Neural networks for electron identification with DAMPE

David DROZ (david.droz@unige.ch), Group Prof. Xin WU, DPNC

In collaboration with ERC PeVSPACE project (prof. Andrii Tykhonov)

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

References
- Chang, J., et al. (2017) Astropart. Phys., vol. 95, pp. 6–24
- Ambrosi, G., et al. (2017). Nature, 552(7683), 63.
- **Droz, D., et al. (2021) J. of Instrumentation 16 P07036**

Software
- Keras: Chollet, F. & al. (2015). https://keras.io
- Theano: Theano dev.team (2016) arXiv:1605.02688

1. DArk Matter Particle Explorer

A cosmic ray space observatory in operations since December 2015. It is equipped with a deep calorimeter (32 X_0) able to detect electrons up to 10 TeV with a 1% energy resolution.
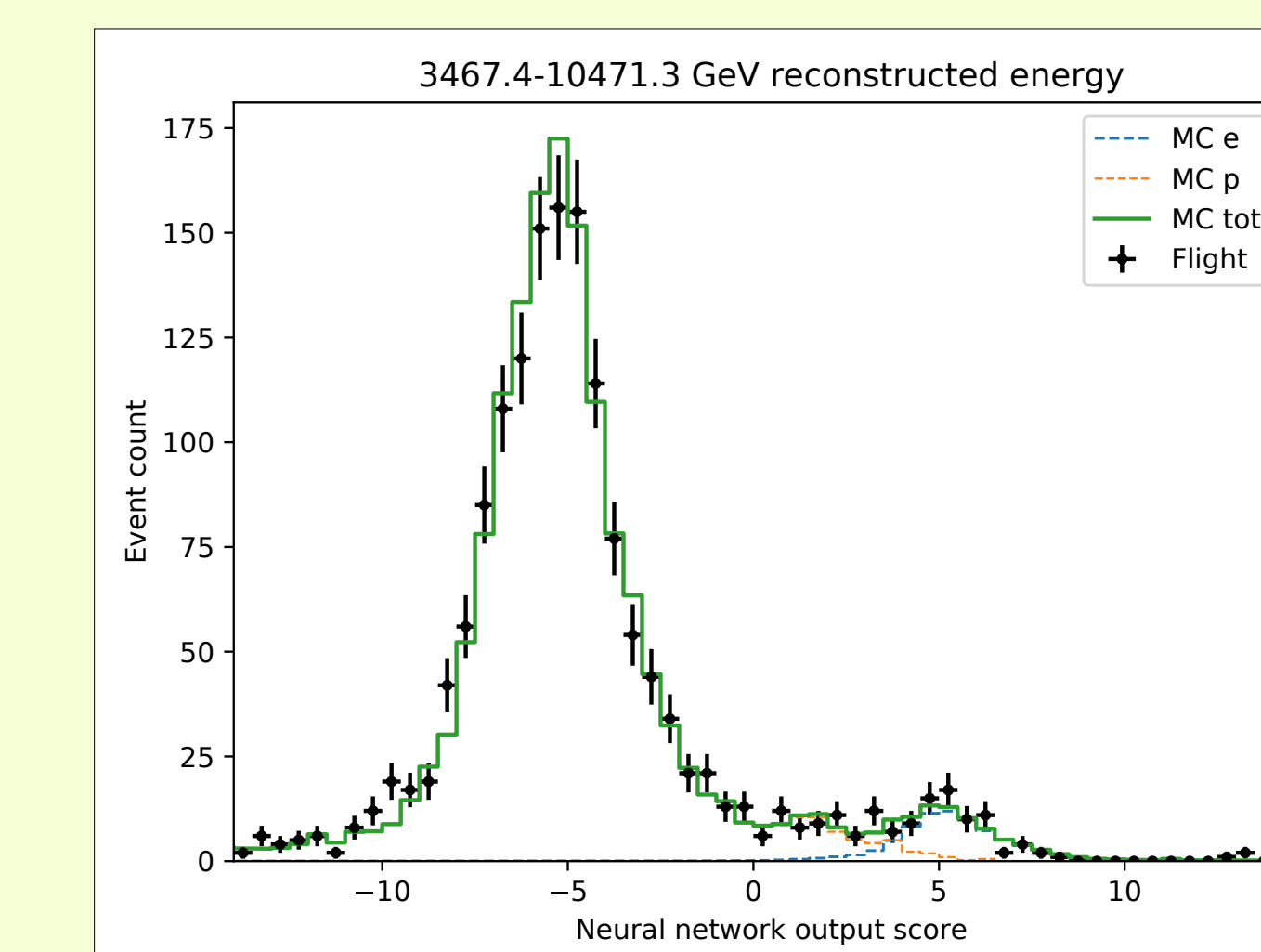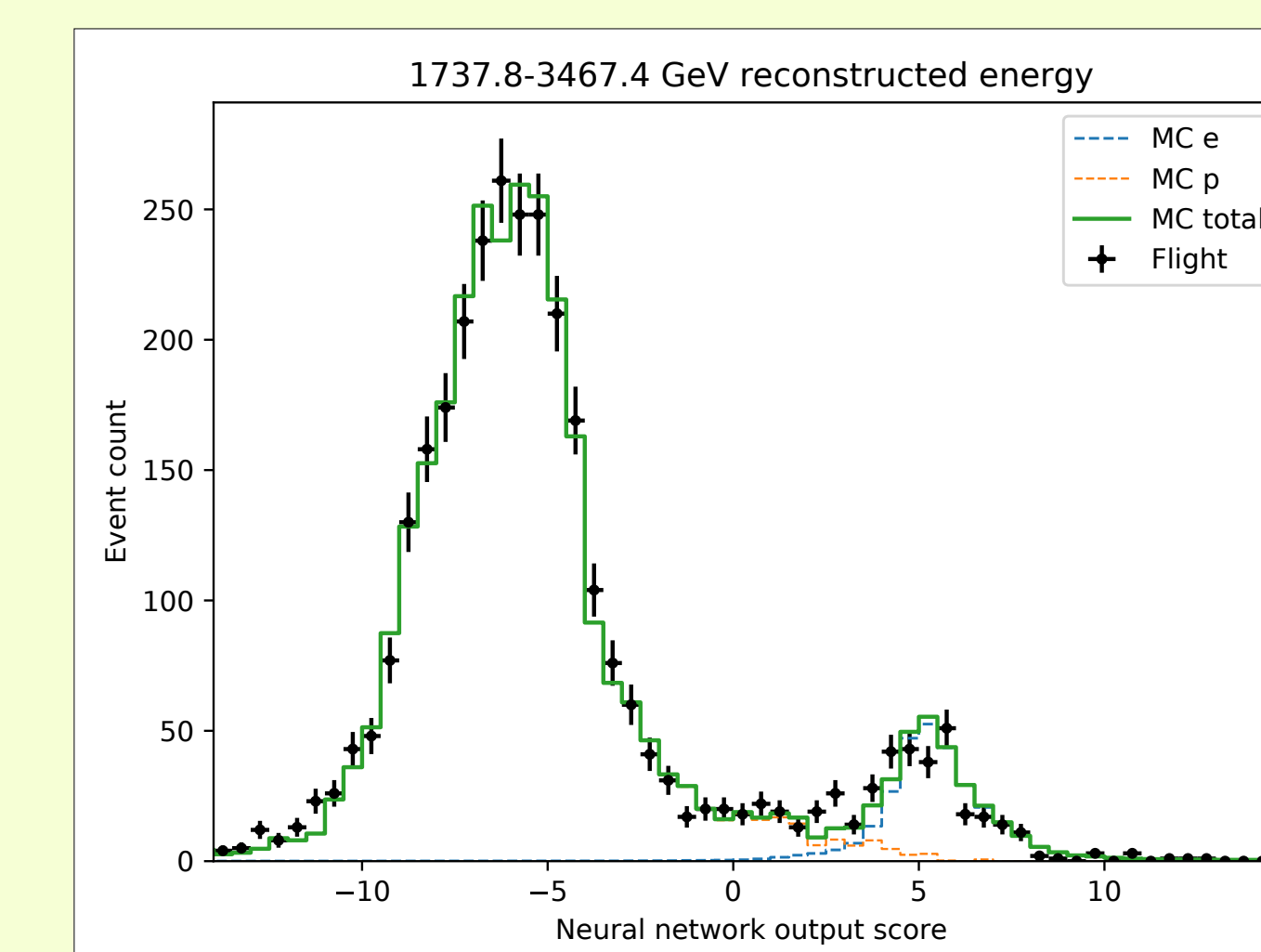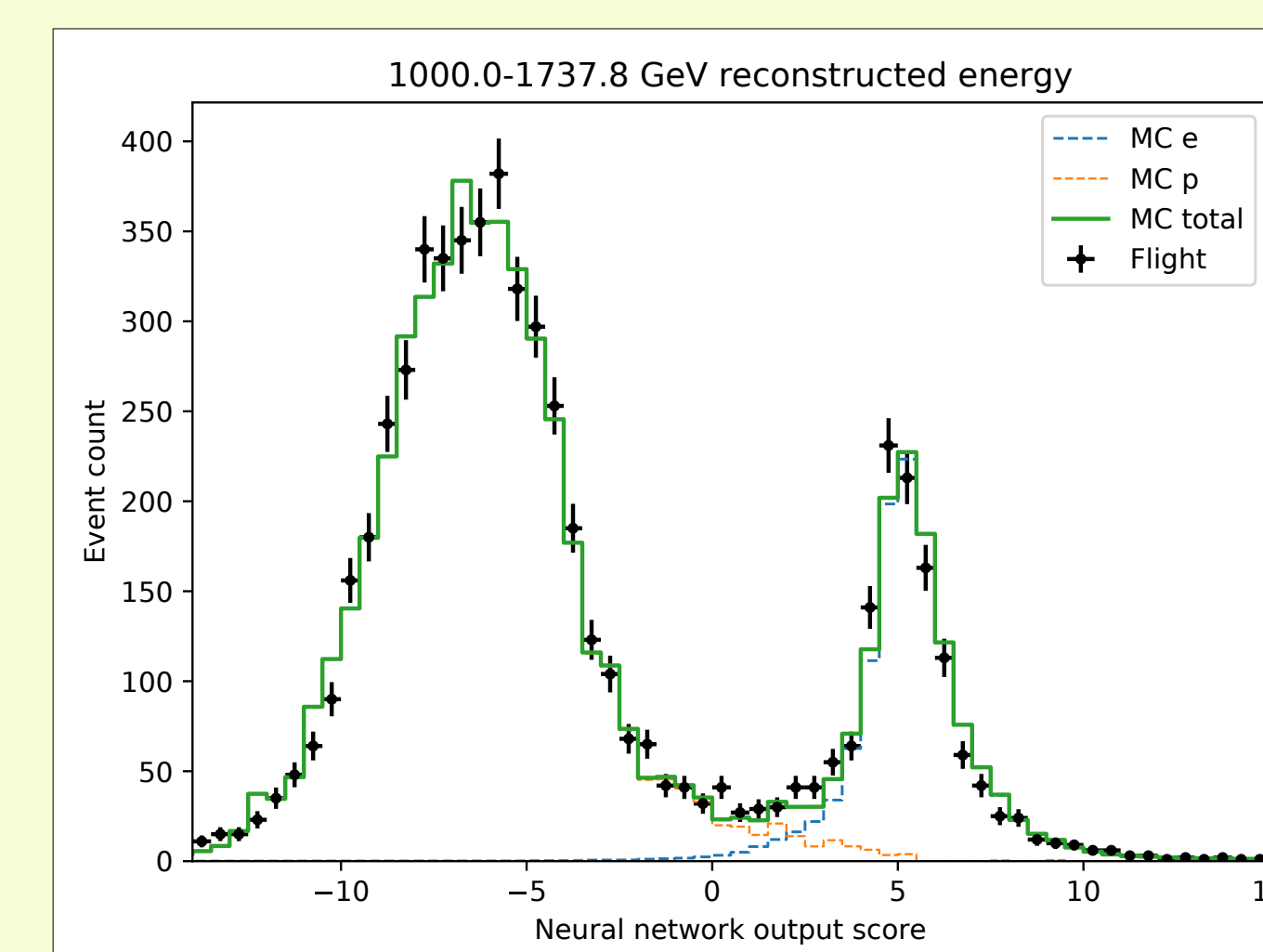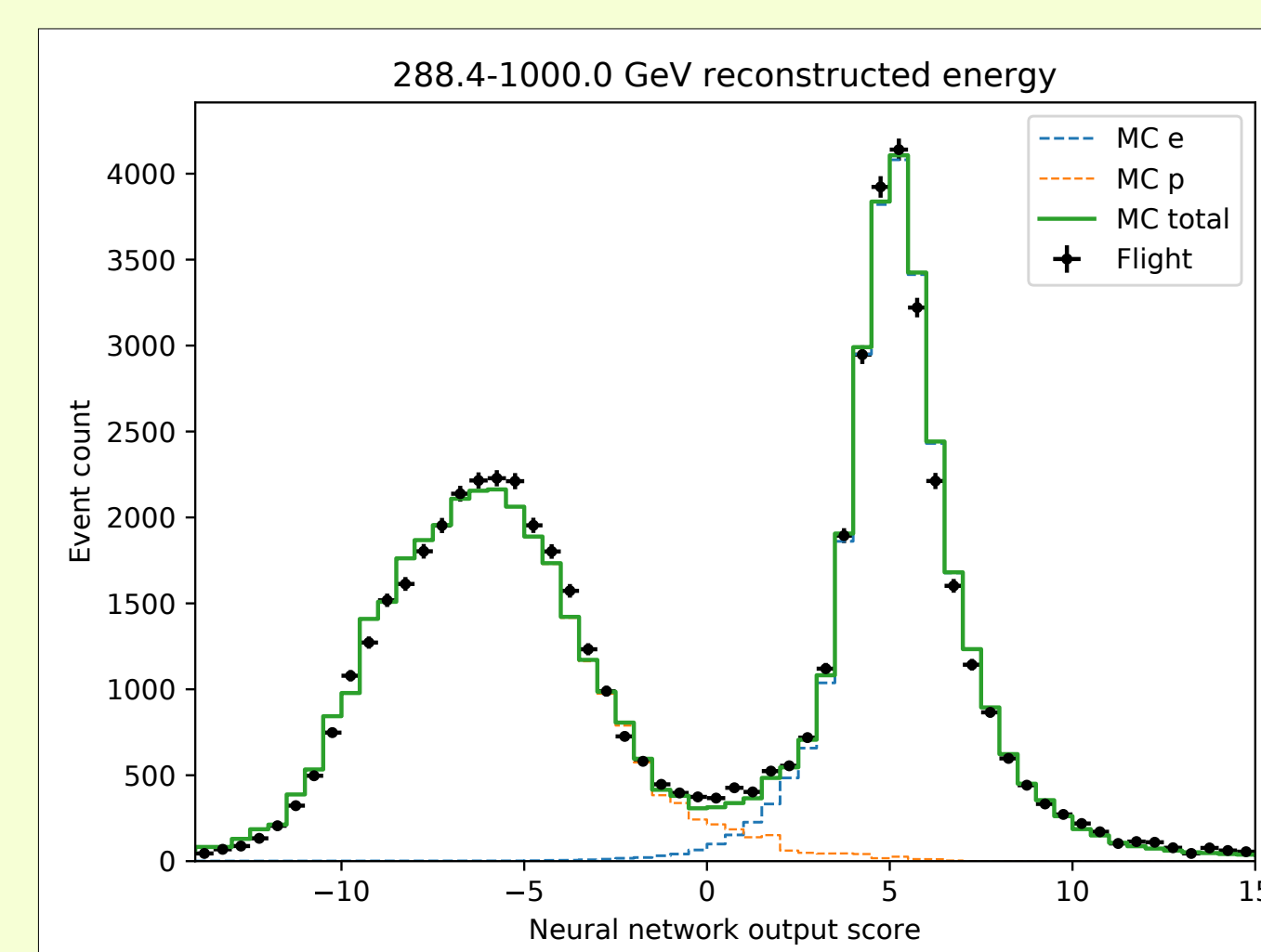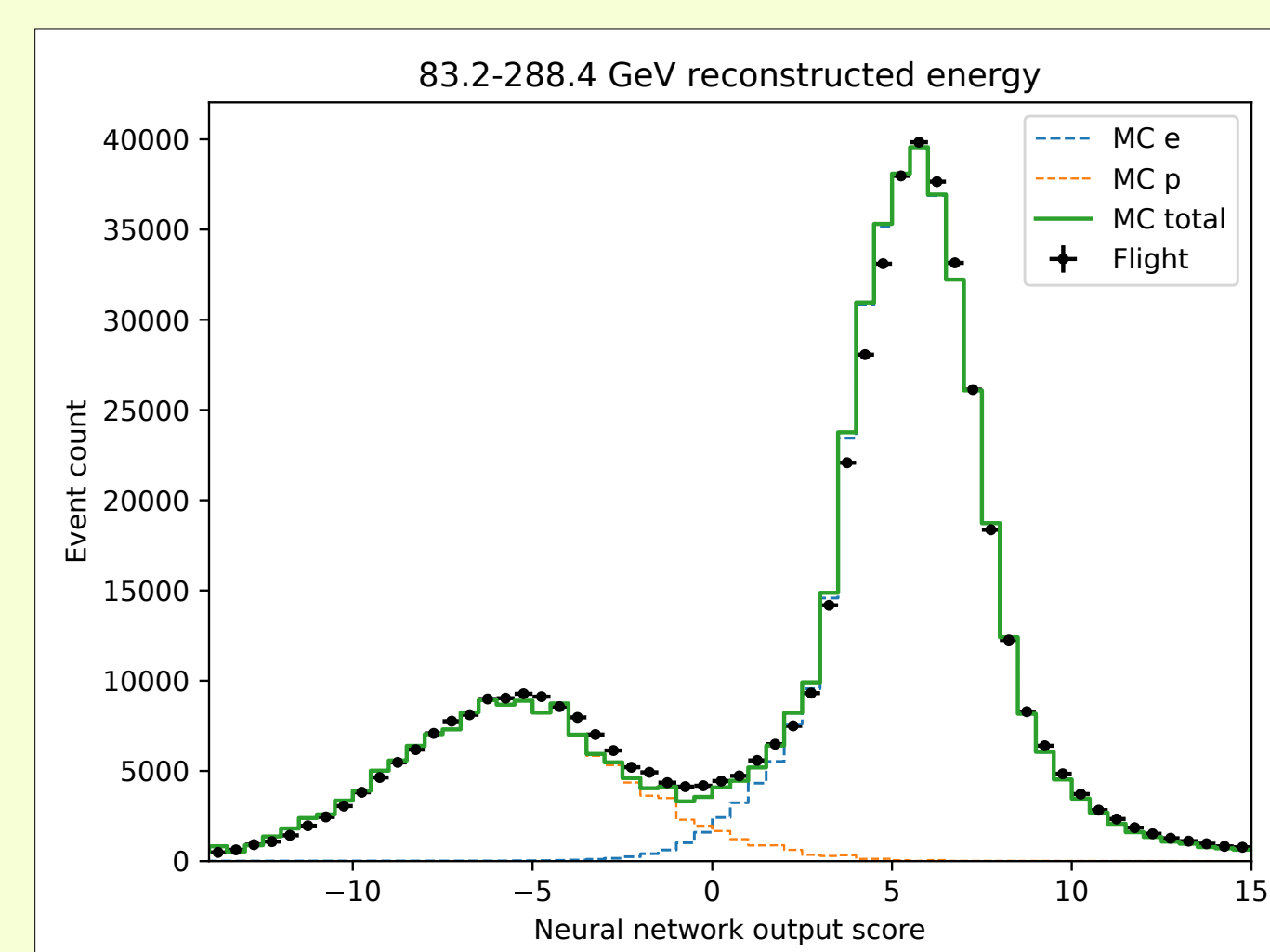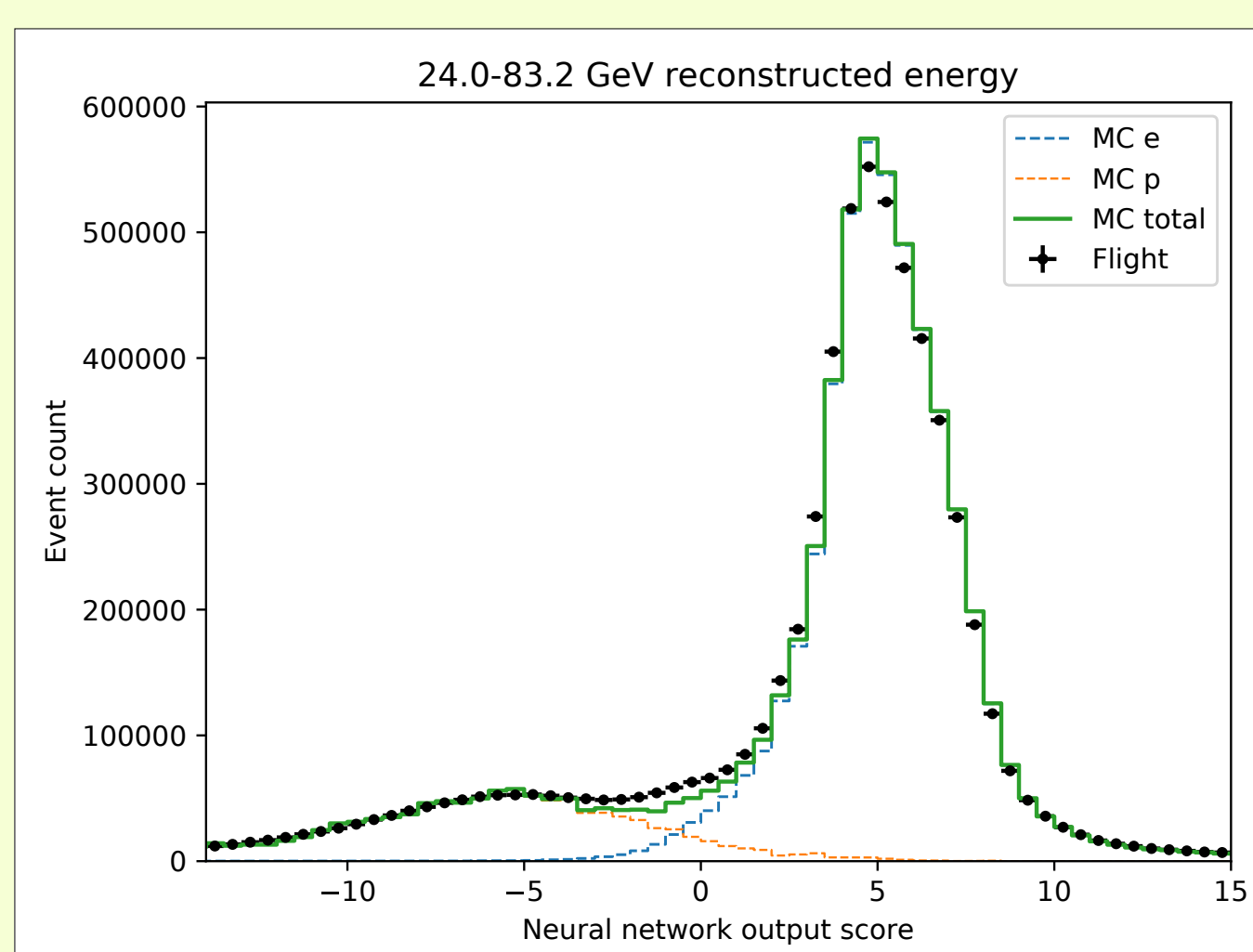
Plastic Scintillator
Tracker
Calorimeter
Neutron Detector
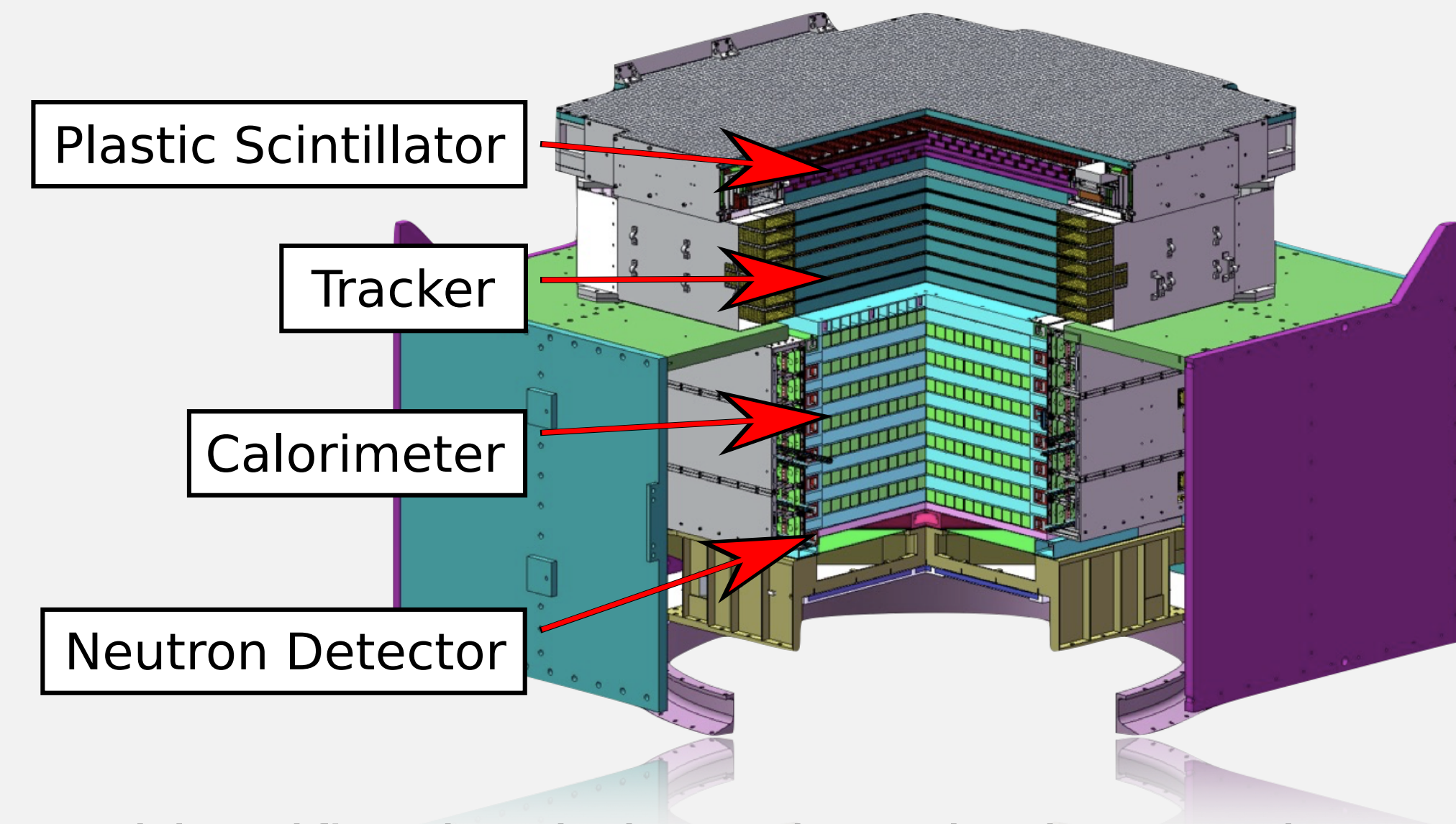
Electron identification is based on the interaction topology. The classical method is to define such observable *[Ambrosi et al.]*

ζ = shower width × shower depth

Electrons have a lower ζ than protons and nuclei.

However ζ is limited at several TeV. A better method is required.
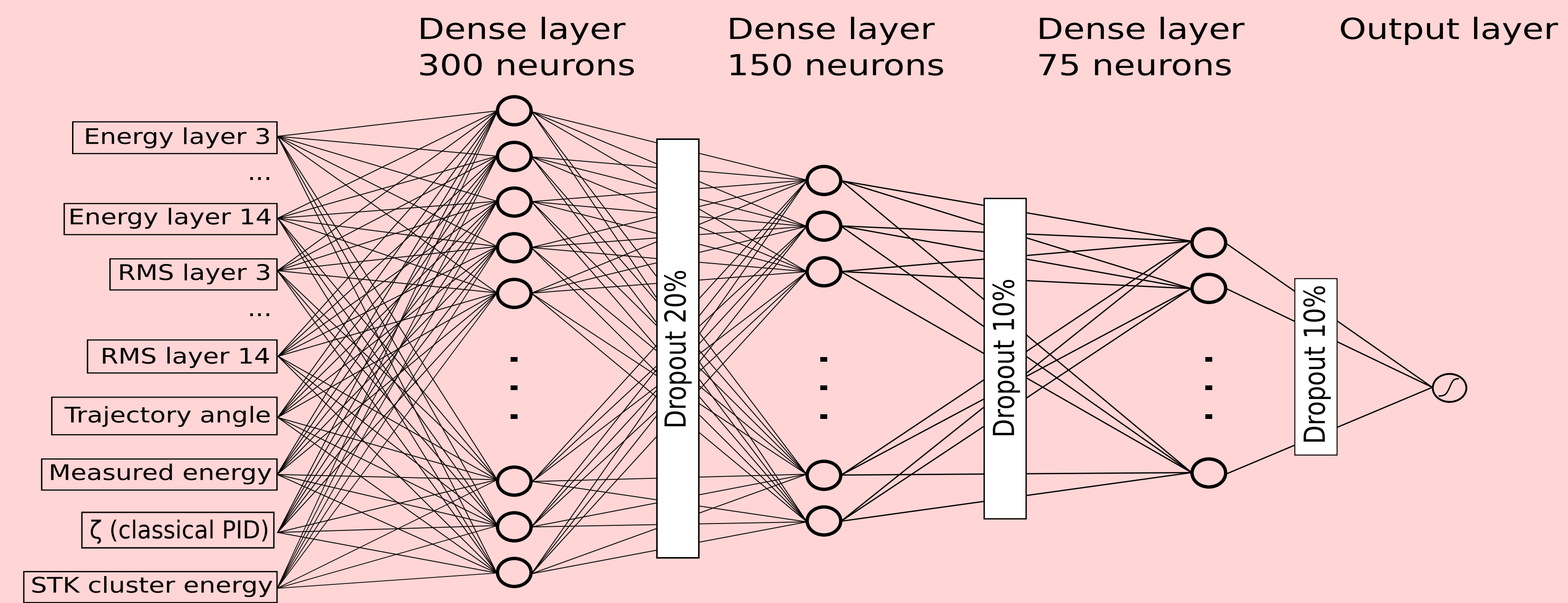
Electron

Proton
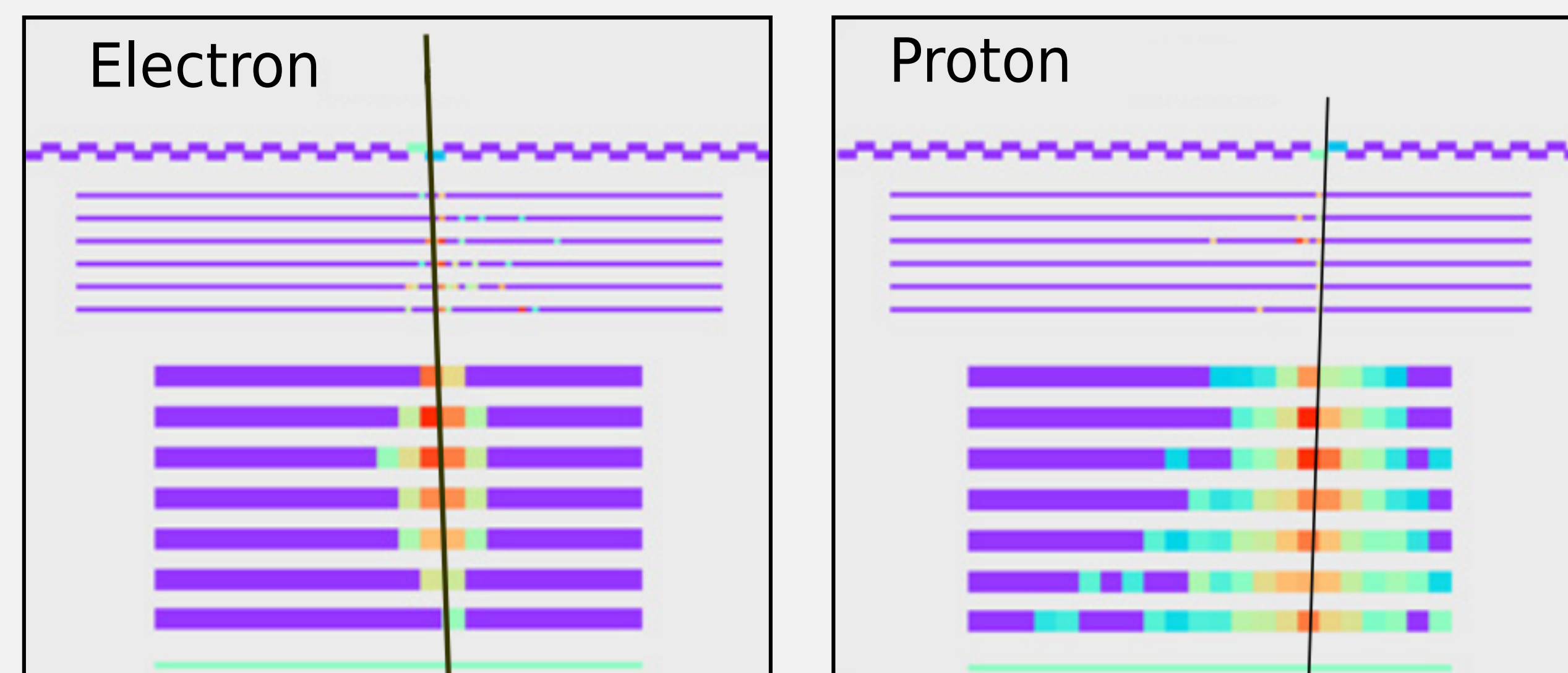
3b. MC validation

MC is scaled to the real data, to verify there are no biases and to confirm the reliability of the method.

2. Neural Network classifier

Dense layer 300 neurons
Dense layer 150 neurons
Dense layer 75 neurons
Output layer

Energy layer 3
Energy layer 14
RMS layer 3
RMS layer 14
Trajectory angle
Measured energy
ζ (classical PID)
STK cluster energy

Dropout 20%
Dropout 10%
Dropout 10%

An artificial neural network is trained on Monte Carlo data, taking as input 28 observables quantifying interaction topology and event characteristics.

In a regular neural network, the output is compressed to the [0;1] range by a sigmoid function. This results in a peak of false-positives in the signal region. Removing the output activation recovers a monotonic distribution, allowing e.g. interpolation methods for background estimation.
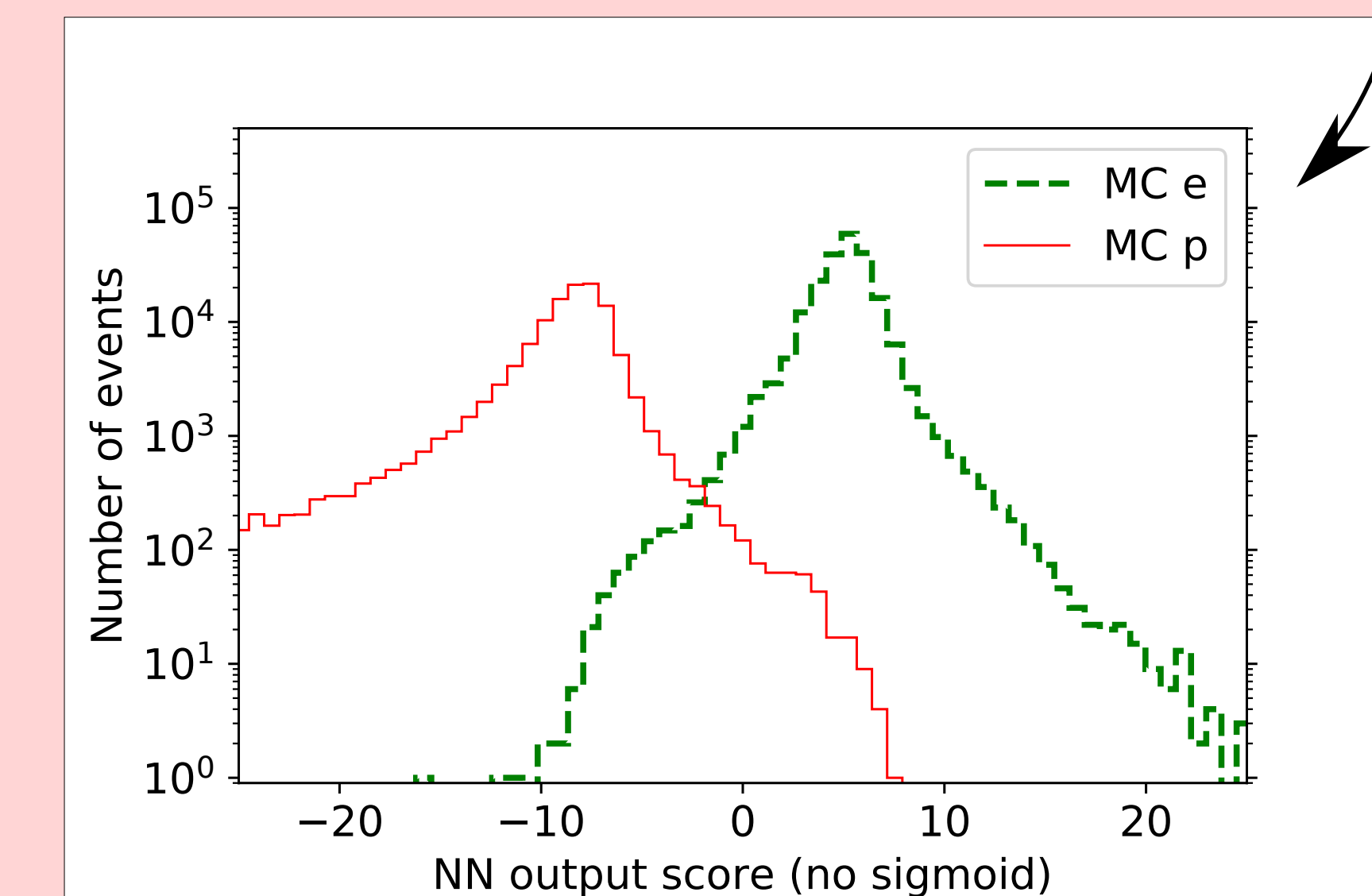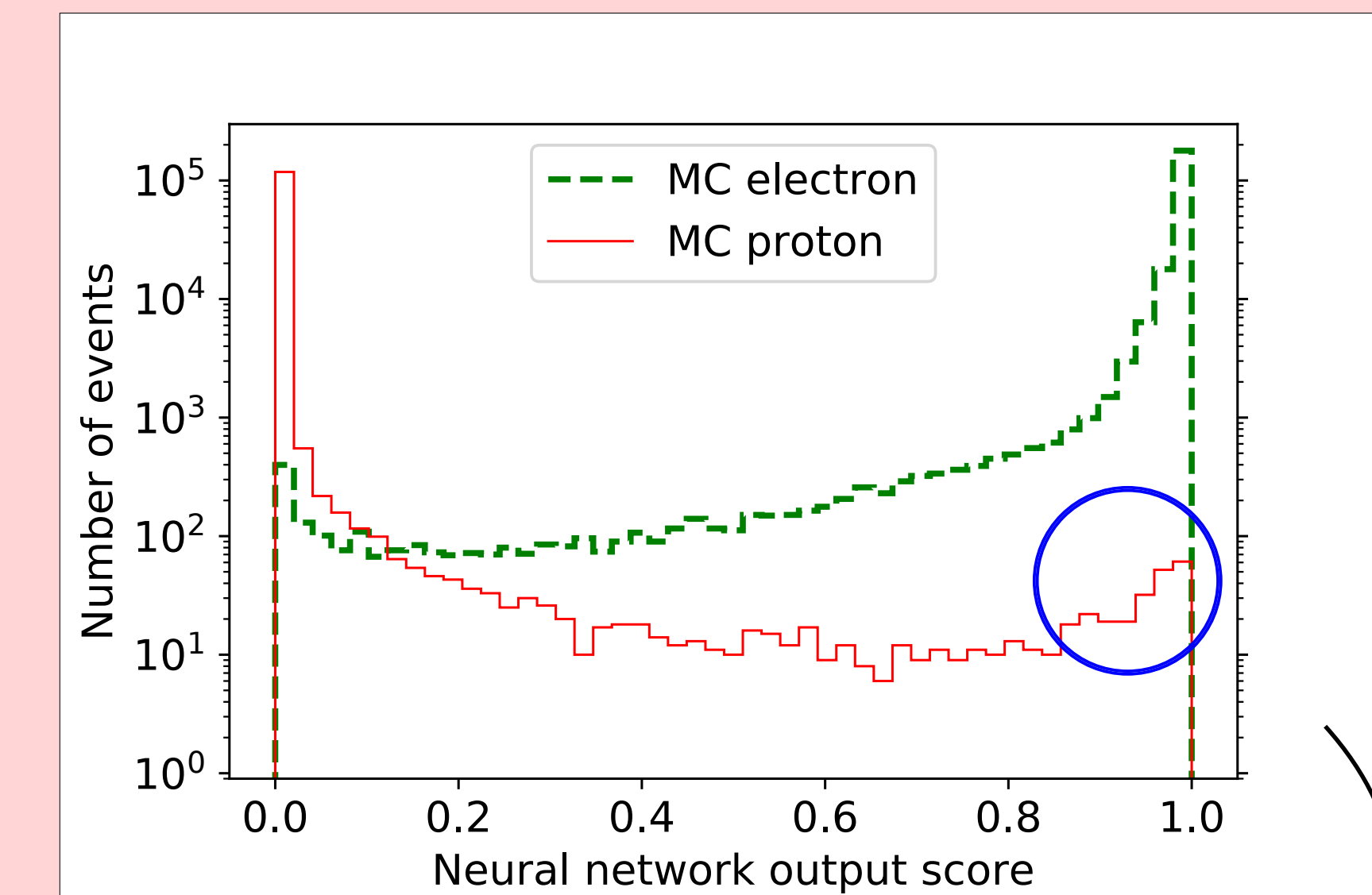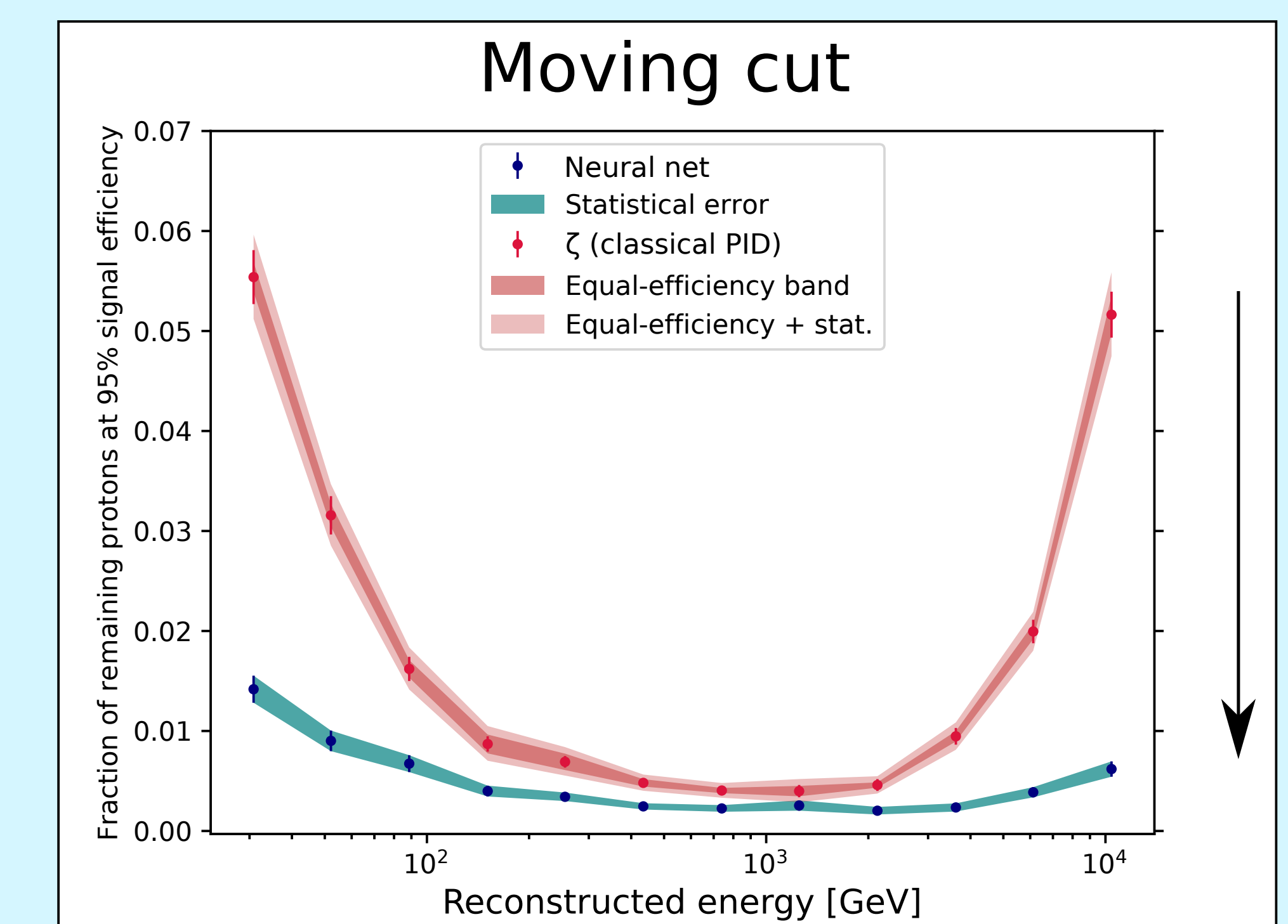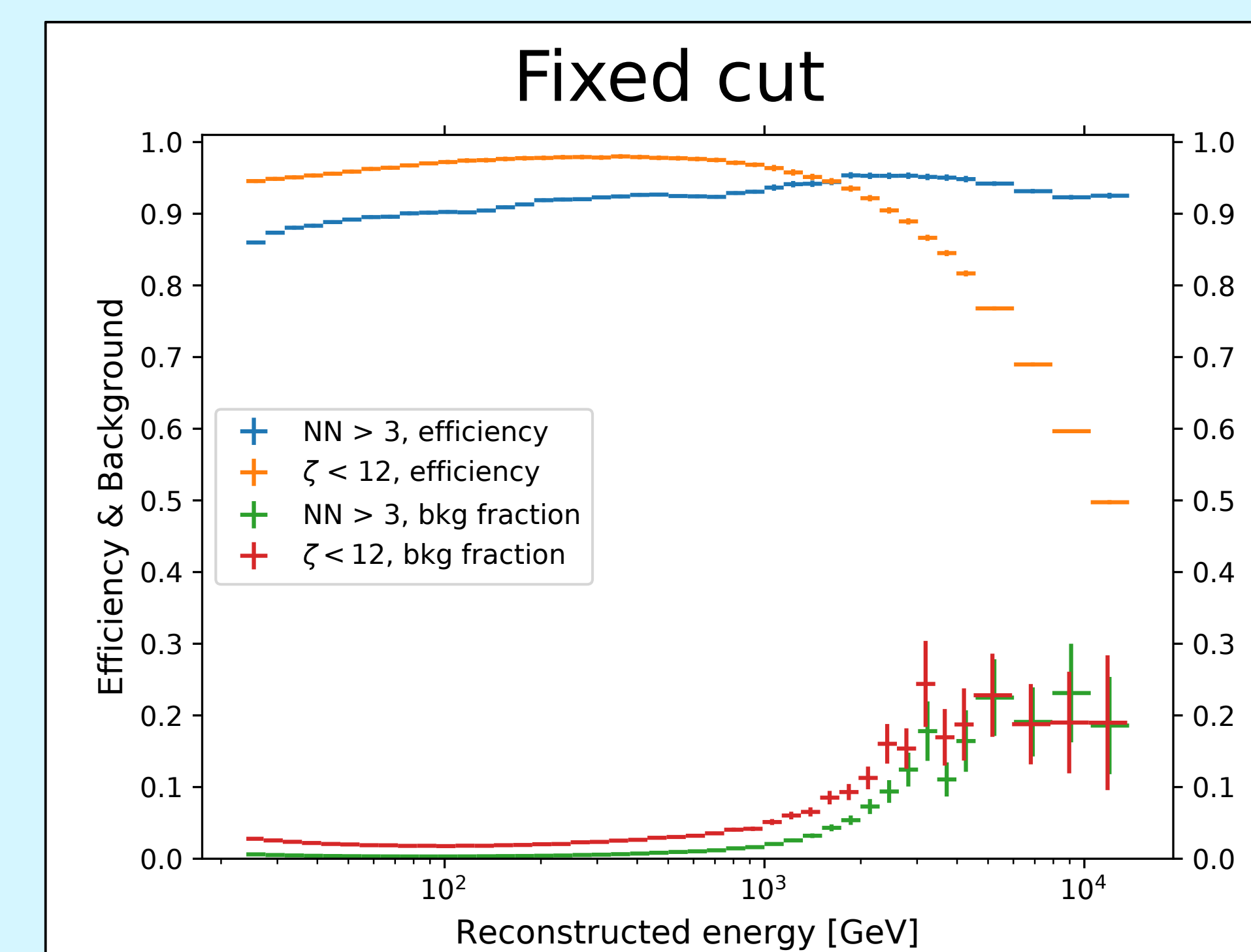
Removing Sigmoid

3a. Performances

The neural network classifier features a much flatter efficiency than the classical method for a fixed cut, yielding at 10 TeV the same contamination for twice the signal efficiency.

For a 1-to-1 comparison, a moving cut is set such that both methods have the same efficiency. With a 95% efficient cut, the background rejection of neural nets is up to 8x better.

Fixed cut

Moving cut

8x fewer false positives