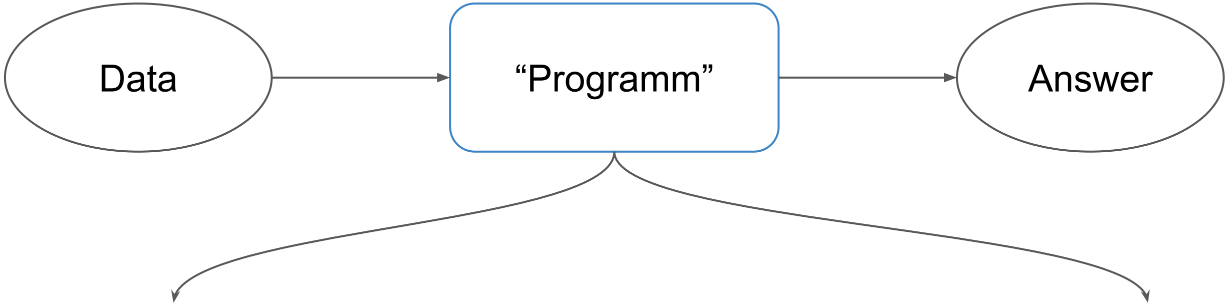
Limits on ultra-high energy photons flux from Telescope Array's data

Quarks 2024, Pereslavl-Zalessky, 21.05.24

1. Neural networks

The idea of machine learning



Standard algorithms:

“Program” is a fixed, created by a human algorithm of solving the problem.

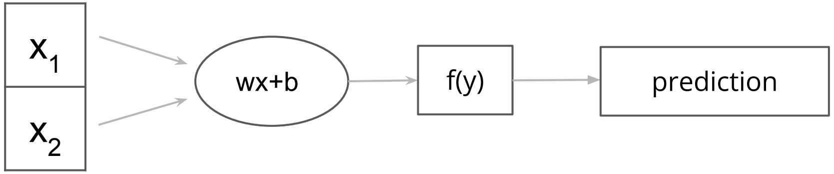
Machine learning:

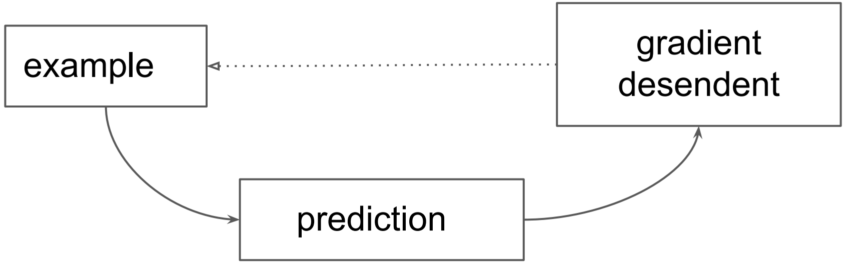
“Programm” - learned on the examples optimal algorithm.

Manually created algorithms

Programs finding optimal algorithms

An illustration



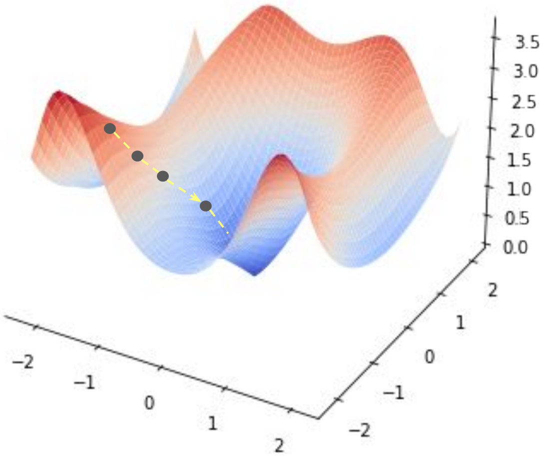


prediction:

[0;1], 0 - insode, 1 - outside

Loss function:

L = | prediction - truth | ≥ 0

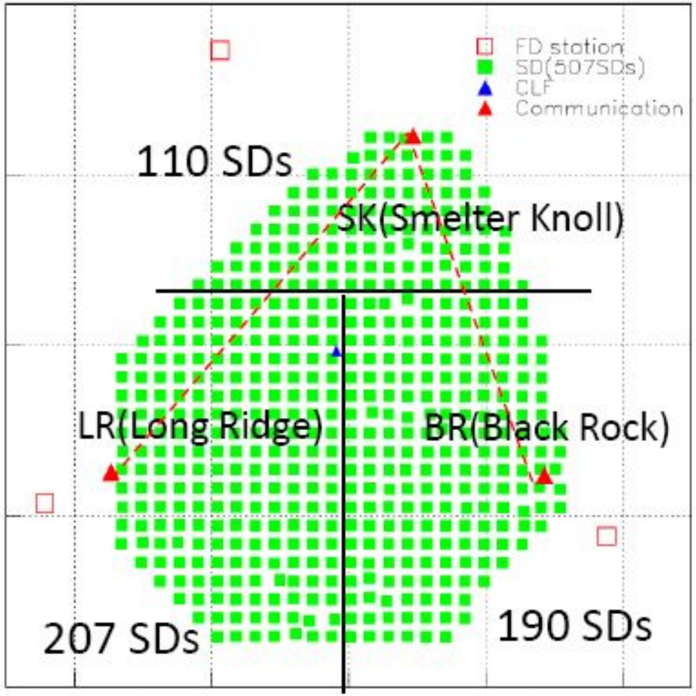


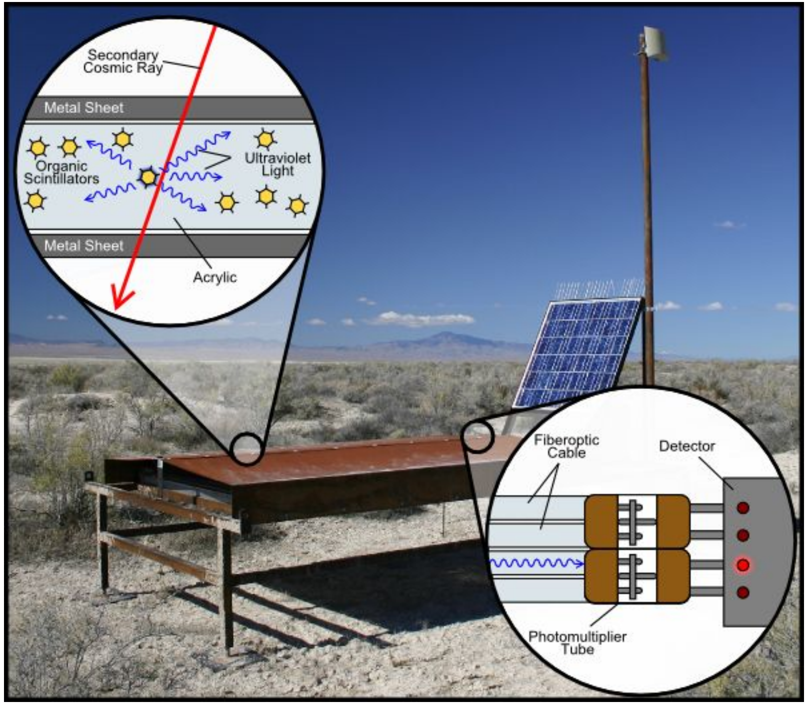
2. Telescope Array

Telescope Array

Telescope Array is the largest cosmic ray observatory in the Northern hemisphere

(700 km^2, 507 surface + 3 fluorescent detectors)



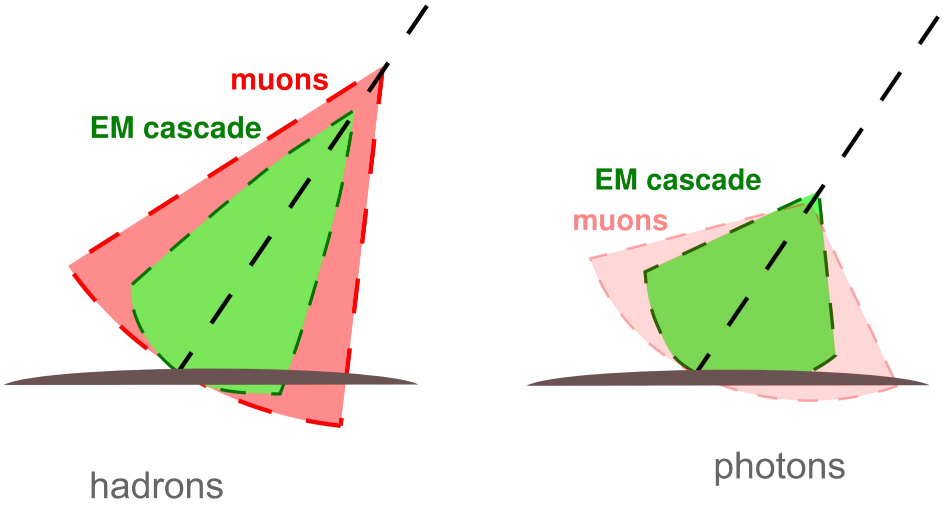


Motivation

Detection of EeV photon-induced air showers would indicate new physics

To search for such events, it is desirable to have as big exposure time as possible

- Surface detectors operate almost at all times.
- To get the best separation of photon- and hadron-induced air showers, employ Neural Networks.



Approach

1. Train a neural network to distinguish photon- and hadron-induced air-showers.
   - Use Monte-Carlo simulated photon- and proton-induced events for training. This provides conservative limits for photon search.

2. Introduce classification threshold.
   - Optimize by requiring best sensitivity to photon candidates in case of the null hypothesis, i.e., no photons in the data.

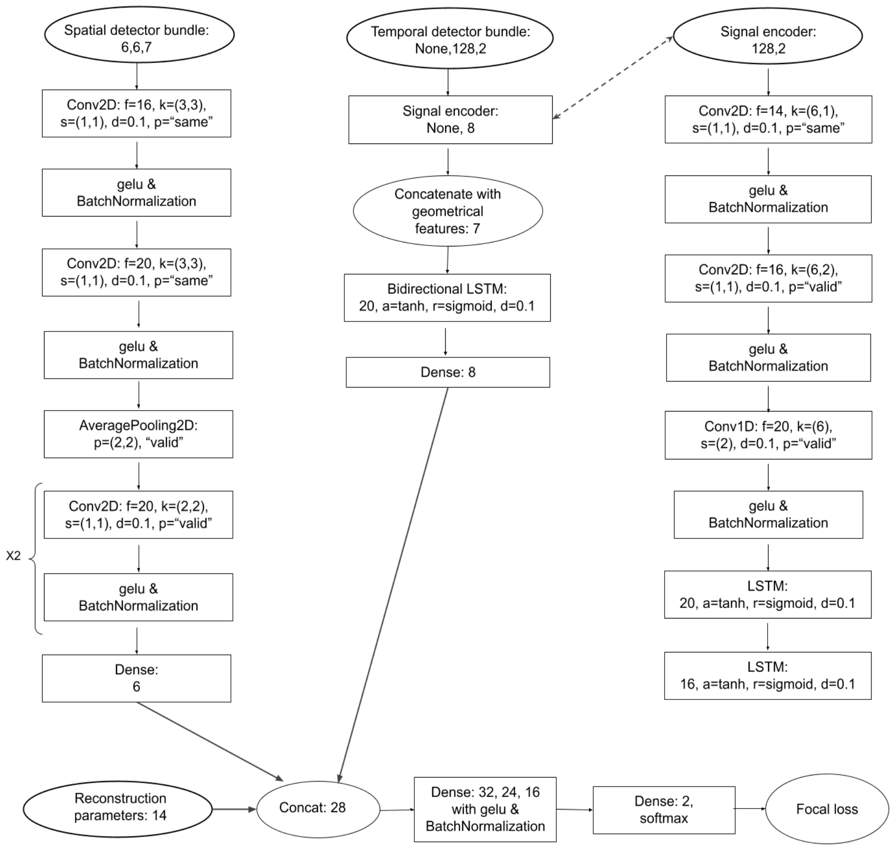
3. Verify applicability to real data analysis

4. Set limits by analyzing the data

Neural network

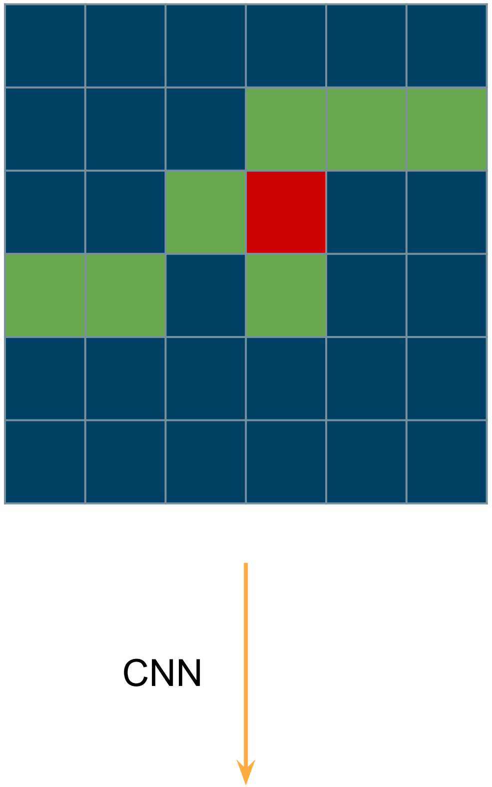
Neural network's blocks:

- Spatial detectors bundle (geometrical features)
- Temporal detector bundle (overall information)
- Reconstruction parameters (high-level information)



Detectors bundle

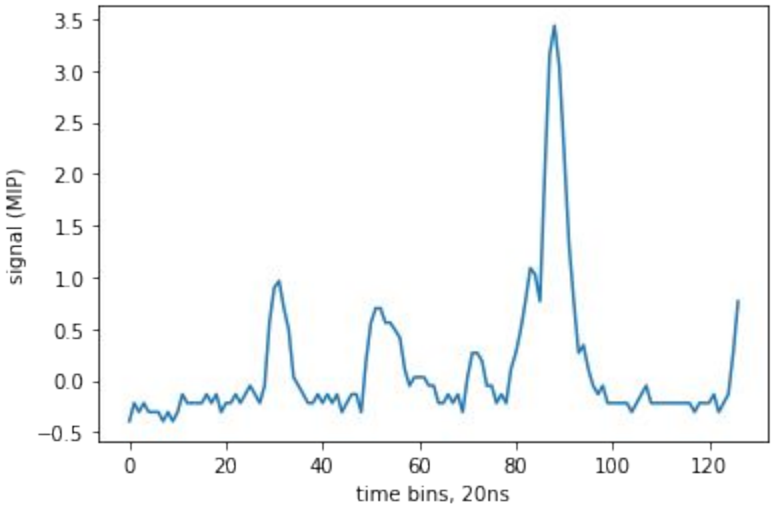
- x, y, and z coordinates of the detector
- Detector’s total signal
- Time of the plane front arrival
- Difference in time between the start of the recorded signal and the wavefront arrival
- Masks:
  - Was triggered?
  - Was saturated?
  - Was excluded from the geometry fit?



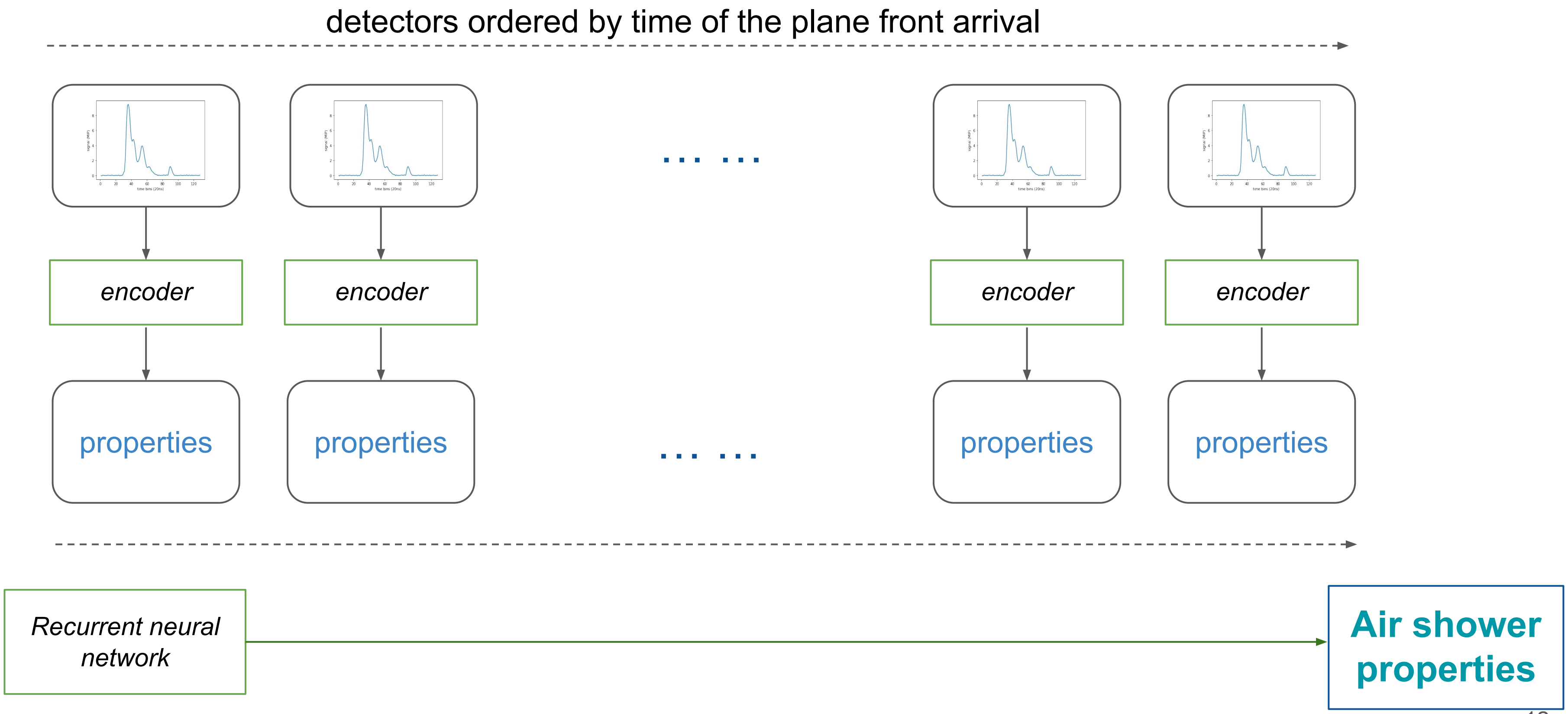
Spatial detectors features

Features from waveforms

- 128 bins (20ns each) of the signal in upper and lower detectors
- Real-time calibration tables, signal measured in “Minimal ionization particle”

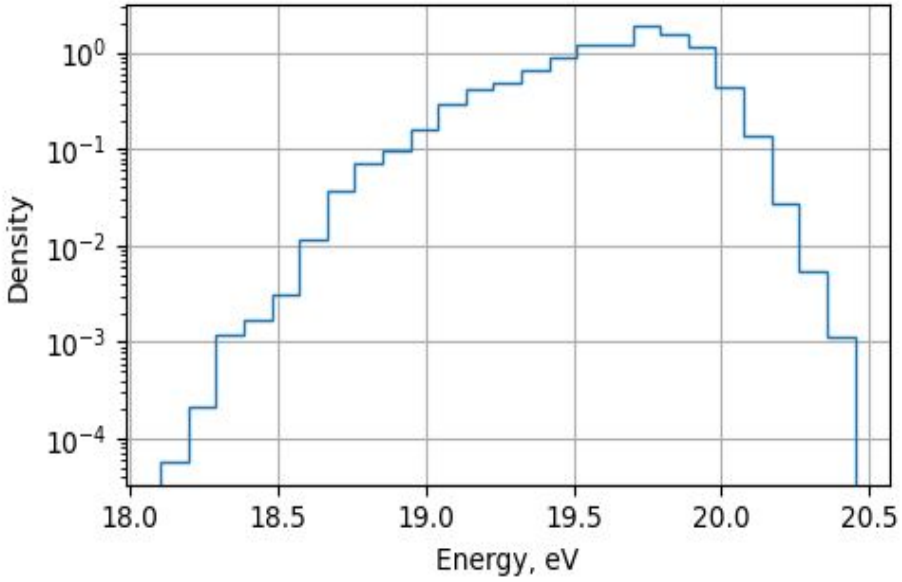


Temporal detector bundle



Dataset

- Full MC simulation of 14 years of TA operation time
- Employ 4 high energy hadronic interaction models: QGSJET-II-03, 04, SIBYLL, EPOS (reduce systematic error).
- Equal-efficiency mixture of proton- and photon-induced events (no unphysical energy related bias).



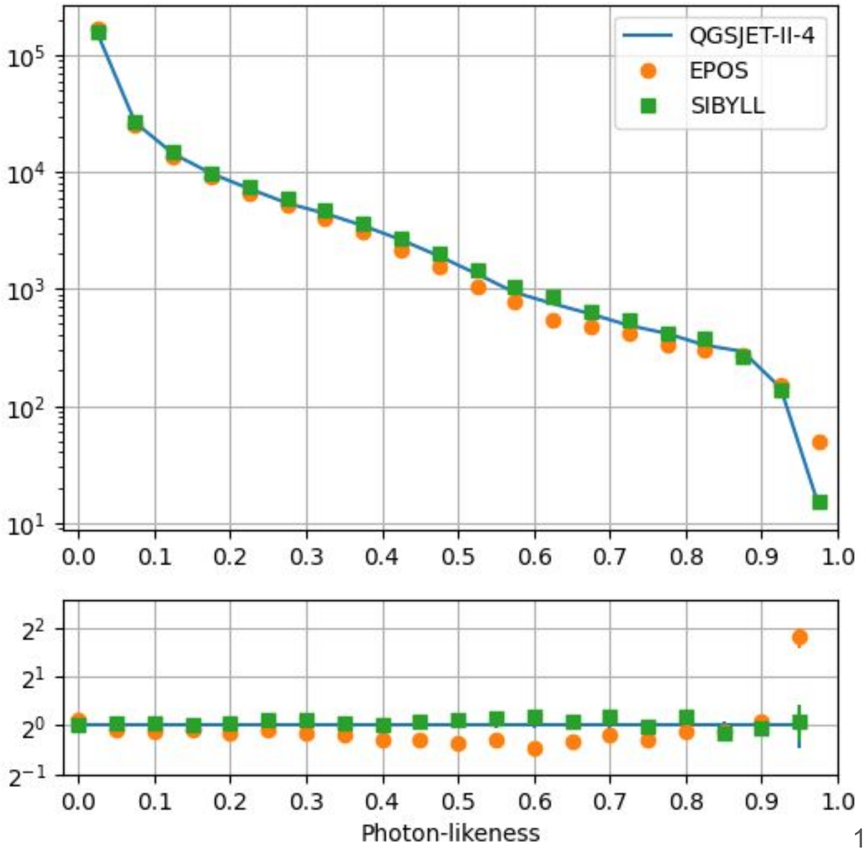
Model dependence

1) Select one of the hadronic models;

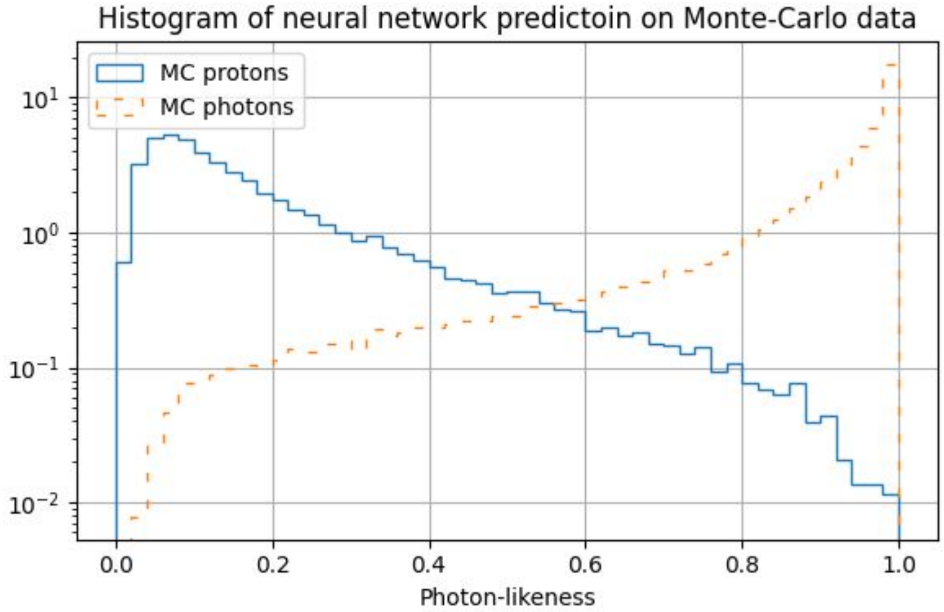
2) Do not use it for training;

3) Compare neural network predictions for the hidden and other hadronic models.

To analyze experimental data, we employ neural network trained on all of the hadronic interaction models.



Training tricks

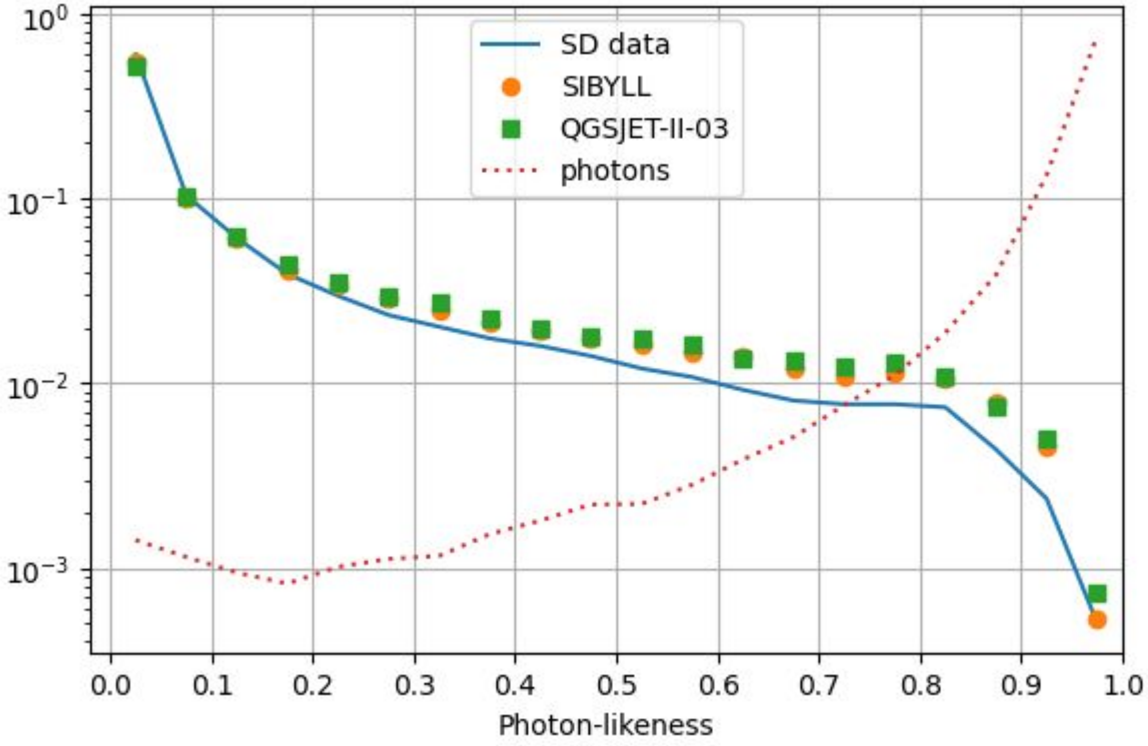


1. Use a special loss function, *focal loss,* forcing neural network to pay more attention to events that are hard to distinguish.

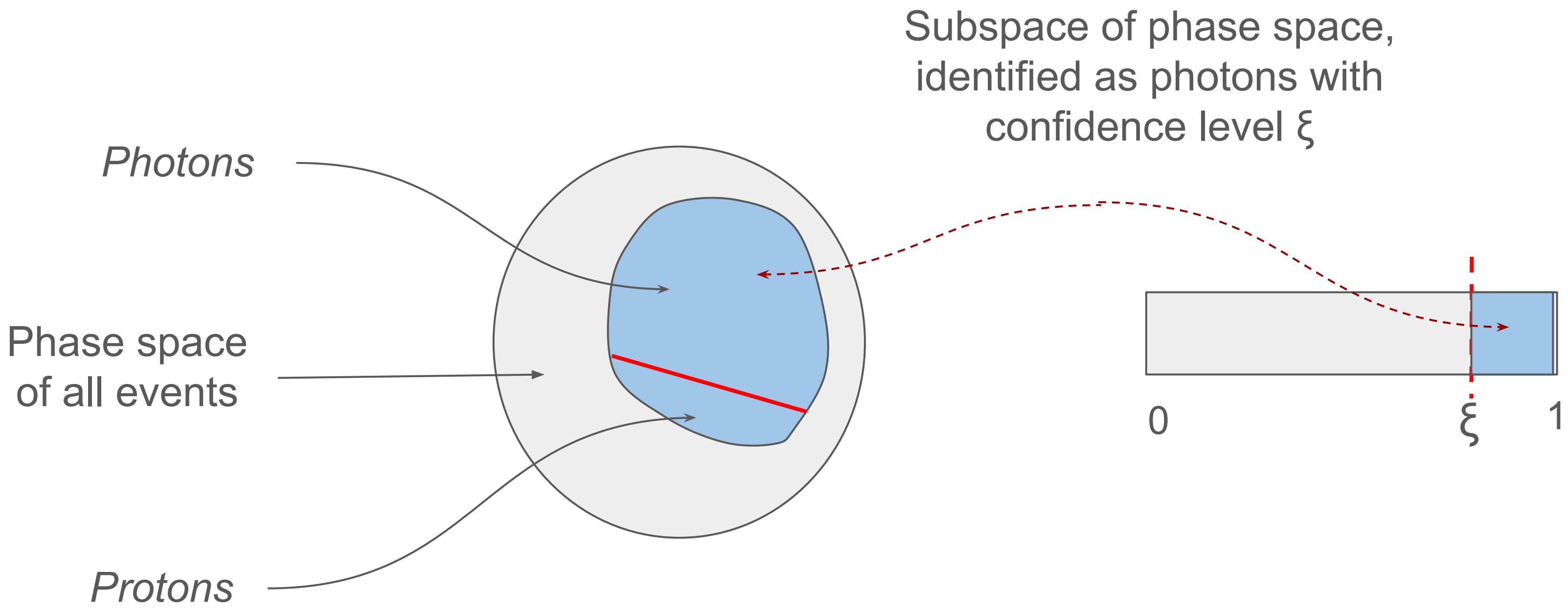
2. Trained an ensemble of neural networks and averaged their predictions. This significantly reduces false-positive rate.

3. Fine-train neural network on experimental data that was classified as protons with high confidence.

Aplicability for analyzing experimental data



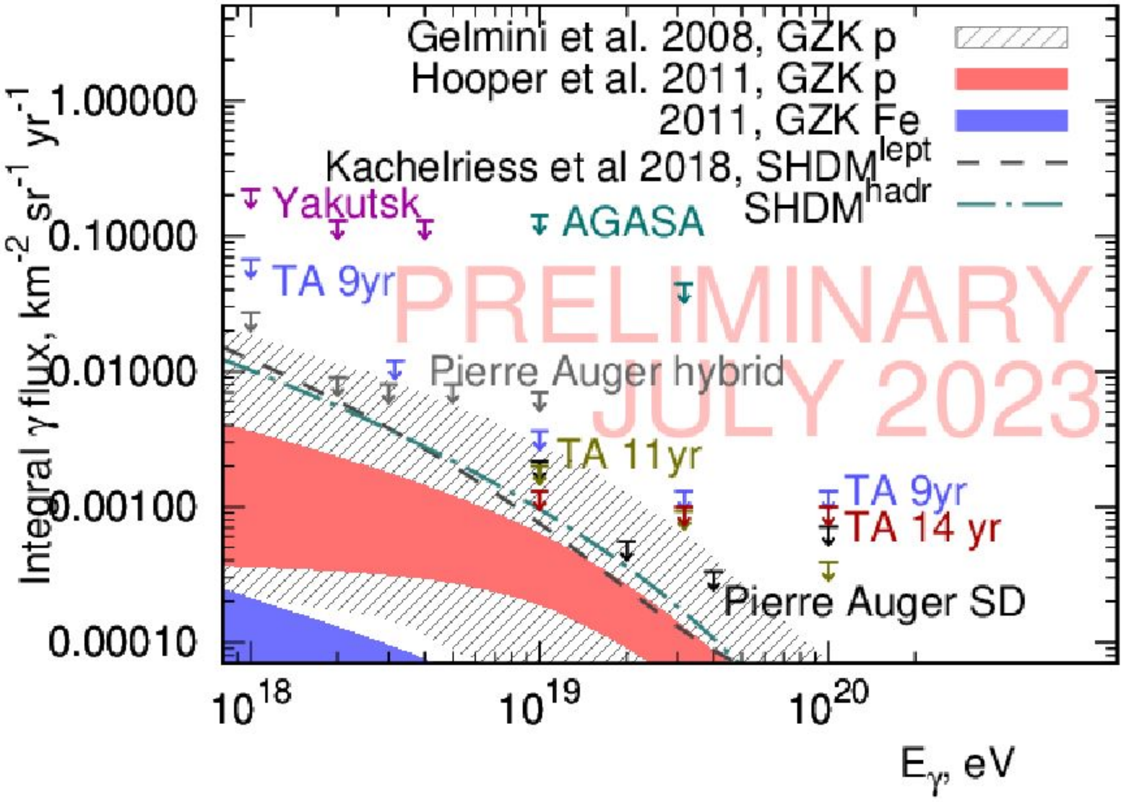
Finding optimal classification threshold



Optimize merit function :

$$L^{95} = \frac{M\left(\sigma^{95}(n_{\text{cand}}(\xi))\right)}{S(\xi)}, \quad S(\xi) = \frac{n_\gamma(\xi)}{n_\gamma^0}$$

Results



Energy, log10 scale, eV	Number of photon candidates
19	0
19.5	0
20	1

Thank you for attention!

This work is supported in the framework of the State project ``Science'' by the Ministry of Science and Higher Education of the Russian Federation under the contract 075-15-2024-541.


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS