Friedrich-Alexander-Universität
Erlangen-Nürnberg

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Gamma Hadron Separation with GNNs for SWGO

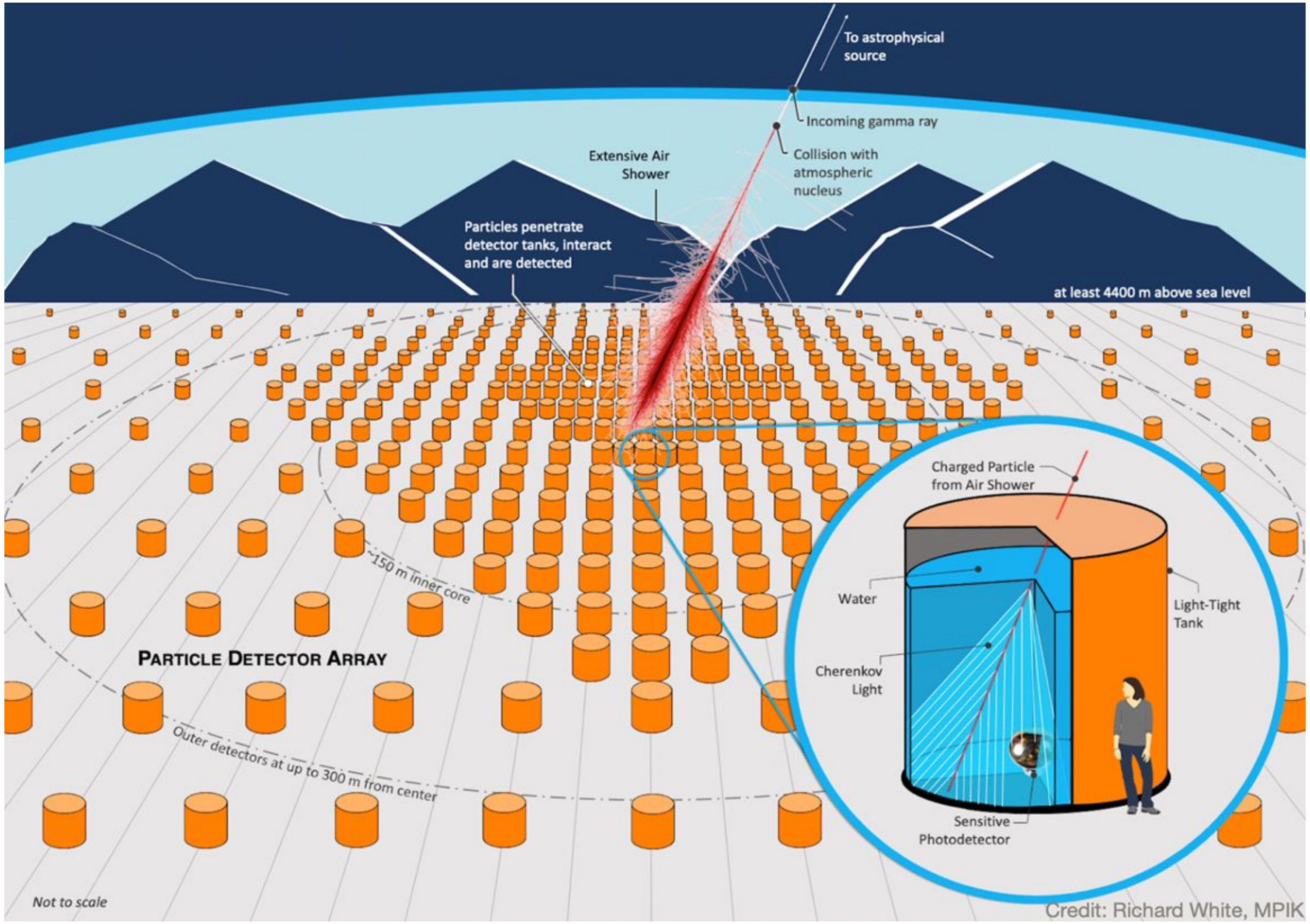
Speaker: Martin Schneider

Prof. Dr. Christopher van Eldik, Dr. Jonas Glombitza

FRANCI Meeting - 13.10.23

Introduction to SWGO

The Southern Wide-field Gamma-ray Observatory

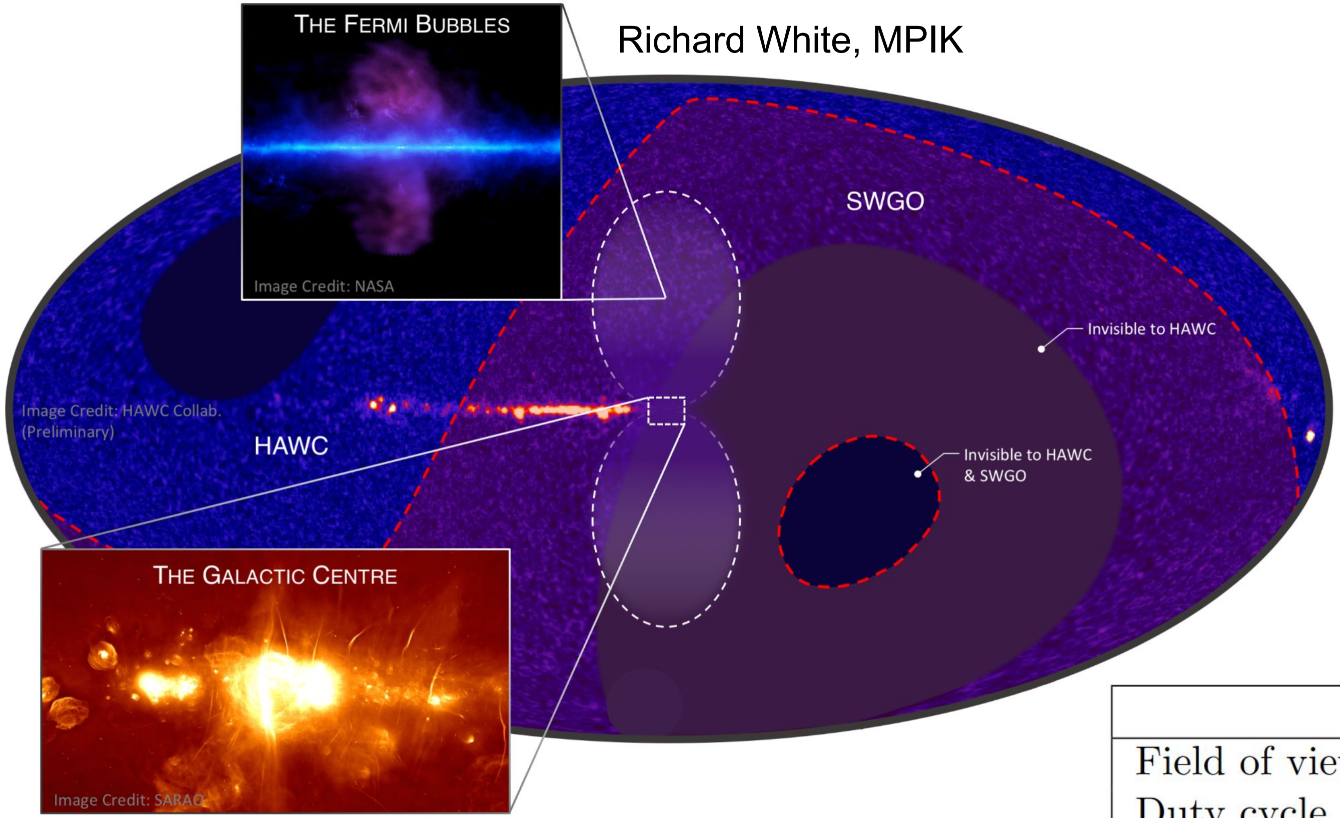


What is SWGO?

- Array of shower particle detectors to measures extensive air showers at ground level
- Will complement IACT based gamma-ray instruments 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 unrivaled)

SGSO 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, TeV Halos, PeVatron sources
- Fermi Bubbles, DM from GC halo
- …

Gamma Hadron Separation

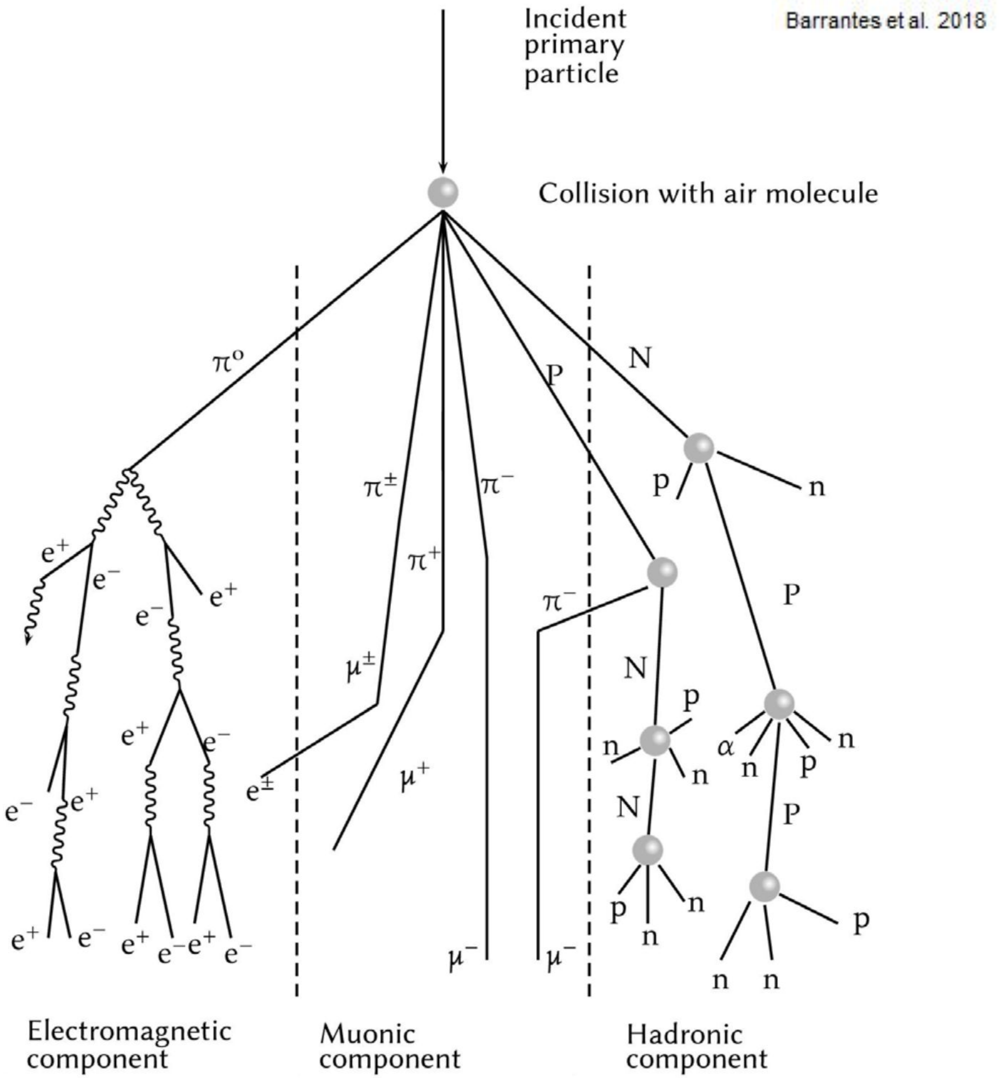
For HAWC like detectors

Common challenge with IACTs:

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

SWGO still in design phase:

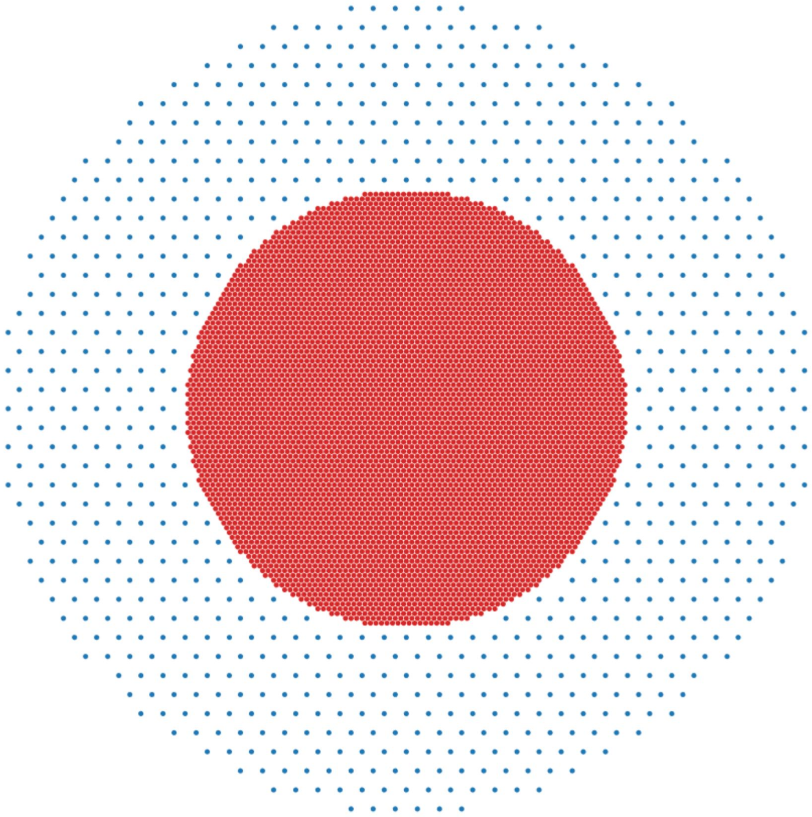
- Muon tagging power (and thus G/H separation) varies by detector design
- Can be improved at moderate additional cost



Why use Graph Neural Networks (GNNs)?

- 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 domains (and perform well on them)
- Easy adaptation to different array layouts and tank designs

Reference Configuration Layout



Inputs and Normalization

Exploit footprint using GNNs

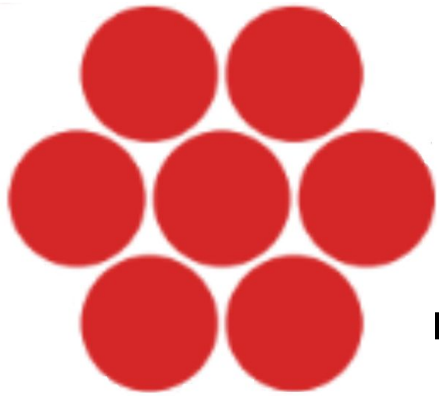
Current inputs:

- Graphs of triggered stations: k-nearest neighbors (kNN) for positions (k = 7)
- Features: x_{pos}, y_{pos}, t_{low}, t_{up}, S_{low}, S_{up}

Normalization:

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

Outer array

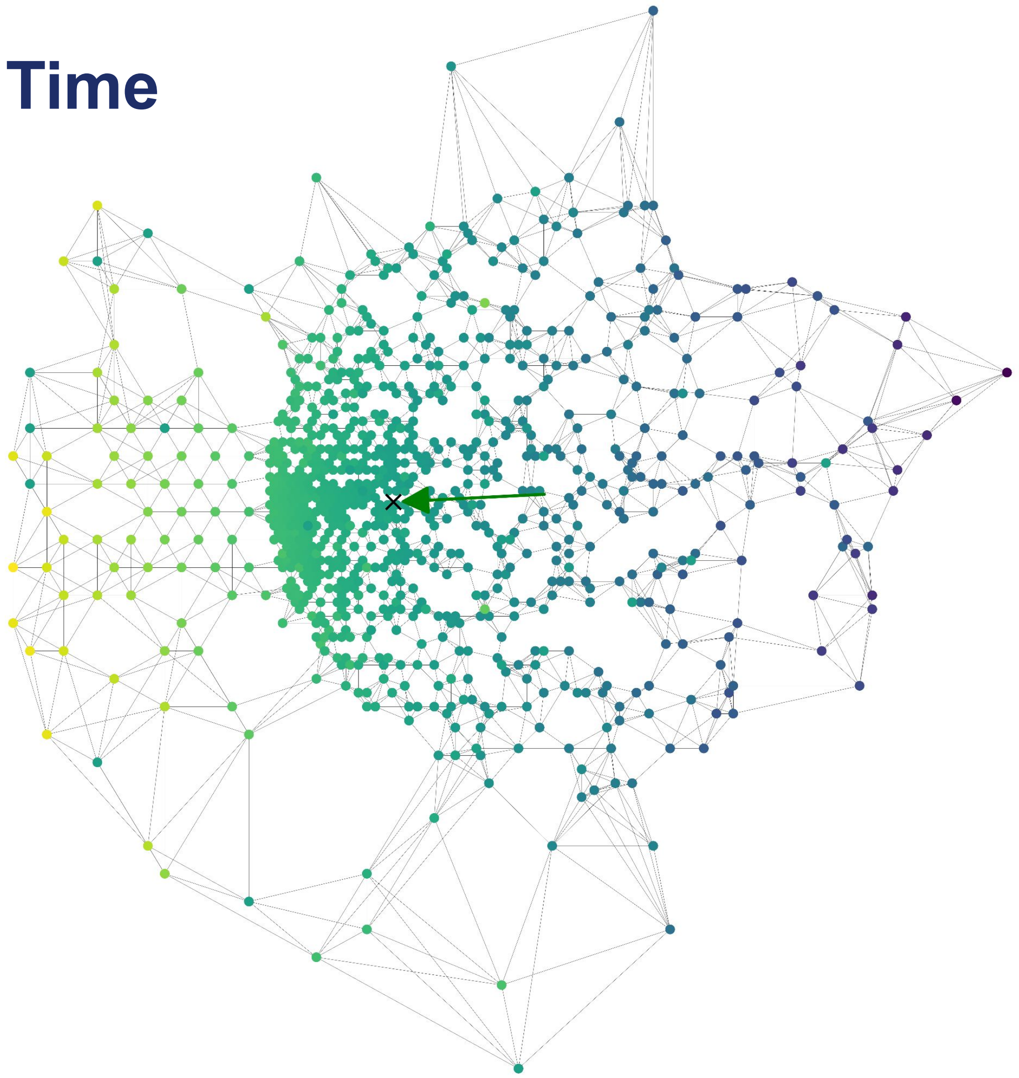


Inner array

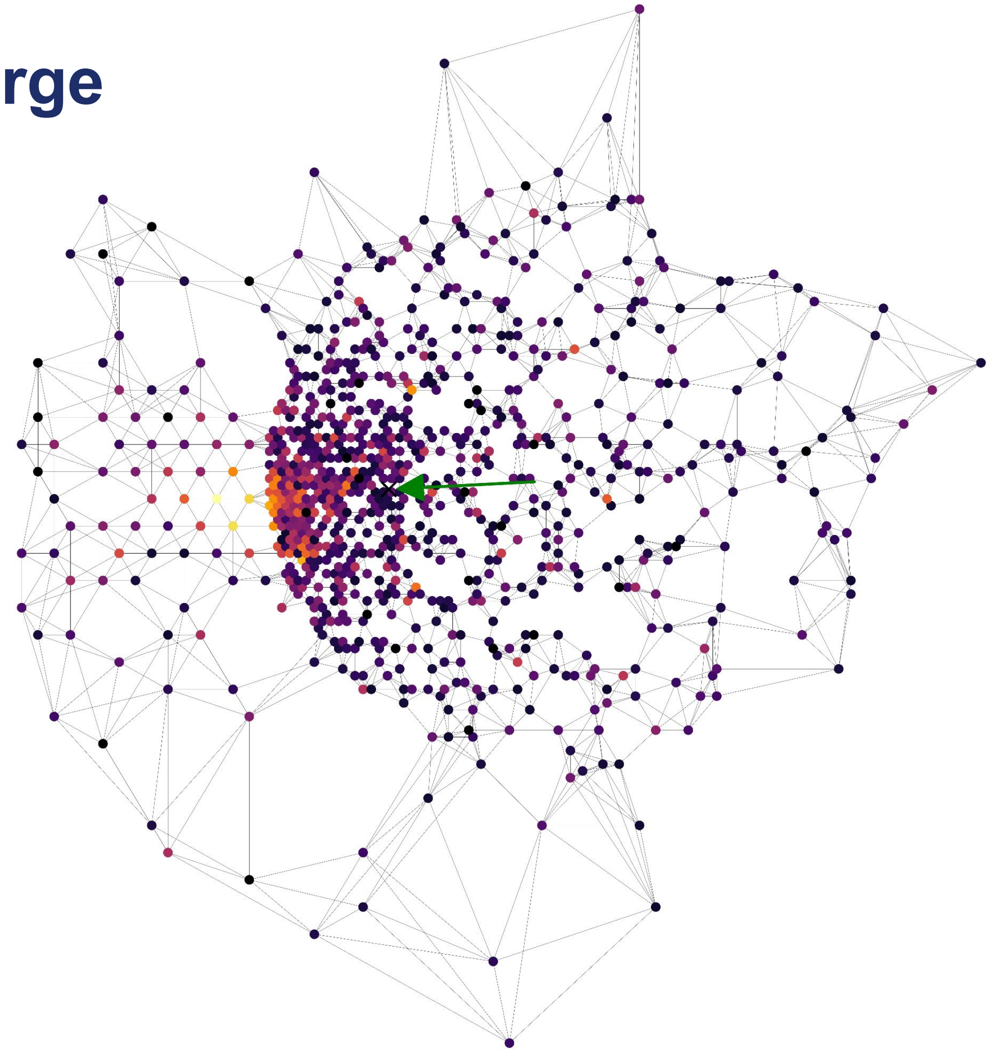
Example of Triggered Graph

Proton | 37.5° zenith angle | 7.5 TeV

Time



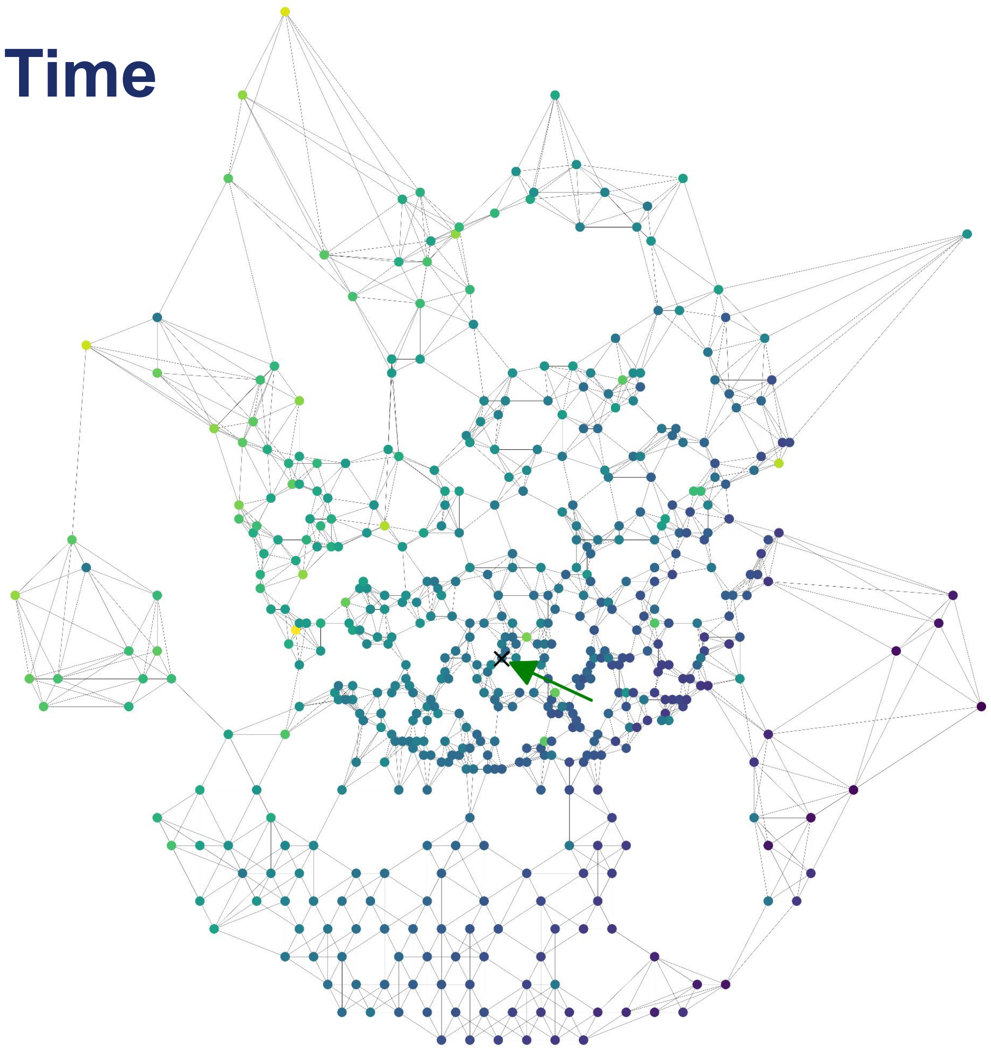
Charge



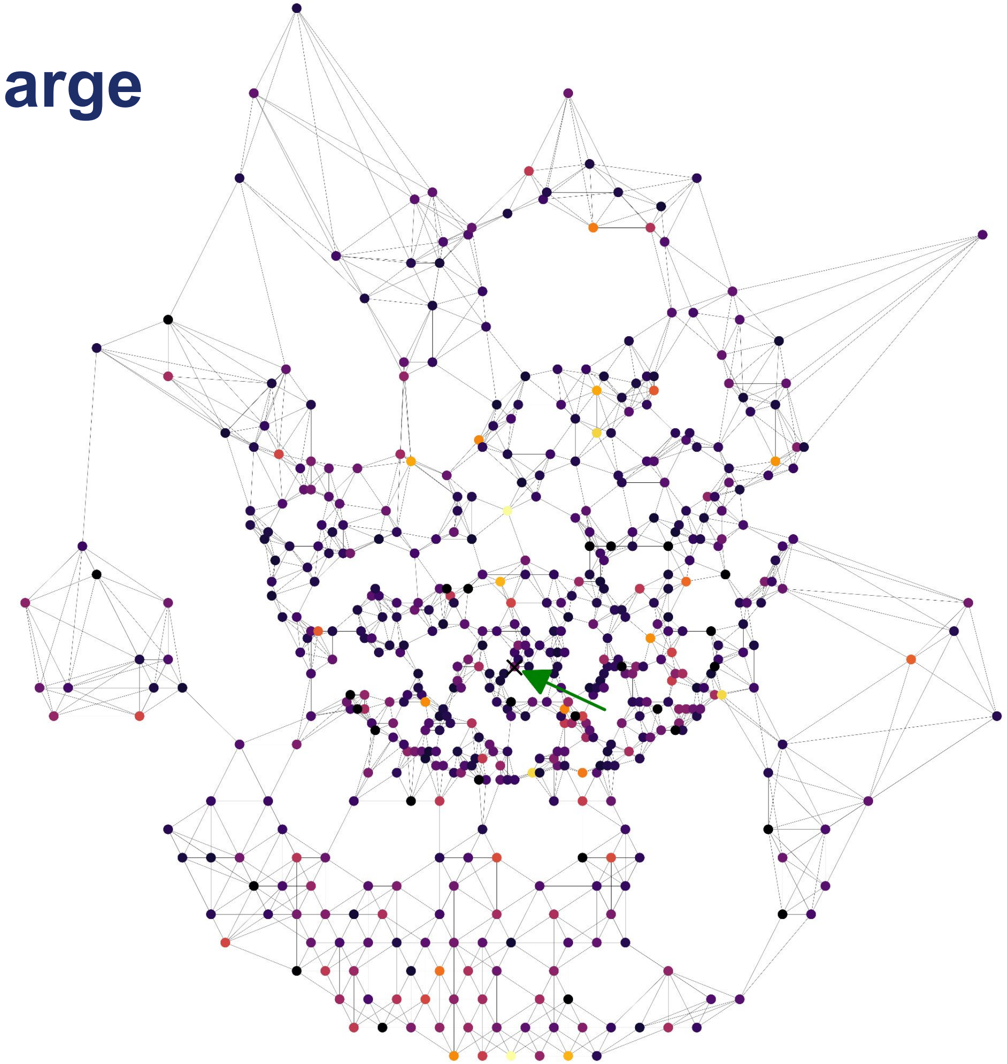
Example of Triggered Graph

Gamma | 28.2° zenith angle | 13.3 TeV

Time



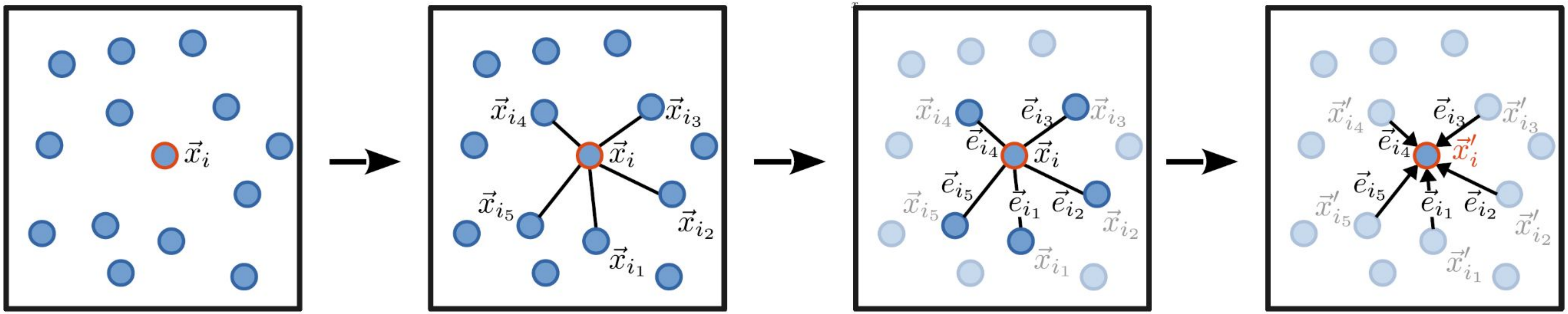
Charge



Convolution in GNNs

EdgeConvolution

Point Cloud → Construct directed Graph → Estimate edge features with kernel function h_θ → Aggregate over neighborhood e.g. sum

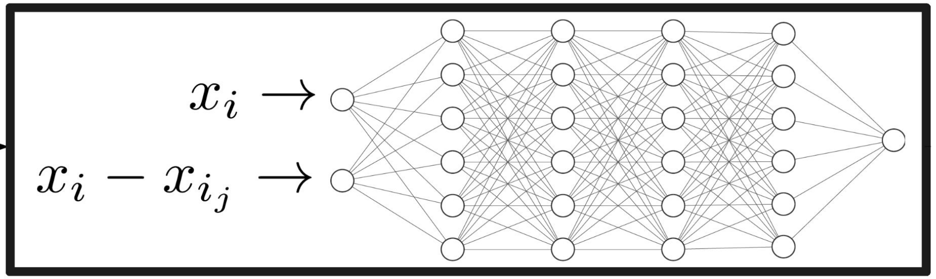


Basic steps of edge convolution:

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

Kernel function is Neural Network!

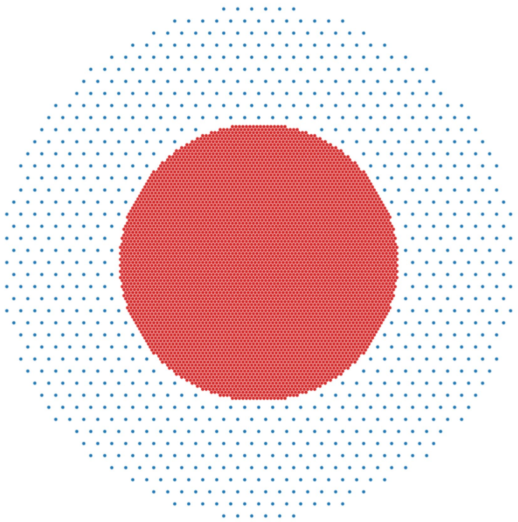
$h_\theta(x_i, x_i - x_{i_j})$

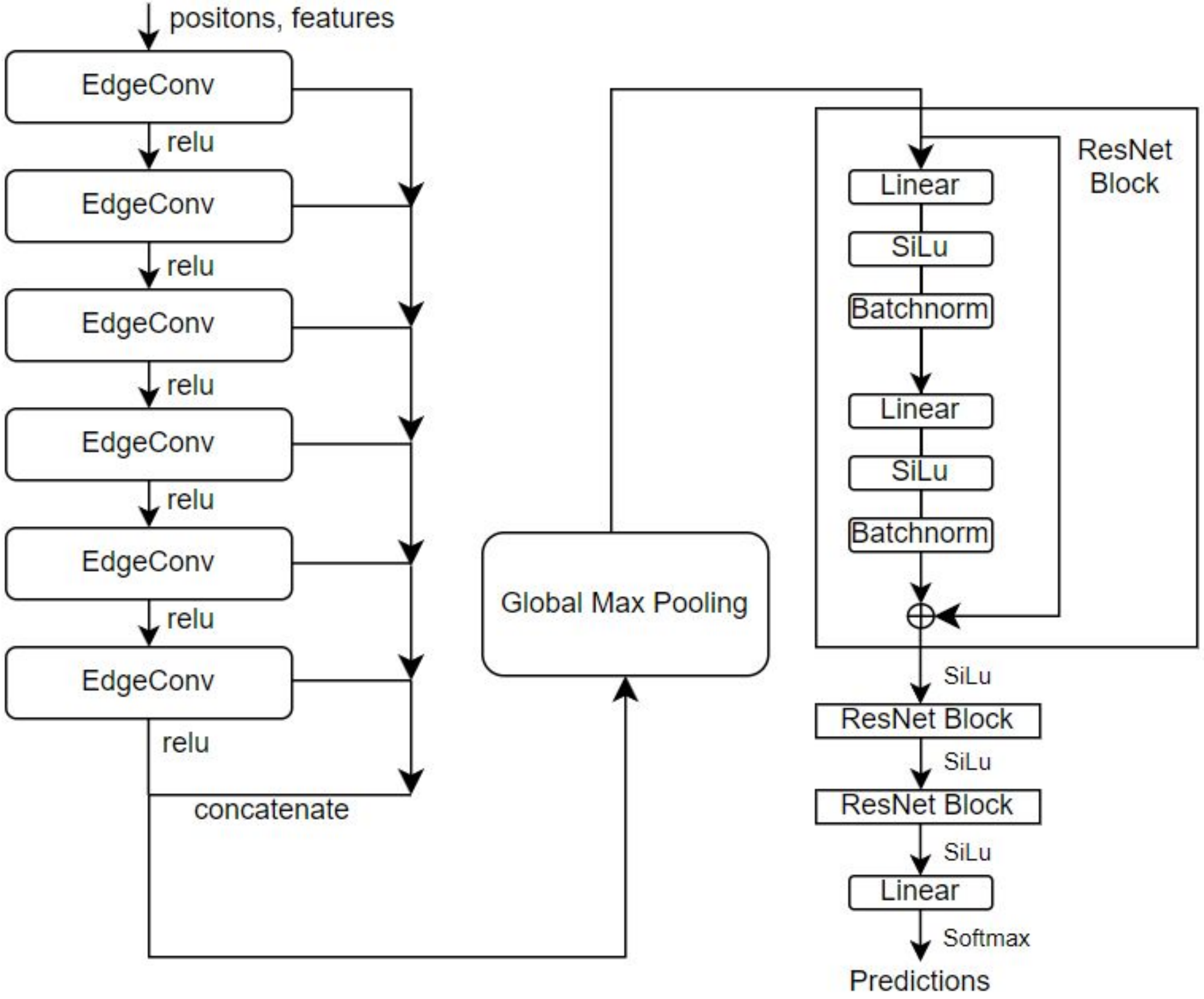


Architecture Sketch

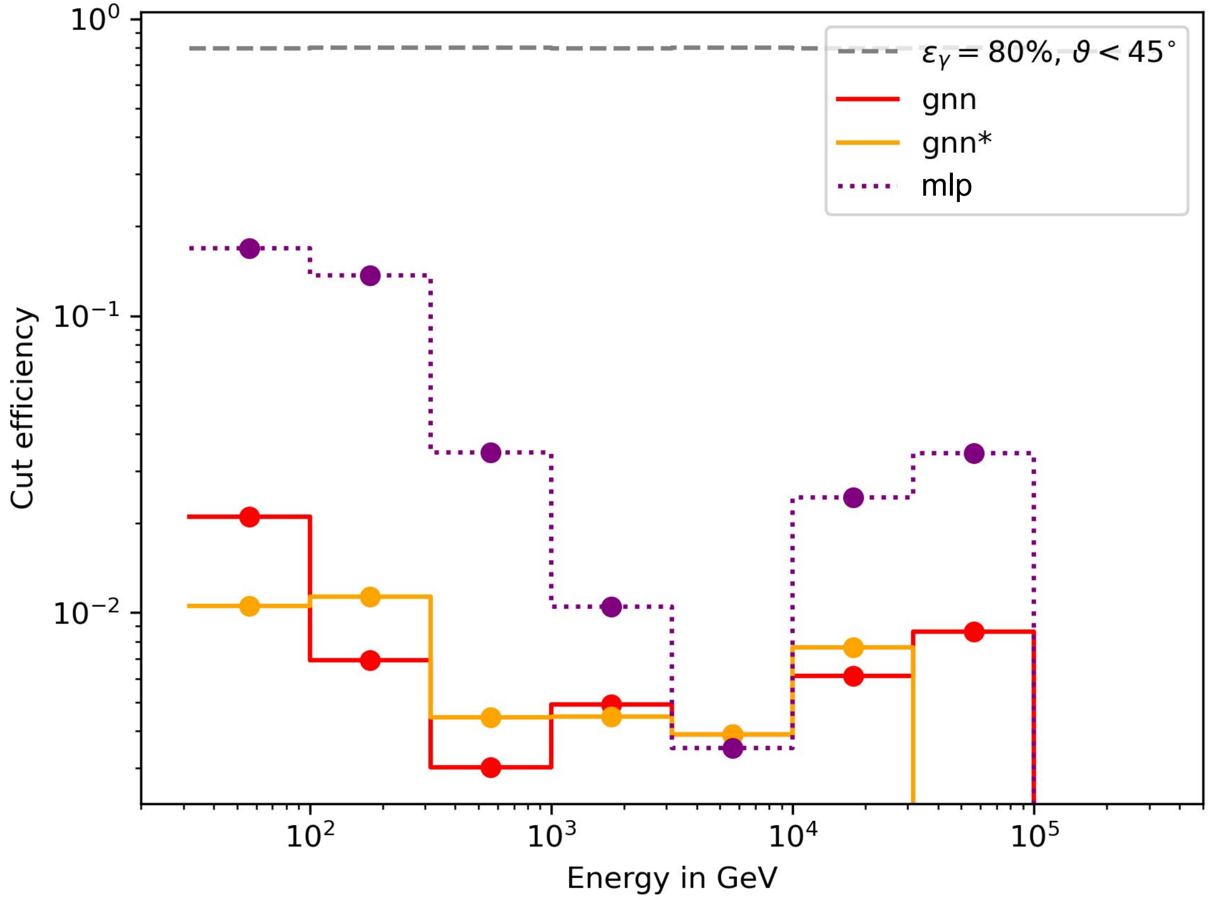
- Train GNN using GPU (Nvidia A100) ~ 1 day
- Implemented using PyTorch_Geometric
- ~500k trainable parameters

Receptive fov estimation:
f(n) ~ $2^{(n+1)} - 1$
→ n = 6 convolutions
for ~100 tank wide array





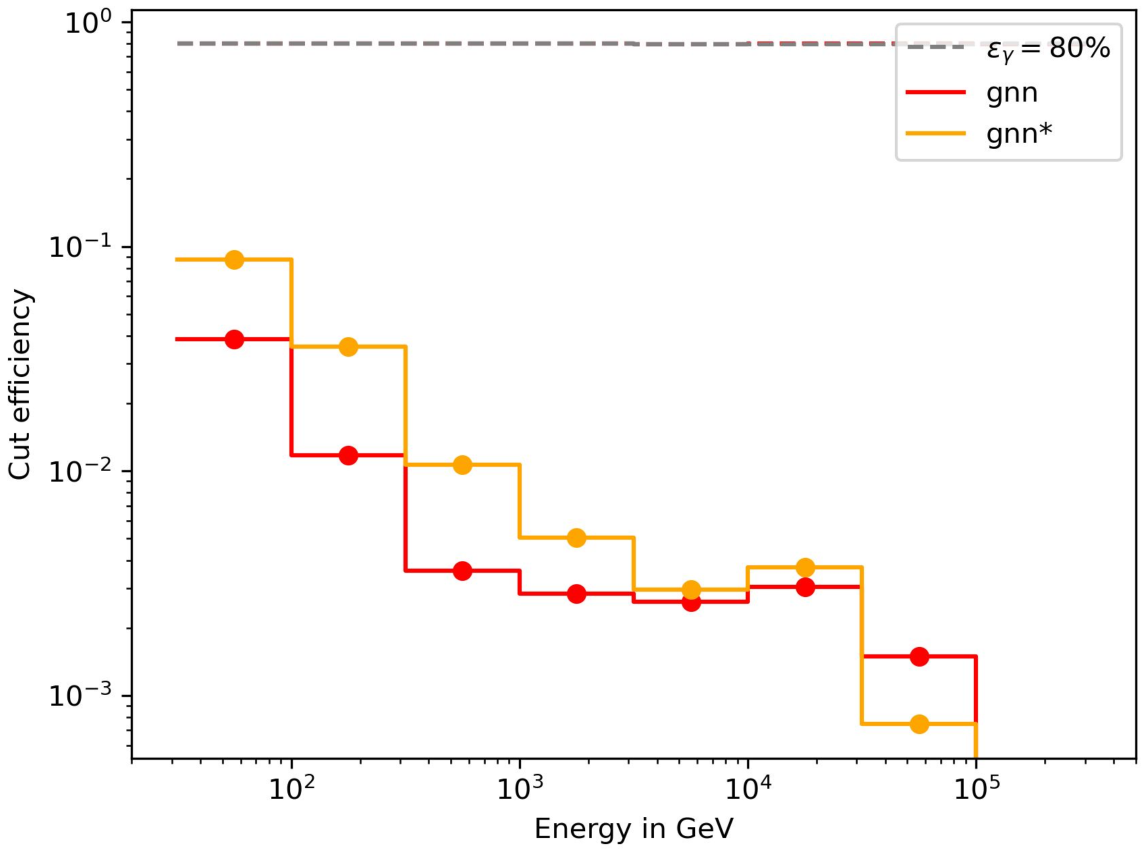
Preliminary Performance vs MLP



For now GNN has been tested for a single configuration:

- Event by Event comparison with quality cuts (needed for MLP)
- GNN outperforms simple MLP implementation (expected)

Preliminary Performance - No quality cuts



For now GNN has been tested for a single configuration

- GNN works well even without quality cuts
- The GNN is seen to effectively utilize the double layer design to improve separation performance

Summary

Develop GNN algorithm for SWGO

- triggered stations interpreted as graphs
- First results for G/H separation promising

Lots of stuff still left to explore with GNNs

- Different graph transformations (e.g. radius based graphs, …)
- Include neighboring non-triggered stations
- Performance studies for different layouts and tank designs

Explore additional task with GNNs

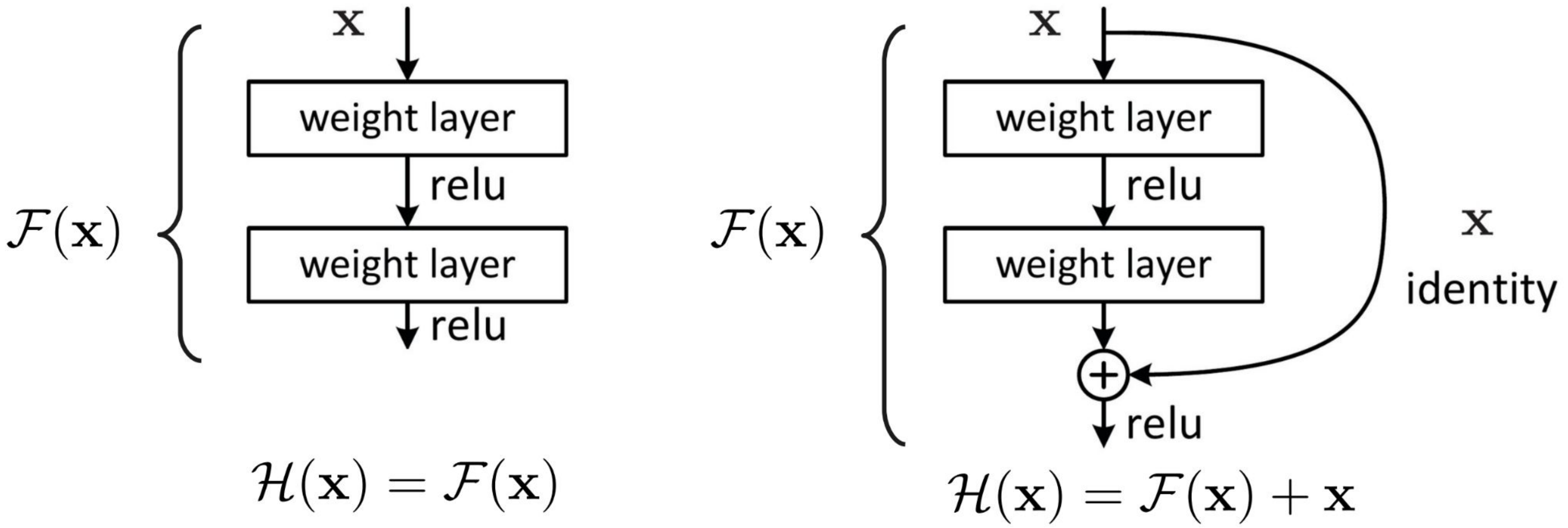
- Apply GNNs to regression tasks e.g. energy/direction reconstruction (Franziska Leitl)
- Explore GNNs in combination with transformer based approaches (Markus Pirke)
- Go deeper using even lower level information

Thank you for your Attention!

Backup

ResNets
Residual Neural Networks

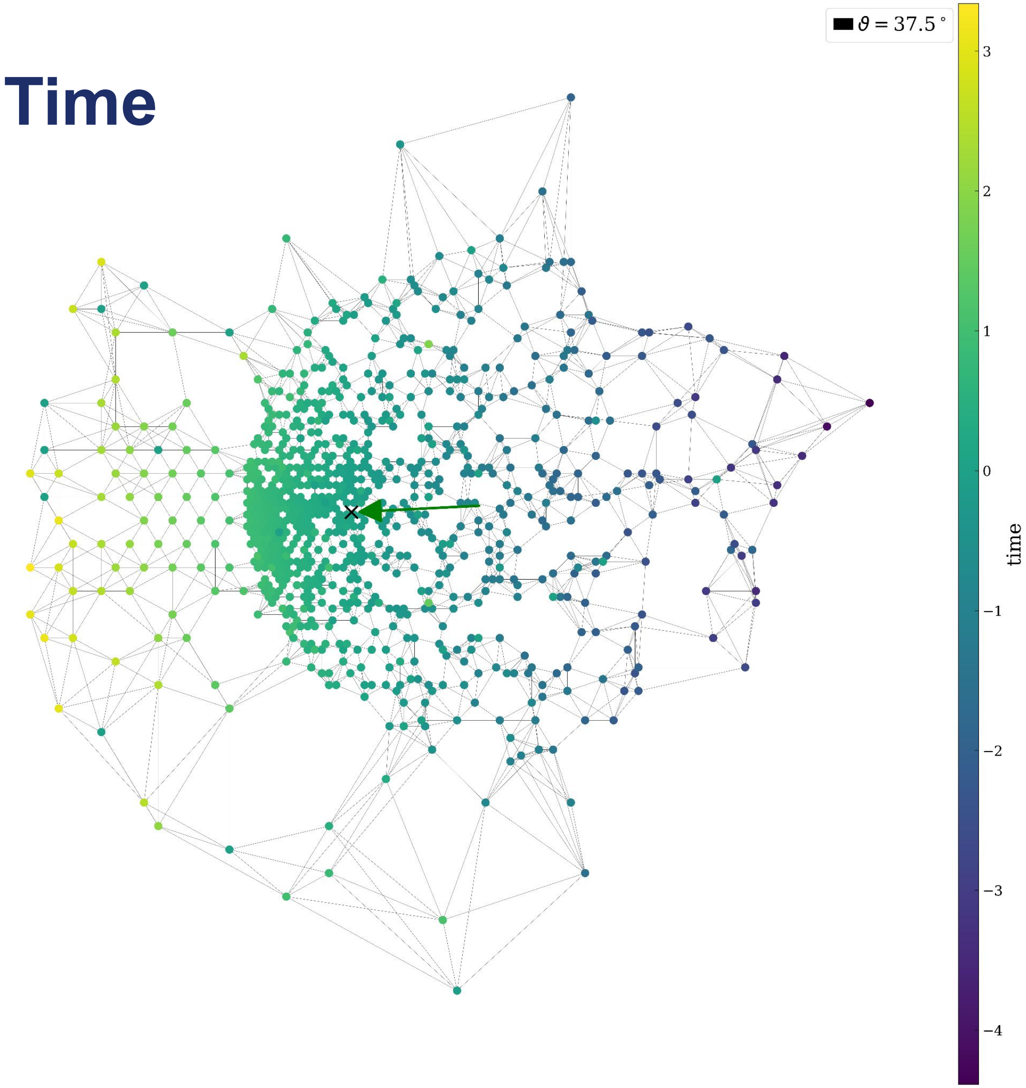
ResNets introduce shortcuts with identity mapping
- Weight block learns residual F(x) instead of learning H(x) directly
- Shortcut allows gradient to propagate easily to earlier layers
- Later layers can easily set weights to zero



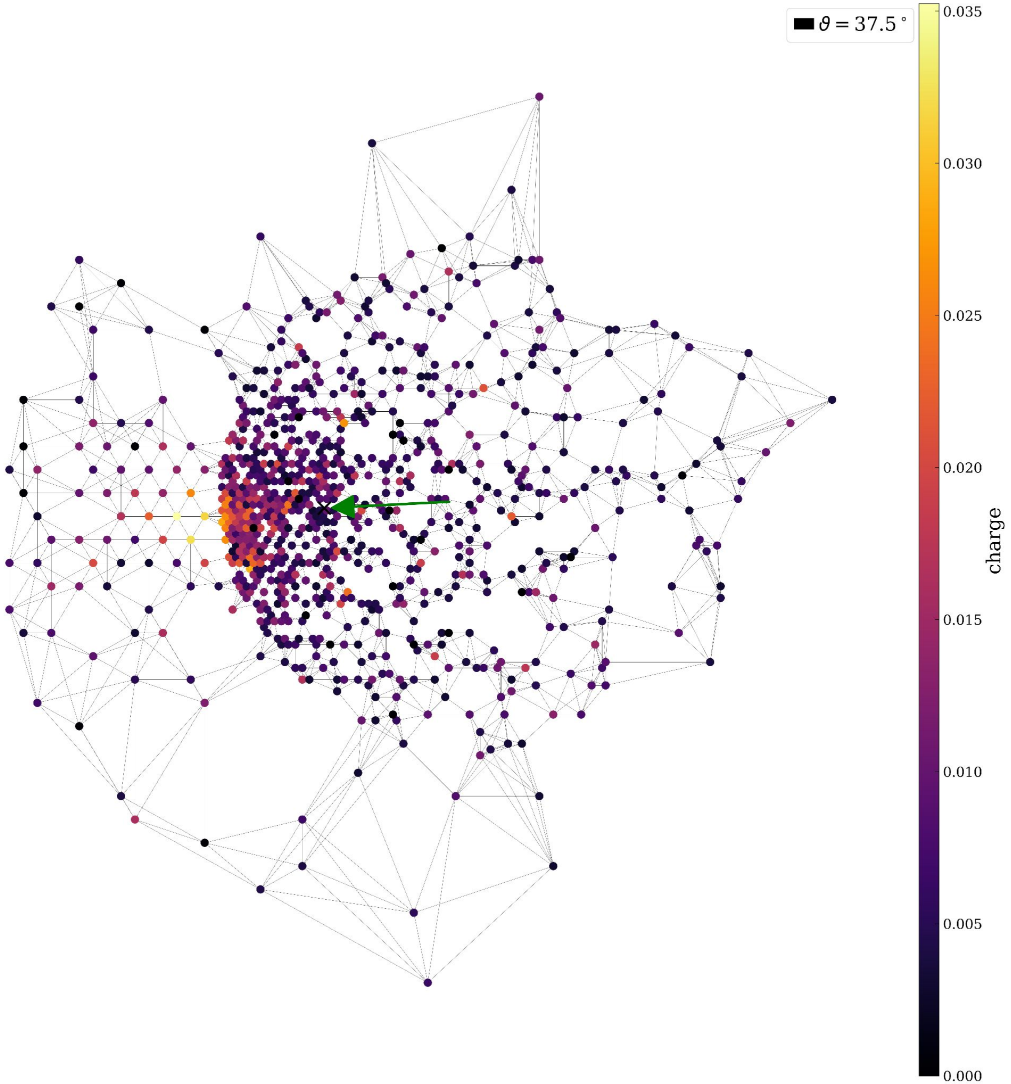
Example of Triggered Graph

Proton | 37.5° zenith angle | 7.5 TeV

Time



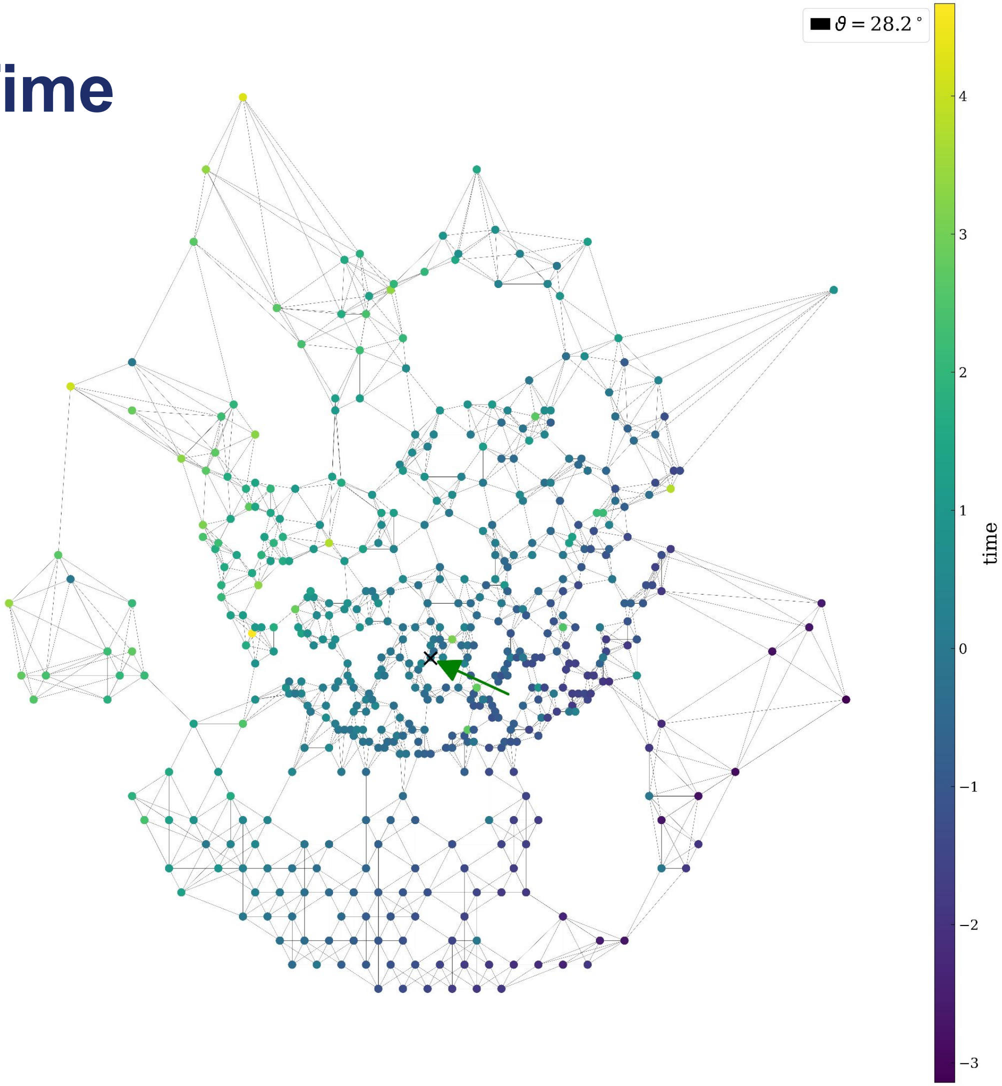
Charge



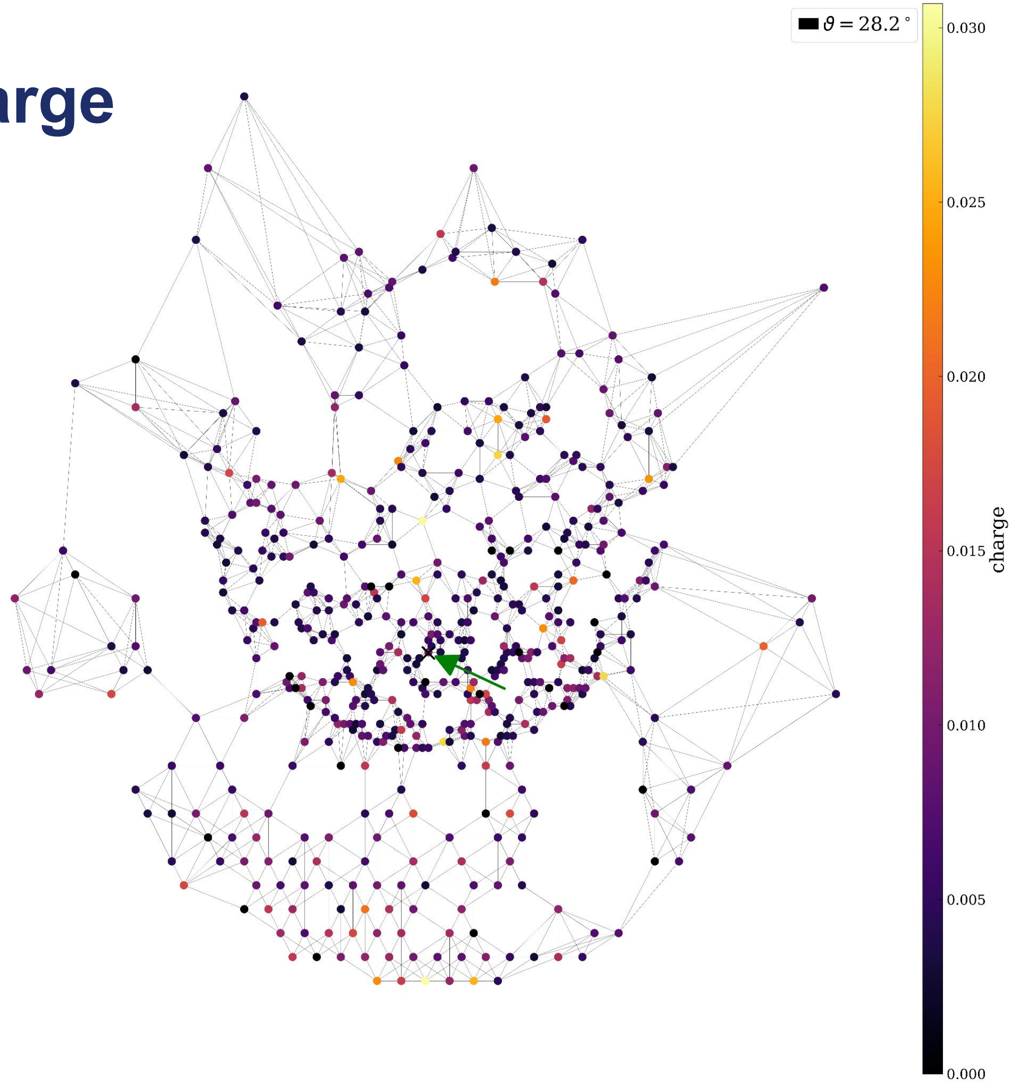
Example of Triggered Graph

Gamma | 28.2° zenith angle | 13.3 TeV

Time



Charge



A1 configuration tank design and layout

A1 configuration tank