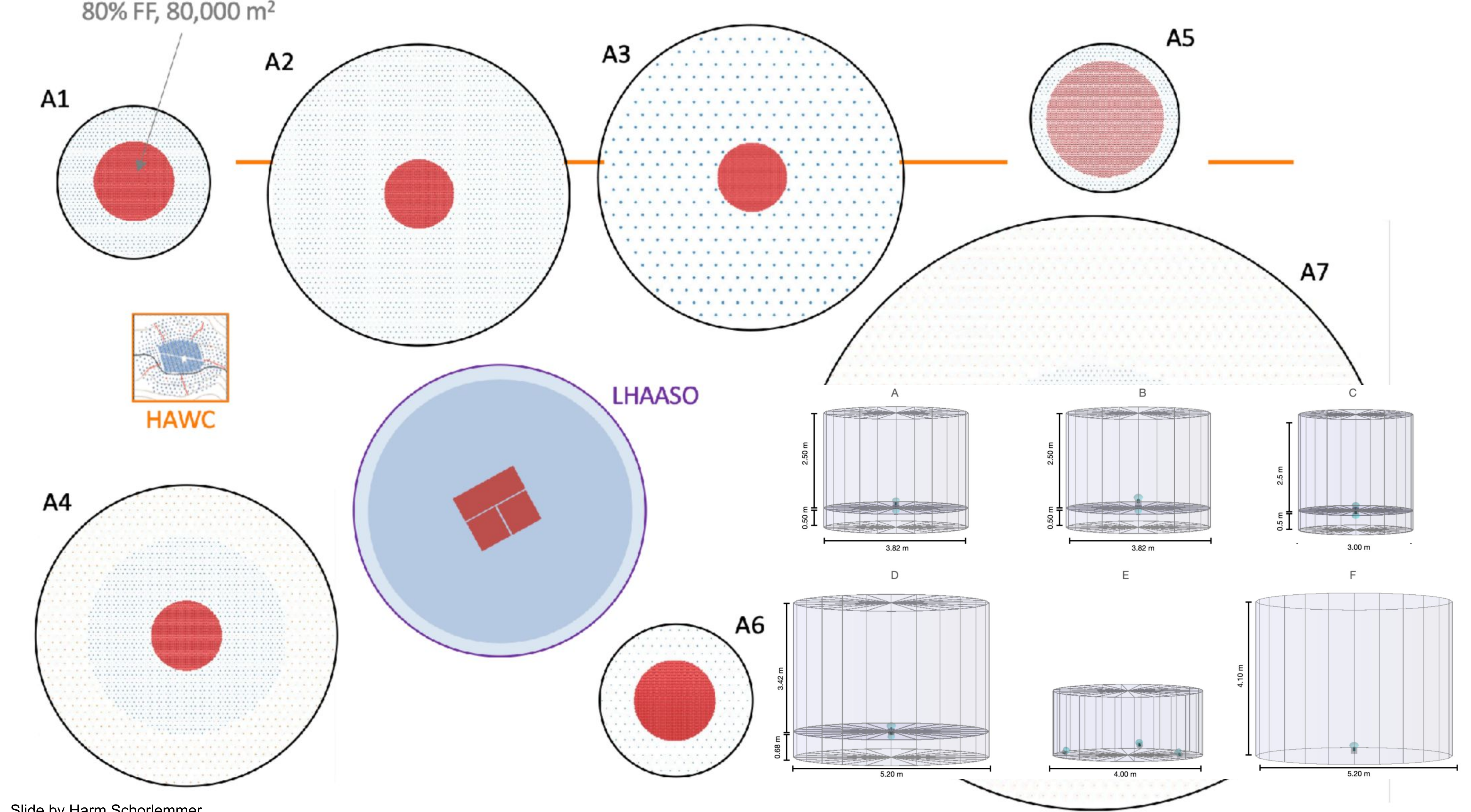
Recap

Status last year in Erlangen



- Deep Learning pipeline setup for SWGO by Jonas
- First promising model trained for a single layout
- Model still relied on the Monte-Carlo shower core
- No comparison plots to the standard method yet

They are ready now, but I'll first walk you through our architecture



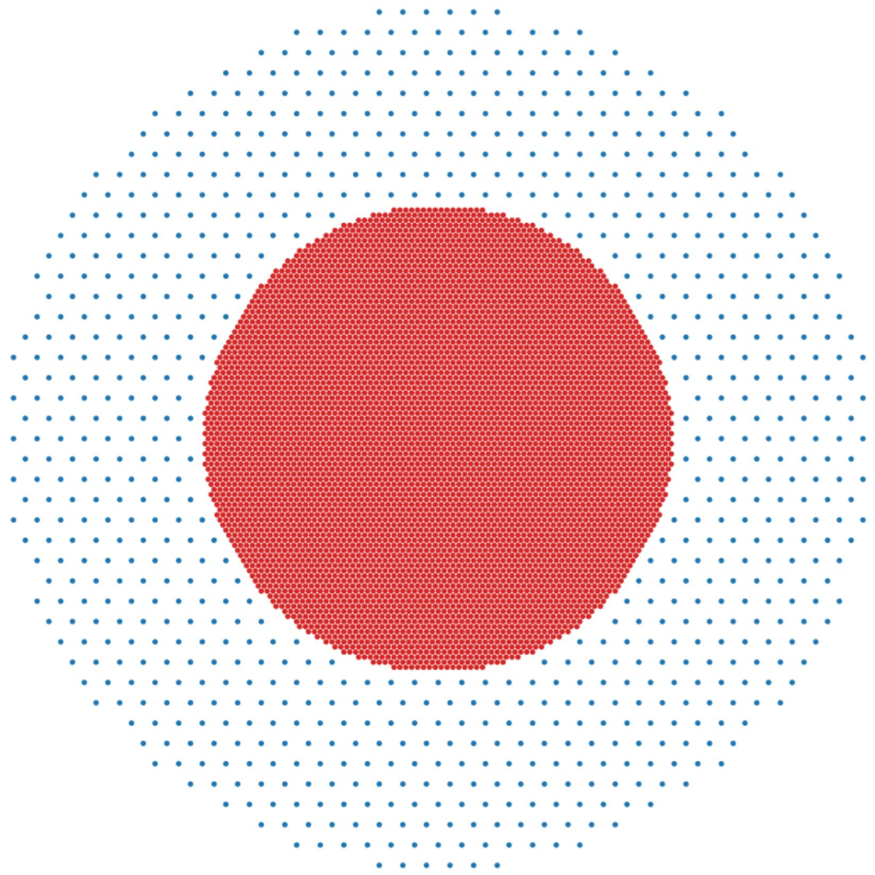
Slide by Harm Schorlemmer

Why use Graph Neural Networks (GNNs)?

Motivation

- Want to improve over standard machine learning methods
- Challenging to exploit underlying symmetry using Convolutional Neural Networks (CNNs)
- Signal footprint is sparse
- Good flexibility as GNNs work on non-regular grids (and perform well on them)
- Easy adaptation to different array layouts and tank designs

Example Layout for SWGO



Inputs and Normalization

Exploit footprint using GNNs

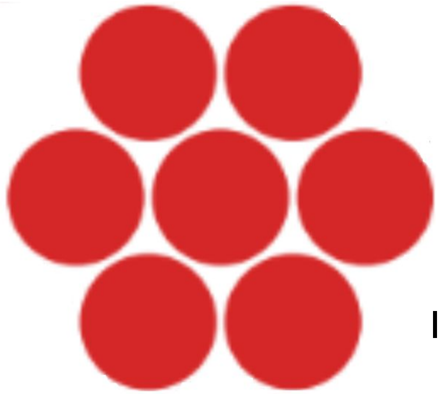
Network inputs:

- Graphs of triggered stations: k-nearest neighbors (kNN) for positions (k = 7)
- Features: Positions x, y, PMT time t, PMT signal S (charge)

Normalization:

- Signals: Logarithmic rescaling $S' = \log_{10}(1 + S) / \sigma$
- Positions (x and y): Simple rescaling $tank' = tank_{pos} / \sigma$
- Time: Z-score normalization $t' = (t - \mu) / \sigma$

Outer array

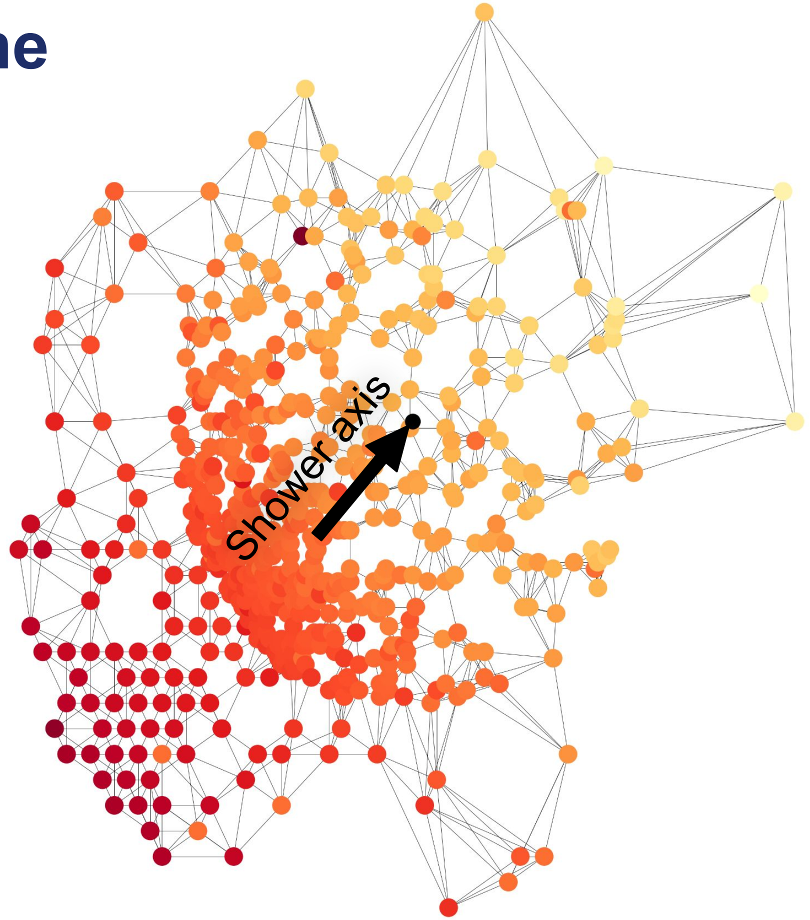


Inner array

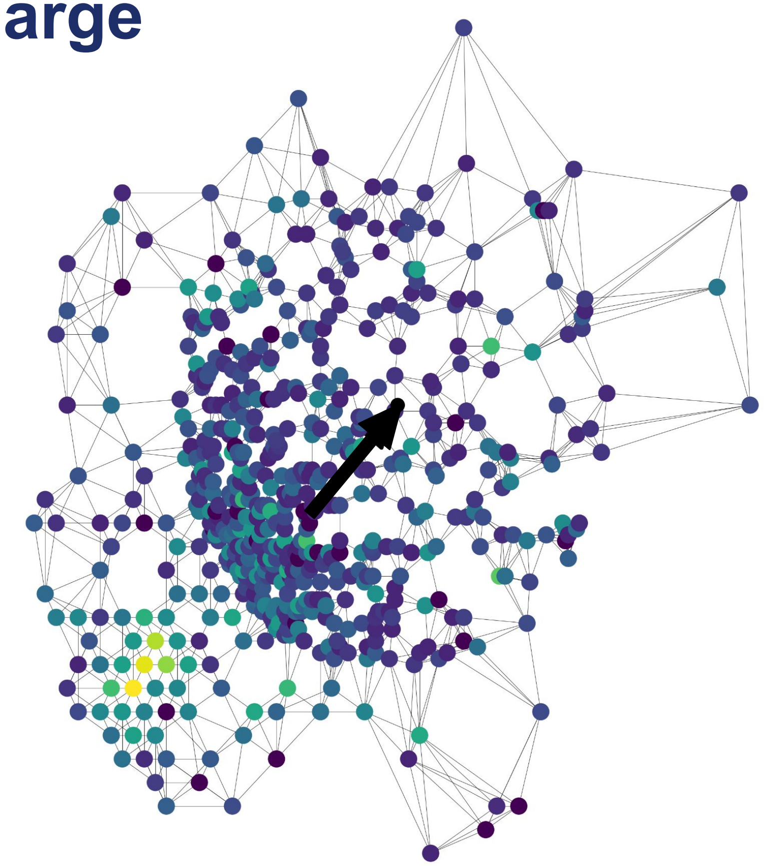
Example of Triggered Graph

Proton | 31.8° zenith angle | 7.5 TeV

Time



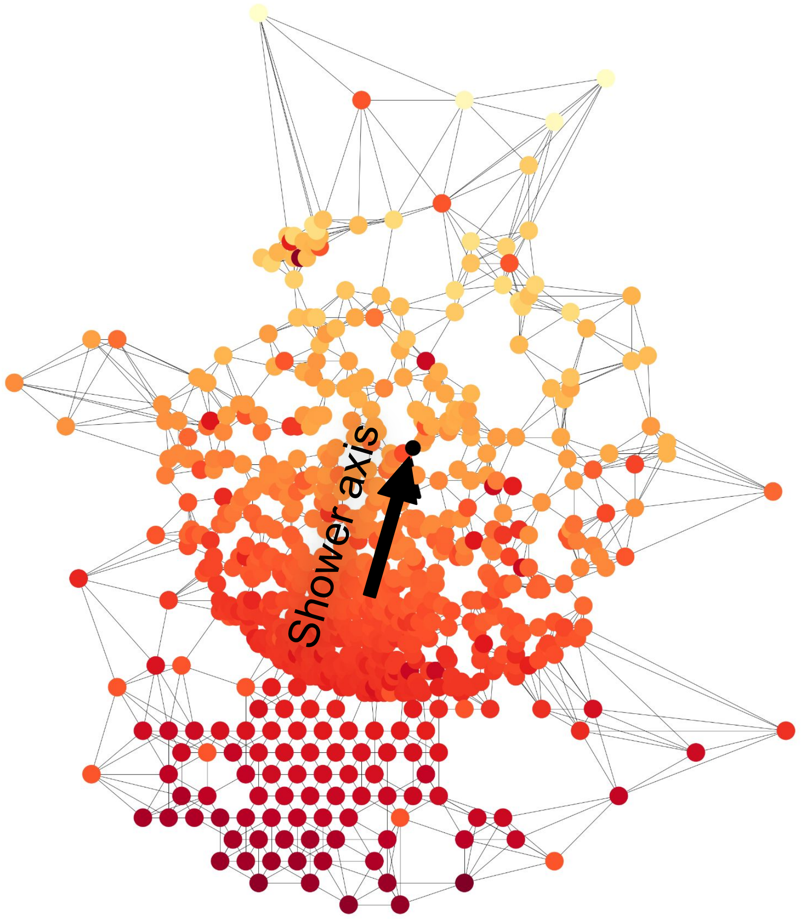
Charge



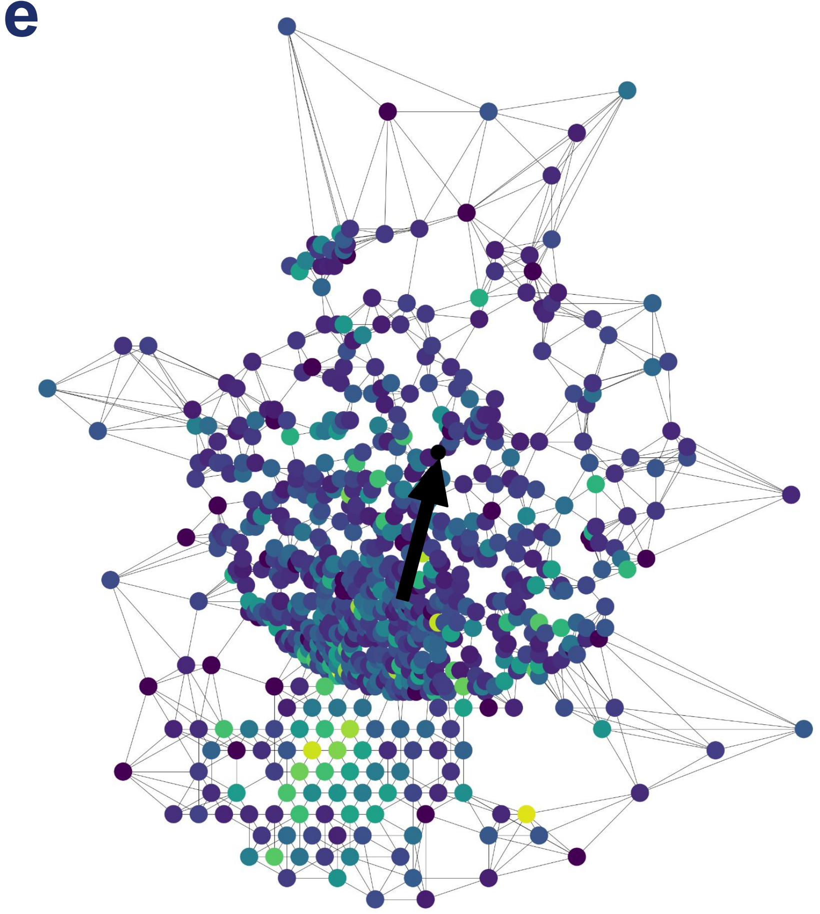
Example of Triggered Graph

Gamma | 27.7° zenith angle | 5.9 TeV

Time

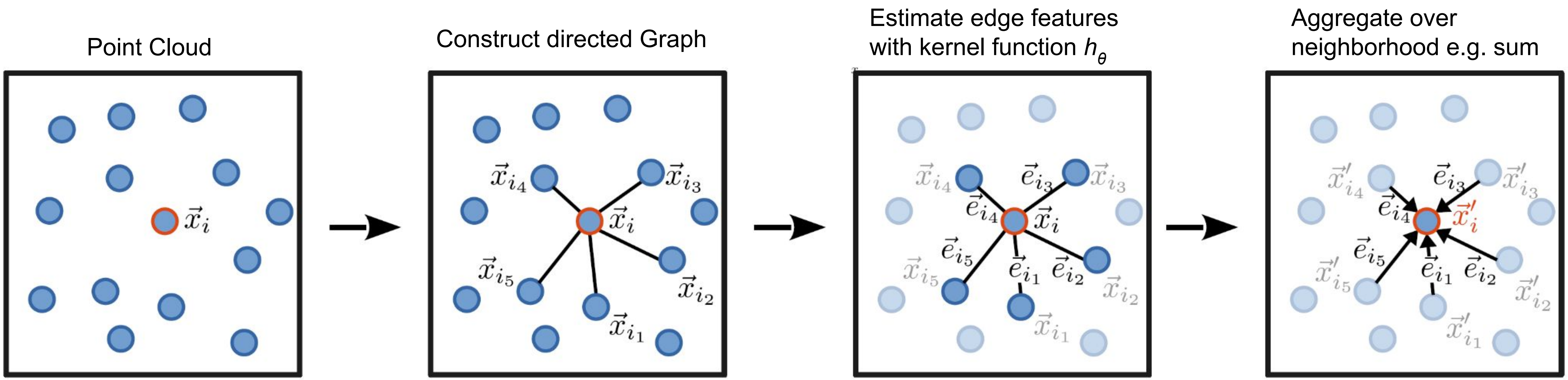


Charge



Convolution in GNNs

EdgeConvolution

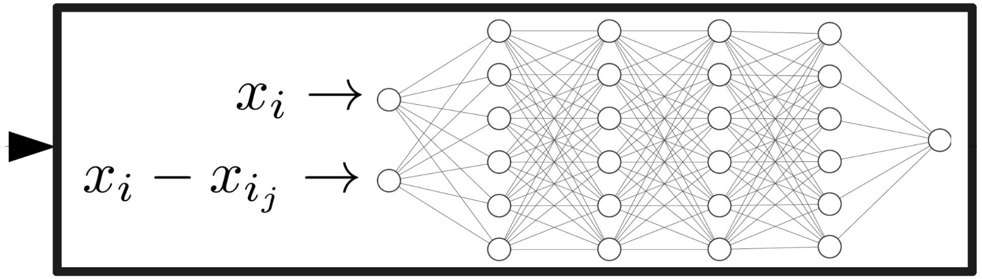


Kernel function is Neural Network!

Basic steps of edge convolution:

- Definition of graph (here with kNN algorithm)
- Estimate edge features by convolving with kernel function h_θ
- Aggregation over the neighborhood

$$h_\theta(x_i, x_i - x_{i_j})$$

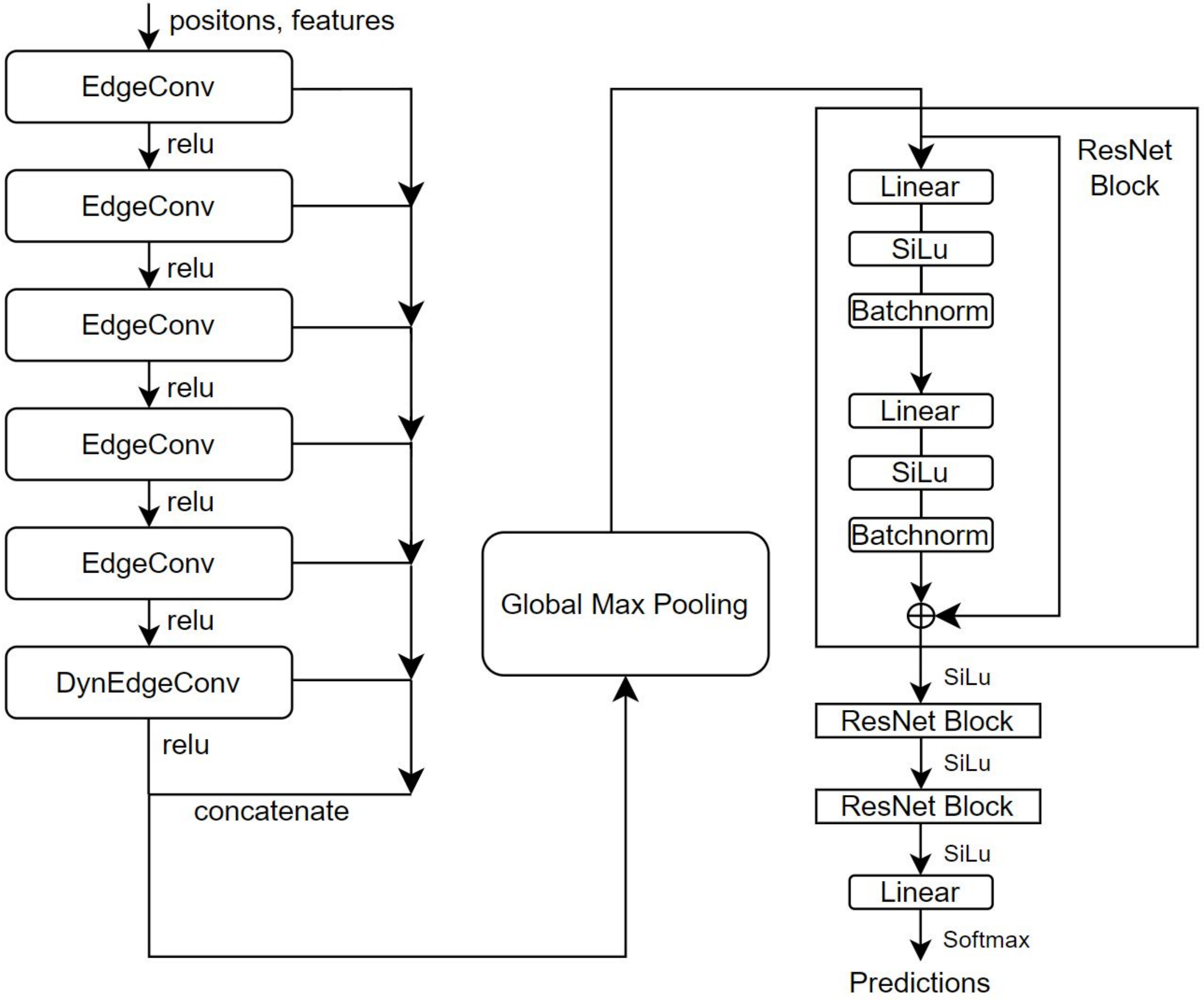


M. Erdmann et al. (2021)

Architecture Sketch

- Train GNN using GPUs (Nvidia A40 / A100)
- Implemented using PyTorch_Geometric
- ~500k trainable parameters

Hyperparameter optimization OPTUNA
(Random search, 70x trainings)
- Learning rate
- Decay factor
- Dropout
- Batchnorm
- ResNet Layers
- EdgeConv Layers
- Kernel features
- DynEdgeConv

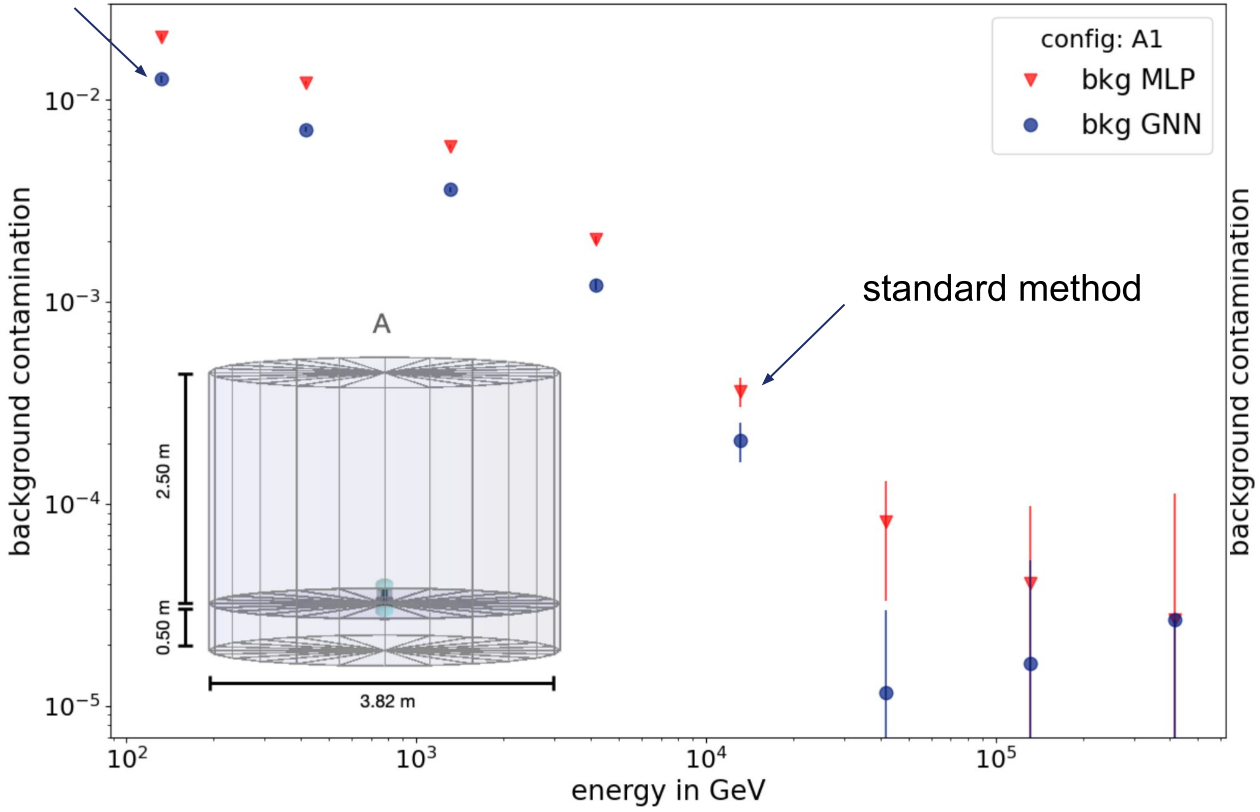


Comparison to regular MLP

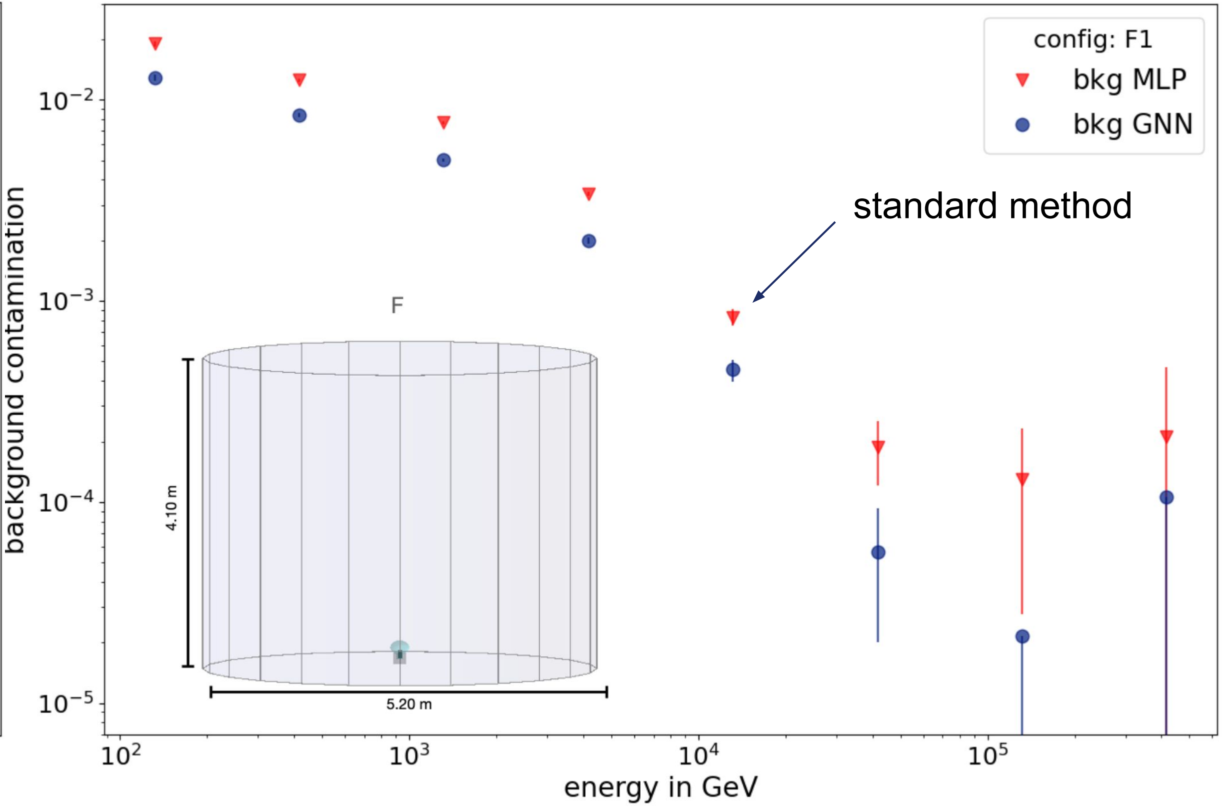
A1 double layer tank, F1 HAWC-like single layer tank

~99% rejection

SWGO reference tank (two layers)



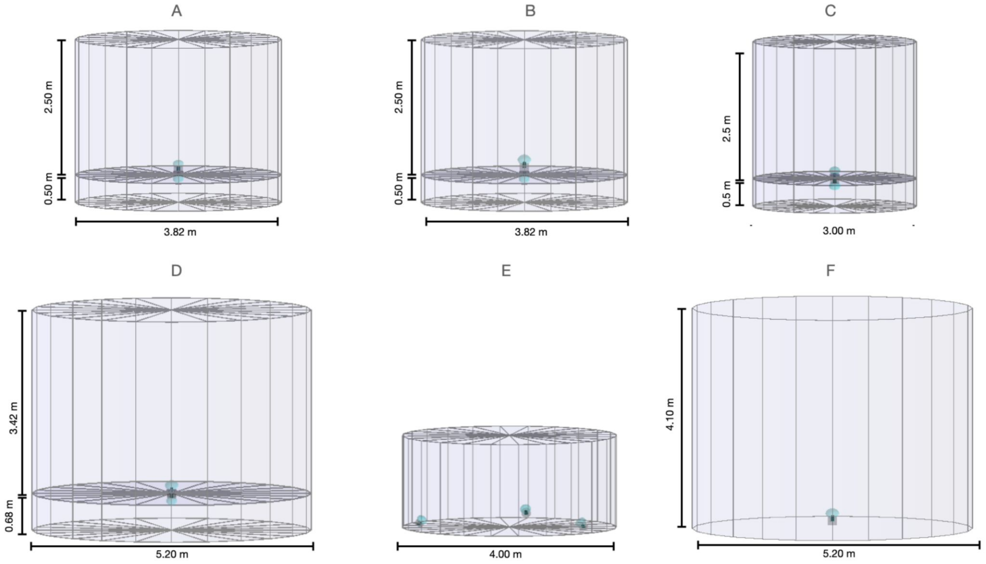
HAWC-like tank (single layer)

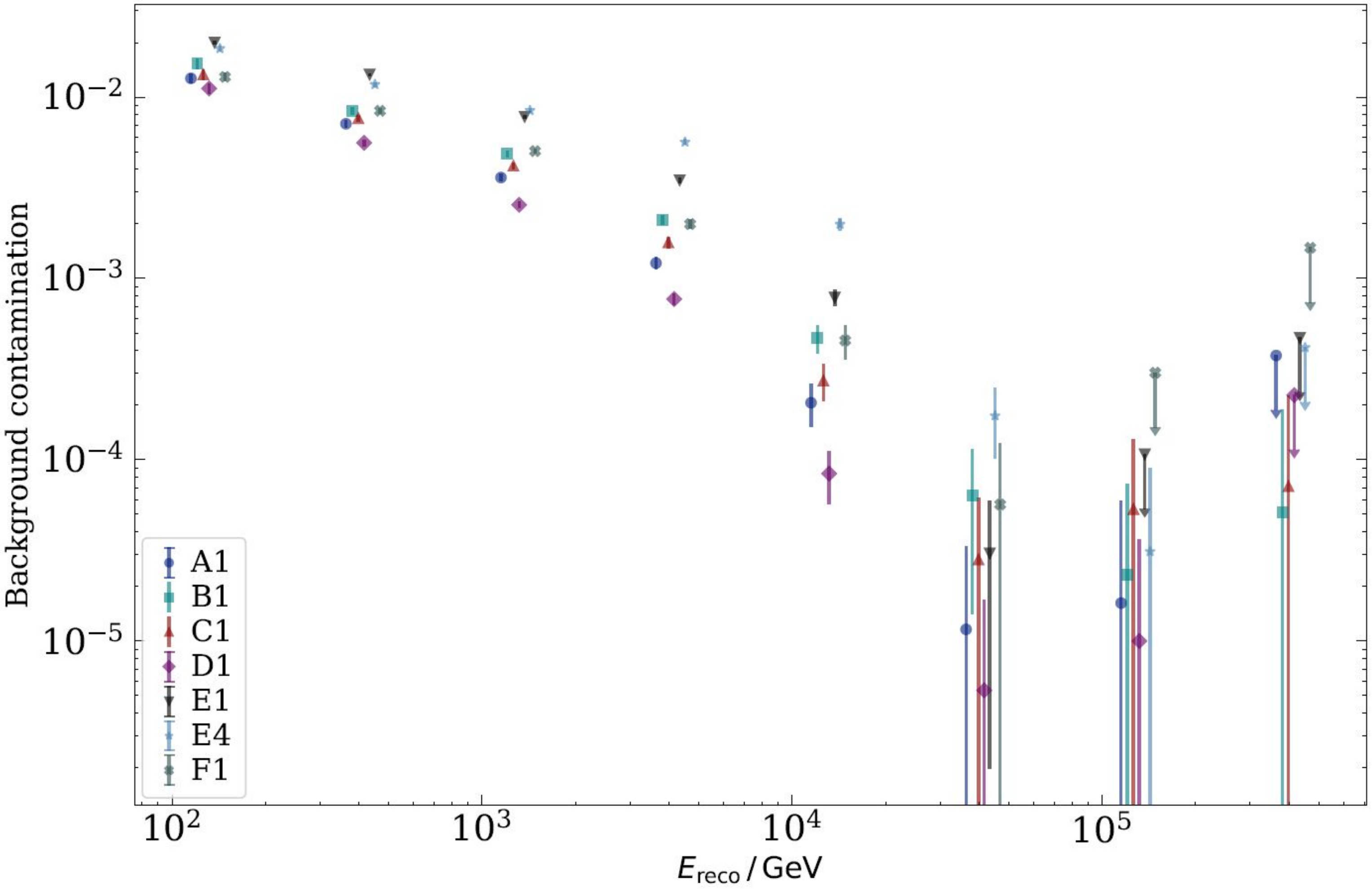


	IACT Arrays	Ground-particle Arrays
Background rejection	>95%	90%–99.8%

Different Tank designs compared

- Large double layer tank offers the best gamma/hadron separation performance using the GNN
- Can have similar separation performance using a smaller double layer tank w.r.t. single layer design
- GNN also works for the small Multi-PMT tanks





Summary

We developed a GNN algorithm for SWGO
- Triggered stations interpreted as graphs
- Promising results for G/H separation

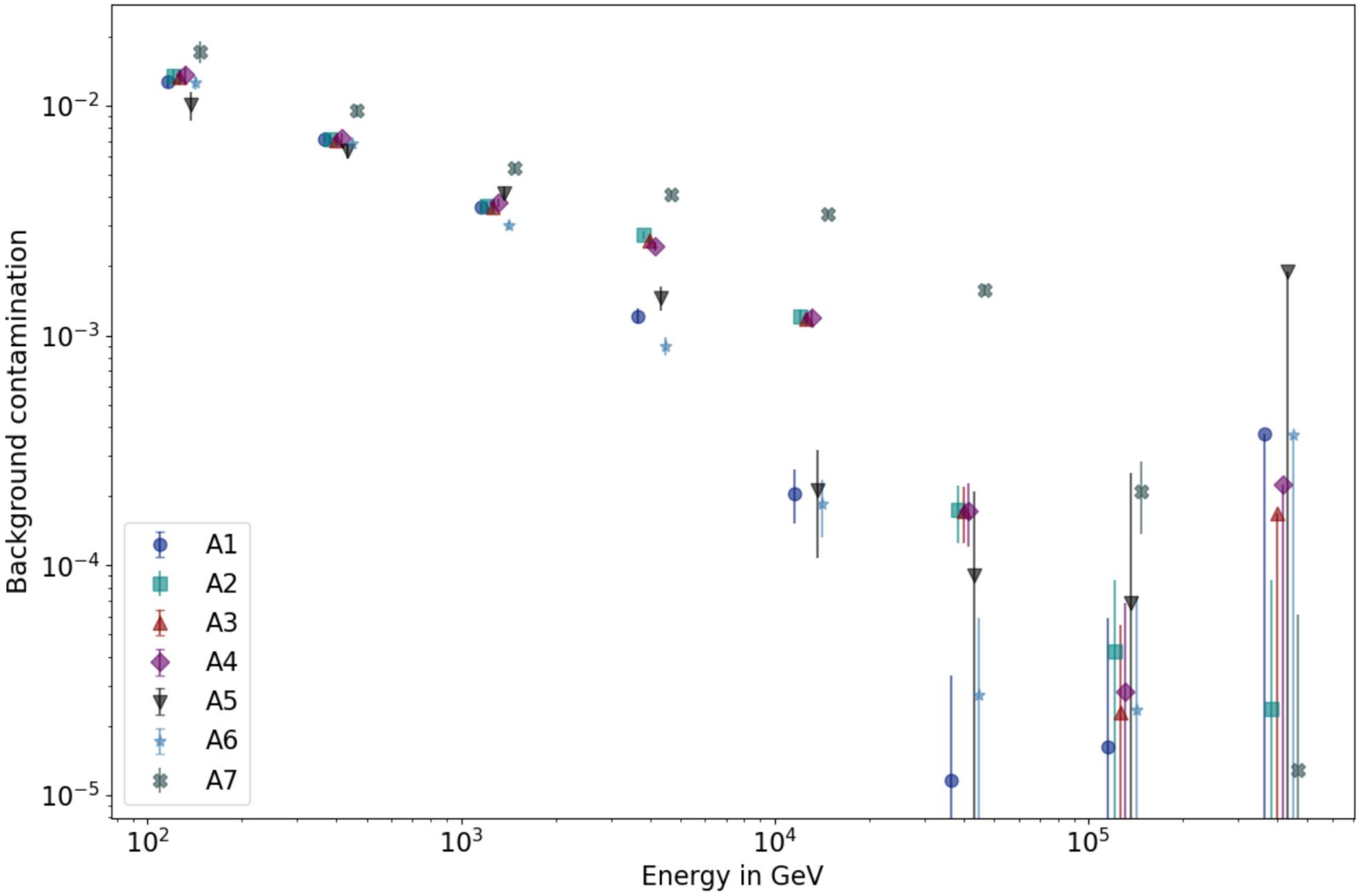
Lots of stuff still left to explore with GNNs
- Different graph transformations (e.g. radius based graphs, …)
- Include neighboring non-triggered stations
- Performance studies with different tanks turned off

Reconstruct other variables using Deep Learning
- GNNs can also solve regression tasks e.g. energy/direction reconstruction
- Explore transformer based approaches (Markus Pirke, Master student at ECAP)
- Go to even lower level information: PMT traces

Thank you for your Attention!

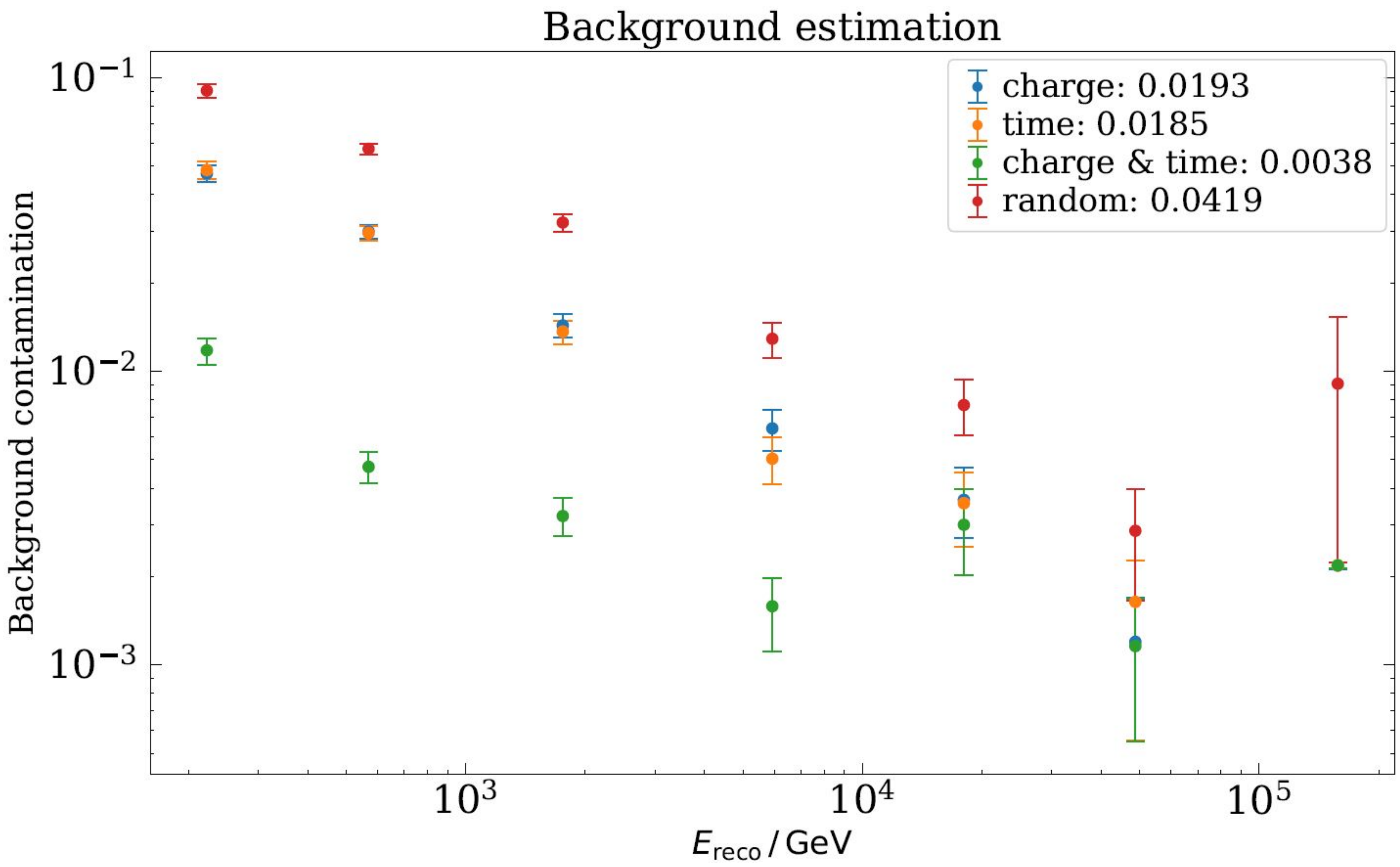
Backup

Different Detector Layouts compared



- As expected the large A7 layout has the worst performance
- Compact layouts like A5 and A6 are good at low energies but lose their advantage at high energies
- A1 layout best across the whole energy range

Different features exploredd



Explore different features

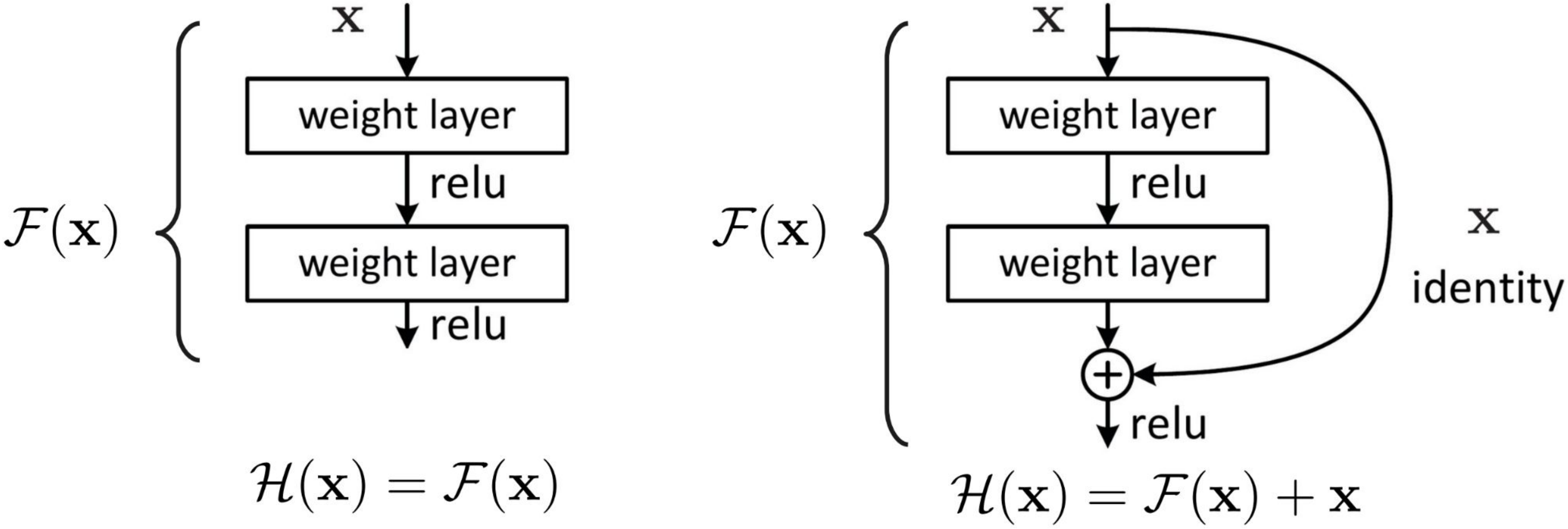
- Network can already separate remarkably well using only positional information
- Charge and time roughly equal in terms of separation power
- Huge performance improvement by combination of time and charge at low energies

ResNets

Residual Neural Networks

ResNets introduce shortcuts with identity mapping

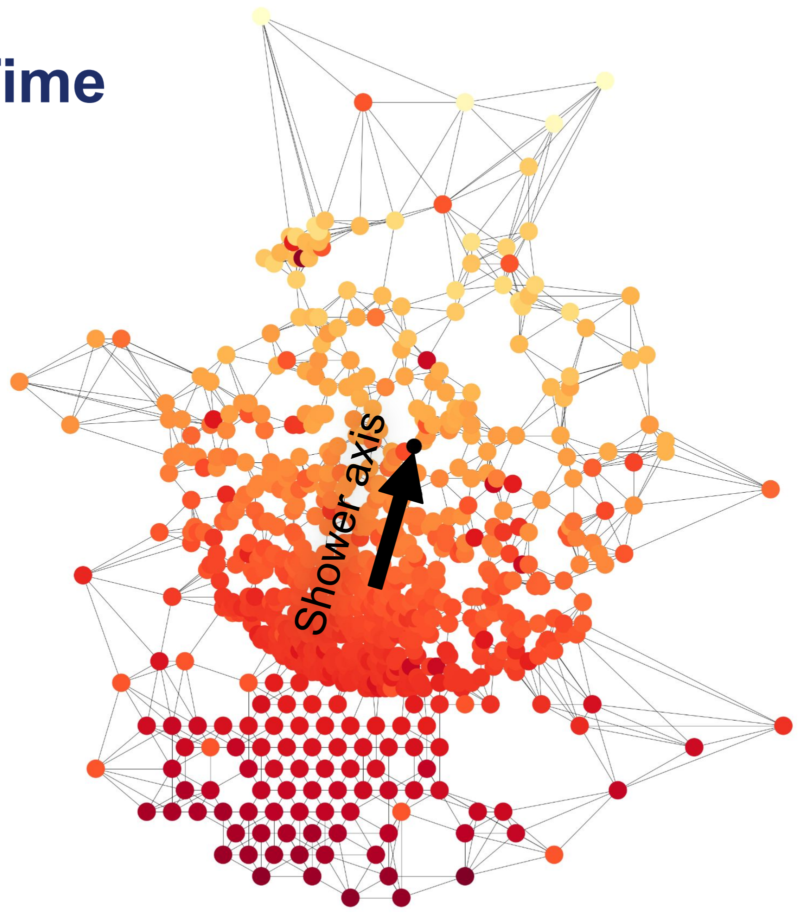
- Weight block learns residual F(x) instead of learning H(x) directly
- Shortcut allows gradient to propagate easily to earlier layers
- Later layers can easily set weights to zero



Example of Triggered Graph

Gamma | 27.7° zenith angle | 5.9 TeV

Time



Charge