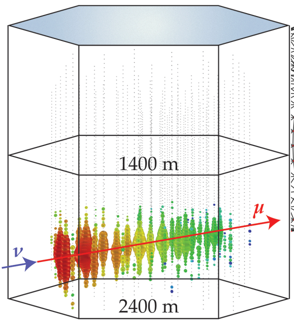
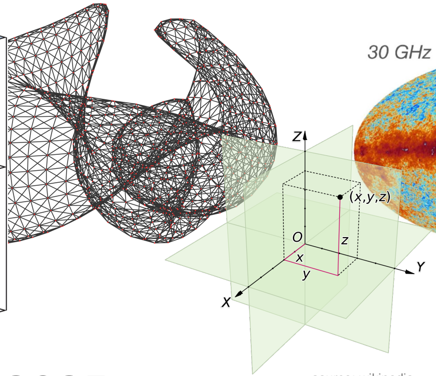
Friedrich-Alexander-Universität
Erlangen-Nürnberg

Deep Learning for Astroparticle Physics

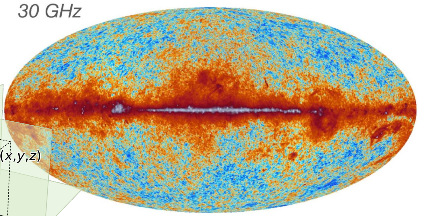
Jonas Glombitza
Erlangen Centre for Astroparticle Physics



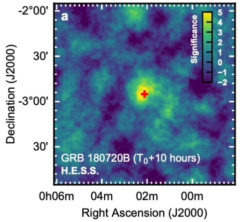
APS/Joan Tycko

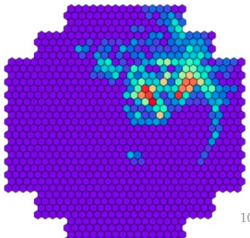


source: wikipedia



Astronomy and Astrophysics 641, p. 1 (2018)





10.1016/j.astropartphys.2018.10.003

April 14, 2025

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

SPONSORED BY THE
Federal Ministry of Education and Research

Machine Learning in Astroparticle Physics

ICRC 1991

OG 4.7.13

SEPARATING GAMMA-RAY SIGNALS BY ČERENKOV IMAGING : NEURAL NETWORK OPTIMIZATION

F. Halzen, R.A. Vazquez, E. Zas

Department of Physics, University of Wisconsin, Madison WI 53706

Abstract

We have performed a systematic study in space and time of air Čerenkov images of photon and proton showers generated by Bartol-Haleakala simulation programs. The rejection power of the azwidth parameter exploited in the TeV discovery of the Crab Nebula is confirmed. We have used a neural net to search for other features discriminating the Čerenkov images of photons and protons and demonstrate how the efficiency of the imaging method can be improved. We also identified differences in (nanosecond) time-image correlations. Although evident, they do not significantly improve proton rejection because of fluctuations. Our analysis and the associated programs are sufficiently general and flexible to be used for computer simulation of the threshold and photon recognition capability of any existing, projected or conceived Čerenkov telescope.

The Artificial Neural Networks as a tool for analysis of the individual Extensive Air Showers data.

1996

Tadeusz Wibig

Experimental Physics Dept., University of Łódź, ul. Pomorska 149/153, PL-90-236 Łódź, Poland

- Dates back to the 90s
- Recently became very popular

ELSEVIER

Astroparticle Physics

Volume 31, Issue 5, June 2009, Pages 383-391

γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analvsis method

S. Ohm, C. van Eldik, K. Egberts

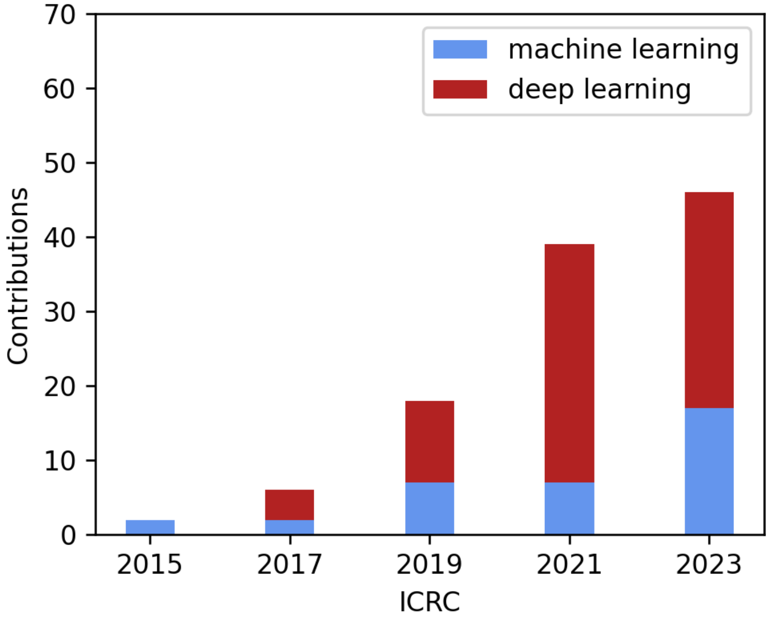
ELSEVIER

Astroparticle Physics

Volume 4, Issue 2, December 1995, Pages 119-132

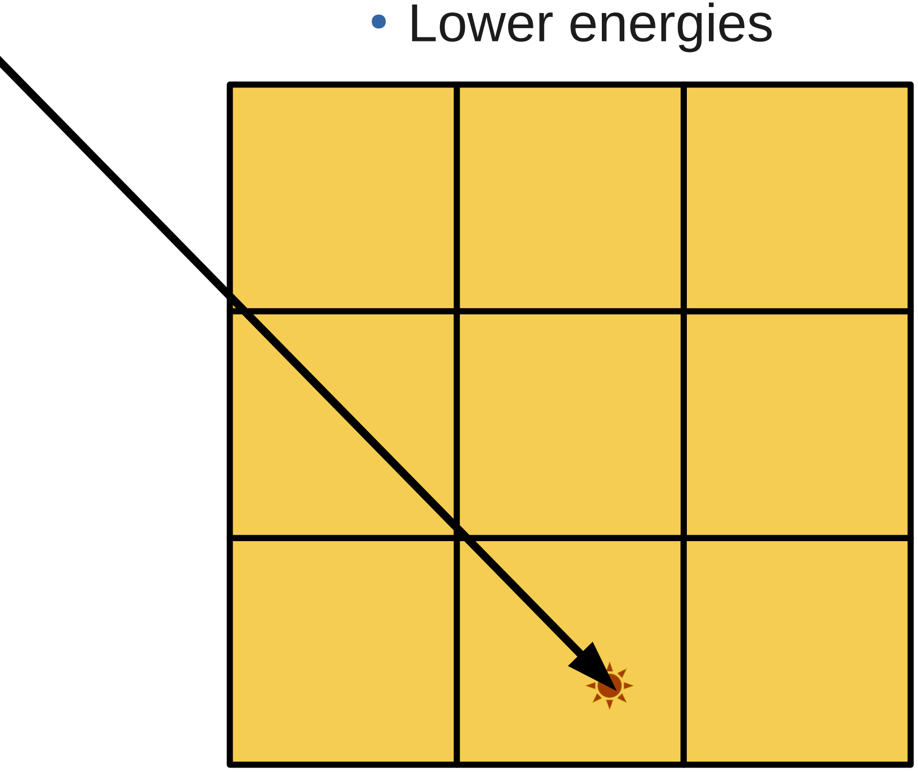
Separating γ- and hadron-induced cosmic ray air showers with feed-forward neural networks using the charged particle information ⋆

S. Westerhoff [a], B. Funk [a], A. Lindner [b], N. Magnussen [a], H. Meyer [a], H. Möller [a], W. Rhode [a], R.N. Sooth [a], B. Wiebel-Sooth [a]



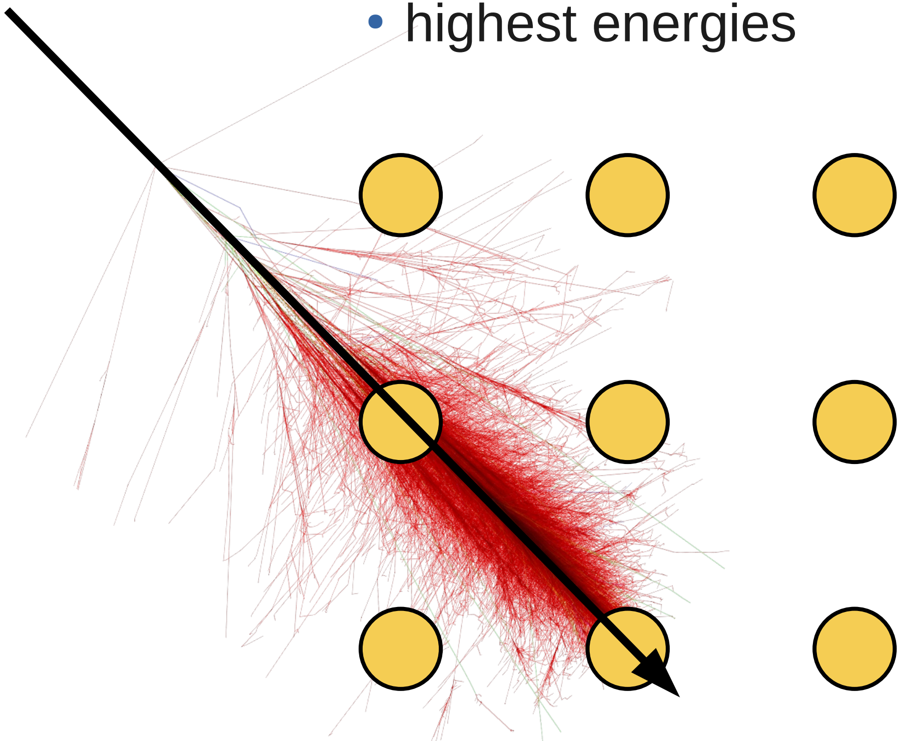
Astronomy at the highest energies

- Lower energies



Direct detection

- highest energies



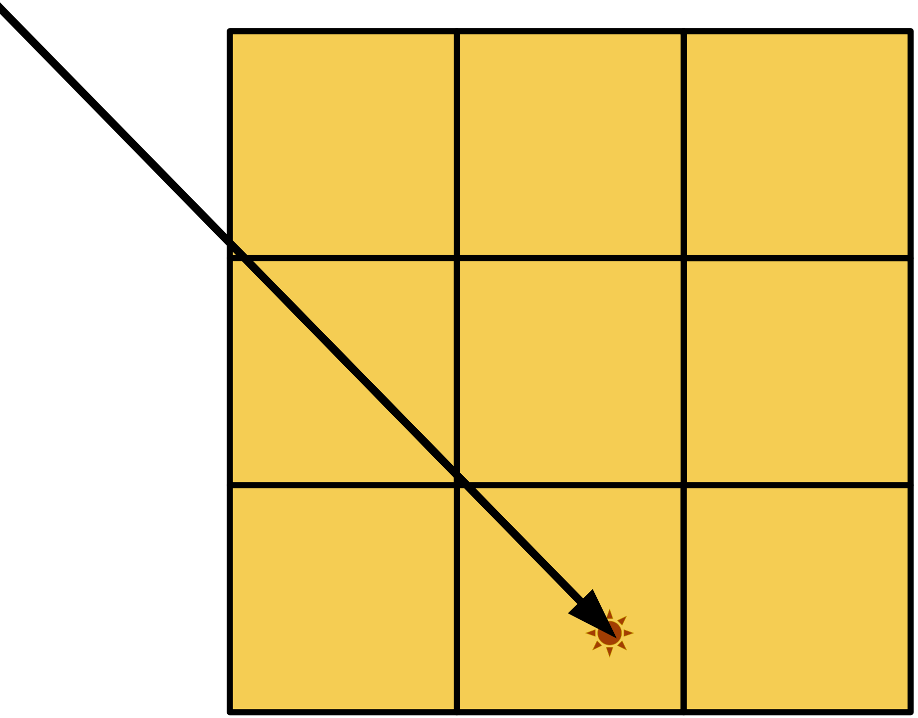
- Low flux & indirect detection
  - Sparsely instrumented detectors
- Complex reconstruction (direction, energy, particle type)

Astronomy at the highest energies

High dimensional data!

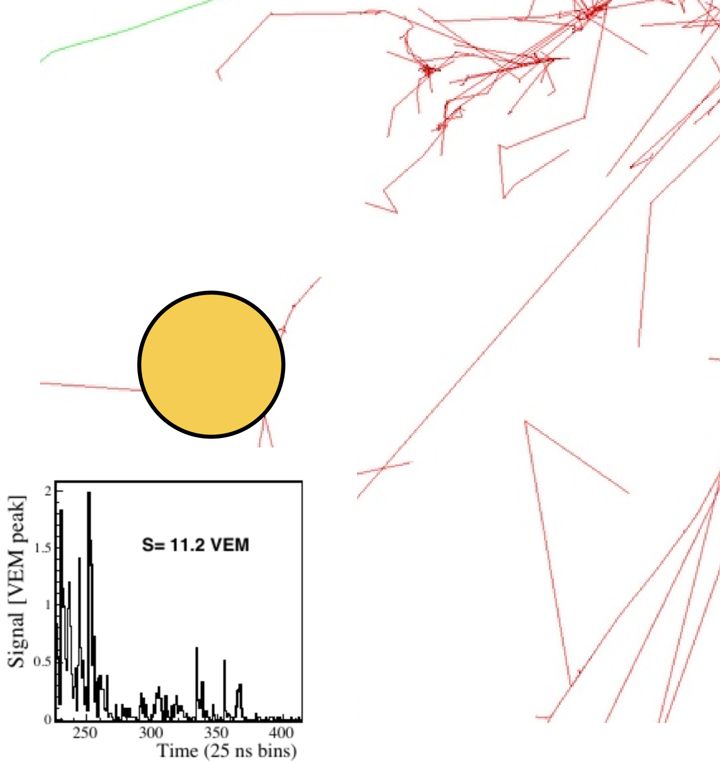
Challenging reconstruction!

- Lower energies



Direct
detection

- highest energies

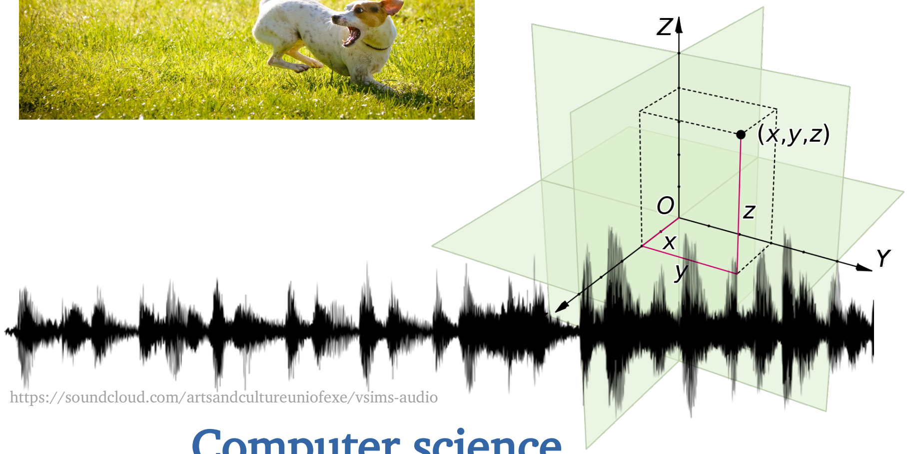


- Single sensors detect
  time resolved signals (per event)

Application in Physics

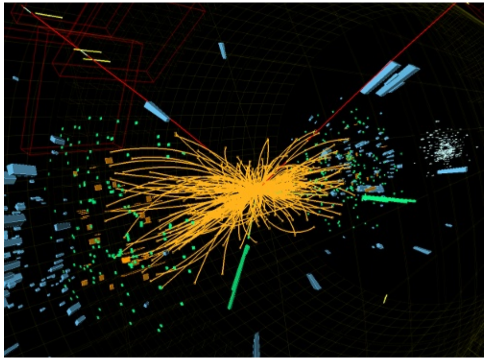
Physics feature different data
Challenge: adapt algorithms from computer science to physics research

source: wikipedia

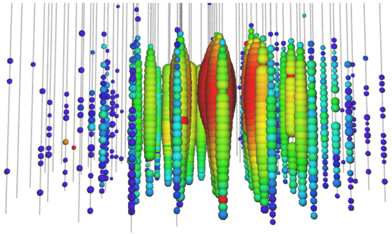


https://soundcloud.com/artsandcultureuniofexe/vsims-audio

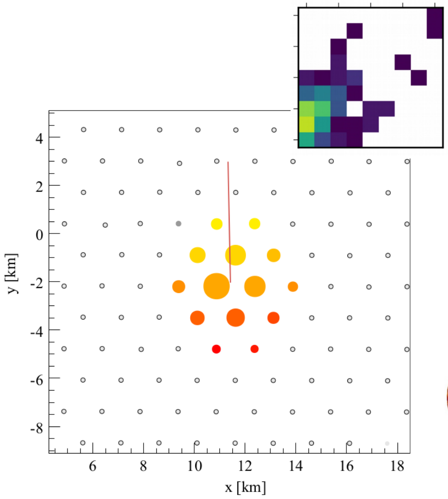
Computer science



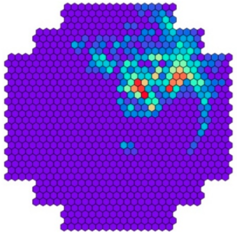
https://cds.cern.ch/record/2711418



https://arxiv.org/abs/1309.7003

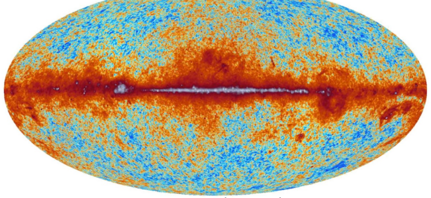


10.1016/j.nima.2015.06.058



10.1016/j.astropartphys.2018.10.003

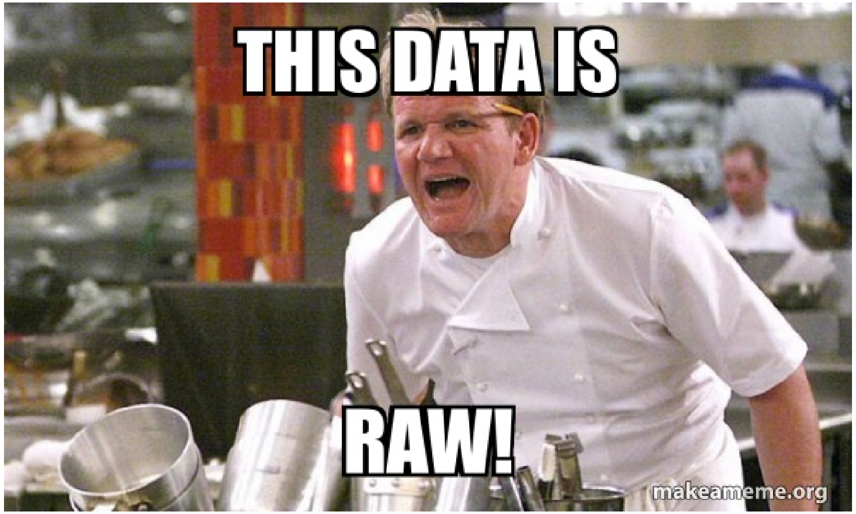
30 GHz



Astronomy and Astrophysics 641, p. 1 (2018)

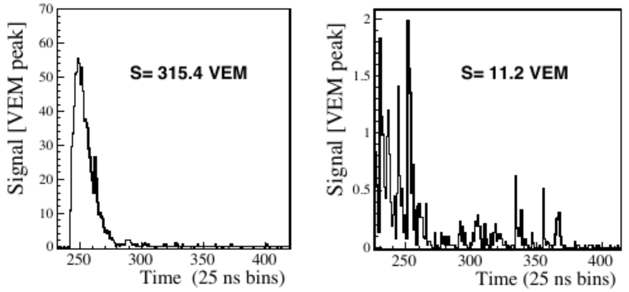
Physics

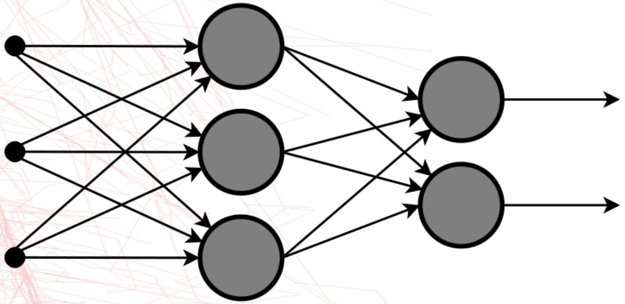
Processing raw signals



Machine Learning to Deep Learning

- Air shower signals measured by surface detectors
  - disentangle muonic and em part at station level





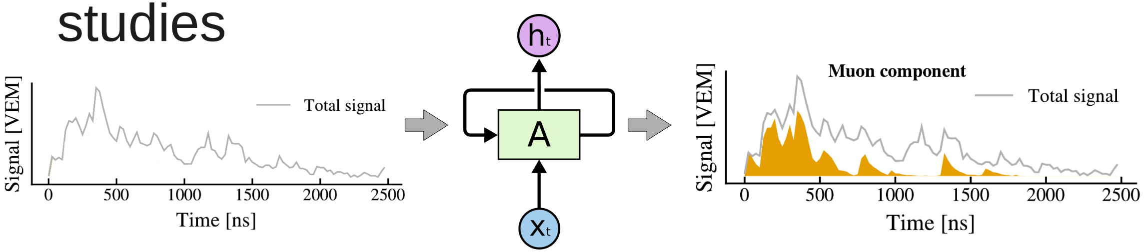
Traditional ML approach

- Extract <u>fraction</u> of muons measured by single station
- Feed physicist observables into a neural network

A. Gulillen et al., 10.1016/j.astropartphys.2019.03.001

Deep learning version

- Use RNN to extract <u>time-dependent signals</u> induced by muons
- Promising results for mass composition studies



Pierre Auger Collaboration, JINST 16 P07016 (2021)

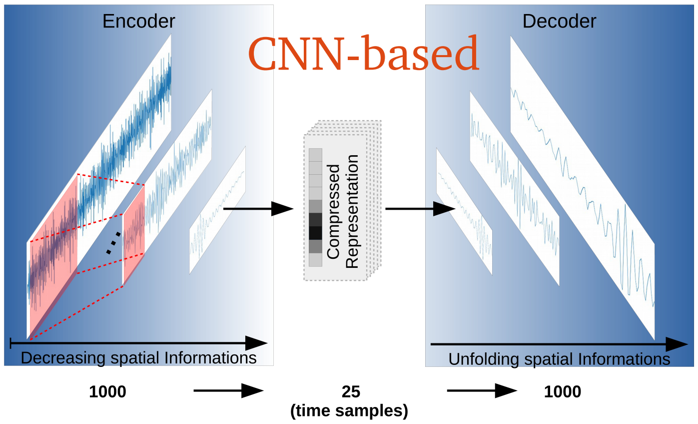
Denoising of Signal Traces (1D)

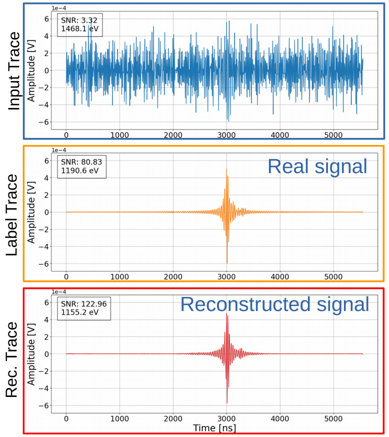
Supervised training of denoising autoencoders

- feature compressed space in between encoder and decoder
- → encodes only relevant information in compressed space

Future application: bringing ML close to the sensor

Denoising of cosmic ray radio signals

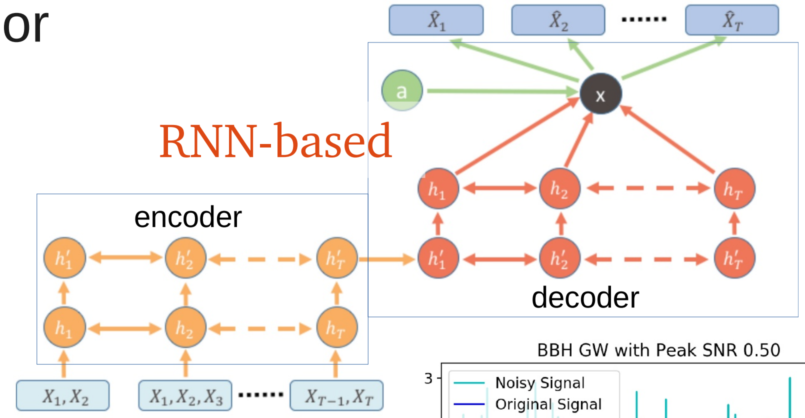


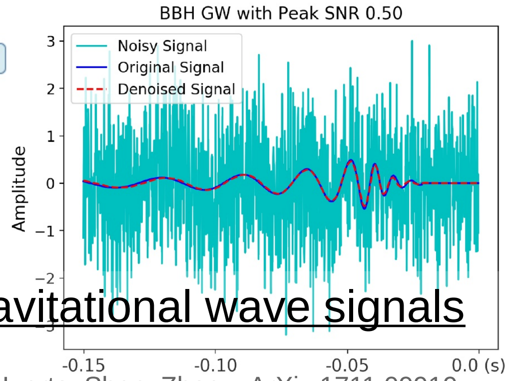


M. Erdmann et al. - 10.1088/1748-0221/14/04/P04005

A. Rehman et al., PoS ICRC2021 417

P. Bezyazeekov et al., ArXiv/2101.02943 & D. Shipilov et al., EPJ (2019) 02003





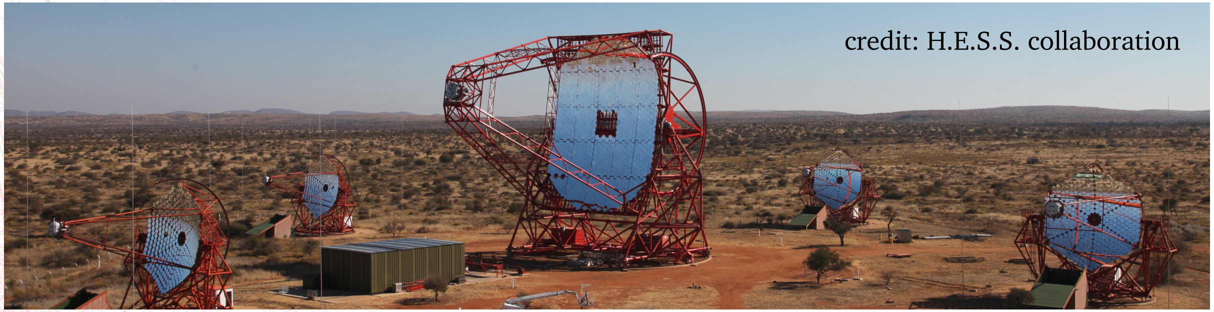
Denoising of gravitational wave signals

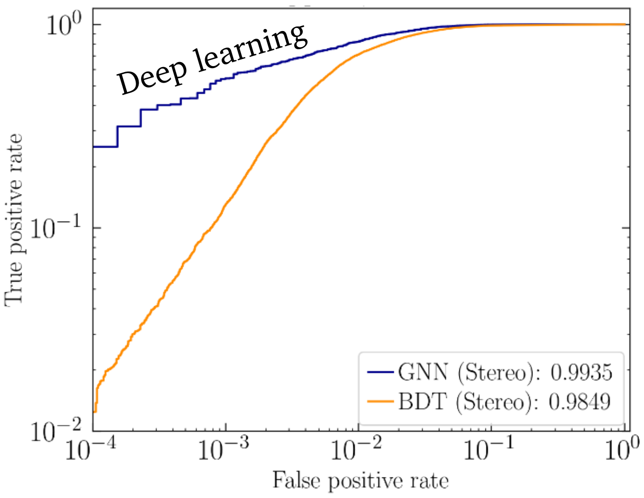
George, Huerta, Shen, Zhao – ArXiv 1711.09919

See talk on Wednesday

Event reconstruction

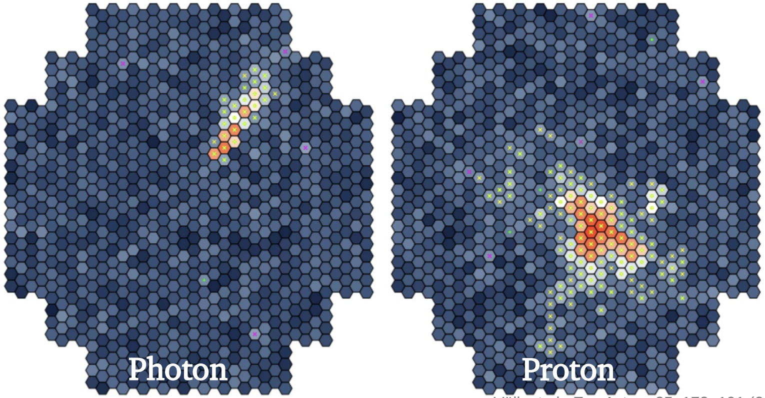
Deep Learning for IACTs





- Gamma ray telescopes in Namibia
- For each photon ~ $10^3 \rightarrow 10^4$ protons
  - Powerful rejection needed
- First promising results on simulations
  - Neural networks outperforms BDTs

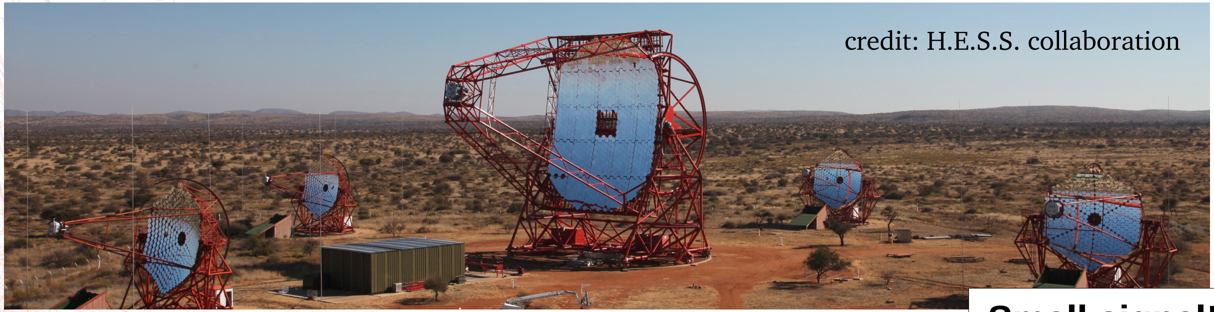
- Currently investigating stereoscopic models
  → exploit telescope-telescope correlations
- Challenge: application to data

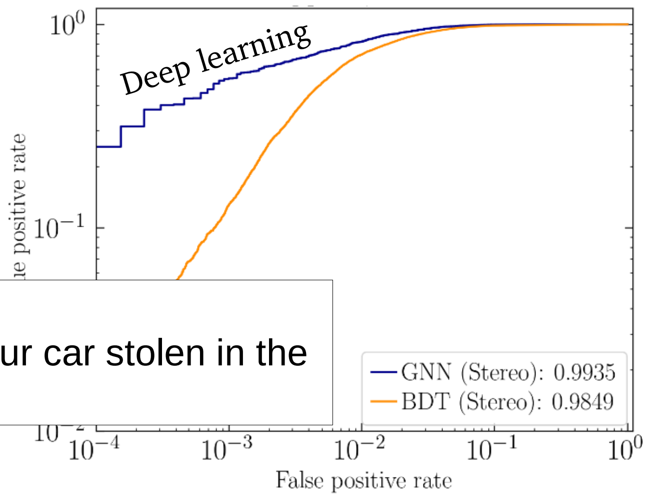


Völk et al., Exp Astron 25, 173–191 (2009)

Shilon et al., 10.1016/j.astropartphys.2018.10.003
Glombitza et al., JCAP11(2023)008, PoS(ICRC2023)715
Jacquemont et. al. arXiv:2105.14927

Deep Learning for IACTs

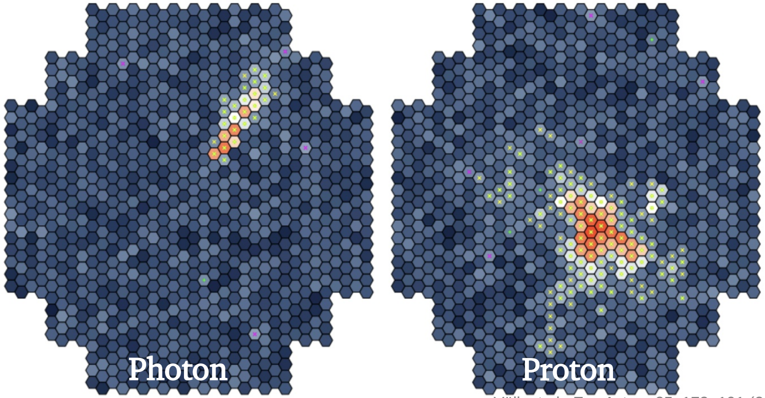




Small signal!
Odds of getting your car stolen in the next year!

- Gamma ray telescopes in Namibia
- For each photon ~ $10^3 \rightarrow 10^4$ protons
  - Powerful rejection needed
- First promising results on simulations
  - Neural networks outperforms BDTs

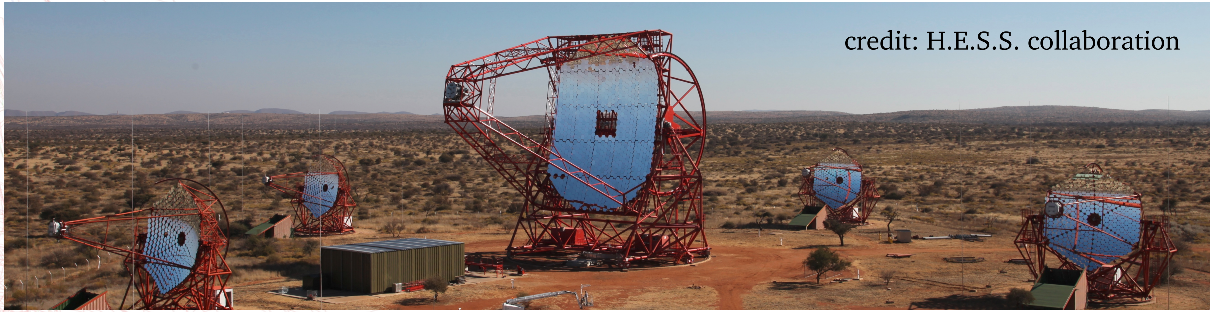
- Currently investigating stereoscopic models
  → exploit telescope-telescope correlations
- Challenge: application to data



Völk et al., Exp Astron 25, 173–191 (2009)

Shilon et al., 10.1016/j.astropartphys.2018.10.003
Glombitza et al., JCAP11(2023)008, PoS(ICRC2023)715
Jacquemont et. al. arXiv:2105.14927

Deep Learning for IACTs

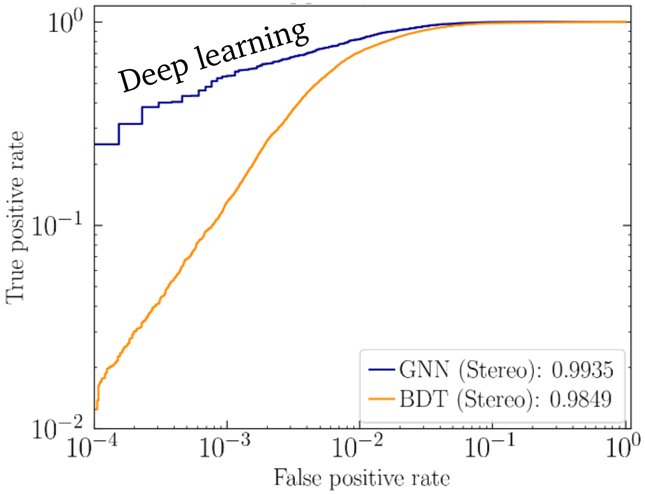


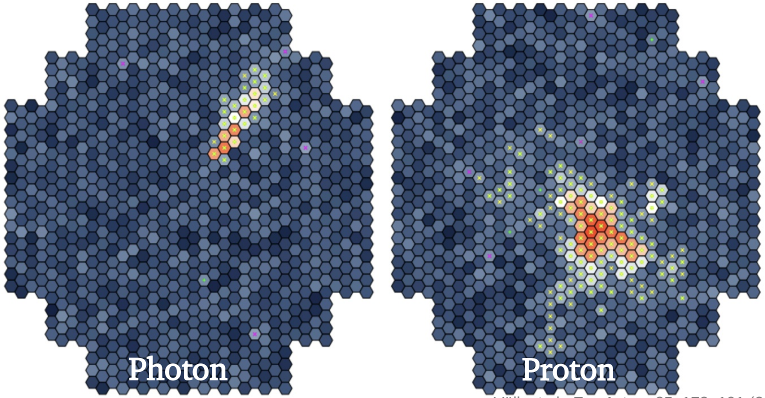
- Gamma ray telescopes in Namibia
- For each photon ~ $10^3 \rightarrow 10^4$ protons
  - Powerful rejection needed
- First promising results on simulations
  - Neural networks outperforms BDTs

- Currently investigating stereoscopic models
  → exploit telescope-telescope correlations
- Challenge: application to data

See talk today

Shilon et al., 10.1016/j.astropartphys.2018.10.003
Glombitza et al., JCAP11(2023)008, PoS(ICRC2023)715
Jacquemont et. al. arXiv:2105.14927





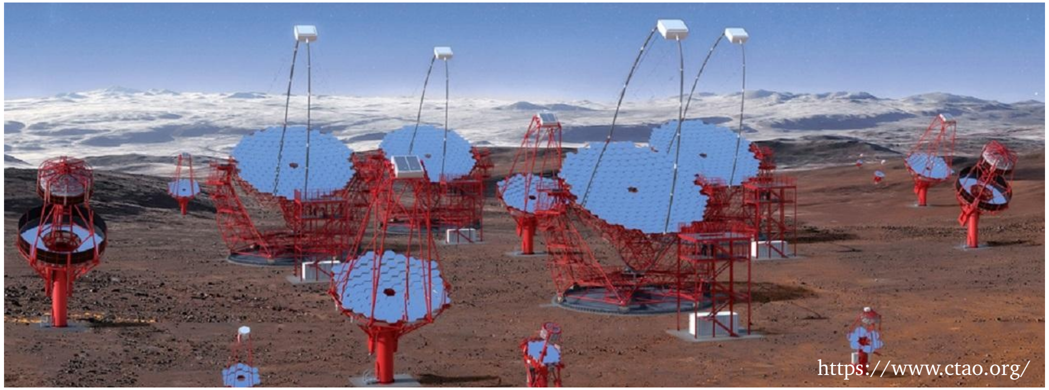
Völk et al., Exp Astron 25, 173–191 (2009)

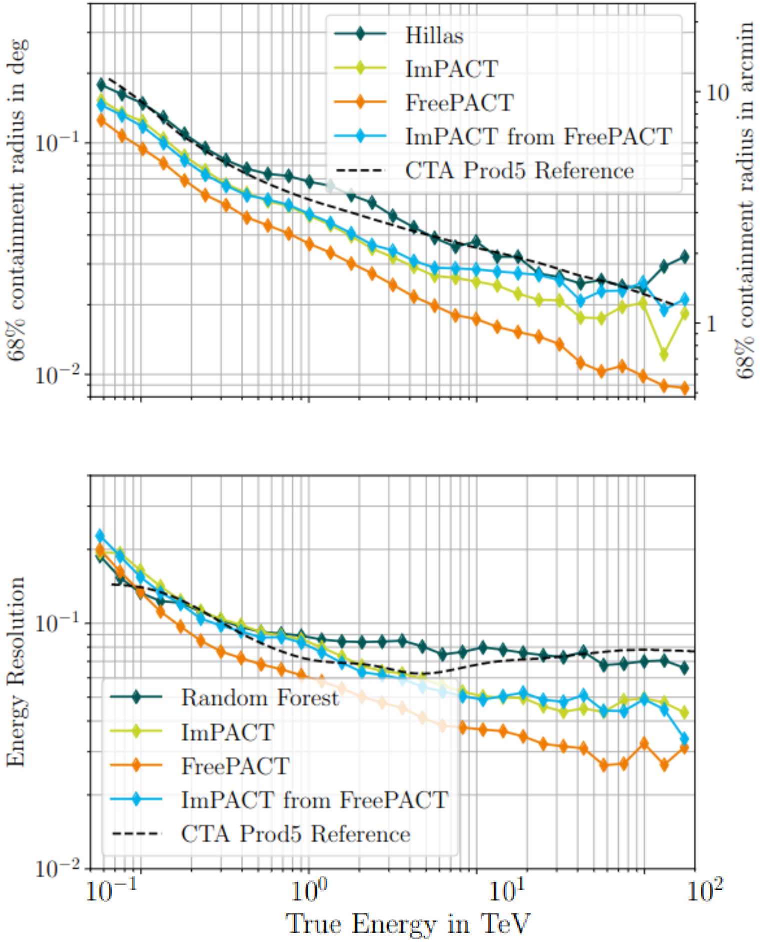
Event reconstruction for CTA

State-of-the-art: template-based reconstruction

Hybrid approach:

- Utilize DNN to approximate charge probability density function for each pixel
- Method outperforms traditional and state-of-the-art approaches on simulations
- Previous works limited to single telescopes
- e.g., T. Miener et al., arXiv:2109.05809, M. Jacquemont et al., arXiv:2105.14927



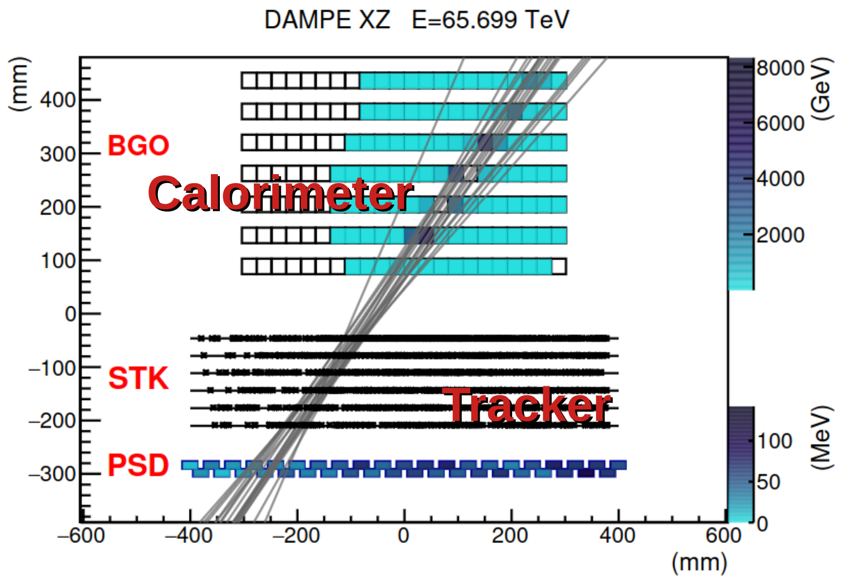


See talk today

G. Schwefer, R. Parsons, J. Hinton, arXiv:2406.17502

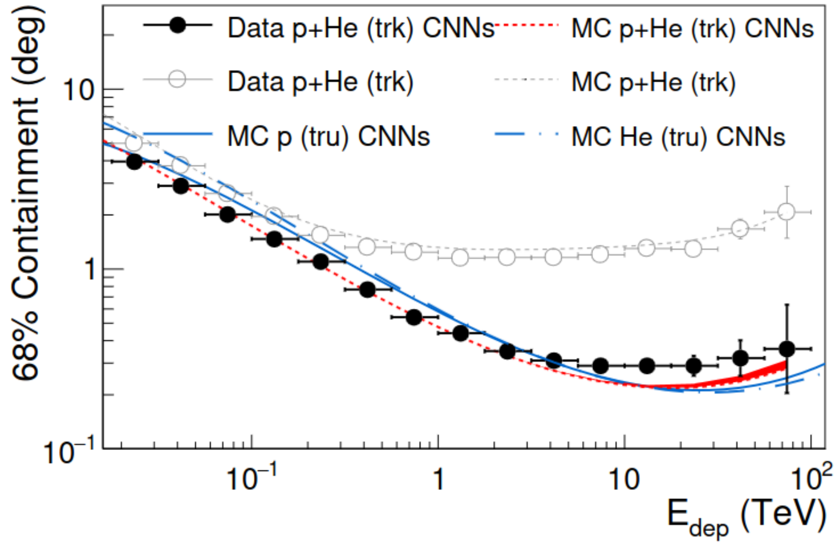
Tracking using DNNs at DAMPE

- DAMPE: cosmic-ray space mission
- Challenge: At high E calorimeter particles back-scatter into tracking
- Use calorimeter data and CNN to perform tracking (+ seed for tracker)



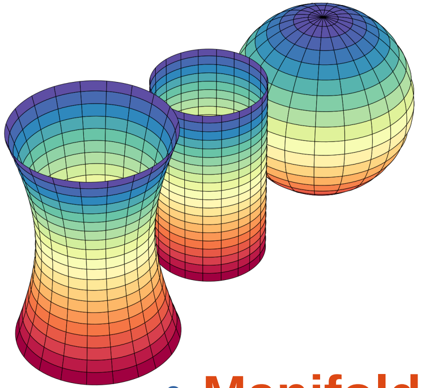
A. Tykhonov et al, Astropart. Phys. 146, 102795 (2023)

- Validation using events with clear tracker
- Significant improvement over classical method
- Increase tracking efficiency using tracker



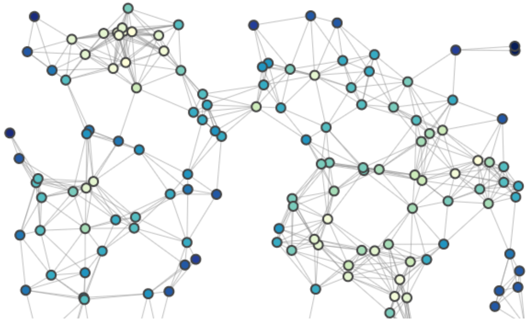
Non-Euclidean Domains

- Defining convolutions, challenging on non-euclidean domains
  - Deformation of filters, changing neighbor relations
  - Non-isometric connections on graphs



- **Manifolds**

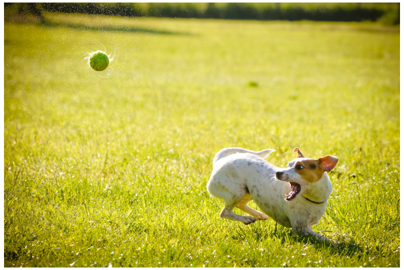
source: wikipedia



- **Graphs**

source: Cody Marie Wild,
Towards Data Science

How can we generalize convolutions?



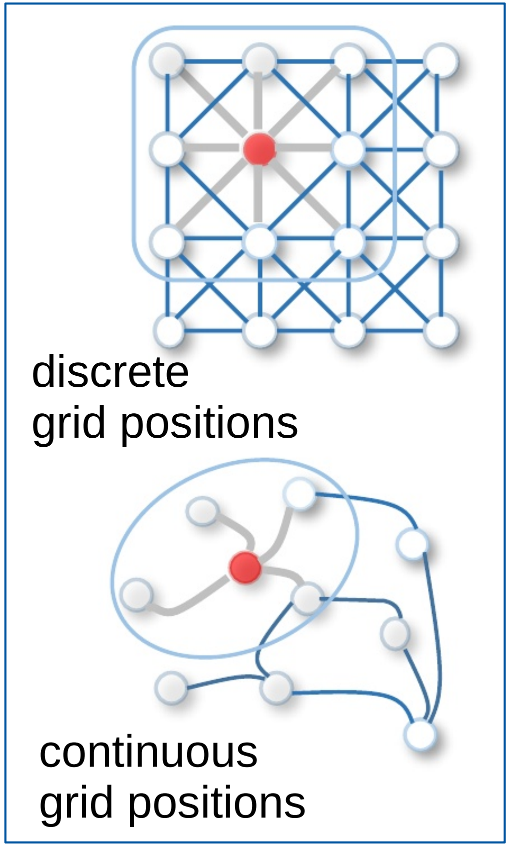
© pxhere.com

Image-like data

- collection of pixels (vector)
- coherent (rarely sparse)
- discrete, regular (symmetric)
- feature euclidean space

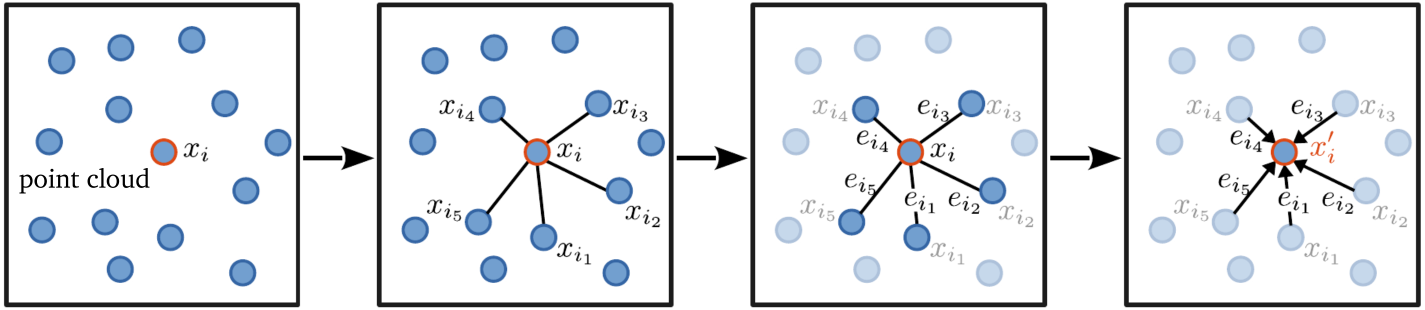
Graph Networks: Edge Convolutions

Y.Wang et al,
https://arxiv.org/abs/1801.07829

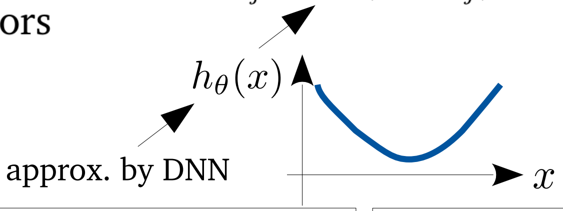


- Define graph/neighborhood → e.g., using kNN
- Apply continuous filter based on distances (filter → DNN)
  - flexible for many settings: irregular structures, point clouds

Erdmann et al., https://doi.org/10.1142/12294



construction of directed graph	estimation of edge features	aggregation over neighborhood
→ search k nearest neighbors	$e_{i_j} = h_\theta(x_i, x_{i_j})$	$x'_i = \square_{j=1}^{k} e_{i_j}$
		e.g. $x'_i = \sum_{j=1}^{k} e_{i_j}$



See talk on Wednesday | See talk on Wednesday | See talk on Thursday | See talk on Thursday

See talk on Thursday

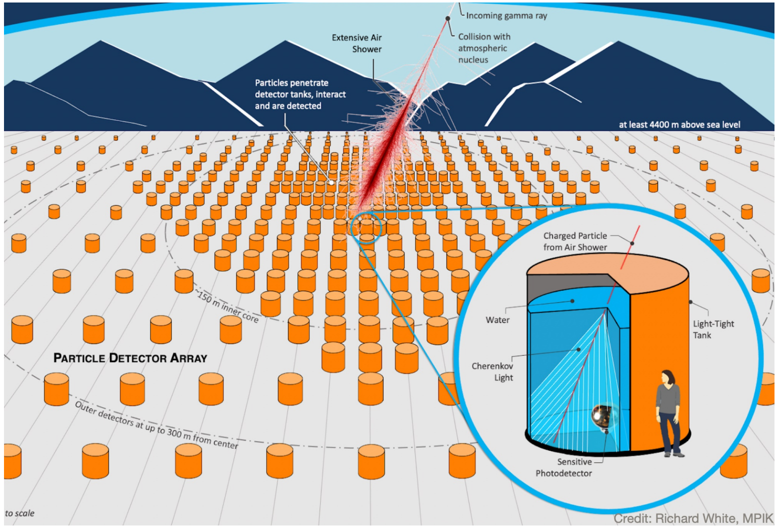
Deep Learning at SWGO

The Southern Wide-field Gamma-ray Observatory

- Surface-detector-based gamma-ray observatory
  - Sensitivity: 100s GeV → PeV scale
- Feature different zones with different fill factors

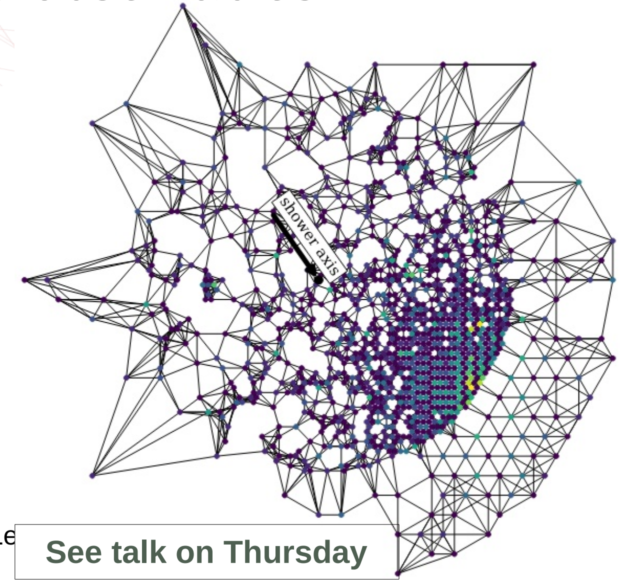
Promising results: GCNs that well handle sparsity

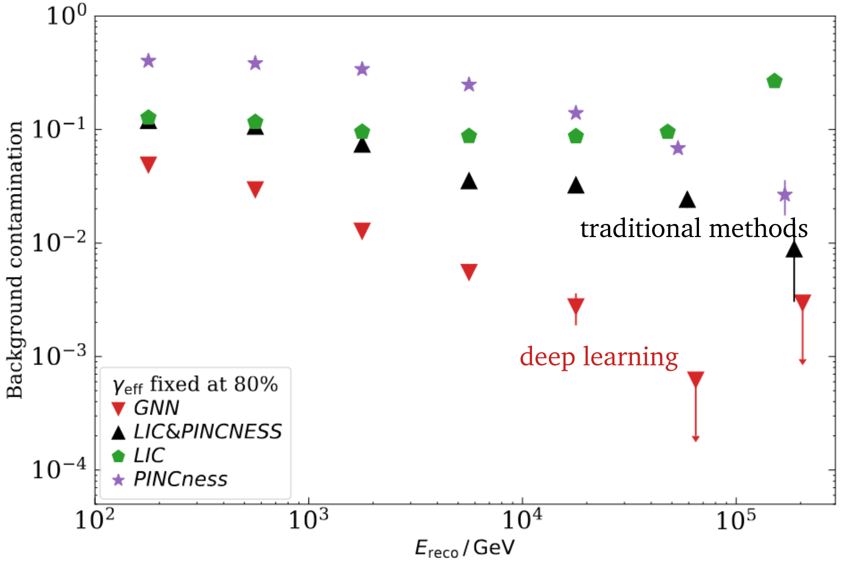
- Superior than ensemble of all previous hand-designed observables



The Southern Wide-field Gamma-ray Observatory

Example signal graph
Proton event
E = 10^4 GeV
Zenith = 35°





J.G., M. Schneider, F. Leitl, S. Funk, C. van Eldik, JCAP02(2025)066

See talk on Thursday

Transformers

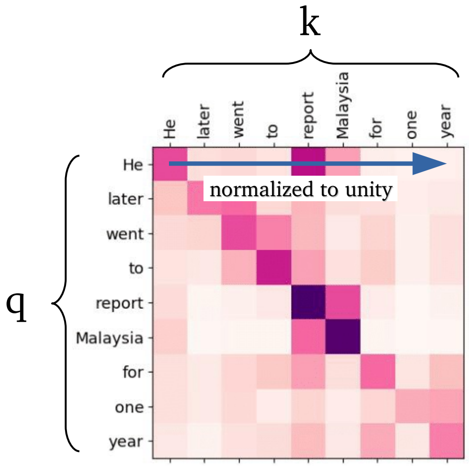
- Transformers are **backbone of latest breakthroughs**: LLMs / Stable Diffusion
- ➔ Building blocks: DNNs with attention mechanism →noise robust
  - Which parts of sequence semantically correlated → analyze together

Analyze *sequences* (arbitrary lengths):

- $(x_1, x_2, x_3, x_4, x_5, \ldots, x_n)$
  - single element called *token* (e.g, word)

Attention: (in a nutshell) extension of fully-connected DNNs

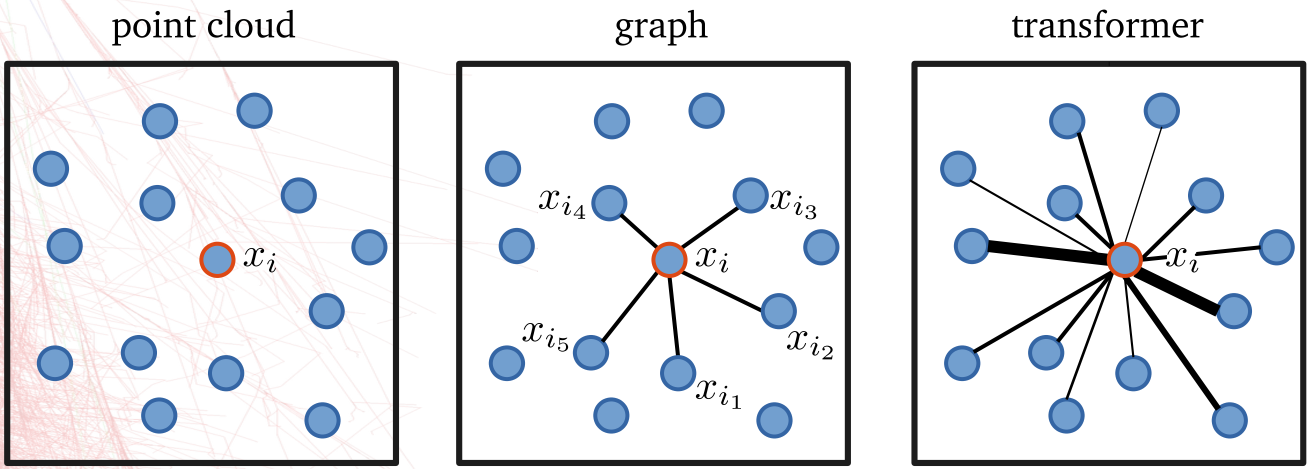
- listen to all inputs, **focus** on most important inputs
- focus (**attention**) given by correlation in feature space
- Independent of sequence length



Visualized attention matrix

Du et al,m 10.3390/fi14030085

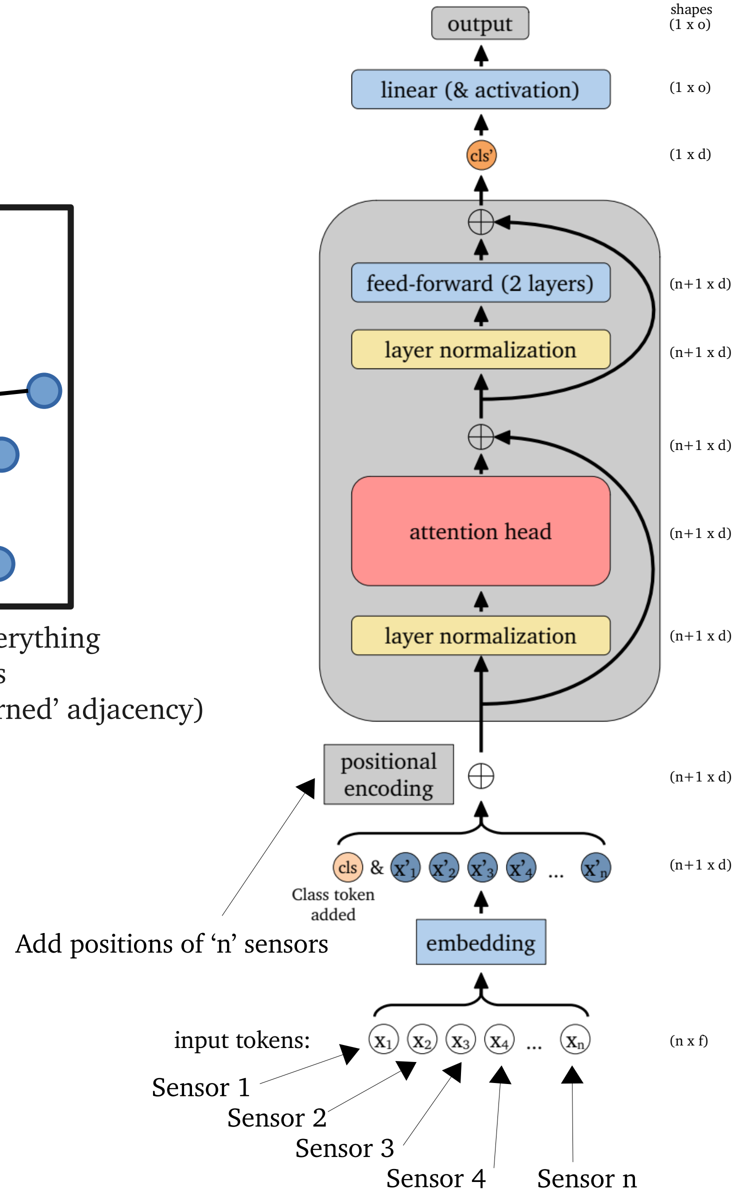
Point cloud transformer



Prior: local correlations
- kNN clustering defines graph
→ defines **adjacency matrix**

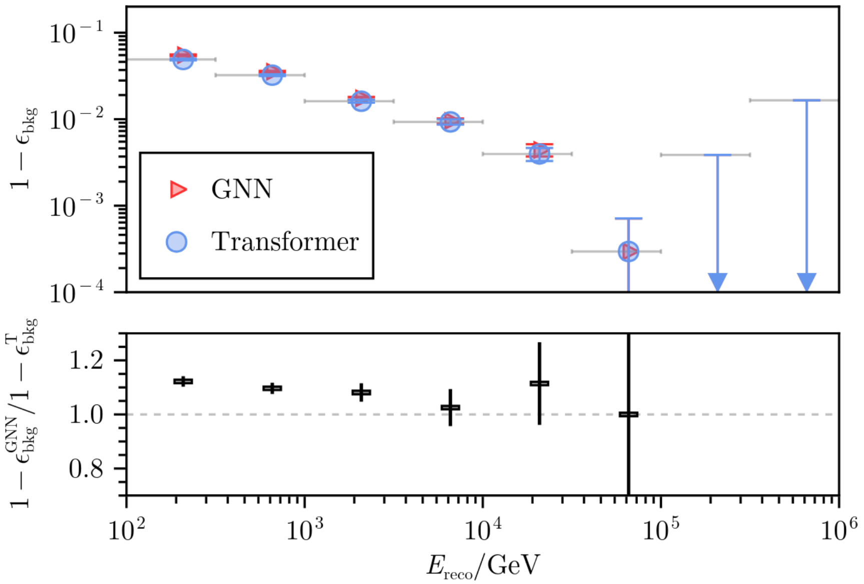
Transformers connect everything
- learns attention weights
→ **attention matrix** ('learned' adjacency)

- Transformers as extensions of graph networks

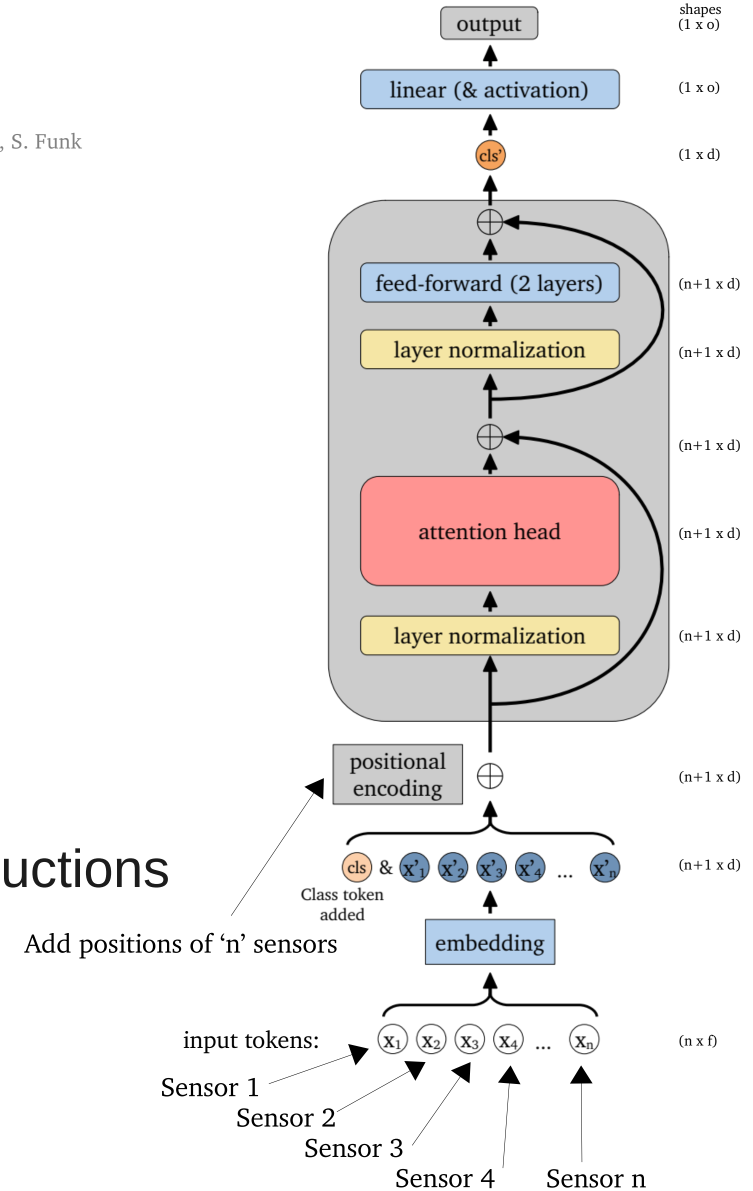


Point cloud transformer

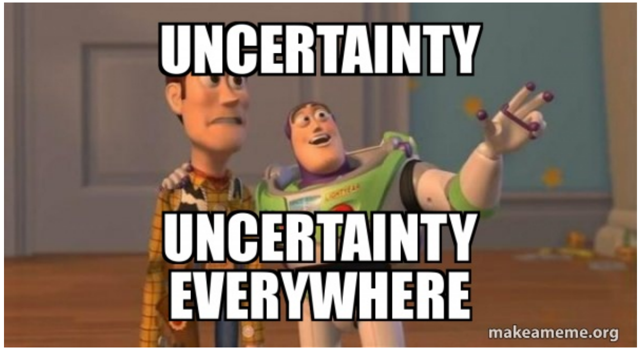
M. Pirke, J.G., F. Leitl, M. Schneider, C.van Eldik, S. Funk



- Transformers as extensions of graph networks
- Additional freedom can lead to improved reconstructions



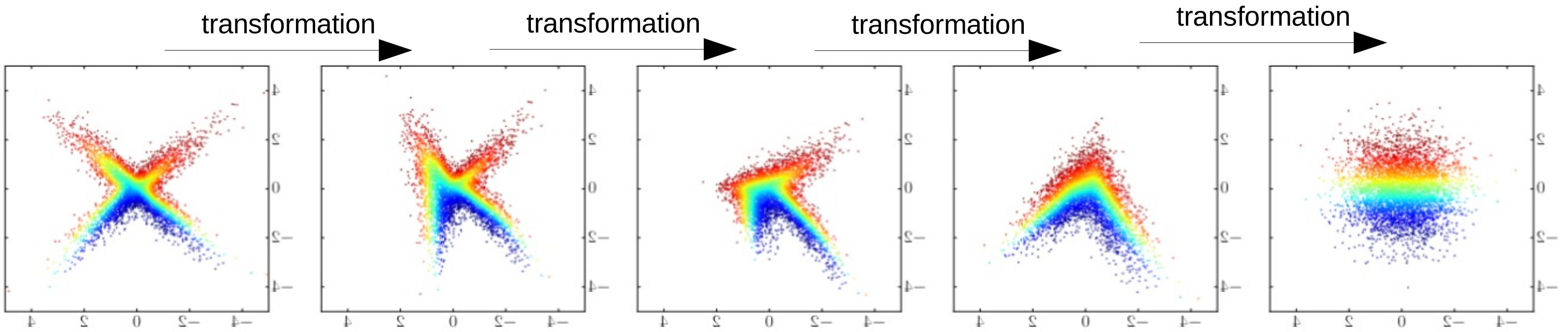
Uncertainty estimation



See talk on Wednesday

Normalizing Flows

Normalizing flows: stack several simple <u>invertible</u> mappings



G. Papamakarios et al., JMLR 22(57):1-64, 2021

<u>training</u>:

complicated distribution (e.g., natural images)

"Fit data distribution to match Gaussian"

→ Direct maximization of Likelihood!

simple distribution (e.g., Gaussian)

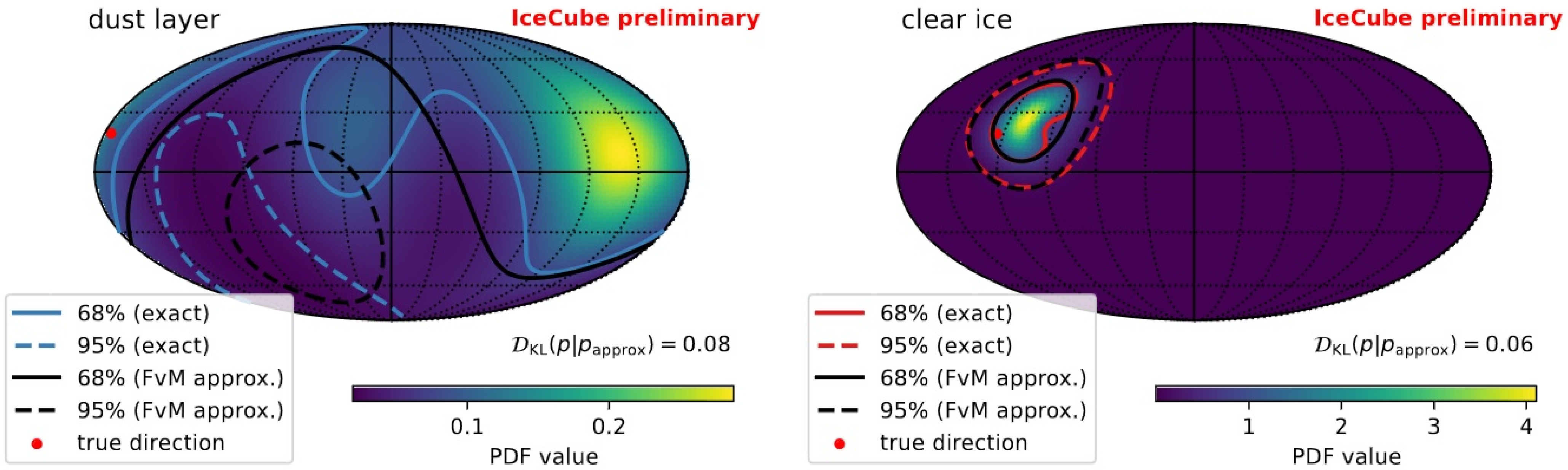
<u>evaluation/ inference</u>:

Since model invertible and distribution normalized

Revert direction → get samples proxy of complicated distribution

enables:
- fast generation of new samples (**direct density estimation**)
- reconstruction of objects, including uncertainty estimate

Normalizing flows at IceCube



- Dust layer can affect reconstruction uncertainty → usually assumed symmetric
- Application of normalization flows: uncertainty of neutrino arrival direction
  - Reconstruction conditions flow that maps to spherical surface → asymmetric uncertainties

T. Glüsenkamp, PoS(ICRC2023)1003
T. Glüsenkamp, Eur. Phys. J. C 84, 163 (2024)

Detector simulations

Generative models

CIFAR10

Learn to generate new samples

generated with stable diffusion

“Albert Einstein using a mobile phone while watching TV”

“A penguin flies in the sky and overtakes other birds. Clouds are seen in the background”

Breakthrough in generative machine learning

- generation of realistic images
- image feature local and global coherence
- realistic image super resolution

Which face is real?

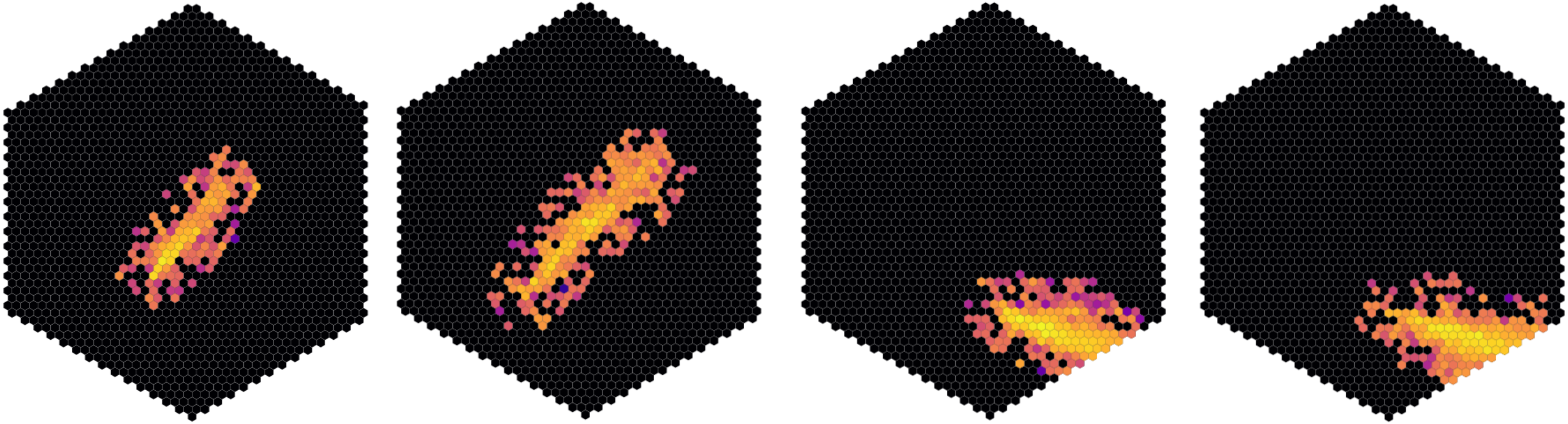


Play the game:
https://www.whichfaceisreal.com

Which generated IACT image is real?

See talk on Thursday

See talk on Wednesday

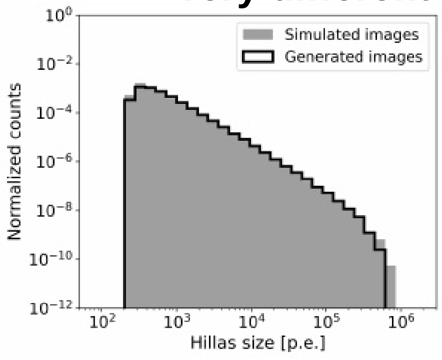


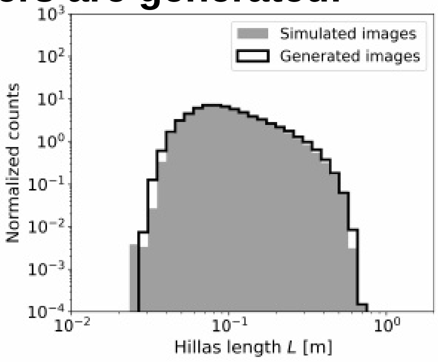
Imaging Air Cherenkov Telescope
Example simulated / generated for the CT5 telescope of the H.E.S.S. array

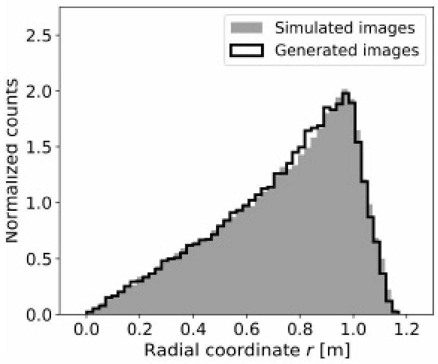
C.Elfein, S.Funk, J.G. JINST 19 P04010 (2024)

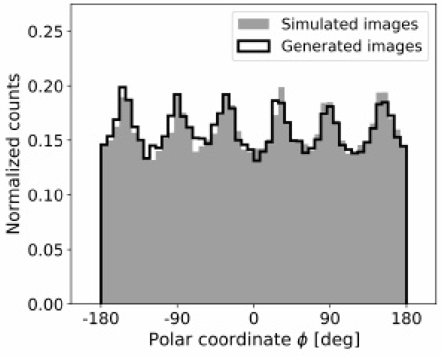
Hillas Parameter

Distributions agree very well → over large range of magnitude!
Very different showers are generated!









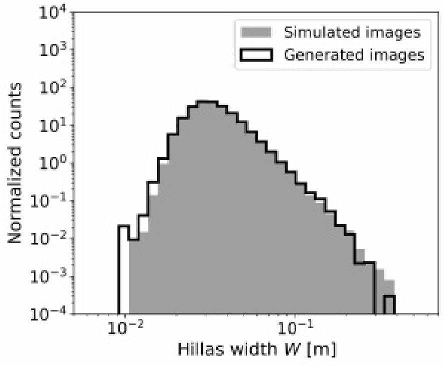
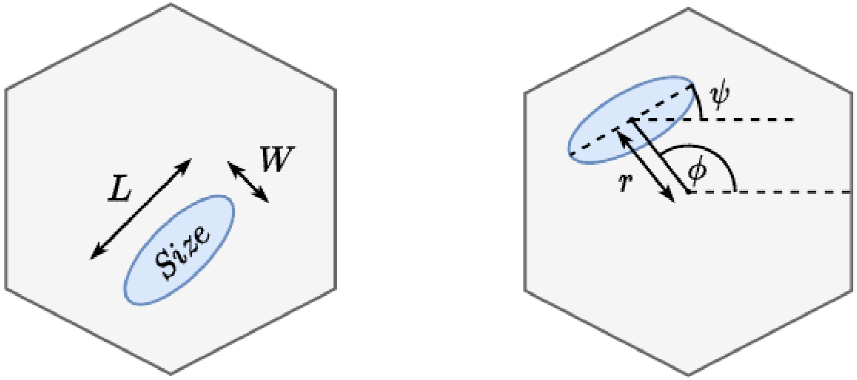
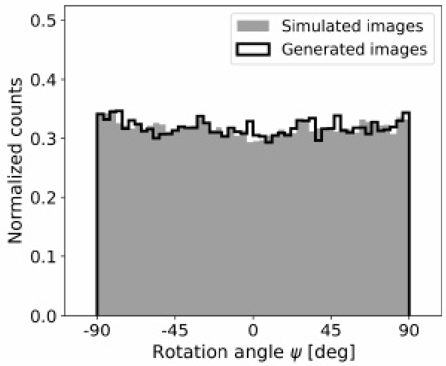


Image shape modeled well!



Investigate Hillas parametrization
→ traditionally used for event reconstruction



Full camera used
→ Very different geometries

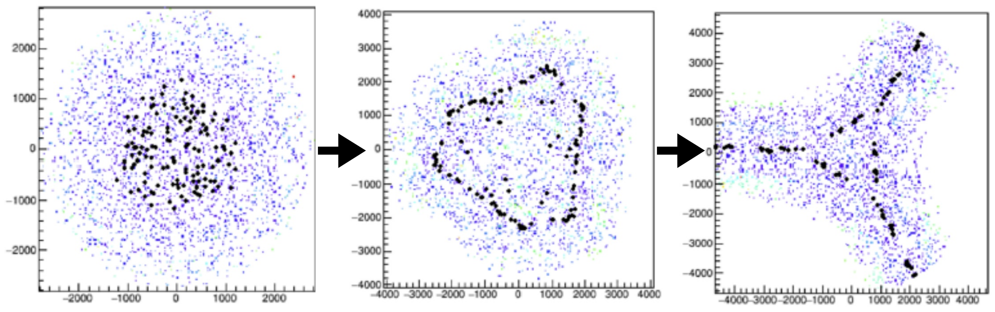
C.Elfein, S.Funk, J.G. JINST 19 P04010 (2024)

Given science requirements → maximize utility function → optimize experiment

T. Dorigo et al, arXiv:2310.01857

Toy example: Gamma ray observatory

- Closed-from parametrization of air shower simulation
- Learn the station placing of a water-Cherenkov gamma-ray observatory

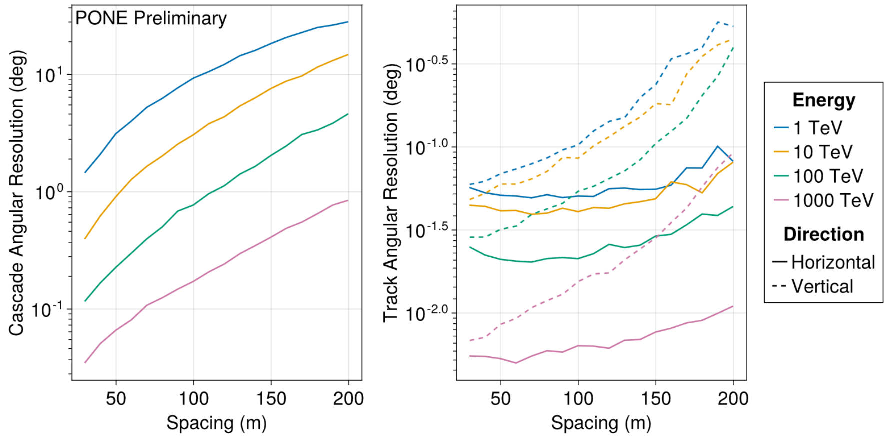


Convergence of station layout

C. Haack, L. Schumacher PoS(ICRC2023)1059

P-ONE: planned neutrino telescope

- Approximate response of single detector for various using surrogate model (NF)
- Estimate stat. limit via Fisher Information



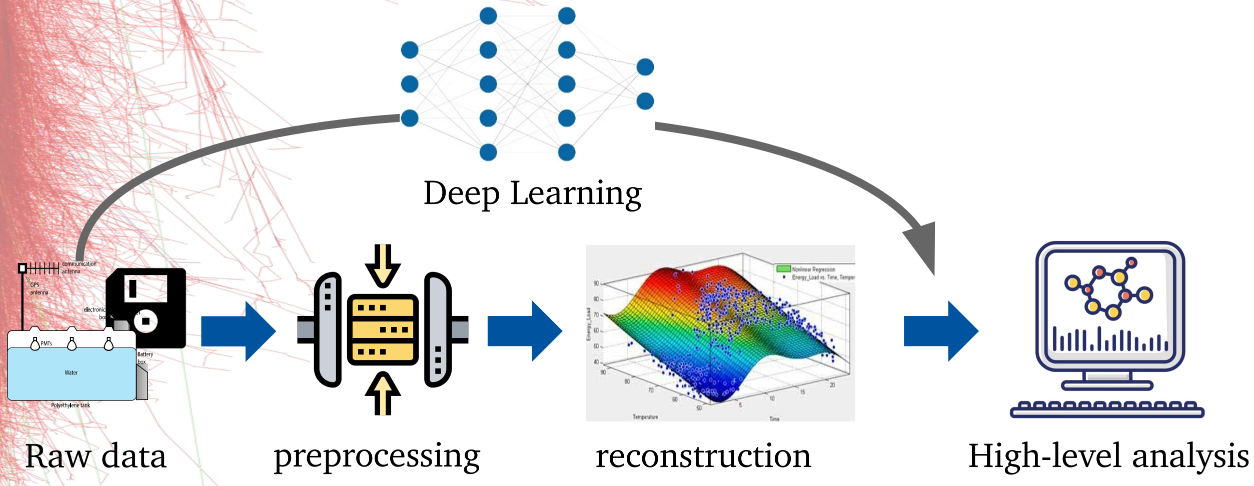
MODE Collaboration

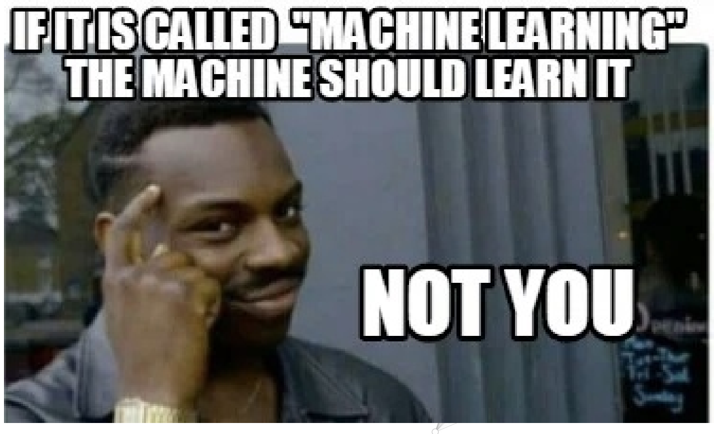
Open collaboration engaging the ML-based design of experiments

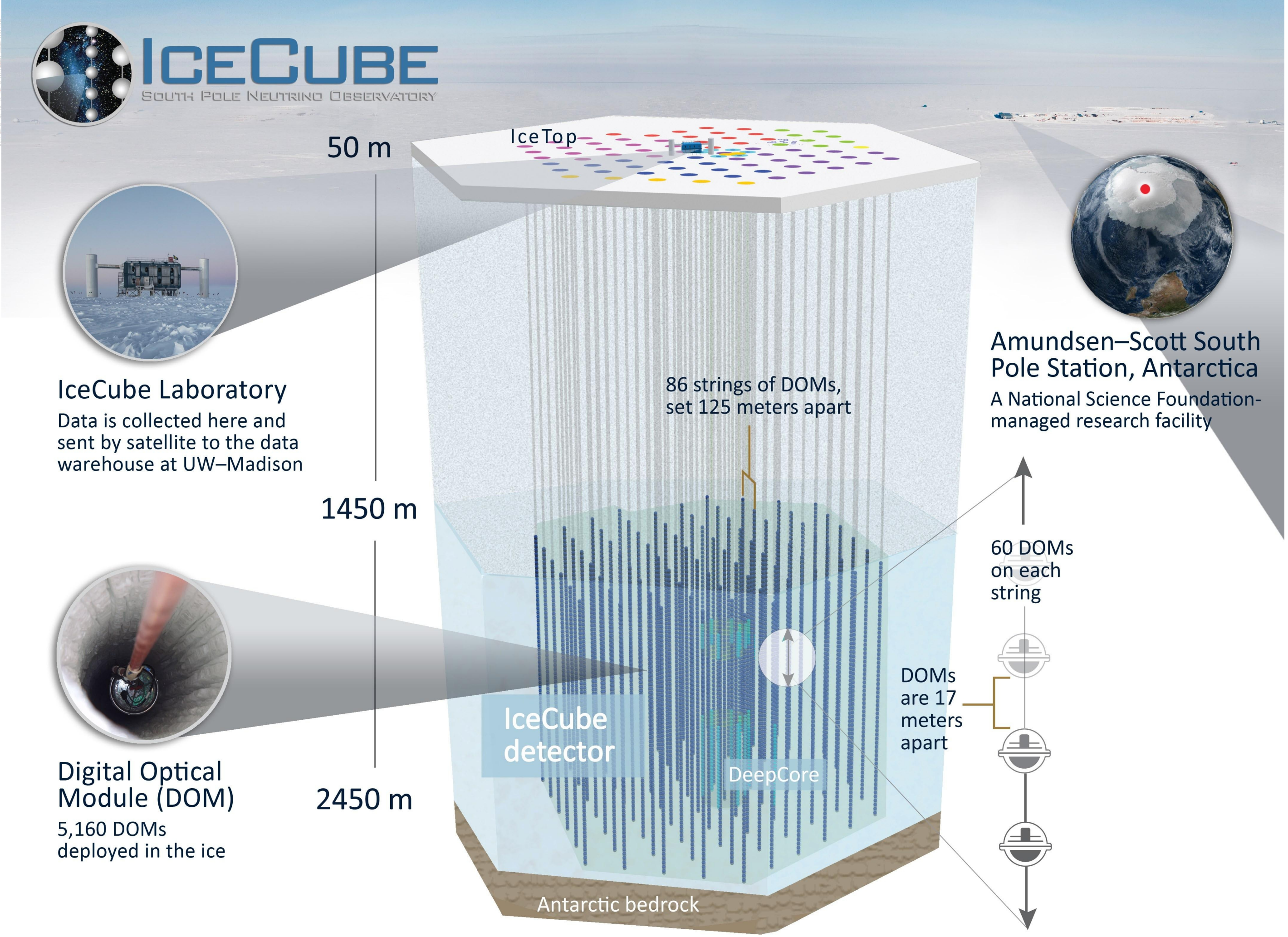
https://mode-collaboration.github.io/

Physics Results
& application to measurement data

Astroparticle physics analysis → based on deep learning







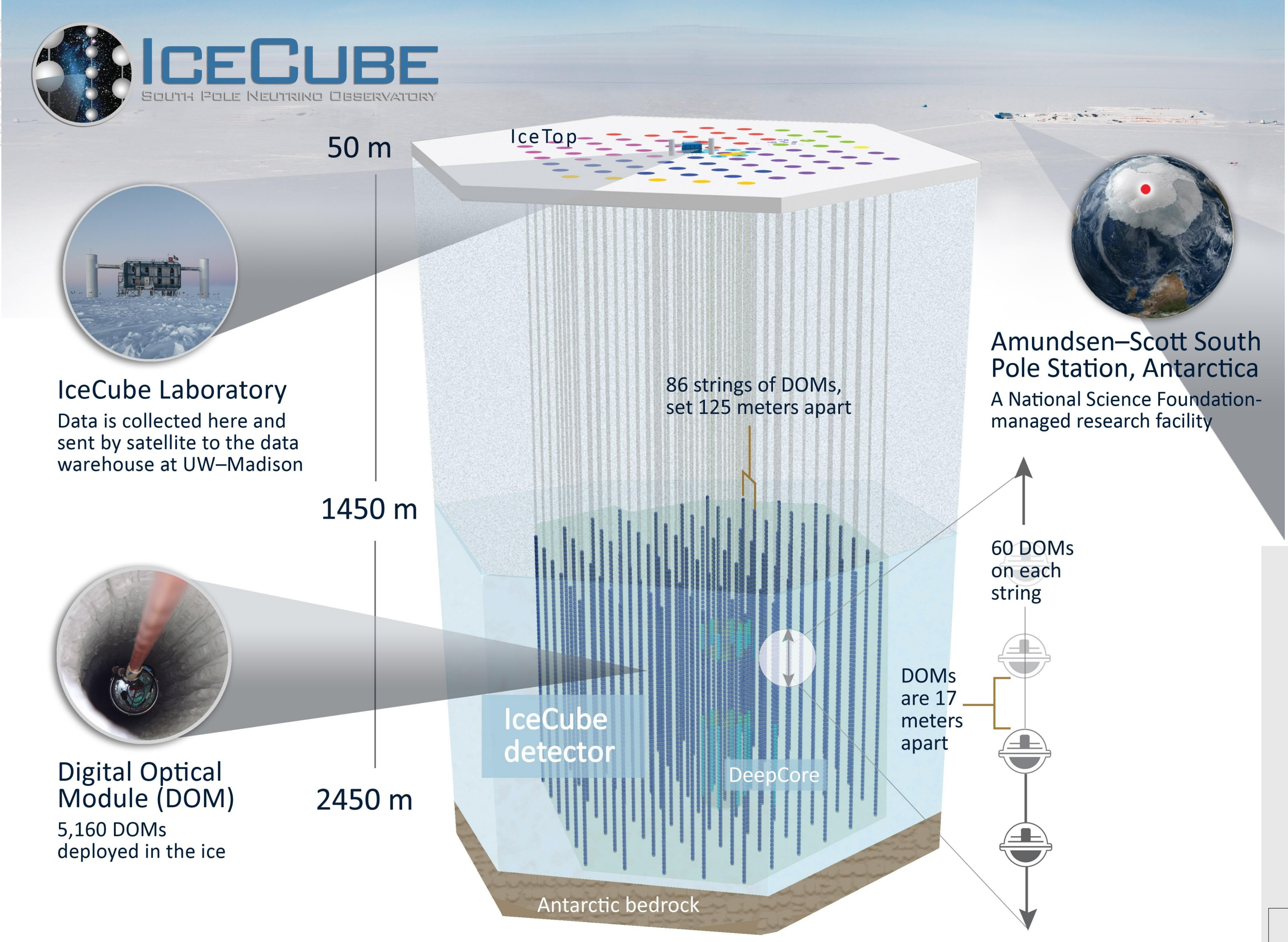
https://icecube.wisc.edu/

- Instrumented km³ of ice
- Detect astrophysical neutrinos (>1TeV)
- DOMs detect time resolved signals (Cherenkov light)

Key findings

- Discovery of astrophysical neutrinos
- Evidence for neutrinos from Blazar, active galaxy, GP
- Indication for astrophysical antineutrinos (Glashow)

Science 361, 147-151 (2018)
Science 378, 538-543 (2022)
Nature 591, 220-224 (2021)
Astrophys.J. 833 (2016) no.1, 3



https://icecube.wisc.edu/

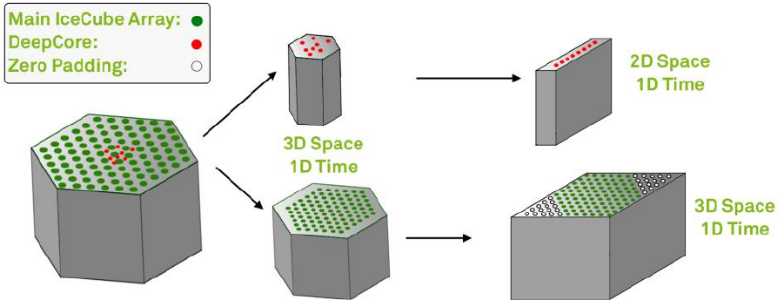
- Instrumented km³ of ice
- Detect astrophysical neutrinos (>1TeV)
- DOMs detect time resolved signals (Cherenkov light)

Challenging background

- Atmospheric muons/neutrinos
- Per single astrophysical neutrino → 10^8 bkg. events

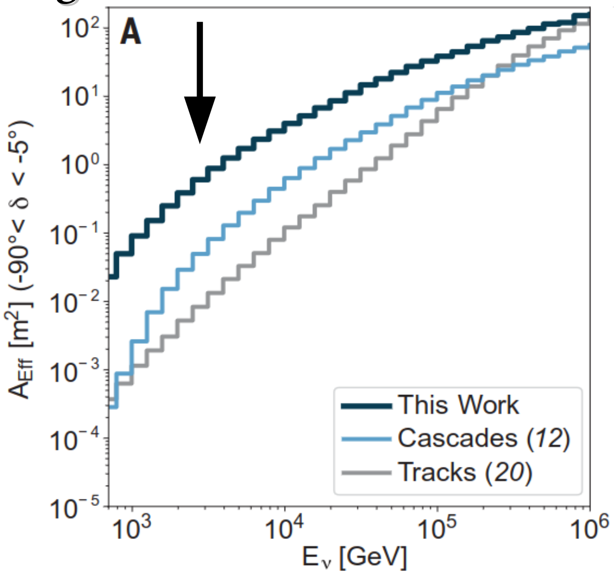
Odds for being killed by a vending machine: 1.2 * 10^8

Improvement: data-driven techniques



Final sample:
87% atmospheric neutrinos
7% astrophysical neutrinos
6% atmospheric muons

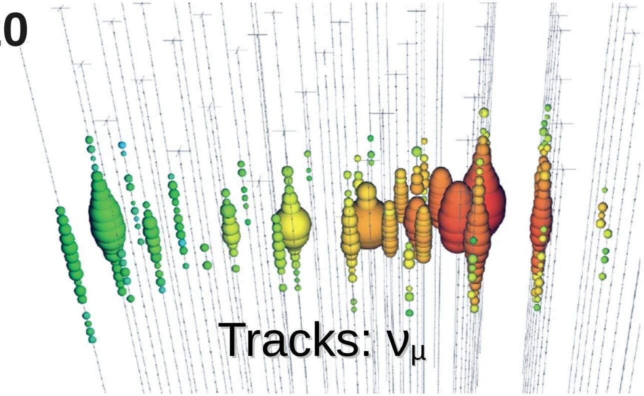
Deep learning: events x20!

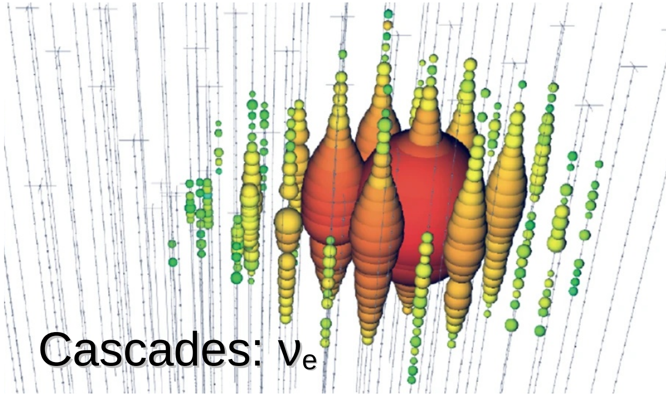


Analysis of cascade events

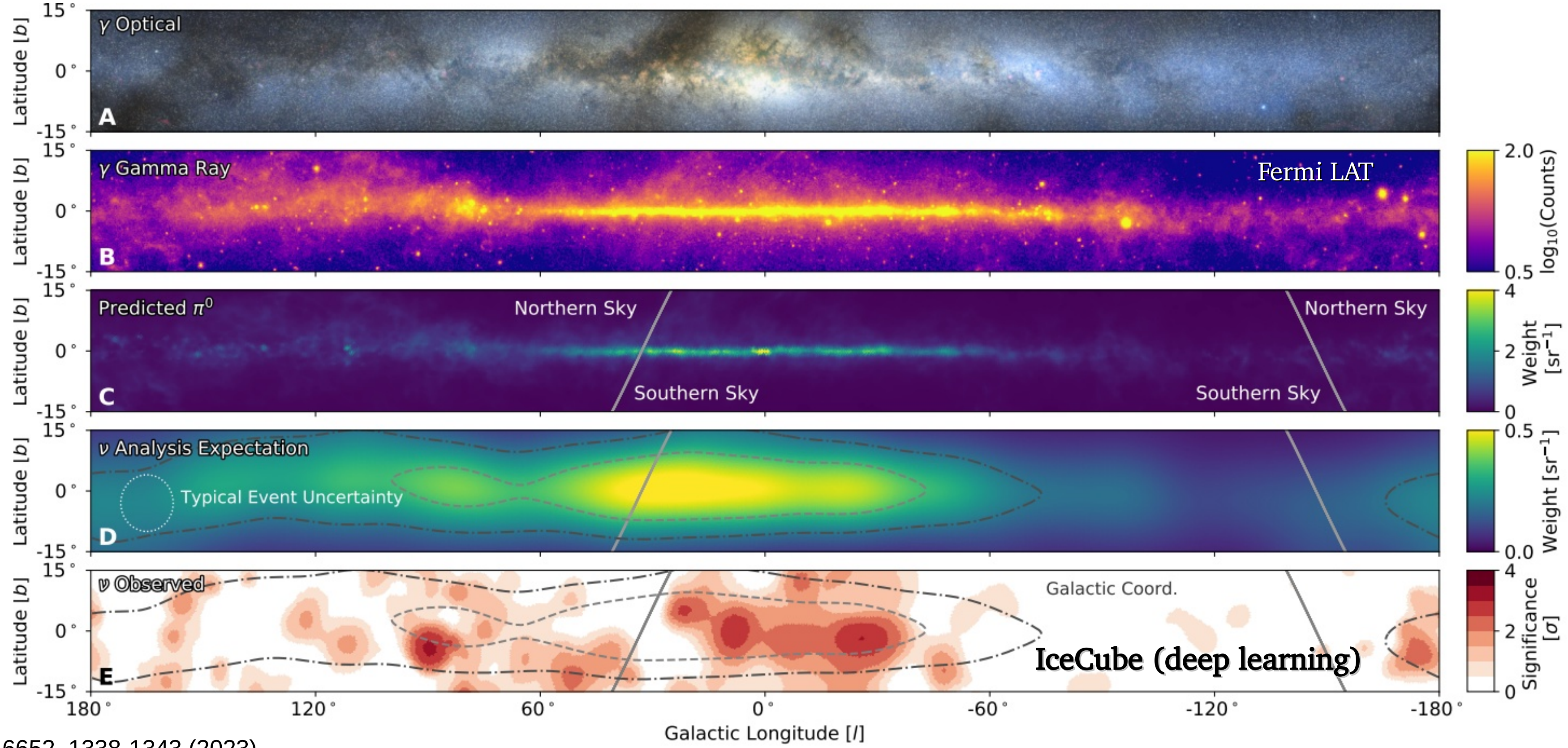
- Improved rejection of atmospheric muons (CNN based)
- Improved reconstruction of cascade events (NN + MLE)
- Reconstruct partially-contained events
- **Statistics increase x20**

[1] M. Hünnefeld et al., PoS(ICRC2017)1057
[2] A. Aiello et al., JINST 15 (2020) P10005
[3] R. Abbasi et al., JINST 16 (2021) P07041
[4] M. Hünnefeld et al., PoS(ICRC2021)1065





The Galactic Plane



Science 380, 6652, 1338-1343 (2023)

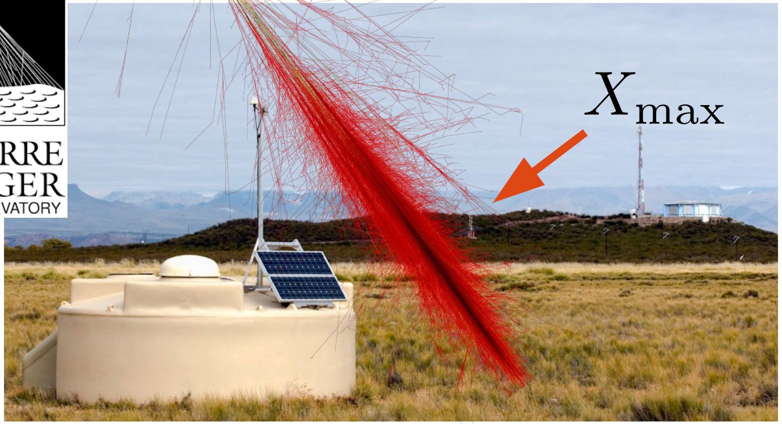
- Comparison to Gamma-ray catalog
- 4.5σ significance (scrambling w. right ascension)

Evidence for diffuse neutrino emission from Galactic plane!

Ultra-high-energy cosmic rays (UHECRs)

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

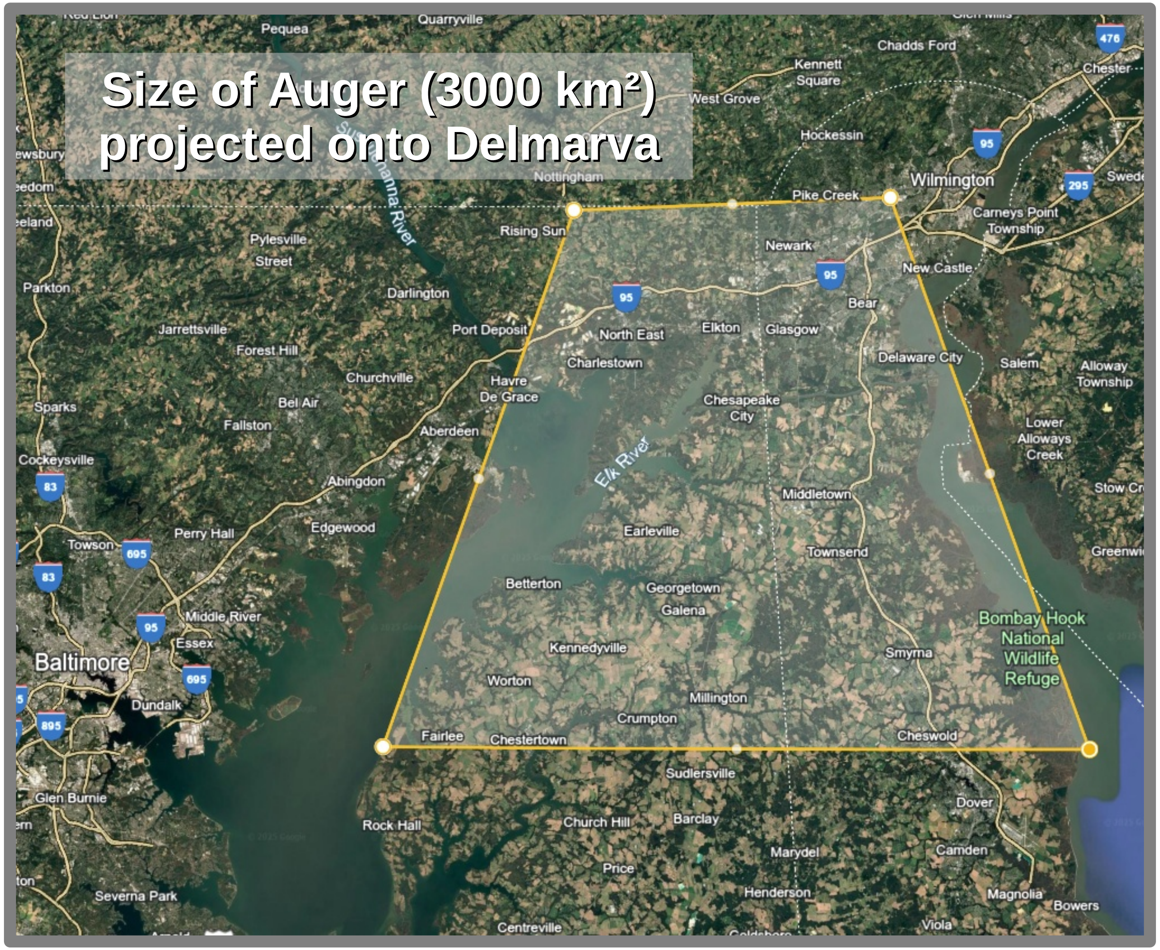
PIERRE AUGER OBSERVATORY



www.auger.org

The Pierre Auger Observatory

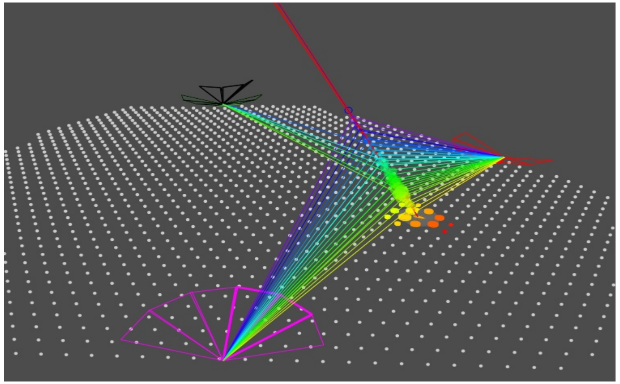
- world’s largest observatory to study ultra-high-energy cosmic rays
- hybrid detection of air showers
  - 1,660 water-Cherenkov detectors
  - 27 fluorescence telescopes
    - can precisely observe Xmax



Air-Shower Reconstruction

The Pierre Auger Collaboration, JINST 16 P07019 (2021)

PIERRE AUGER OBSERVATORY



www.auger.org

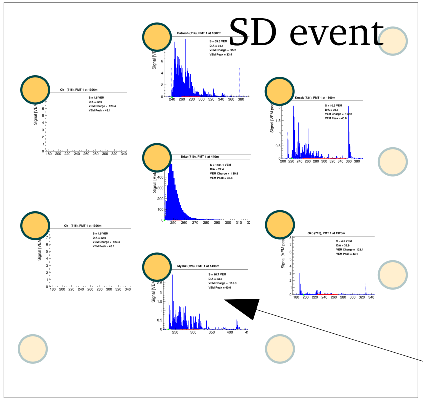
Pierre Auger Observatory

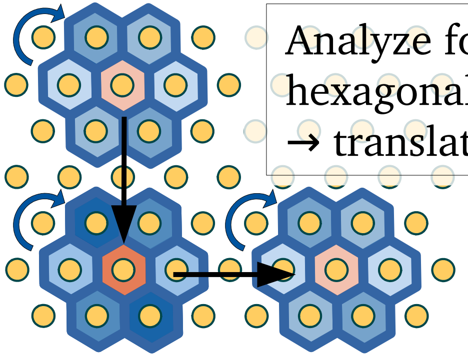
Fluorescence Detector (15% duty cycle)

- direct and precise observation of shower maximum *X*max

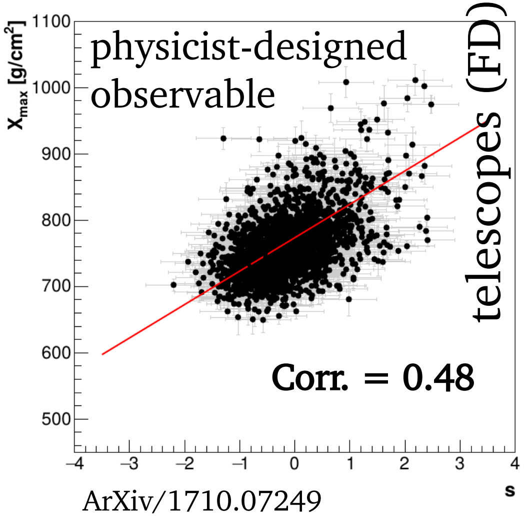
Surface Detector (~100% duty cycle)

- reconstruction of shower maximum using deep learning
- verification with hybrid detection

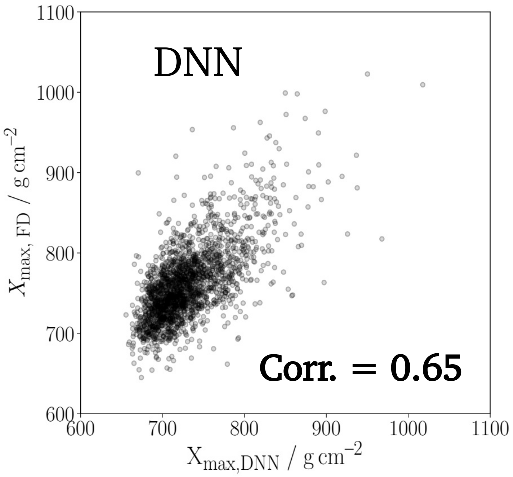




analyze traces with RNNs



ArXiv/1710.07249



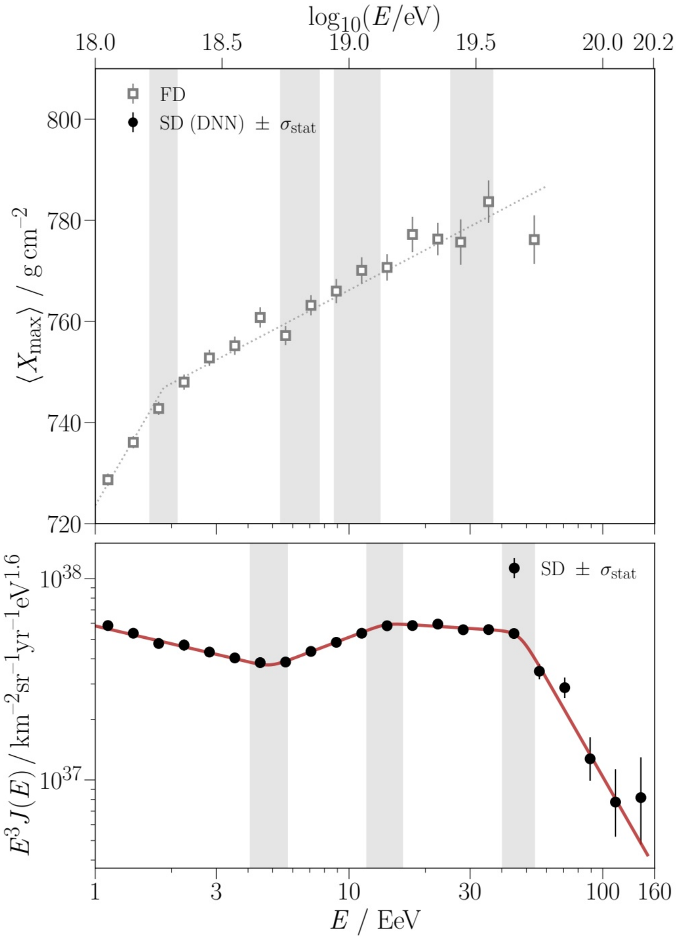
Evidence for breaks in the elongation rate

Critical for understanding astrophysical sources

- Energy spectrum feature (deviations from simple power law)
- Evolution of mass composition

Telescope-based measurements:

- Linear model describes transition from light to heavy



Current interpretation:

- Ankle: transition from galactic to extra galactic
- Cut-off: maximum injection energy accelerator & propagation?

A. Abdul Halim (Pierre Auger Collaboration) et al., Phys. Rev. Lett. 134, 021001 (2025)

A. Abdul Halim (Pierre Auger Collaboration) et al., Phys. Rev. D 111, 022003 (2025)

Evidence for breaks in the elongation rate

Critical for understanding astrophysical sources

- Energy spectrum feature (deviations from simple power law)
- Evolution of mass composition

Telescope-based measurements:

- Linear model describes transition from light to heavy

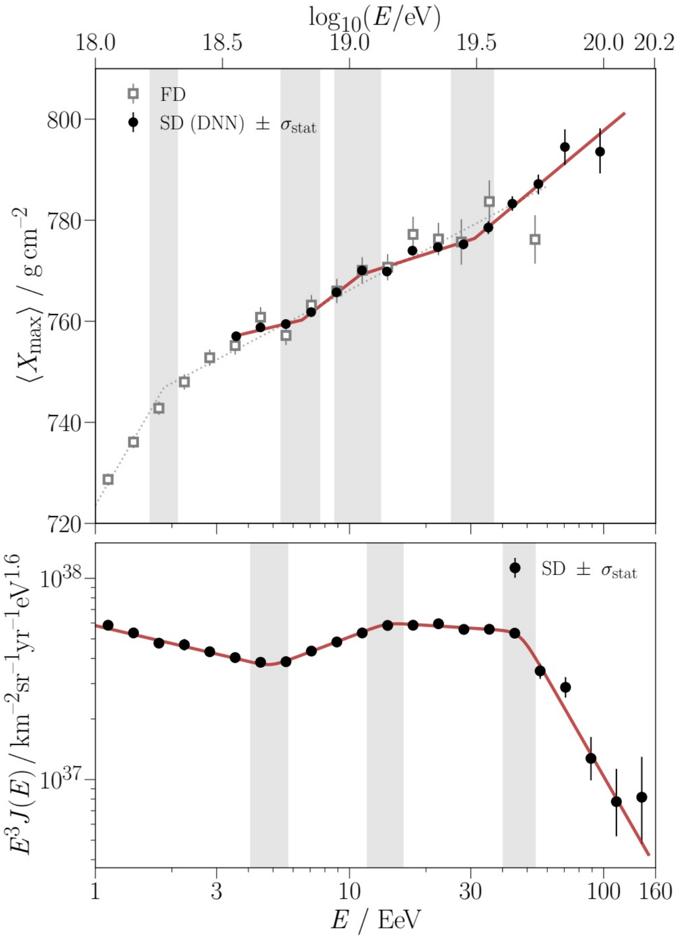
Surface-detector based (utilizing **deep learning**): statistics x10

- Evidence for three breaks, in proximity of spectrum features

  same statistic: telescopes would need to operate for 150 years!

Current interpretation:

- Ankle: transition from galactic to extra galactic
- Cut-off: maximum injection energy accelerator & propagation?



A. Abdul Halim (Pierre Auger Collaboration) et al., Phys. Rev. Lett. 134, 021001 (2025)

A. Abdul Halim (Pierre Auger Collaboration) et al., Phys. Rev. D 111, 022003 (2025) cartoon credit: Doug Neill

*Note: There is no free lunch → ”Older algorithms” are not necessarily worse

Past, Present, and Future – Deep Learning in Astroparticle Physics

I. Classic ML
- Published physics analyses using high-level observables, BDTs, RFs

II. Proof of concept
- First SAL publications of applying DL at low- & high level data (MC)

III. Verified reconstruction mechanisms
- First publications by Collaborations, e.g., Pierre Auger, IceCube, KM3Net ...

IV. Exploiting symmetries
- Incorporating symmetries into architectures
- increase robustness

V. Full Physics analyses
- Publications by Collaborations
- Application to data
- Extensive study of systematics

‘Unsupervised era’
- exploiting measured data
- refinement of simulations
- AI-based detector design

DL close to sensors
On-site application of ML algorithms

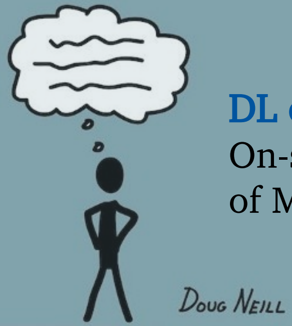
Physics with LLMs
- use pre-trained models
- try “to teach” physics

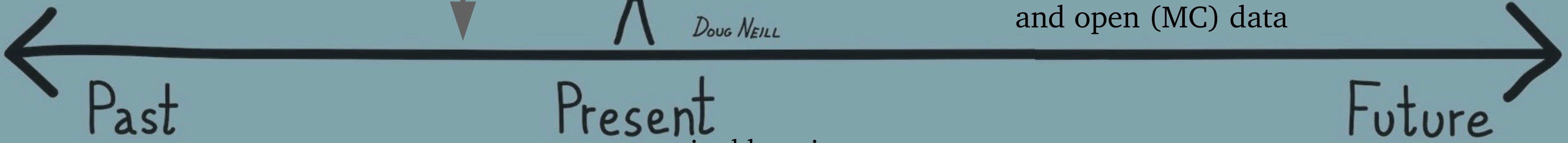
Interpretability
- introspection & causality
- Distilling physics laws from DNNs

AGPI?
Artificial general Physics Intelligence

Multi-experiment DL
Application of ML methods to open data

Open data
Large, complete and open (MC) data



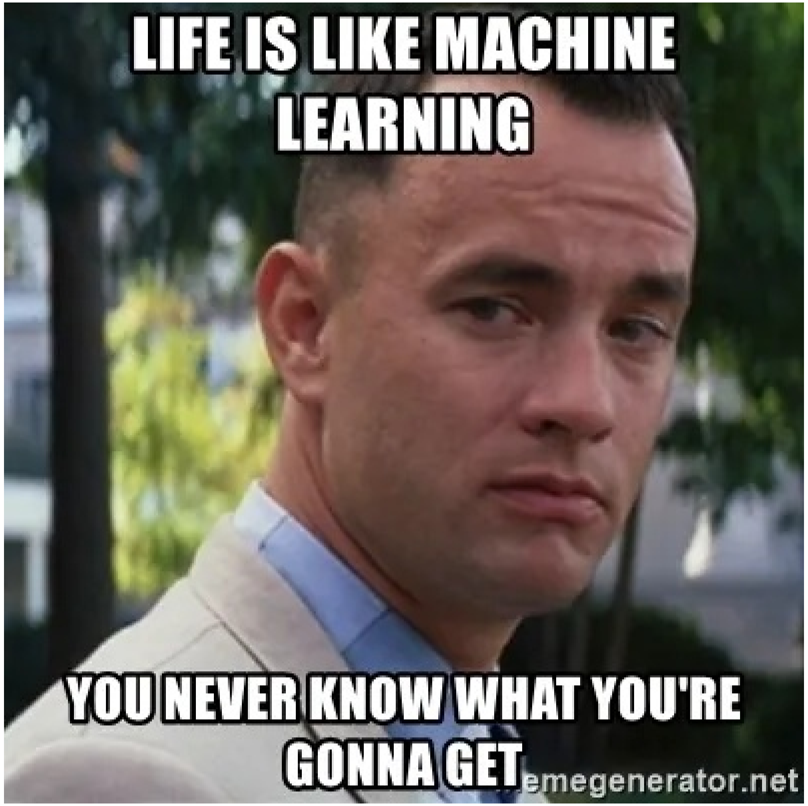


supervised learning →

unsupervised learning →

self-supervised learning →

BACKUP


LIFE IS LIKE MACHINE LEARNING
YOU NEVER KNOW WHAT YOU'RE GONNA GET
emegenerator.net