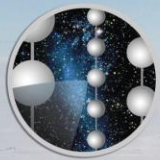


Online Event Reconstruction in IceCube Using Deep Learning Techniques

Mirco Huennefeld for the IceCube Collaboration*
mirco.huennefeld@tu-dortmund.de

Connecting The Dots 2018
Seattle – March 21th, 2018



ICECUBE

SOUTH POLE NEUTRINO OBSERVATORY



IceCube Laboratory

Data is collected here and sent by satellite to the data warehouse at UW-Madison



Digital Optical Module (DOM)

5,160 DOMs deployed in the ice

50 m

Ice Top

1450 m

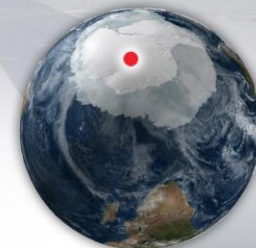
2450 m

IceCube detector

86 strings of DOMs,
set 125 meters apart

DeepCore

Antarctic bedrock



Amundsen-Scott South Pole Station, Antarctica

A National Science Foundation-managed research facility

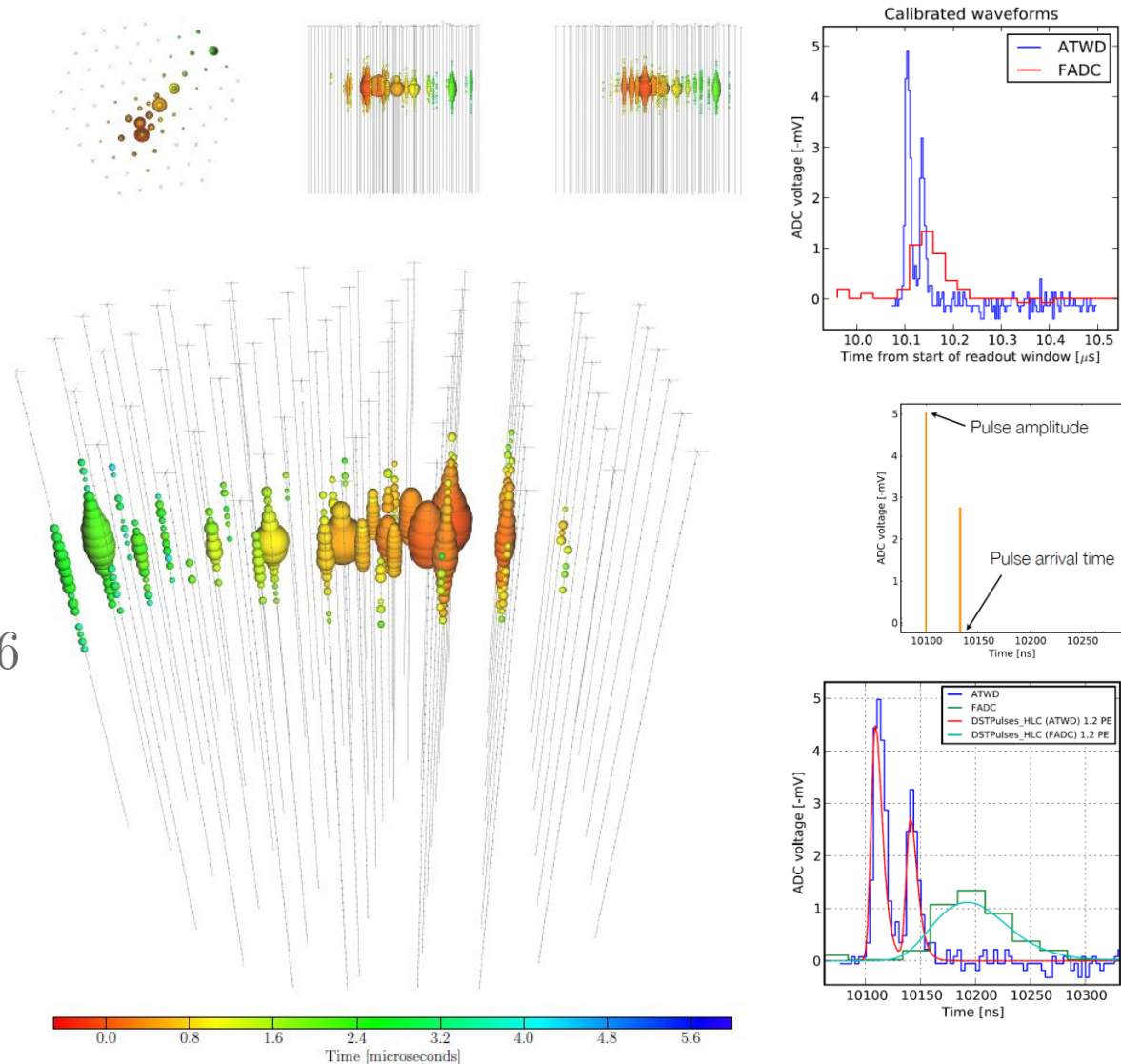
60 DOMs on each string

DOMs are 17 meters apart



IceCube – Data Format

- 5160 DOMs
- 128 bins per waveform
- Arbitrary number of waveforms per DOM
- $5160 \cdot 128 \cdot n = 660480 \cdot n$
 MNIST: $28 \cdot 28 = 784$
 CIFAR: $32 \cdot 32 = 1024$
 ImageNet: $256 \cdot 256 = 65536$ (cropped)
- Readout window dependent on trigger
- 4-dimensional, highly variable data



Deep Learning for Online Event Reconstruction

Online Analysis Pipeline:



- Used for real-time analyses and follow-up programs
- Main goal: discover neutrino sources
- Powerful reconstructions desired
- Challenges:
 - Limited resources
 - High data rate
 - Low level data

Can we further improve the online analysis pipeline
with modern techniques?

Presentation Outline

- Introduction & Motivation
- Neural Network Introduction
- Network Architecture
- Runtime
- Performance
- Loss Function Bias
- Uncertainty Estimation
- Summary & Outlook

Goal: Improvement of online reconstructions

Deep Learning – Deep Neural Networks

- Neural network defines a function:

$$f_{\theta} : I \rightarrow O$$

θ : Free parameters defined by model architecture

I: Input data

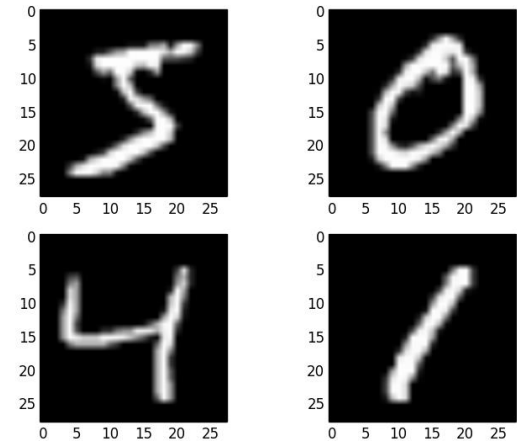
Greyscale values of pixels (image recognition)

Pulse information of DOMs (IceCube)

O: Output

Digit (image recognition)

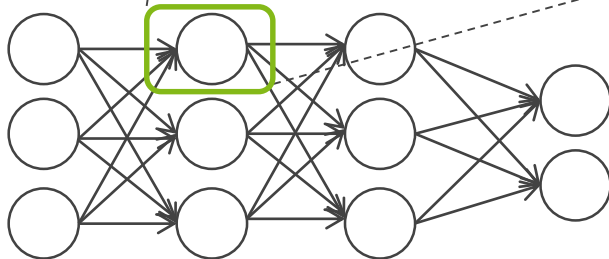
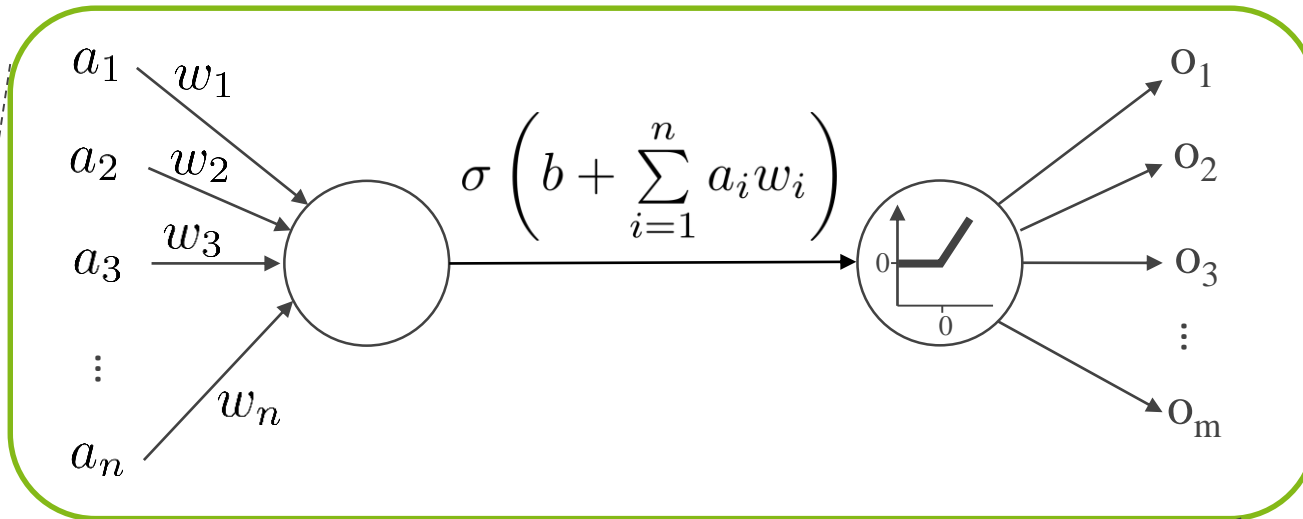
Energy/direction of particle (IceCube)



MNIST Dataset

Deep Learning – Deep Neural Networks

Artificial Neuron and fully connected layer



Weights and bias

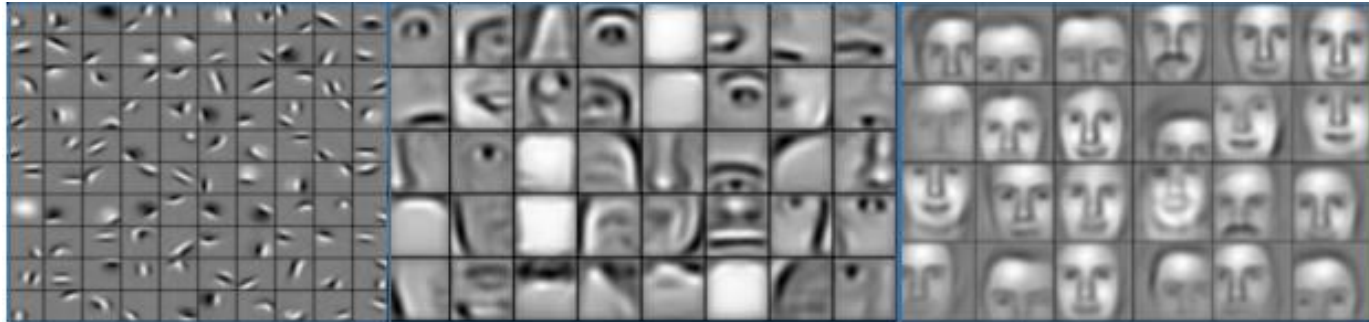
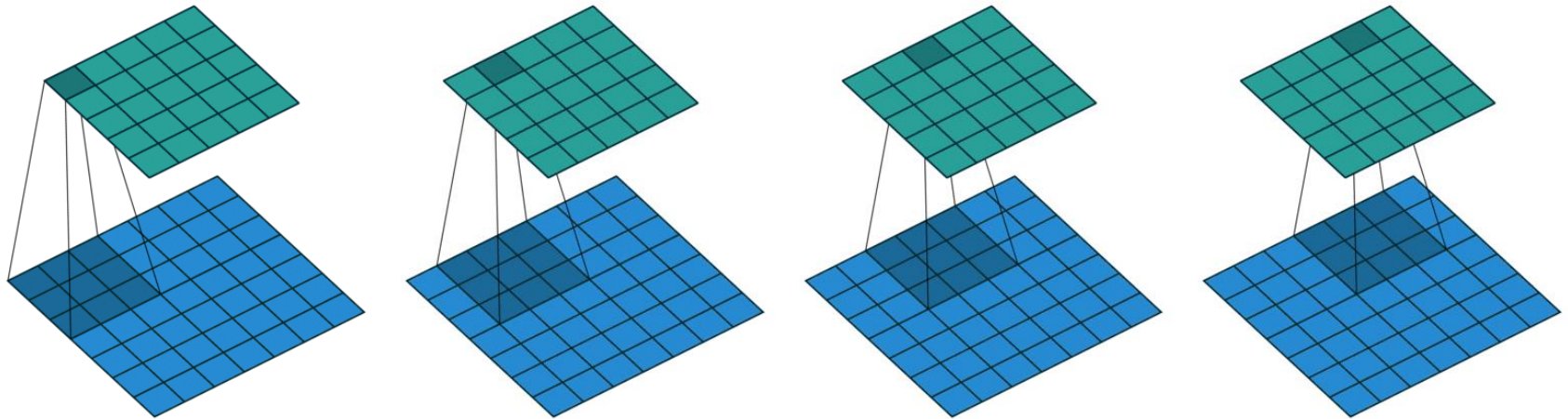
$n + 1$ free parameters per neuron

Nonlinear activation function e.g. ReLU

0 up to a fixed threshold

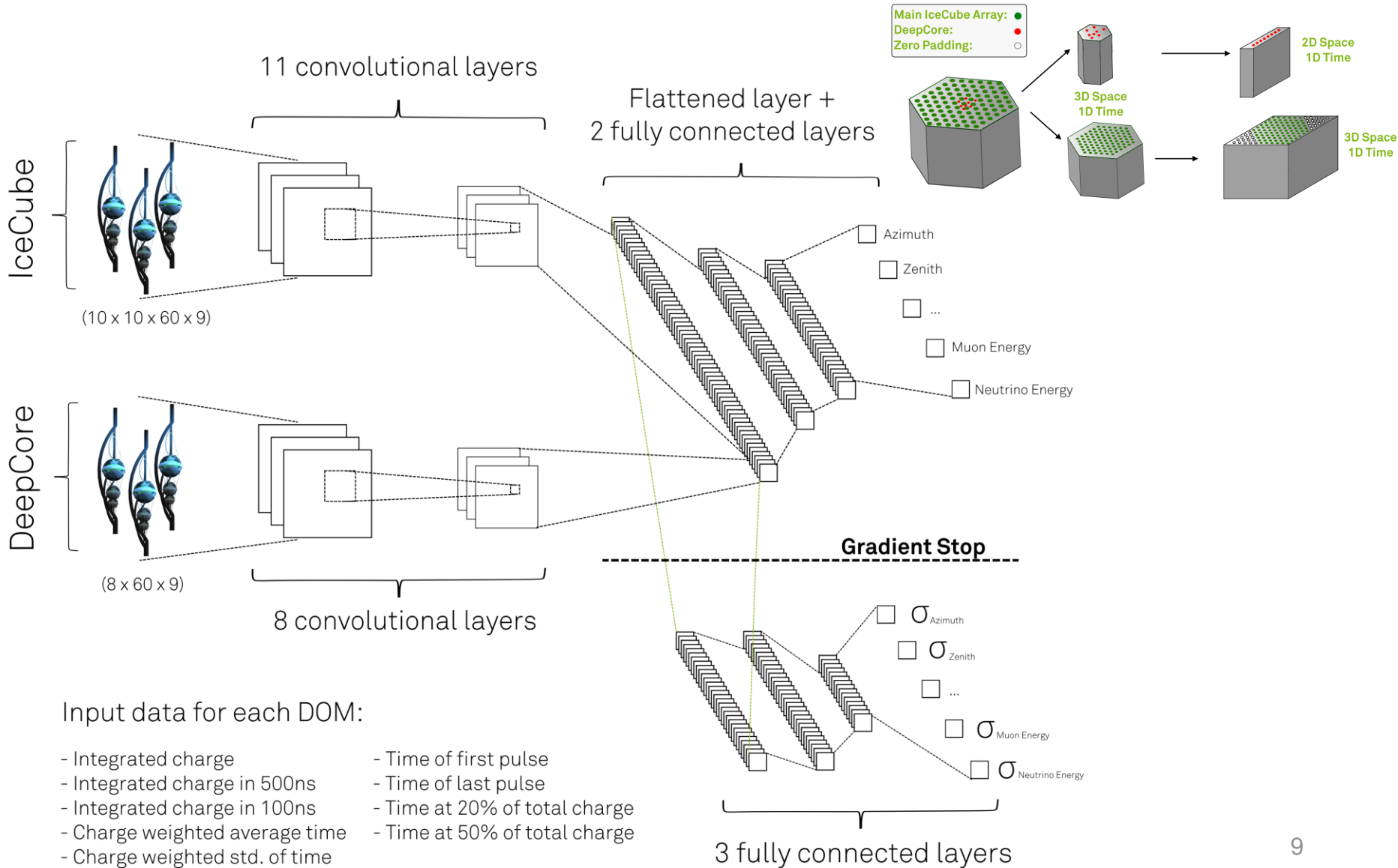
Deep Learning – Convolutional Neural Nets

Convolutional Layer



- Only neighboring neurons are connected
- Kernel weights are shared
- Greatly reduces number of free parameters

Network Architecture



Input data for each DOM:

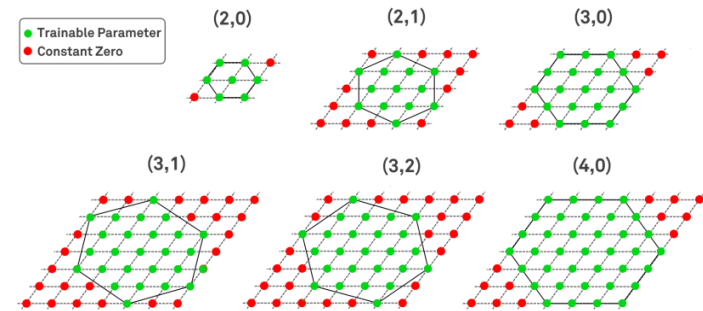
- Integrated charge
- Integrated charge in 500ns
- Integrated charge in 100ns
- Charge weighted average time
- Charge weighted std. of time
- Time of first pulse
- Time of last pulse
- Time at 20% of total charge
- Time at 50% of total charge

Network Architecture

- Residual additions:
$$\text{output} = \text{input} + \underbrace{f(\text{input})}_{\text{residual}}$$

DOI: 10.1109/CVPR.2016.90

- Hexagonally shaped convolution kernels
- Normalization of input and labels to mean 0 and variance 1
- Variance control in layers



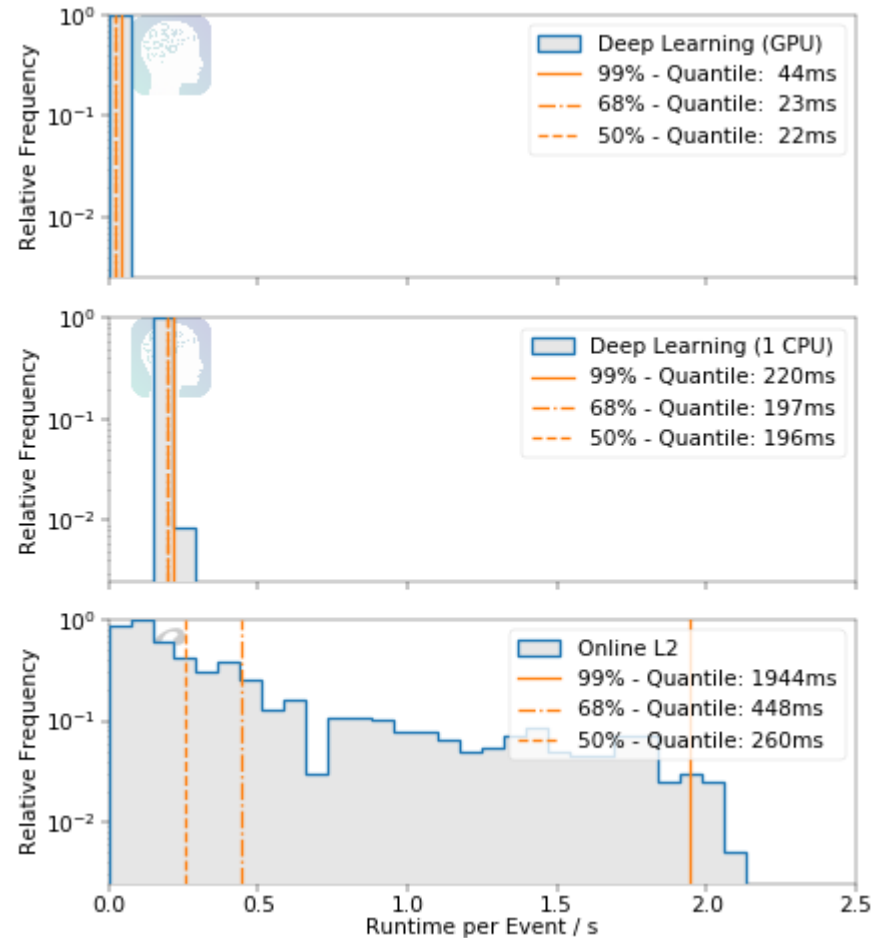
- Assuming input into a layer is normalized:
 - Ensure that output is normalized as well
- At initialization: random output is as good as predicting based on the label distribution
- Multilabel loss function
 - Adaptive factors for each label according to predefined importance
 - Ensures that labels are learnt at same speed

Runtime

- Fast and stable runtime:
 - Runtime is independent of event
 - Set number of mathematical operations are performed
 - No extreme outliers

- Slight dependence on event added due to preprocessing of input data

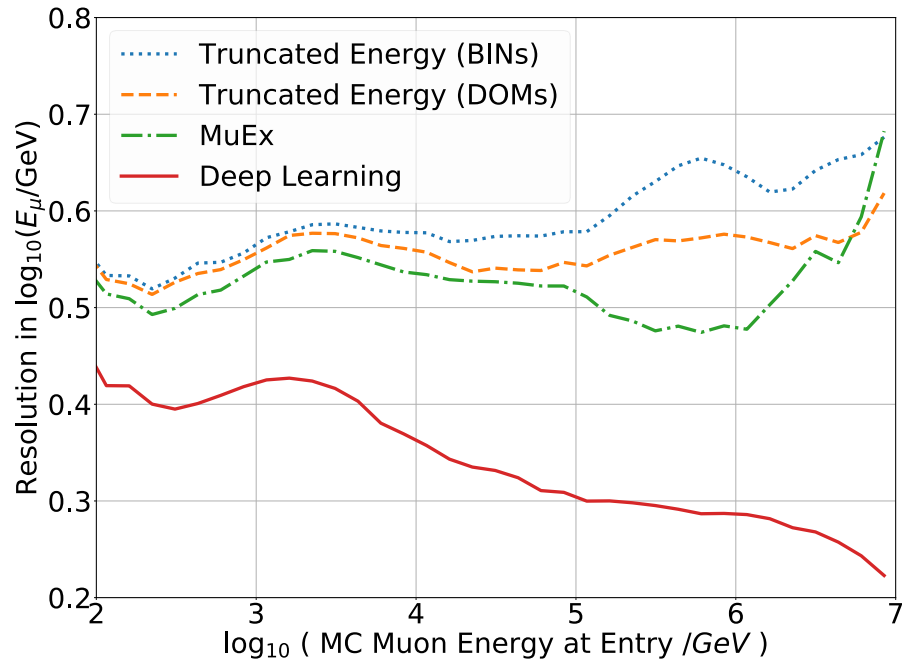
- Can be added to online filter without noticeably increasing runtime



Performance

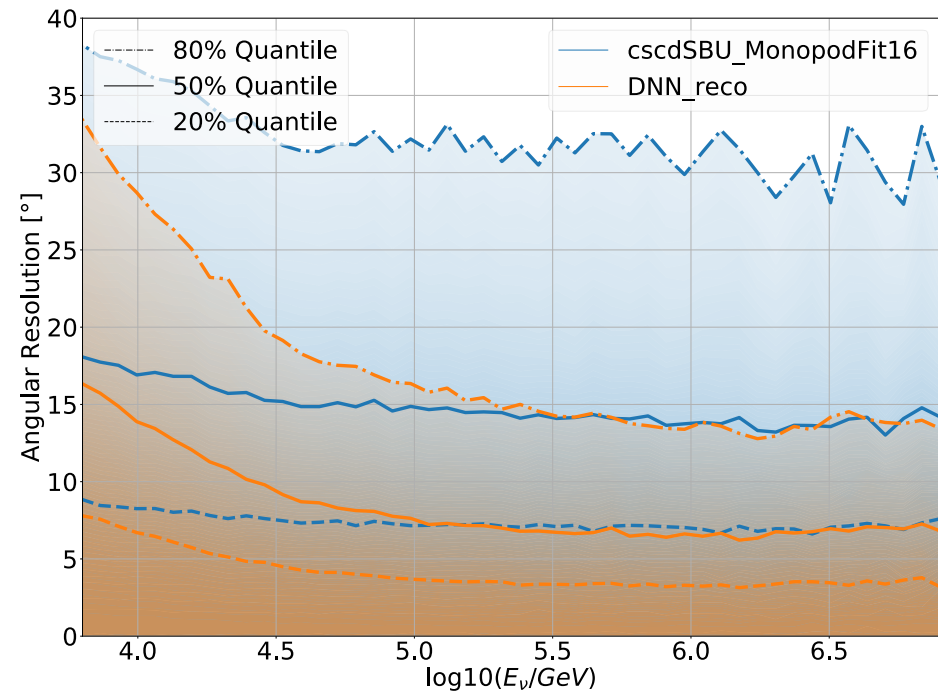
Track-like Events

Muon Energy Resolution



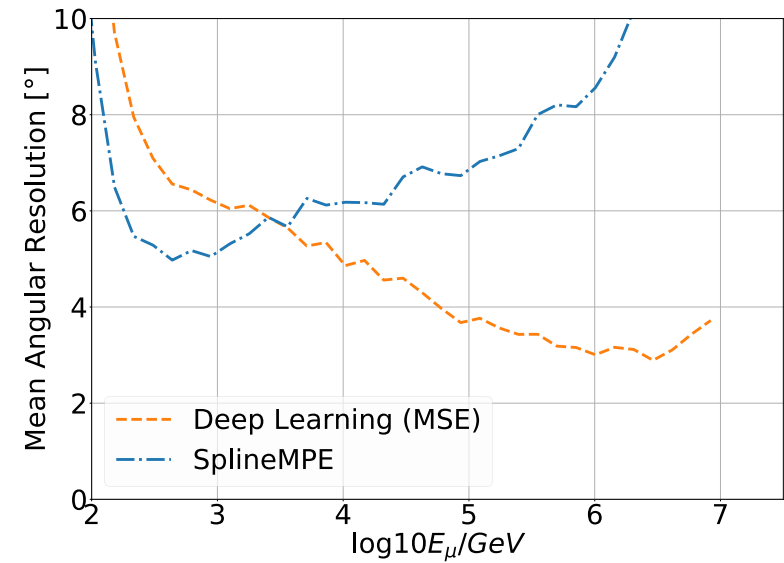
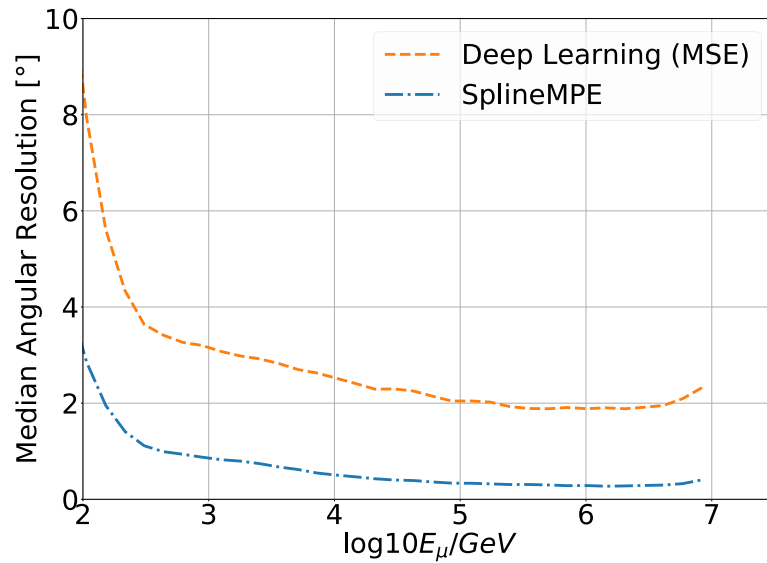
Cascade-like Events

Energy dependent Angular Resolution



Loss Function Bias

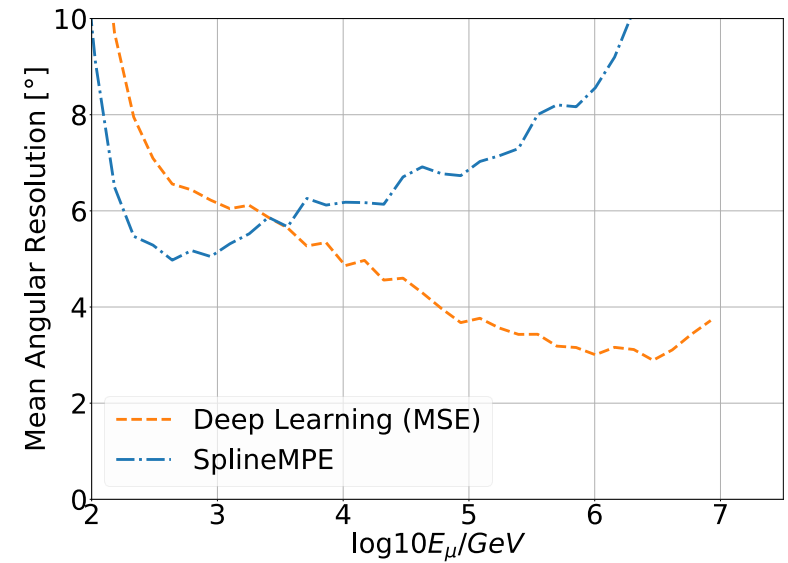
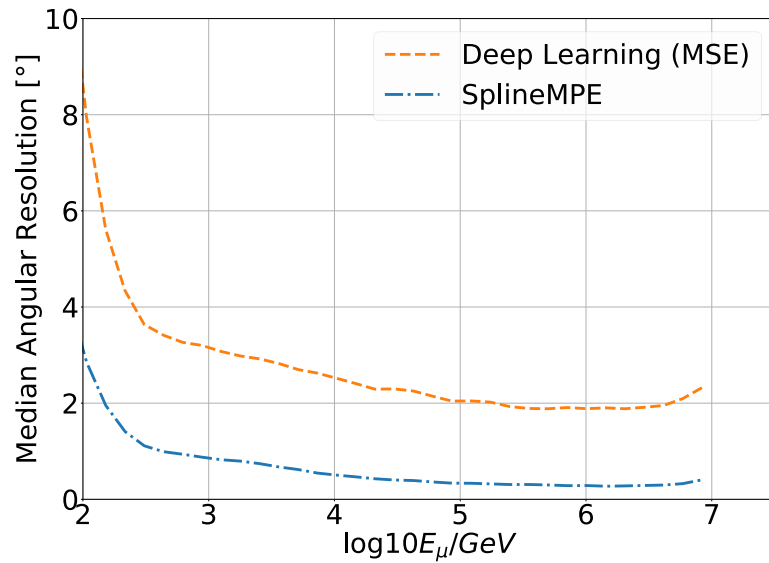
Angular Resolution of Track-like Events



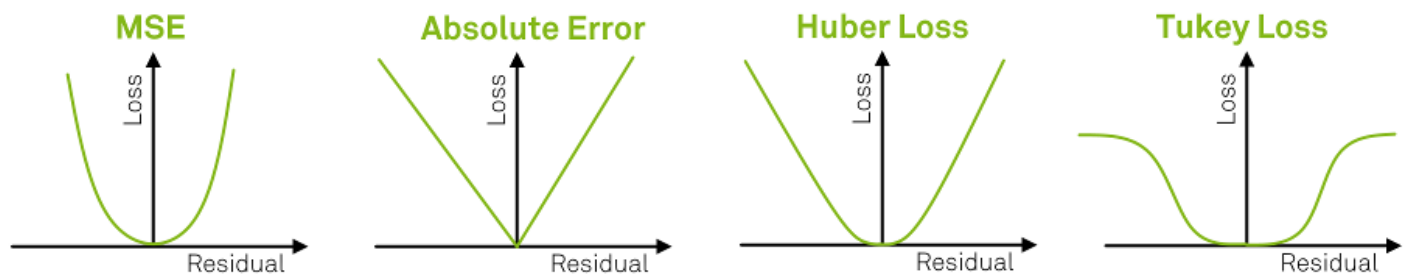
- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution

Loss Function Bias

Angular Resolution of Track-like Events

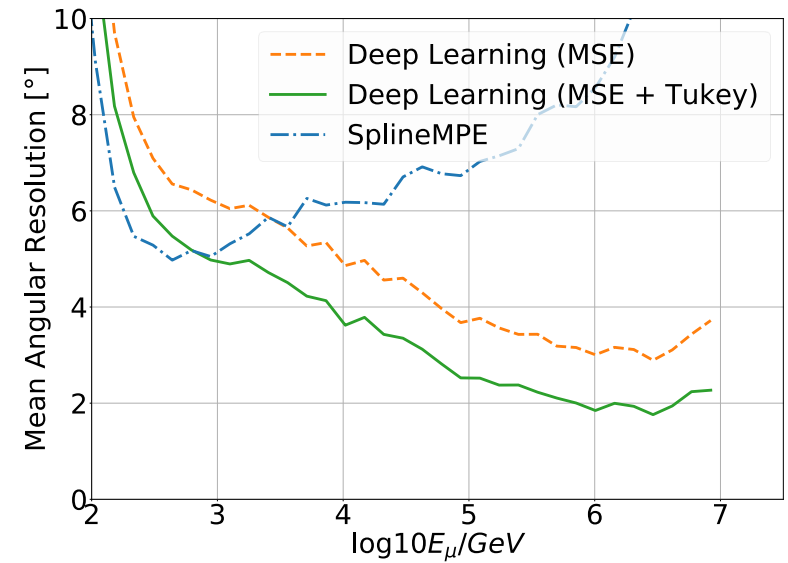
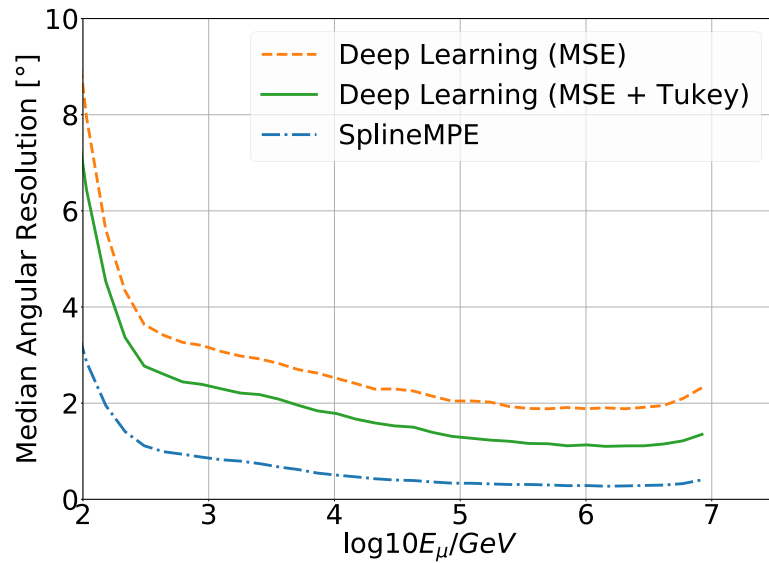


- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution

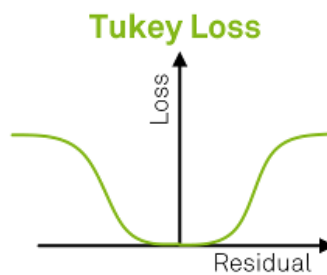
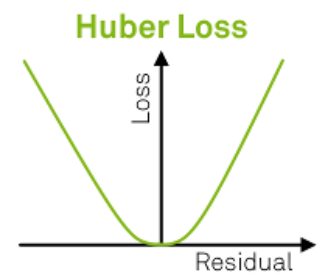
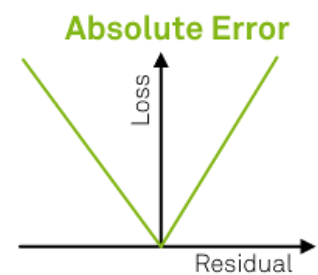
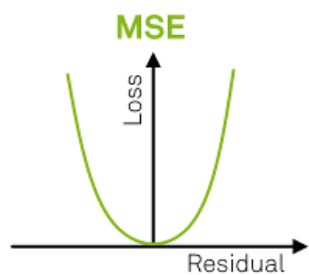


Loss Function Bias

Angular Resolution of Track-like Events

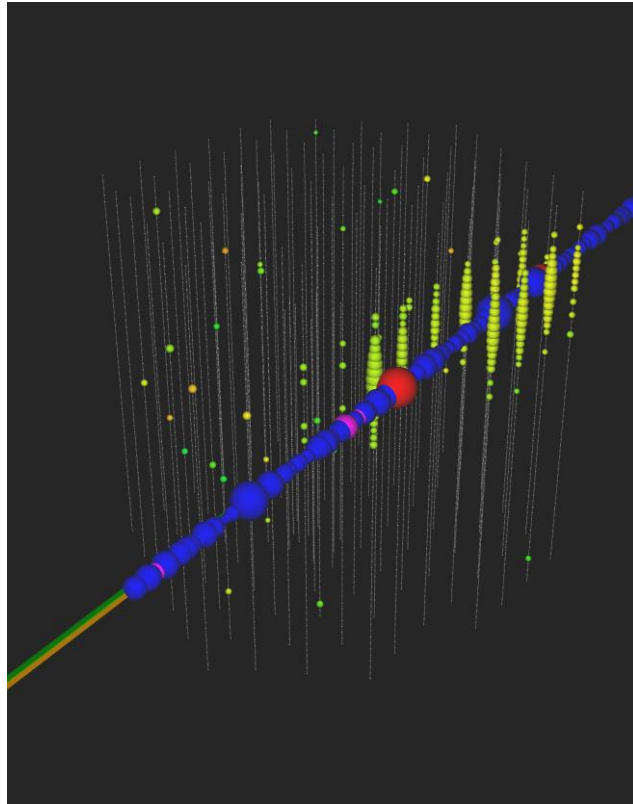


- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution

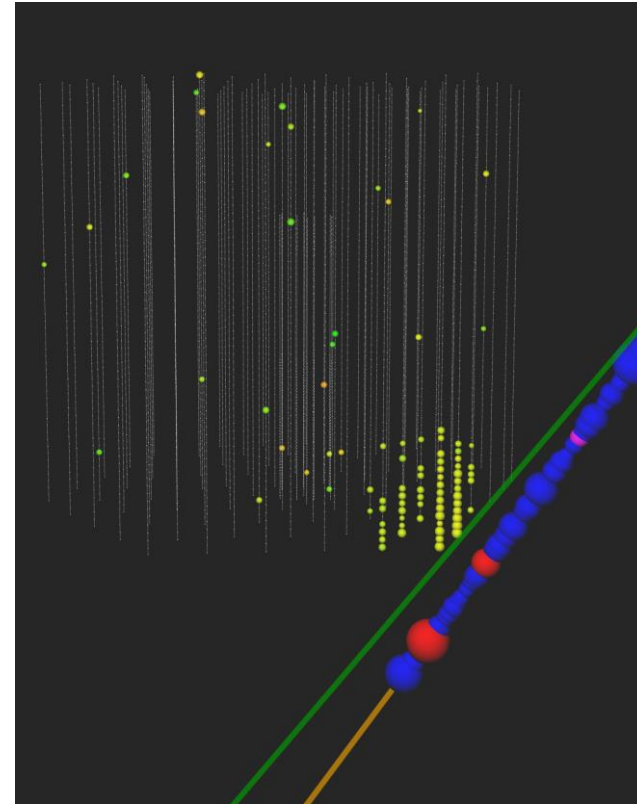


Loss Function Bias

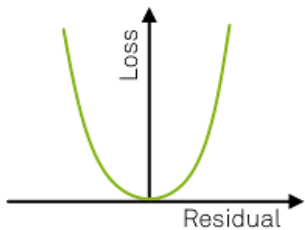
Angular Error: 0.5°



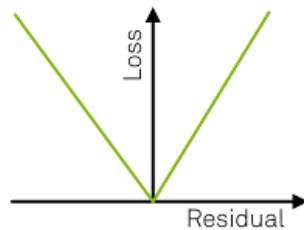
Angular Error: 15°



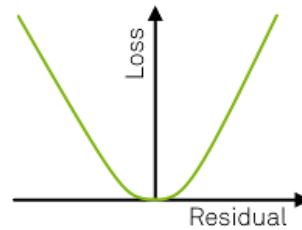
MSE



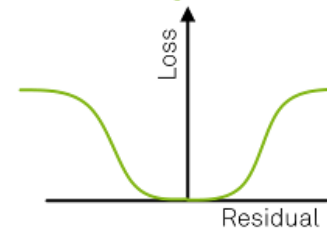
Absolute Error



Huber Loss

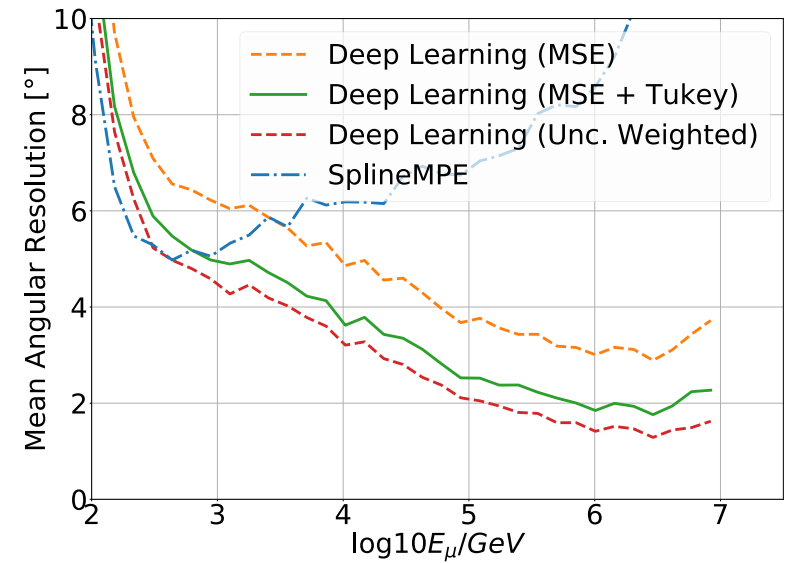
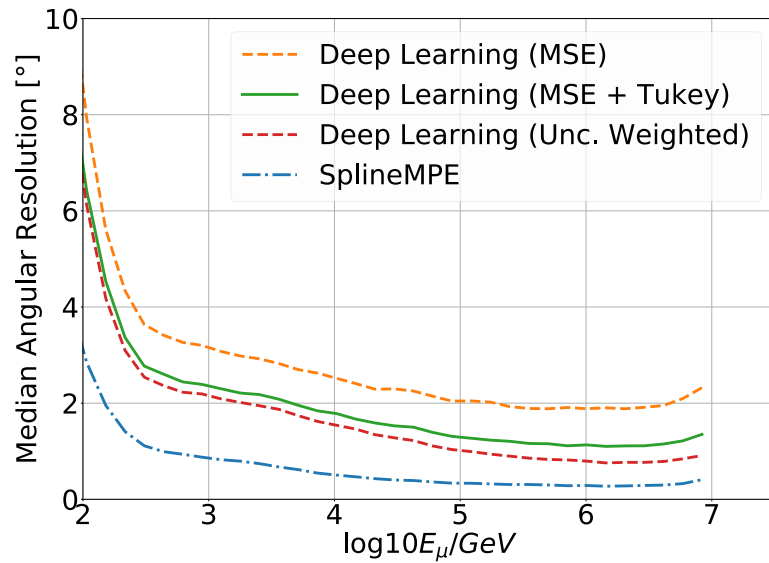


Tukey Loss

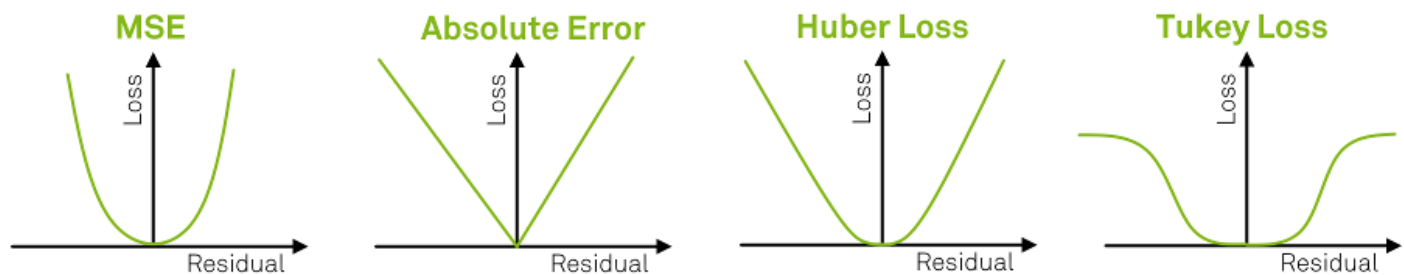


Loss Function Bias

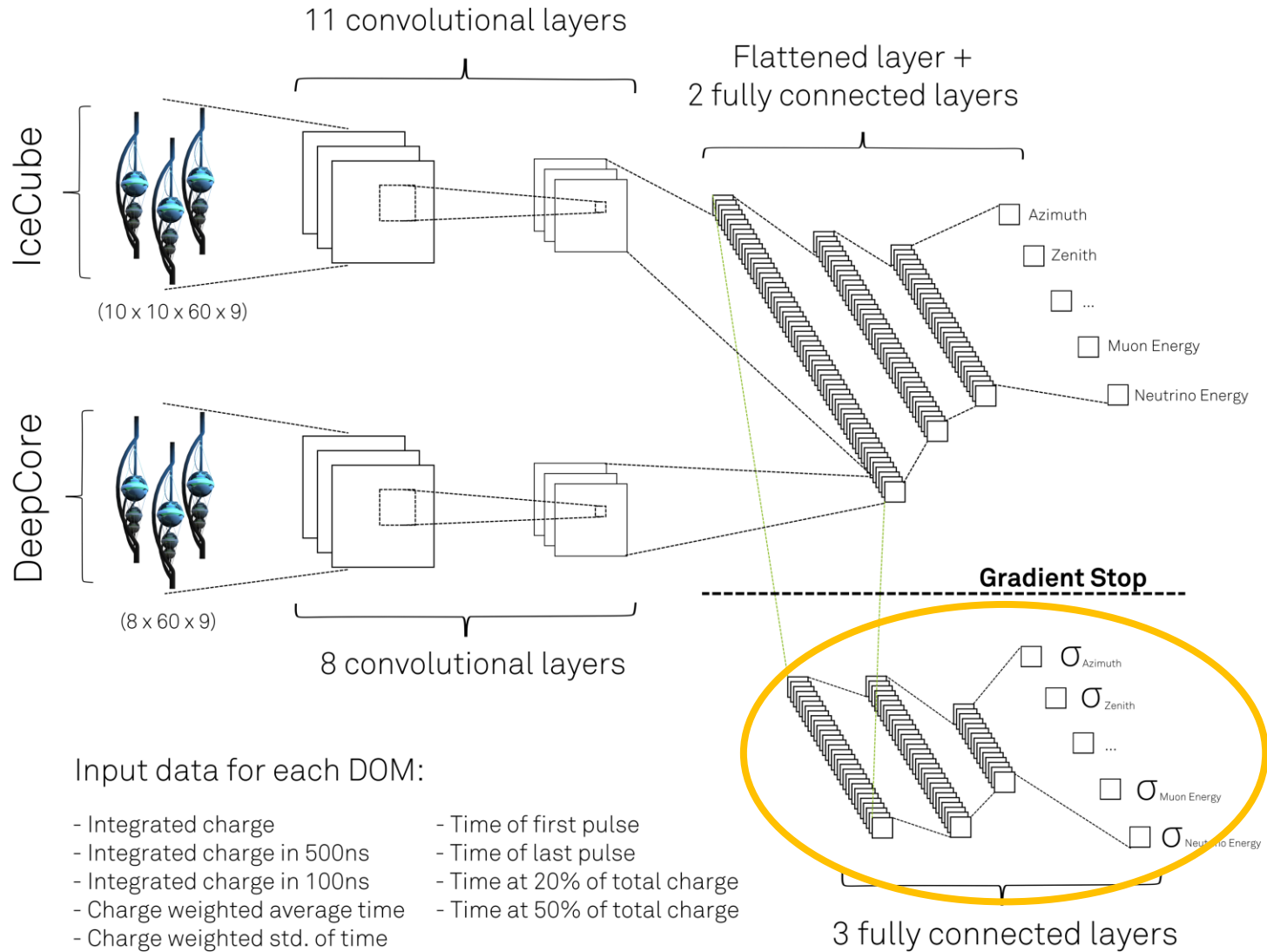
Angular Resolution of Track-like Events



- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution



Uncertainty Estimation



Input data for each DOM:

- Integrated charge
- Integrated charge in 500ns
- Integrated charge in 100ns
- Charge weighted average time
- Charge weighted std. of time
- Time of first pulse
- Time of last pulse
- Time at 20% of total charge
- Time at 50% of total charge

Uncertainty Estimation

- Mean squared error (MSE) of absolute residuals:

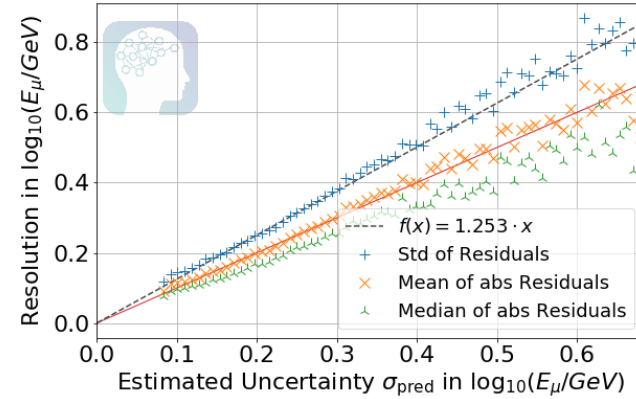
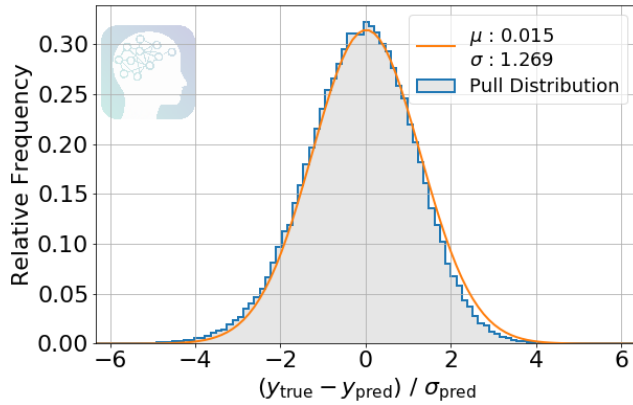
$$\text{loss} = \frac{1}{n} \sum_{i=1}^n (\sigma_{\text{pred}} - |y_{\text{true}} - y_{\text{pred}}|)^2$$

- Network is trained to correctly estimate absolute residual on average: $\sigma_{\text{pred}} \approx \langle y_{\text{abs}} \rangle$
 - If residuals are Gaussian-distributed, the mean absolute residual can be converted to the standard deviation: $\sigma = \sqrt{\frac{\pi}{2}} \cdot \langle y_{\text{abs}} \rangle \approx 1.253 \cdot \langle y_{\text{abs}} \rangle$
- Gaussian Likelihood Minimization:

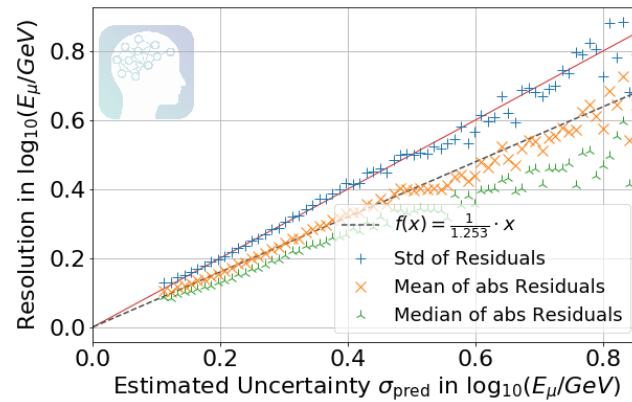
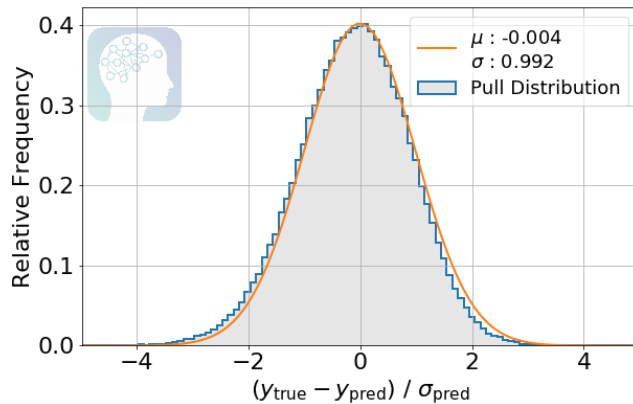
$$\text{loss} = \frac{1}{n} \sum_{i=1}^n \left(\ln \left(\sigma_{\text{pred}_i}^2 \right) + \frac{(y_{\text{pred}_i} - y_{\text{true}_i})^2}{\sigma_{\text{pred}_i}^2} \right)$$

Uncertainty Estimation

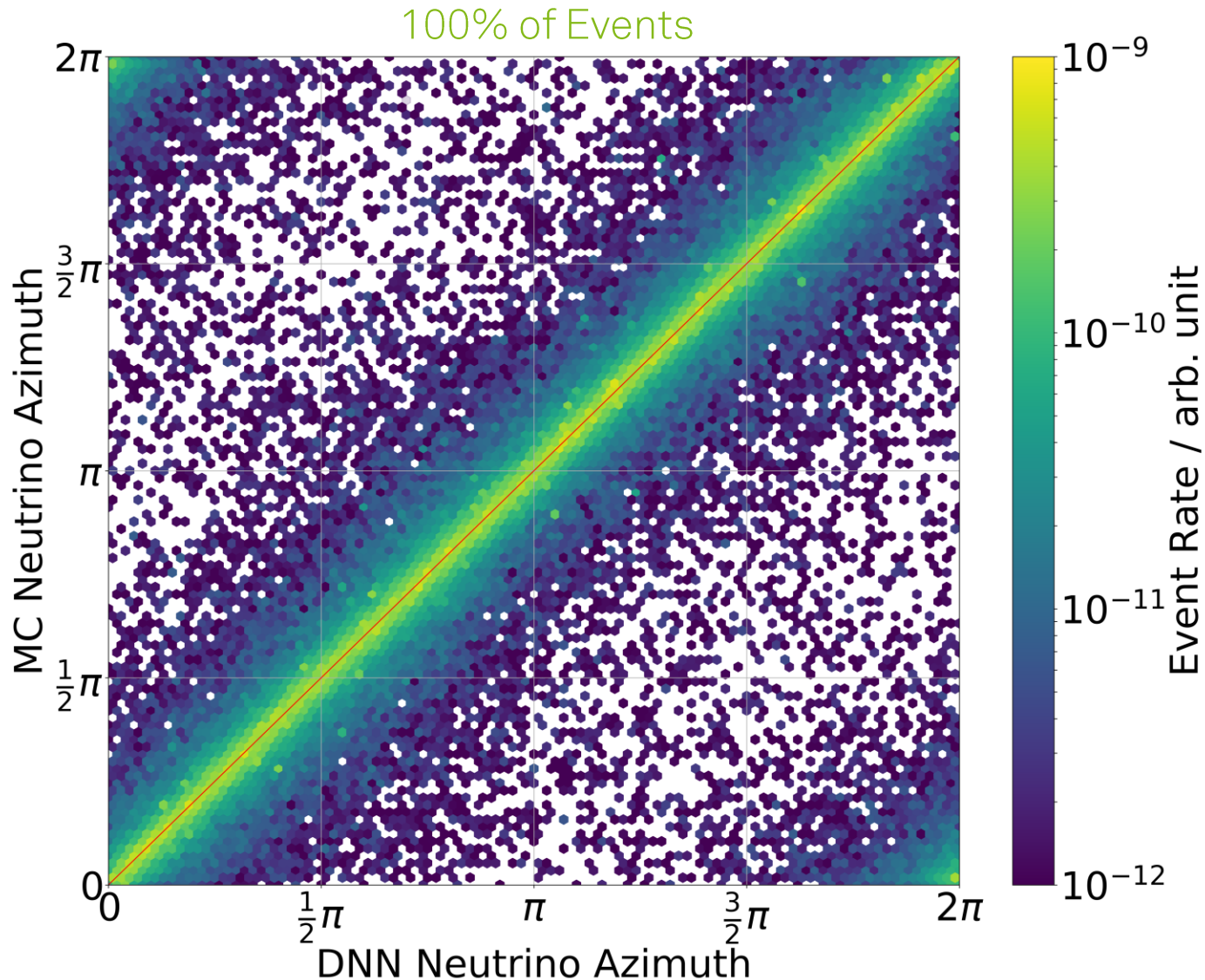
MSE of absolute Residuals



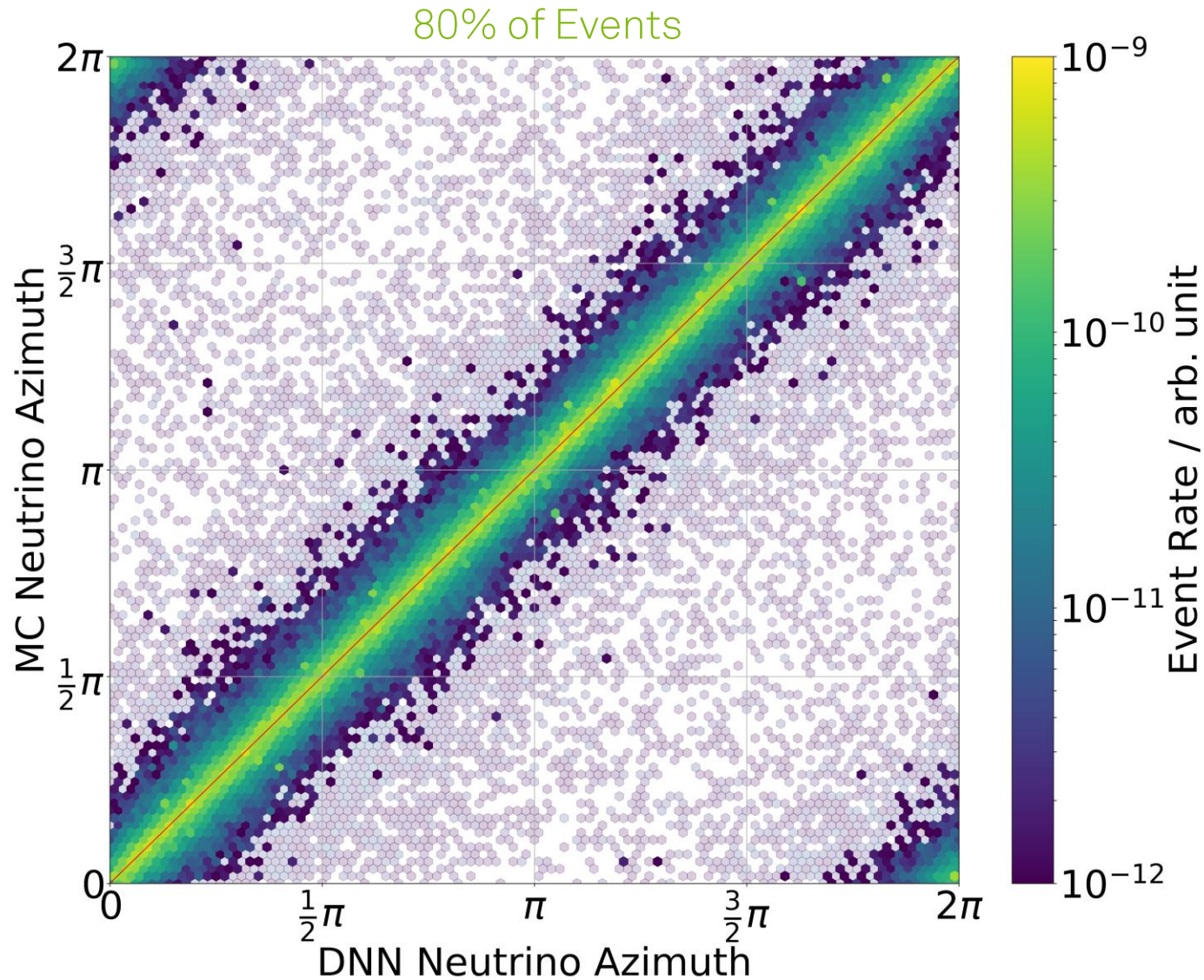
Gaussian Likelihood



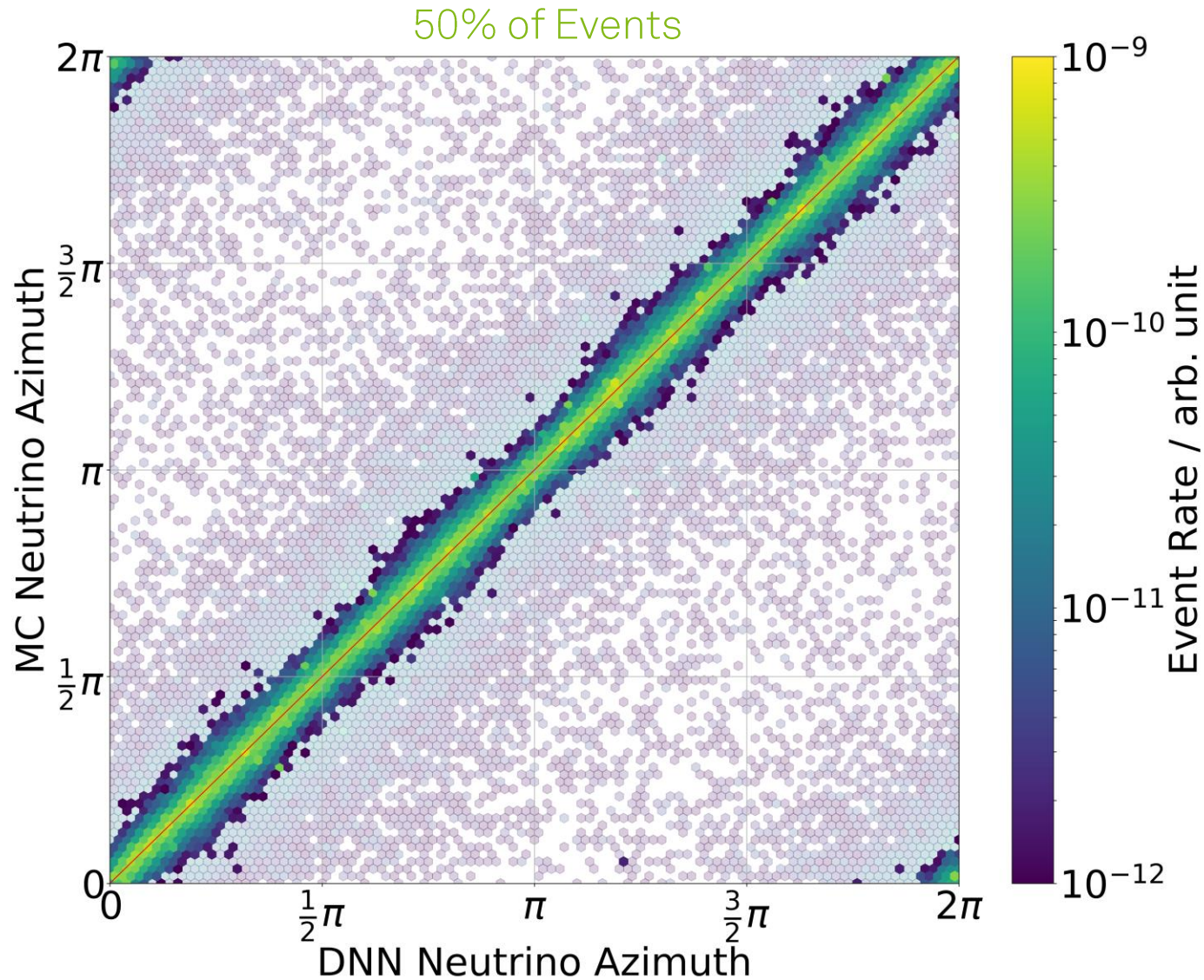
Uncertainty Estimation



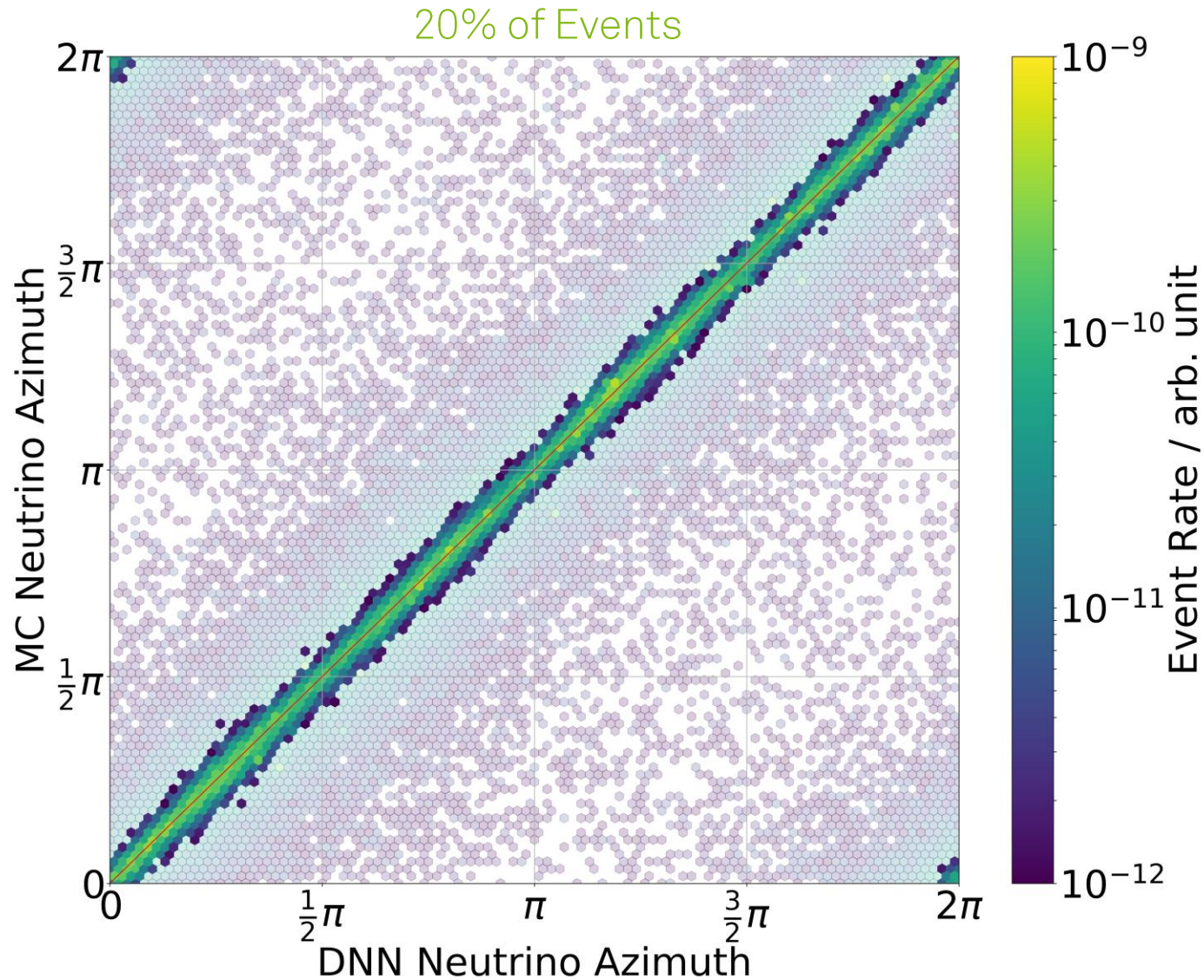
Uncertainty Estimation



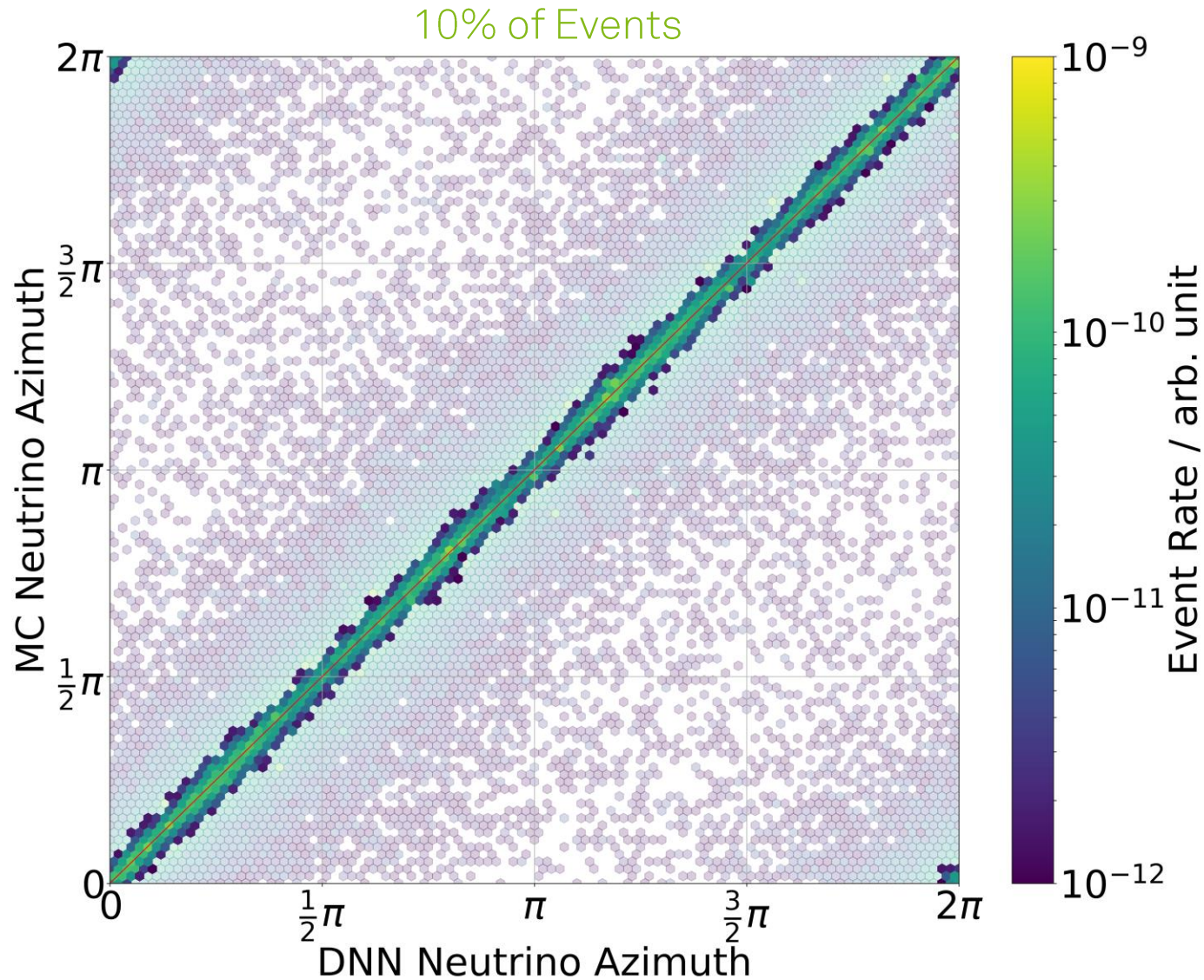
Uncertainty Estimation



Uncertainty Estimation



Uncertainty Estimation



Conclusions

Can we further improve the online analysis pipeline
with modern techniques?

Conclusions

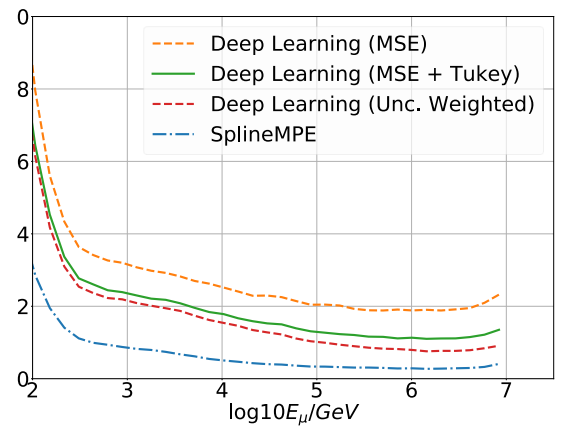
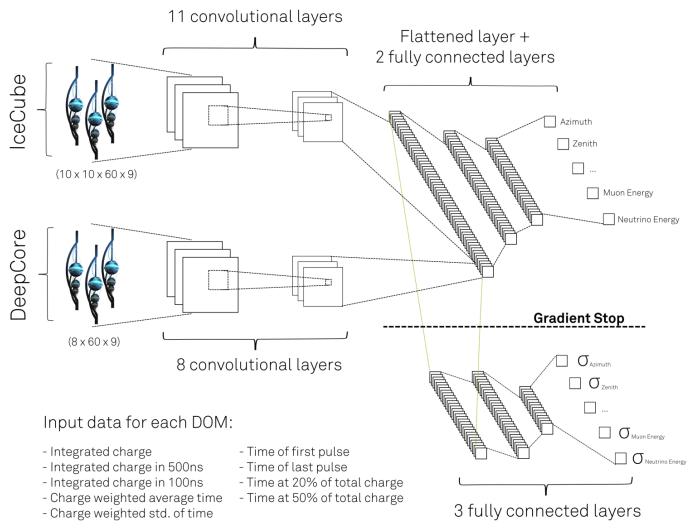
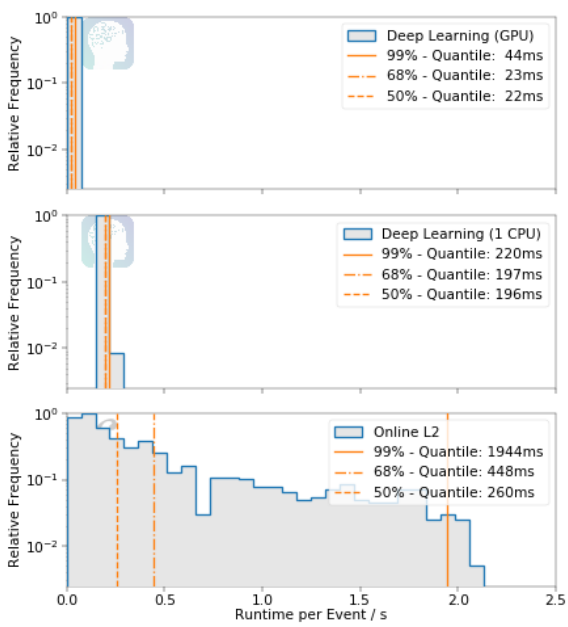
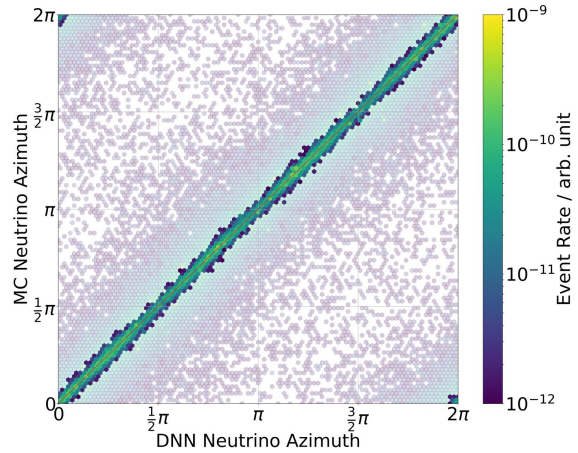
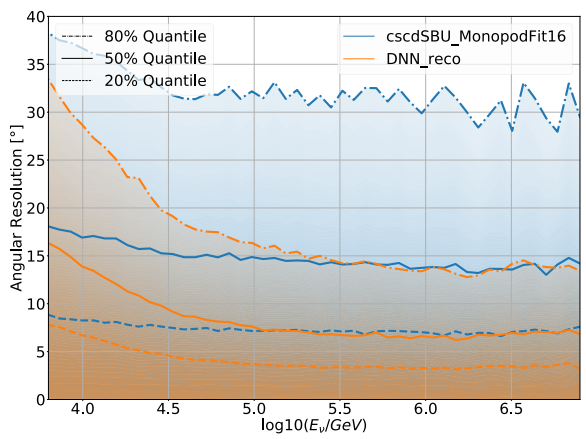
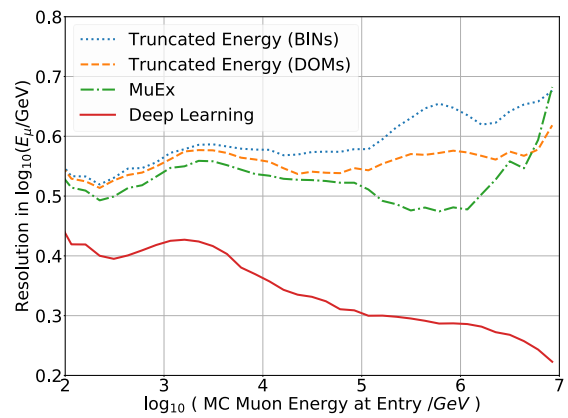
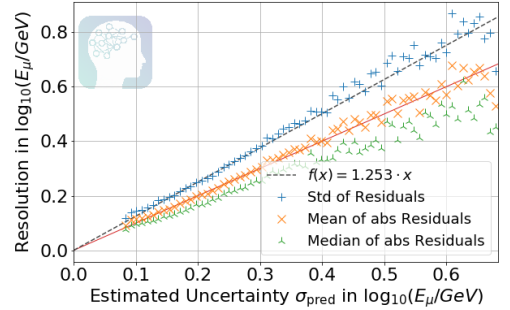
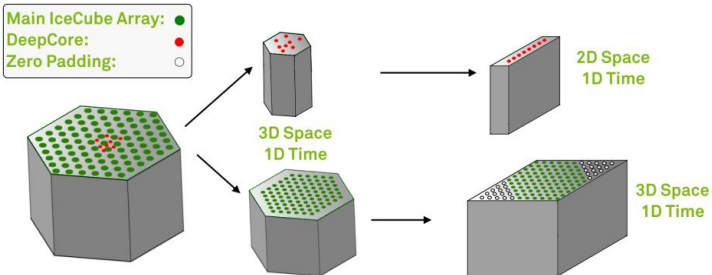
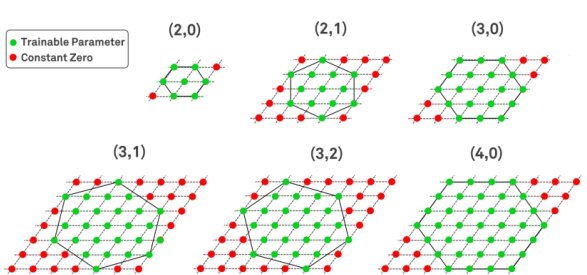
Can we further improve the online analysis pipeline with modern techniques? **Yes!**

- Fast and stable runtime
- Drastic improvement of energy and directional resolution
- Reliable uncertainty estimate

Conclusions

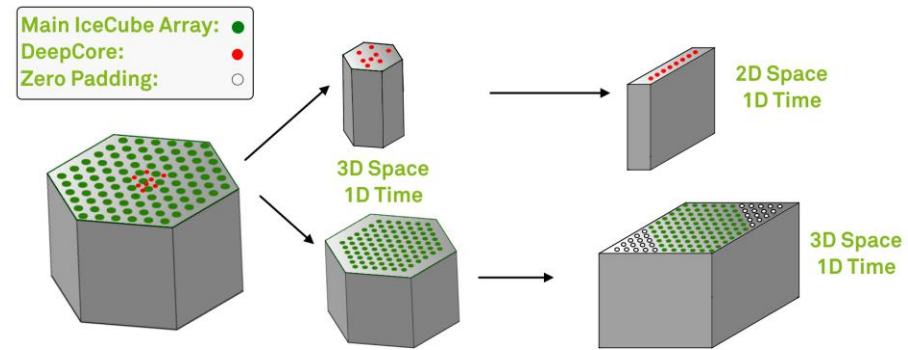
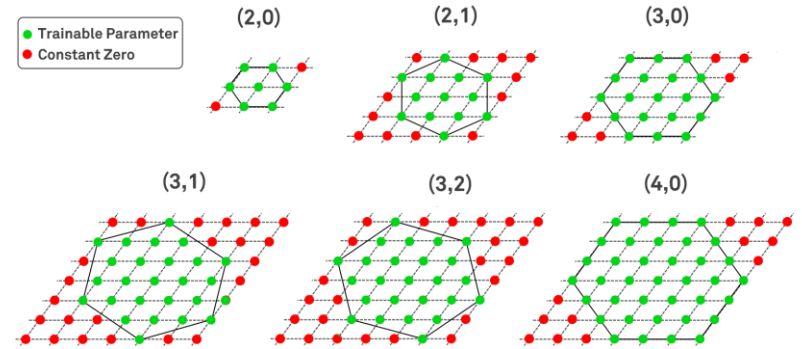
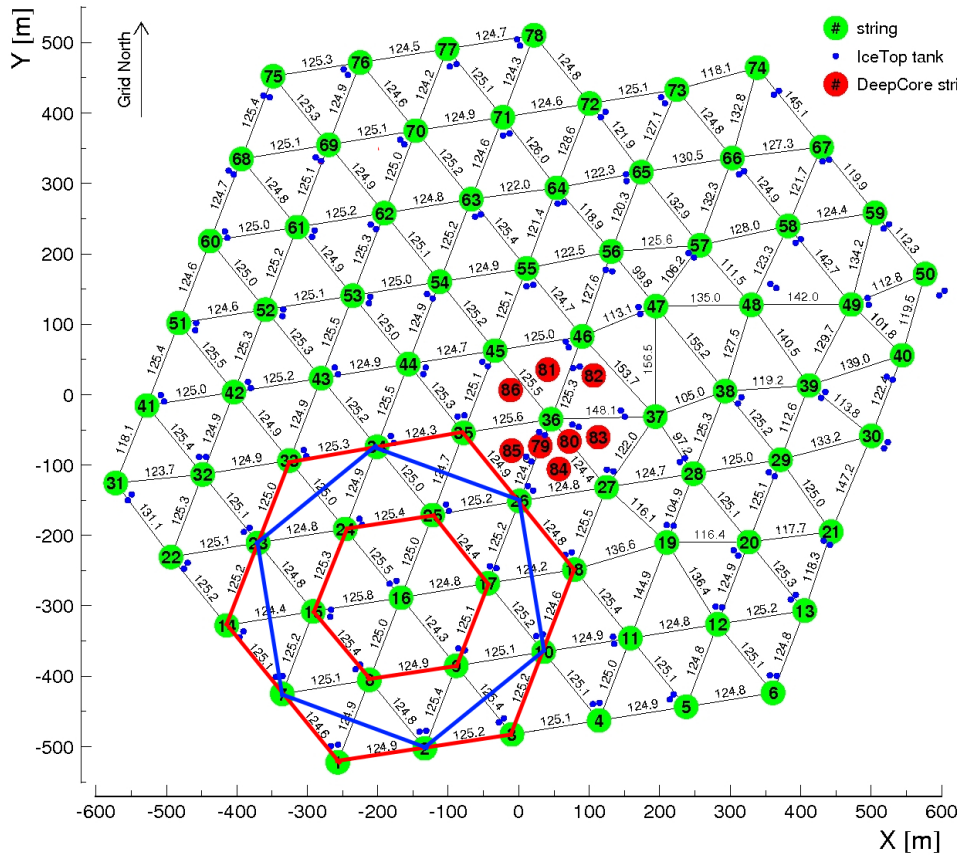
Can we further improve the online analysis pipeline with modern techniques? **Yes!**

- Fast and stable runtime
- Drastic improvement of energy and directional resolution
- Reliable uncertainty estimate
- More research and new methods/ architectures are needed:
 - Inclusion of symmetries, invariances, and prior knowledge into the network architecture
 - Combining strengths of likelihood-based reconstructions and DL
 - Limitations of network architecture
 - Numerical limits and regression specific problems

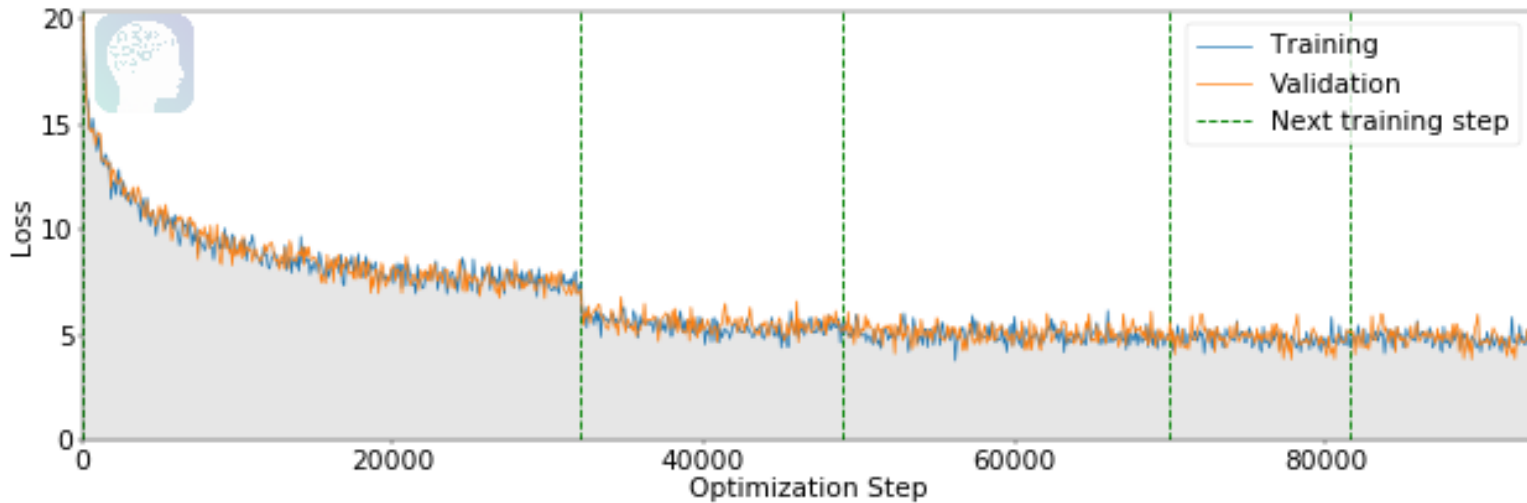


Appendix

Network Architecture



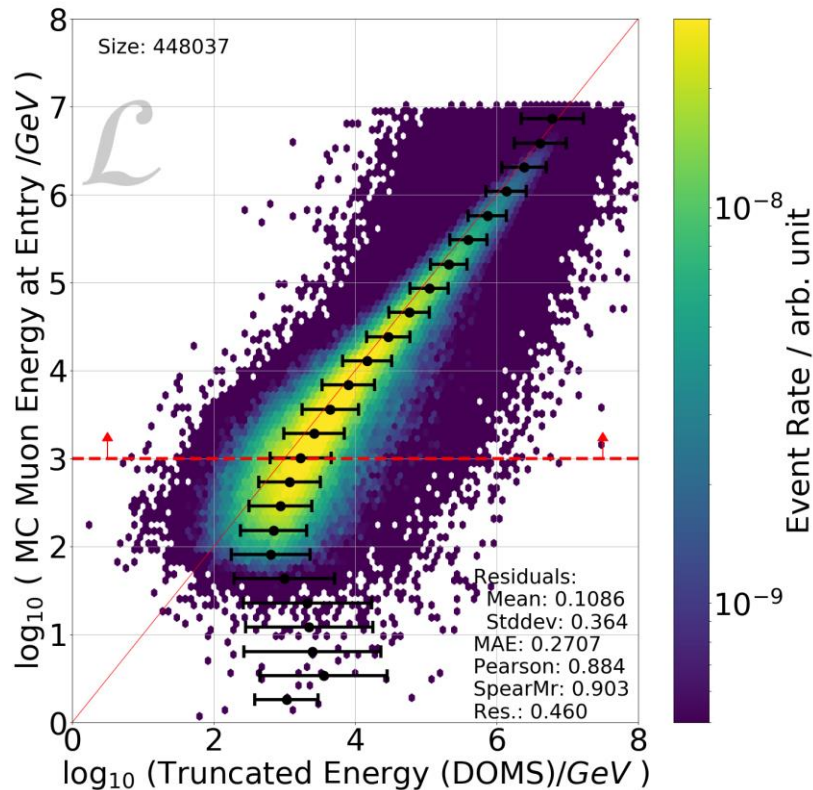
Deep Learning – Training



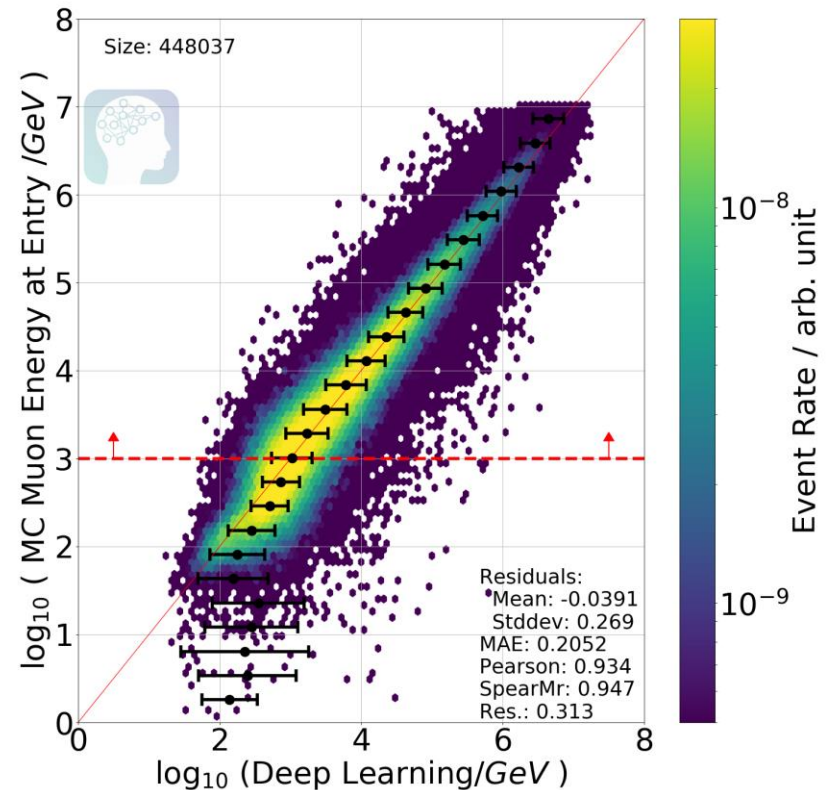
Energy Reconstruction – Muon Energy at Entry

OnlineL2 Muon Filter – CC events

Truncated Energy (DOMS)



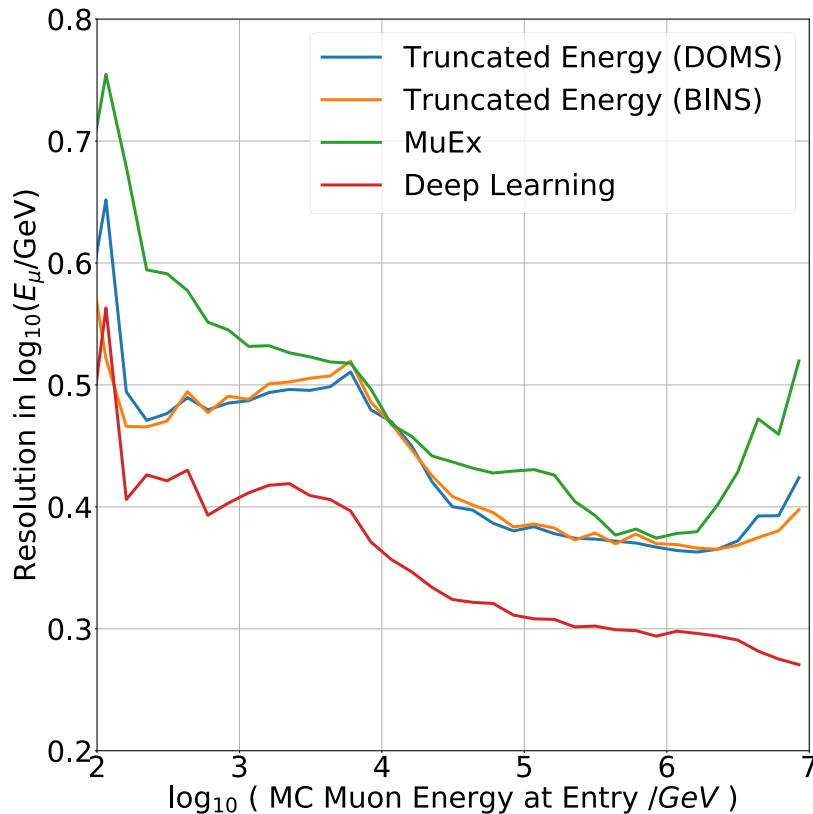
Deep Learning



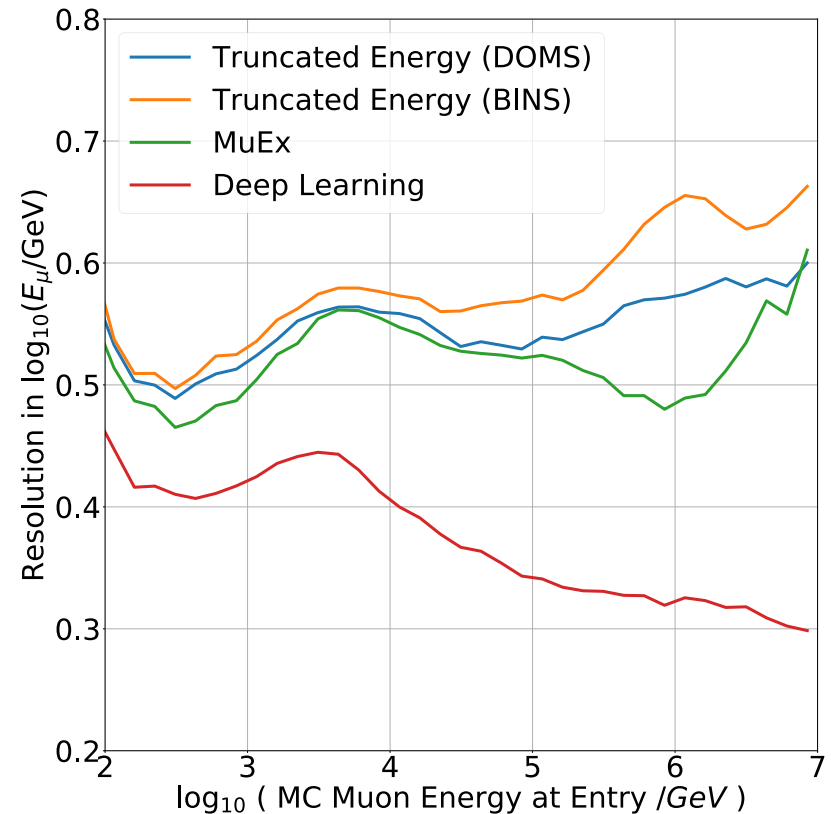
Energy Reconstruction – Muon Energy at Entry

OnlineL2 Muon Filter – CC events

$$\gamma = -1.0$$



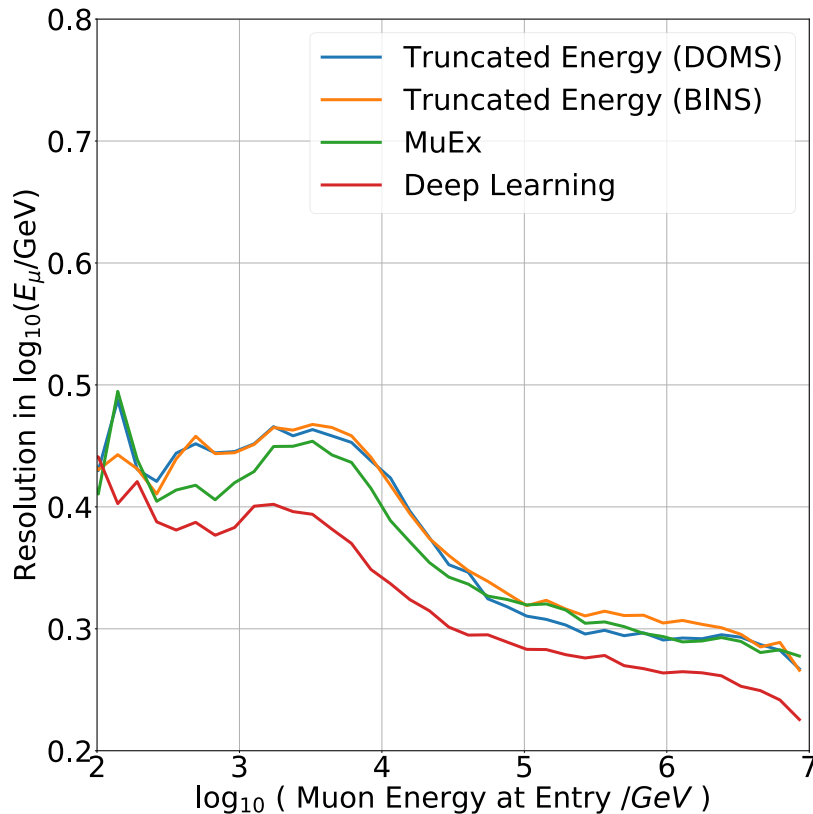
$$\gamma = -2.19$$



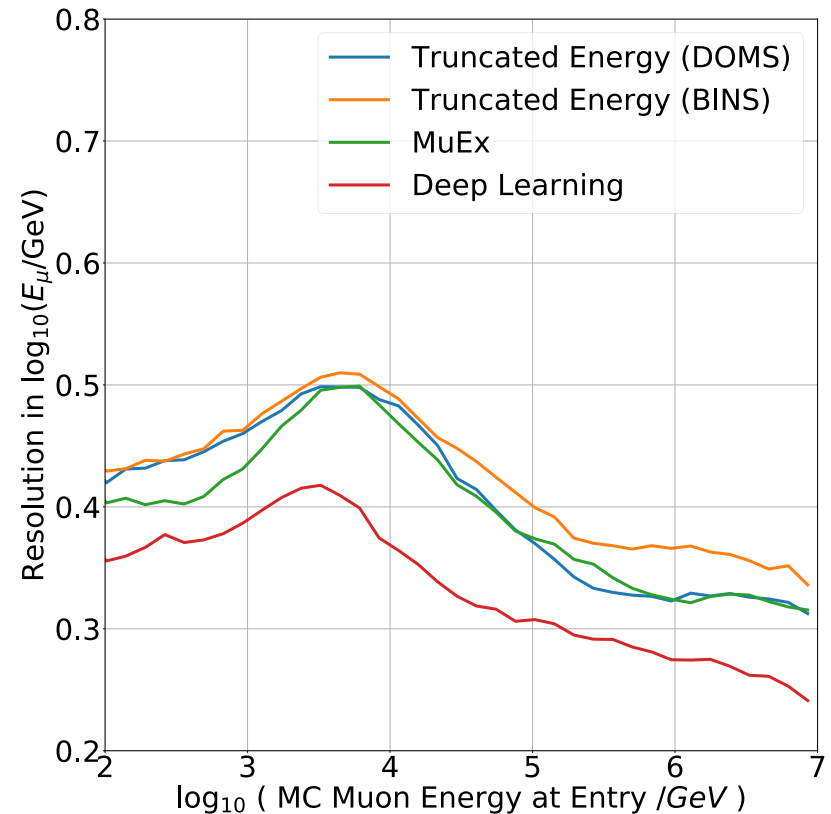
Energy Reconstruction – Muon Energy at Entry

GFU Filter – CC events

$$\gamma = -1.0$$



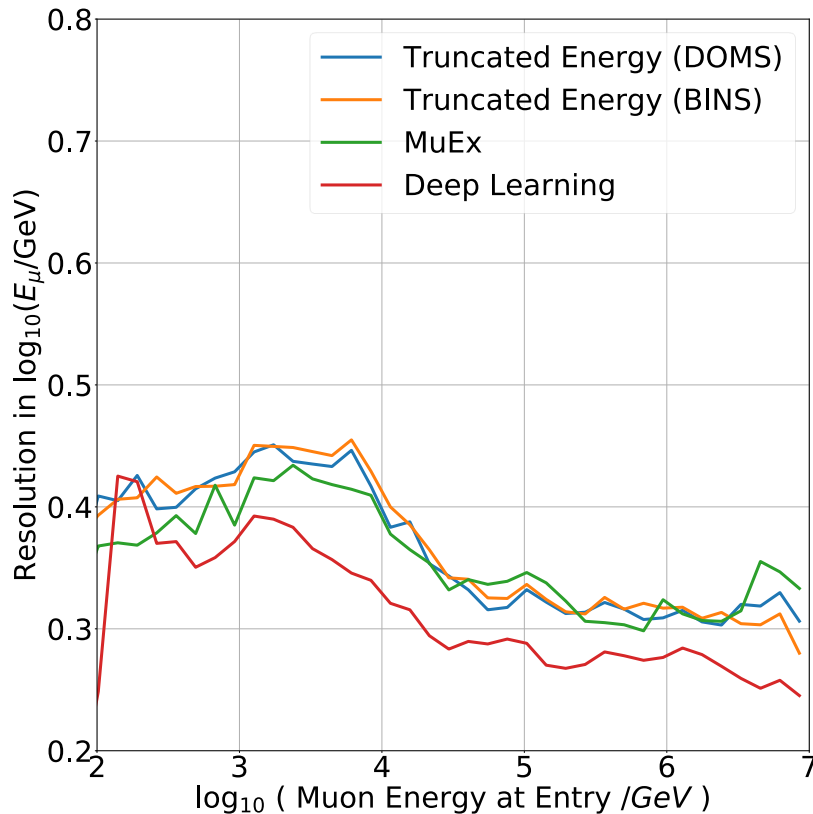
$$\gamma = -2.19$$



Energy Reconstruction – Muon Energy at Entry

Point source final level– CC events

$$\gamma = -1.0$$



$$\gamma = -2.19$$

