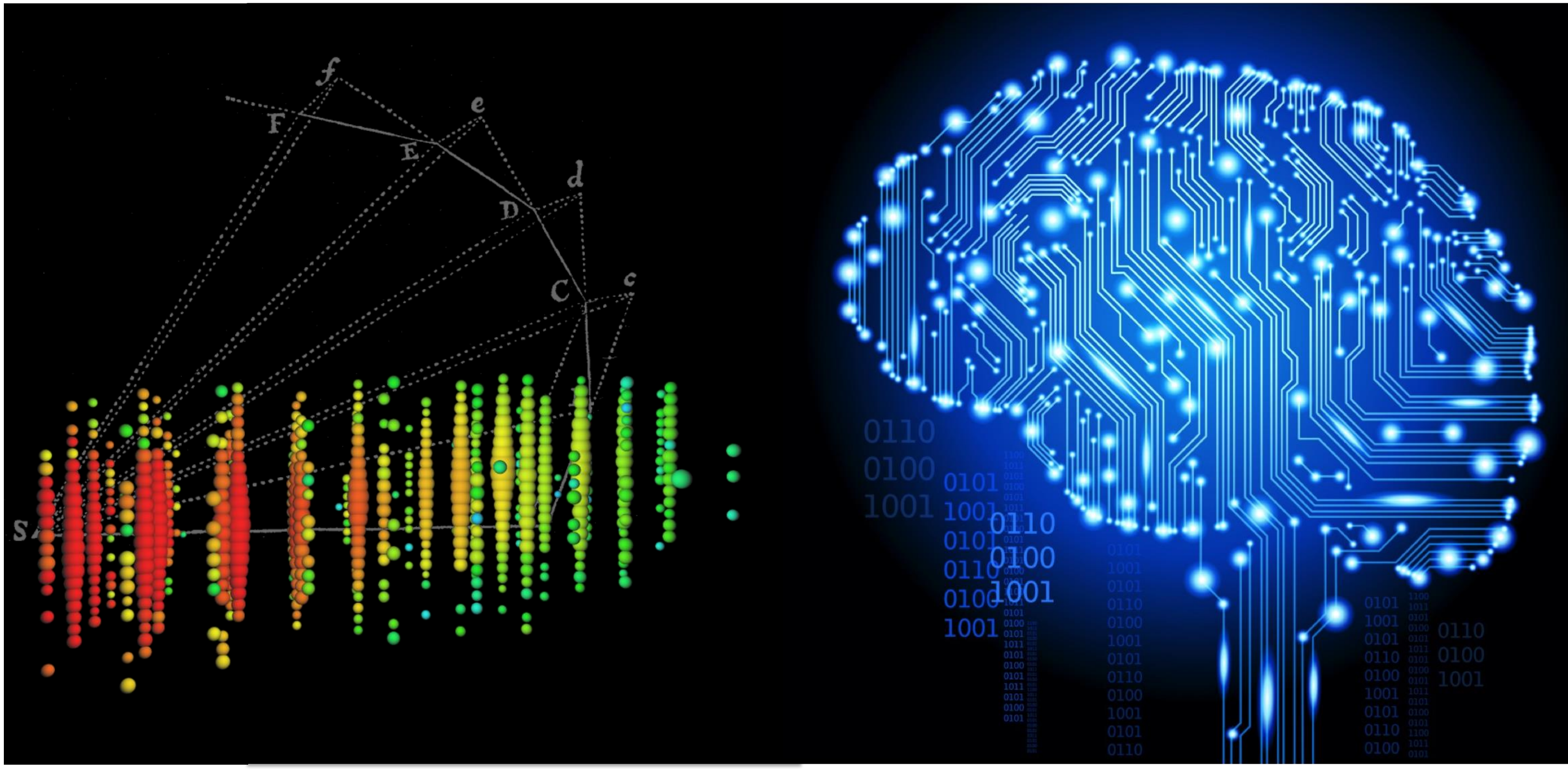
technische universität dortmund



Event Reconstruction in IceCube using Deep Learning Techniques

Mirco Huennefeld
mirco.huennefeld@tu-dortmund.de

February 20th, 2018

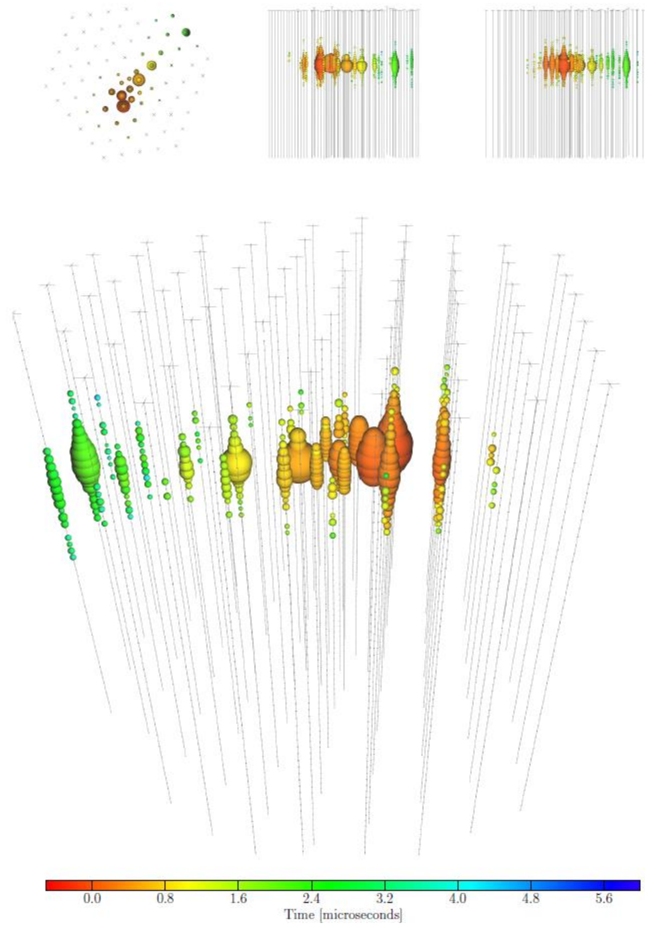
IceCube – Data Format

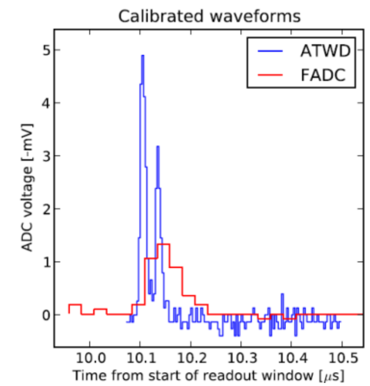
- 5160 DOMs
- 128 bins per waveform
- Arbitrary number of waveforms per DOM
- $5160 \cdot 128 \cdot n = 660480 \cdot n$

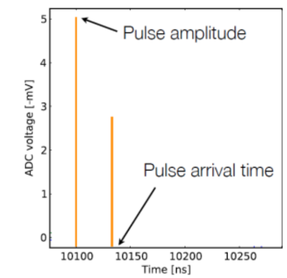
  MNIST: $28 \cdot 28 = 784$

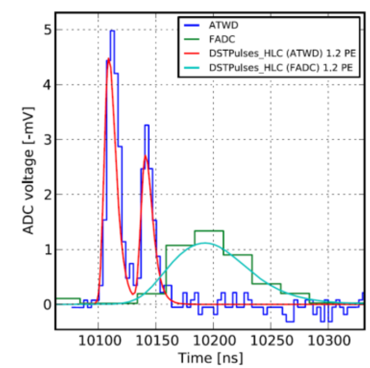
  CIFAR: $32 \cdot 32 = 1024$

  ImageNet: $256 \cdot 256 = 65536$ (cropped)
- Readout window dependent on trigger
- 4-dimensional, highly variable data



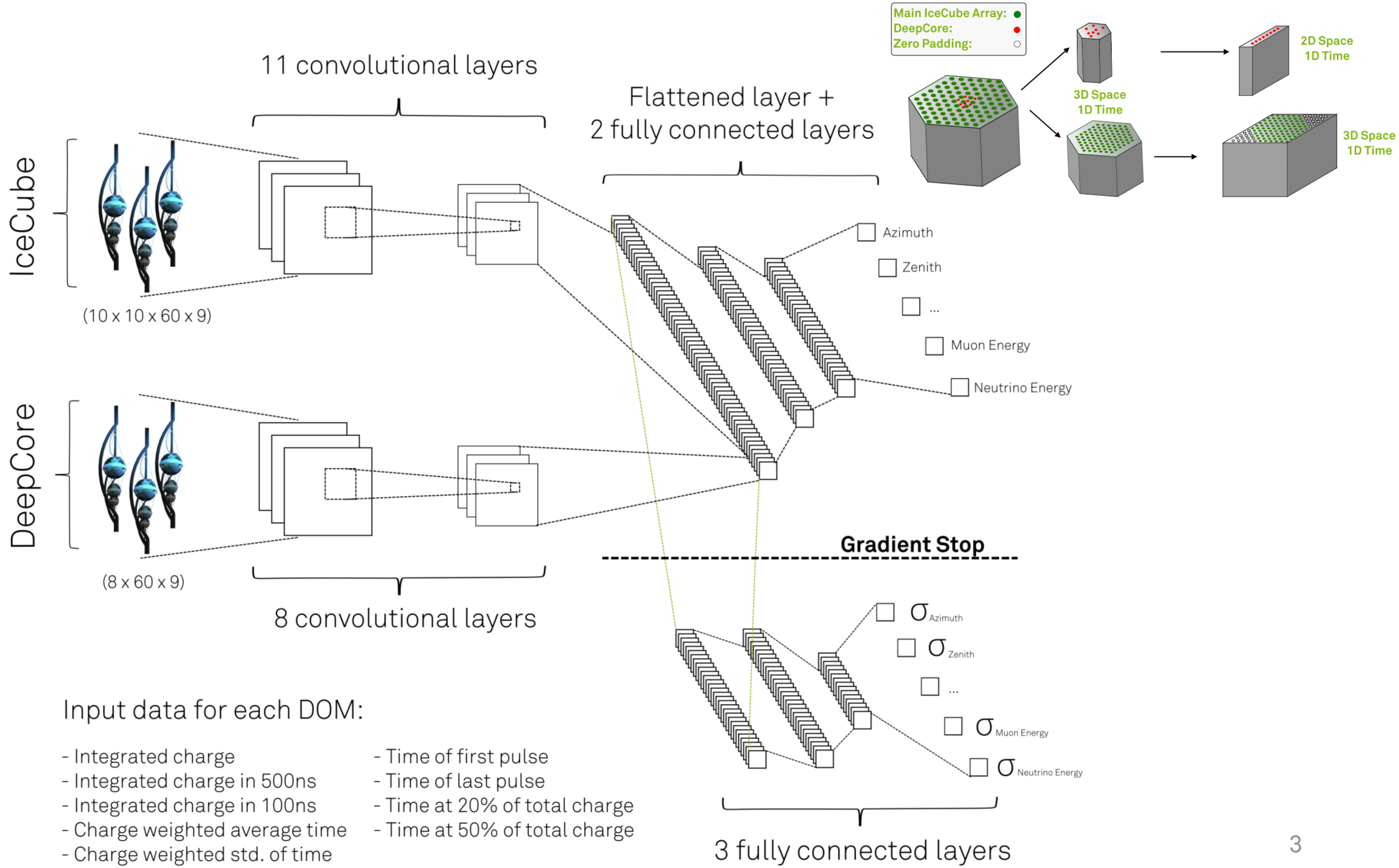






Goal: online reconstruction of particle track and energy

Network Architecture



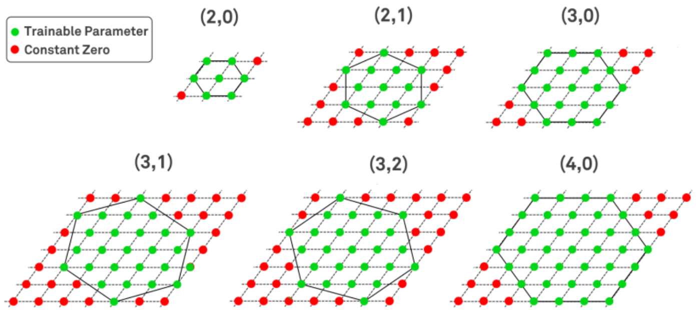
Input data for each DOM:

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

- Hexagonally shaped convolution kernels

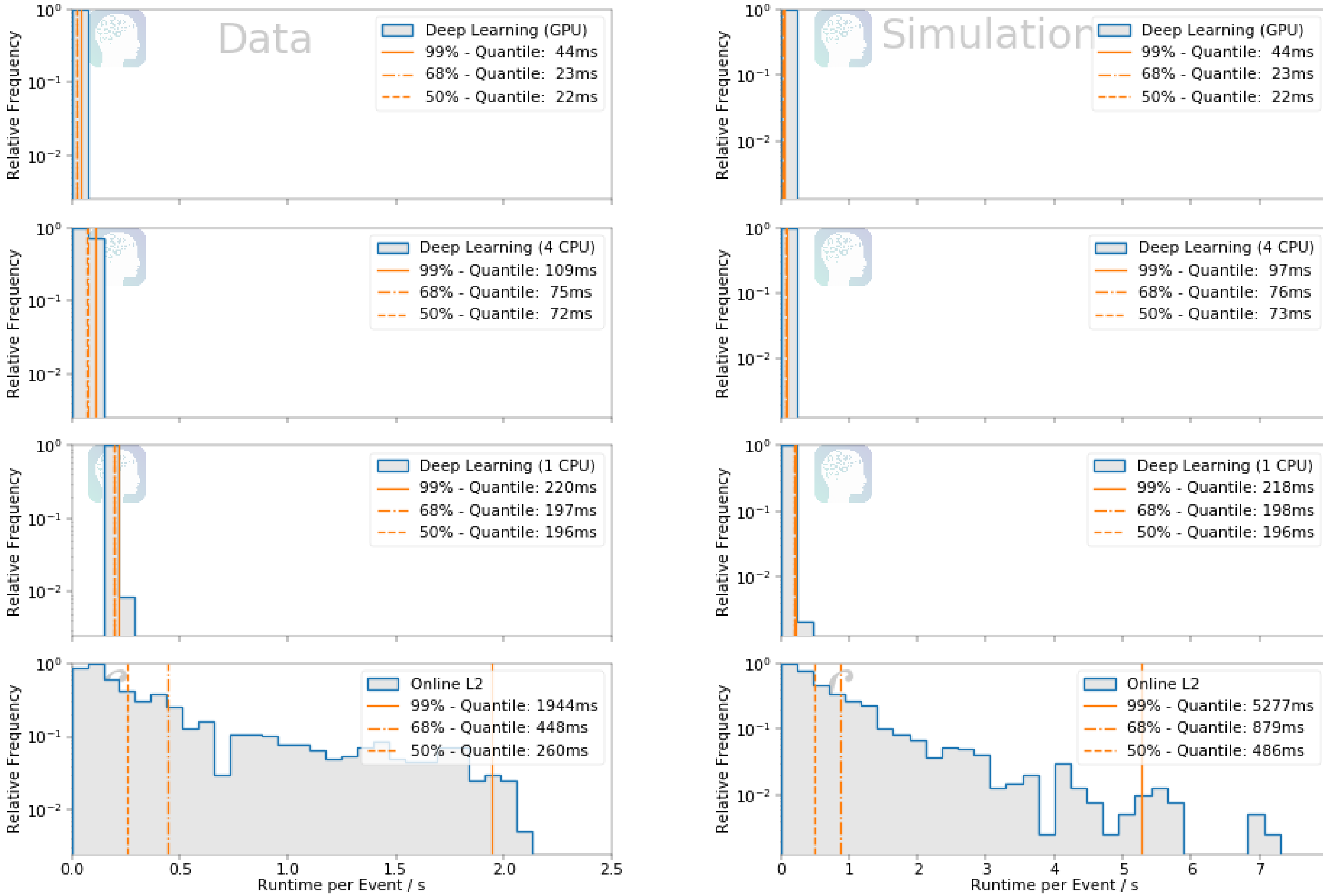


- Normalization of input and labels to mean 0 and variance 1

- 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

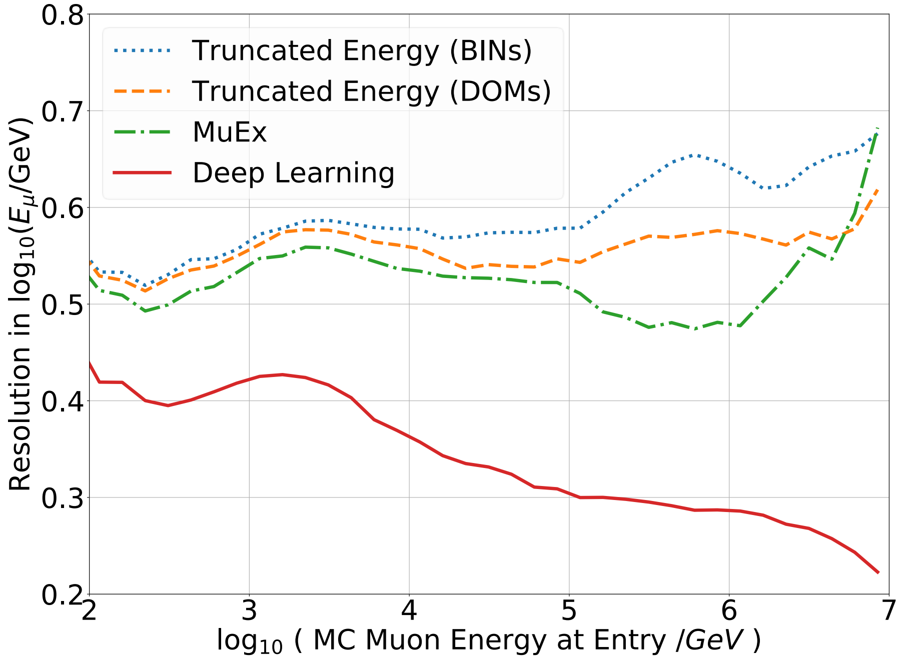
Runtime


Data
Simulation
Relative Frequency
Deep Learning (GPU)
99% - Quantile: 44ms
68% - Quantile: 23ms
50% - Quantile: 22ms
Deep Learning (GPU)
99% - Quantile: 44ms
68% - Quantile: 23ms
50% - Quantile: 22ms
Deep Learning (4 CPU)
99% - Quantile: 109ms
68% - Quantile: 75ms
50% - Quantile: 72ms
Deep Learning (4 CPU)
99% - Quantile: 97ms
68% - Quantile: 76ms
50% - Quantile: 73ms
Deep Learning (1 CPU)
99% - Quantile: 220ms
68% - Quantile: 197ms
50% - Quantile: 196ms
Deep Learning (1 CPU)
99% - Quantile: 218ms
68% - Quantile: 198ms
50% - Quantile: 196ms
Online L2
99% - Quantile: 1944ms
68% - Quantile: 448ms
50% - Quantile: 260ms
Online L2
99% - Quantile: 5277ms
68% - Quantile: 879ms
50% - Quantile: 486ms
Runtime per Event / s
Runtime per Event / s

Performance

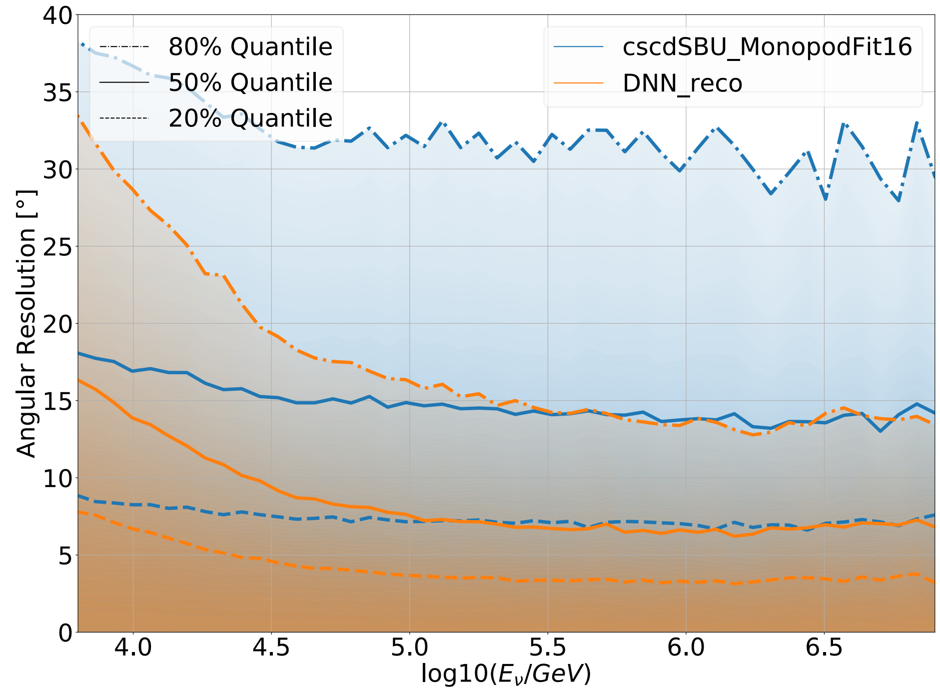
Track-like Events

Muon Energy Resolution

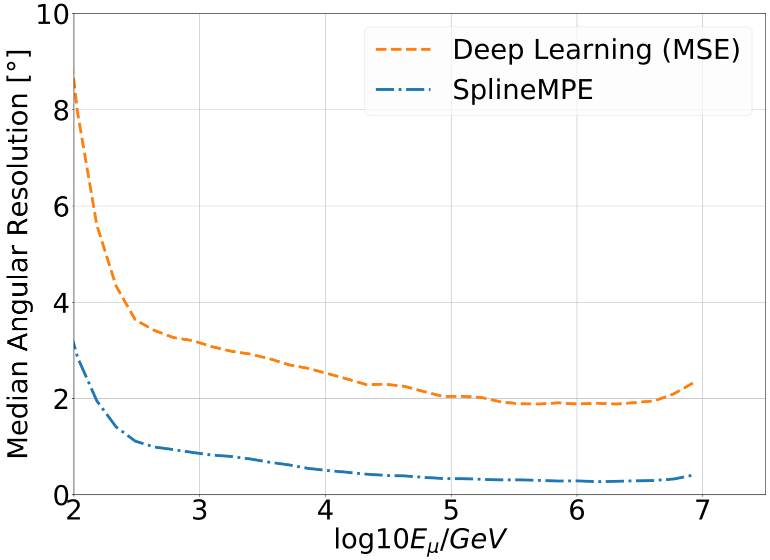


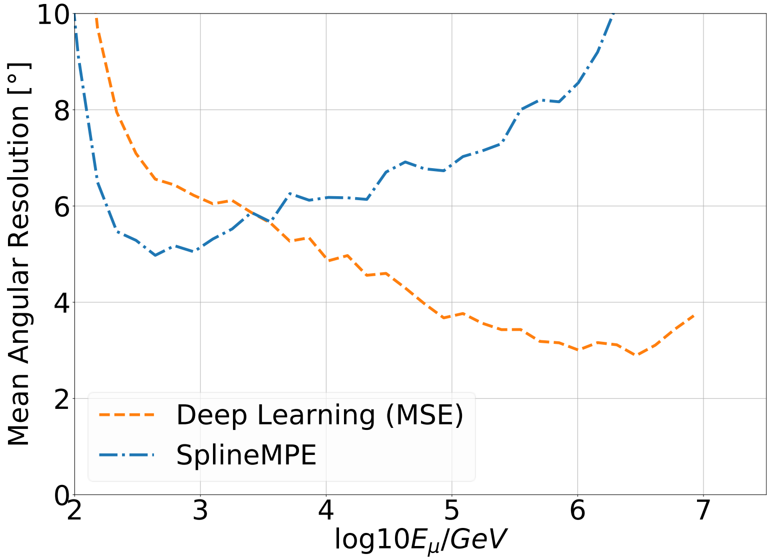
Cascade-like Events

Energy dependent Angular Resolution



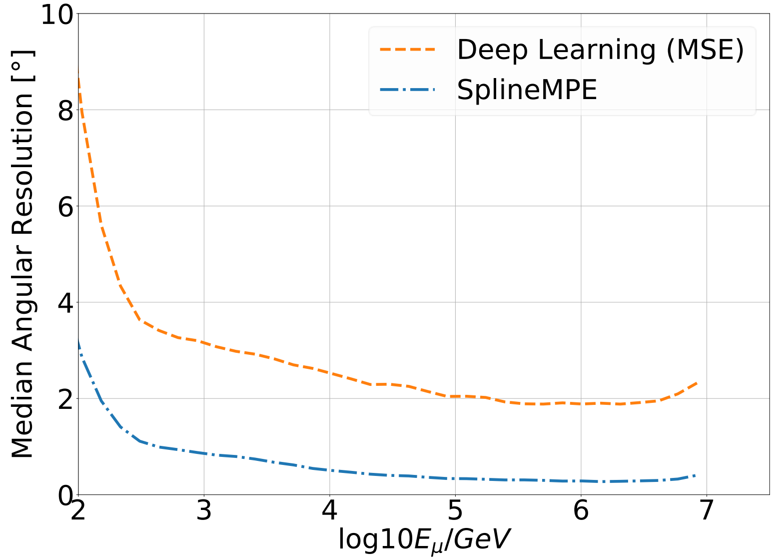
Loss Function Bias

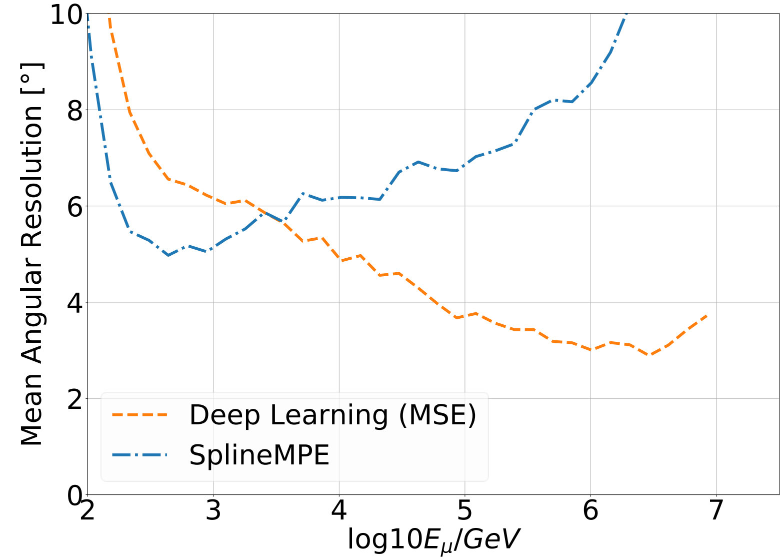




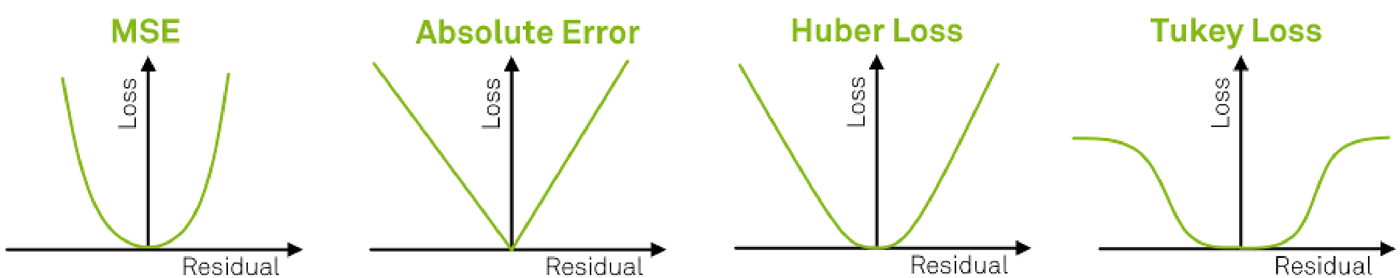
- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution

Loss Function Bias



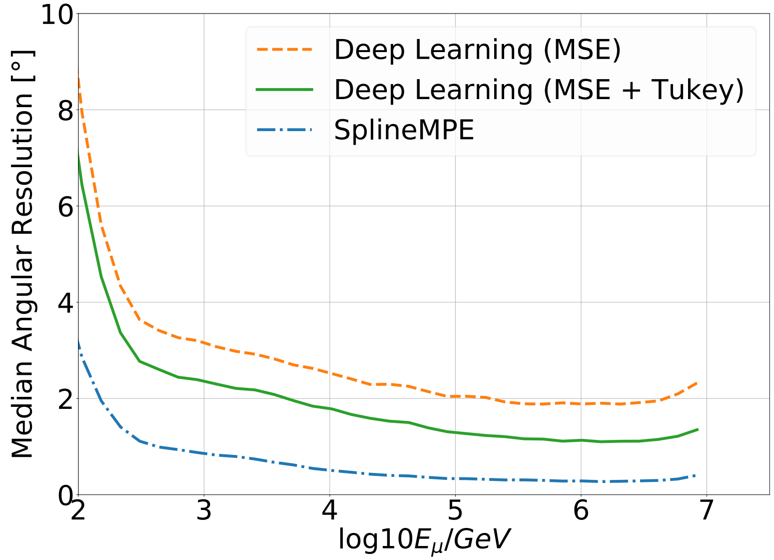


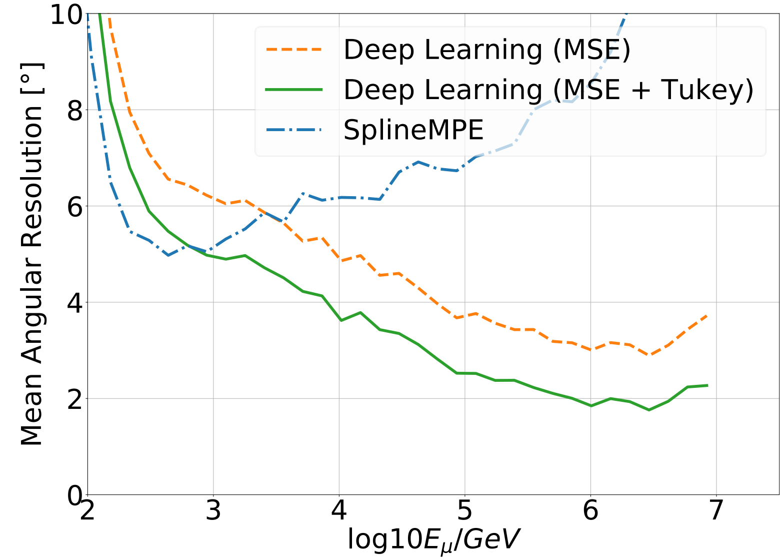
- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution



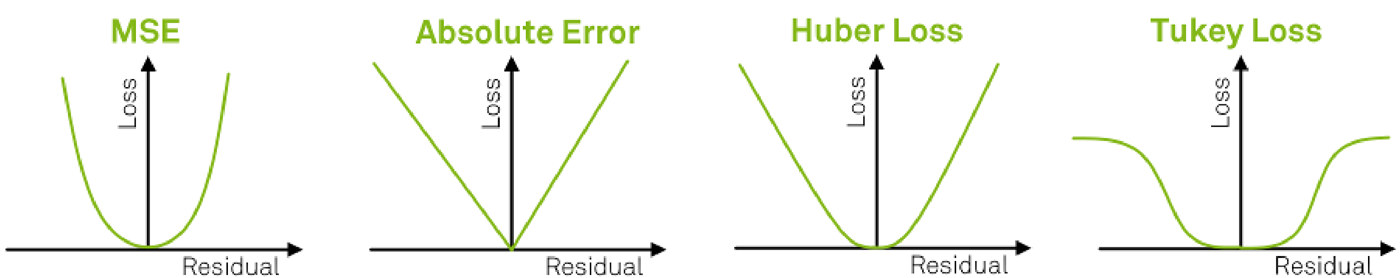
DOI: 10.1109/ICCV.2015.324

Loss Function Bias





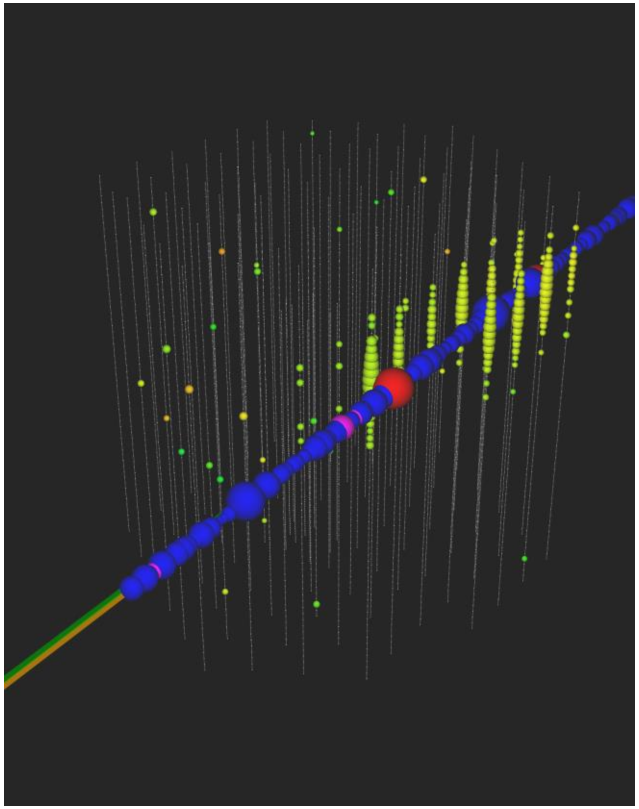
- Deep Learning approach can greatly reduce outliers
- Likelihood-based approach better when looking at median angular resolution



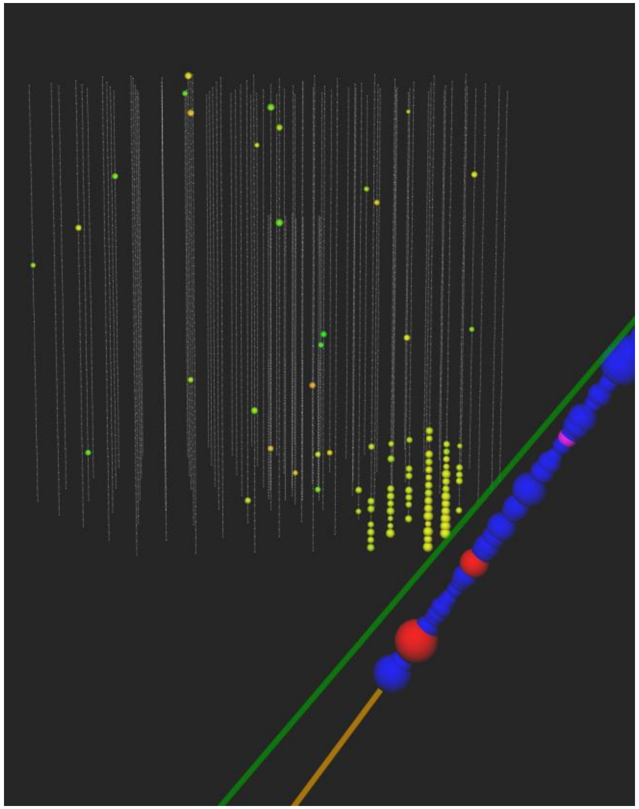
DOI: 10.1109/ICCV.2015.324

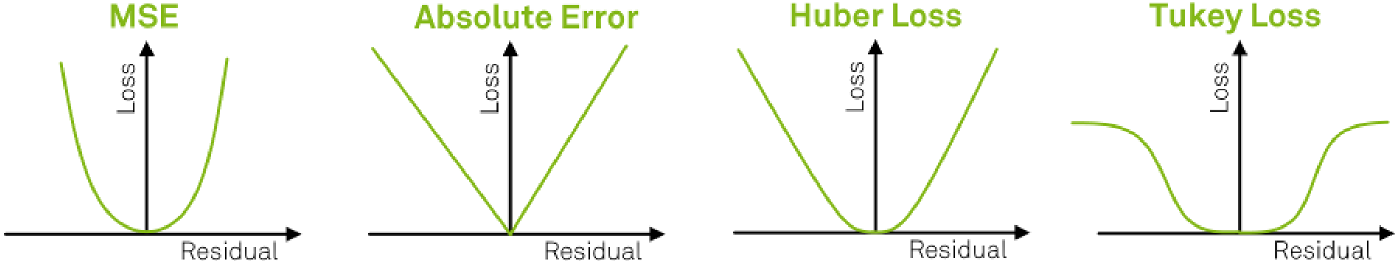
Loss Function Bias

Angular Error: 0.5°



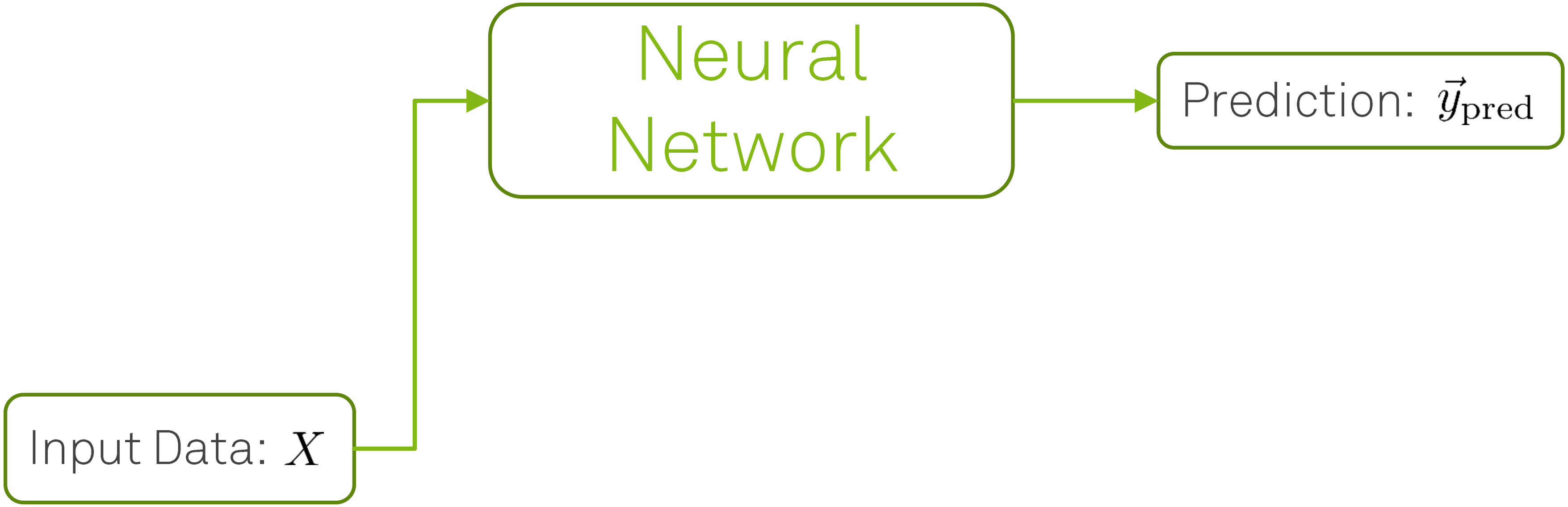
Angular Error: 15°



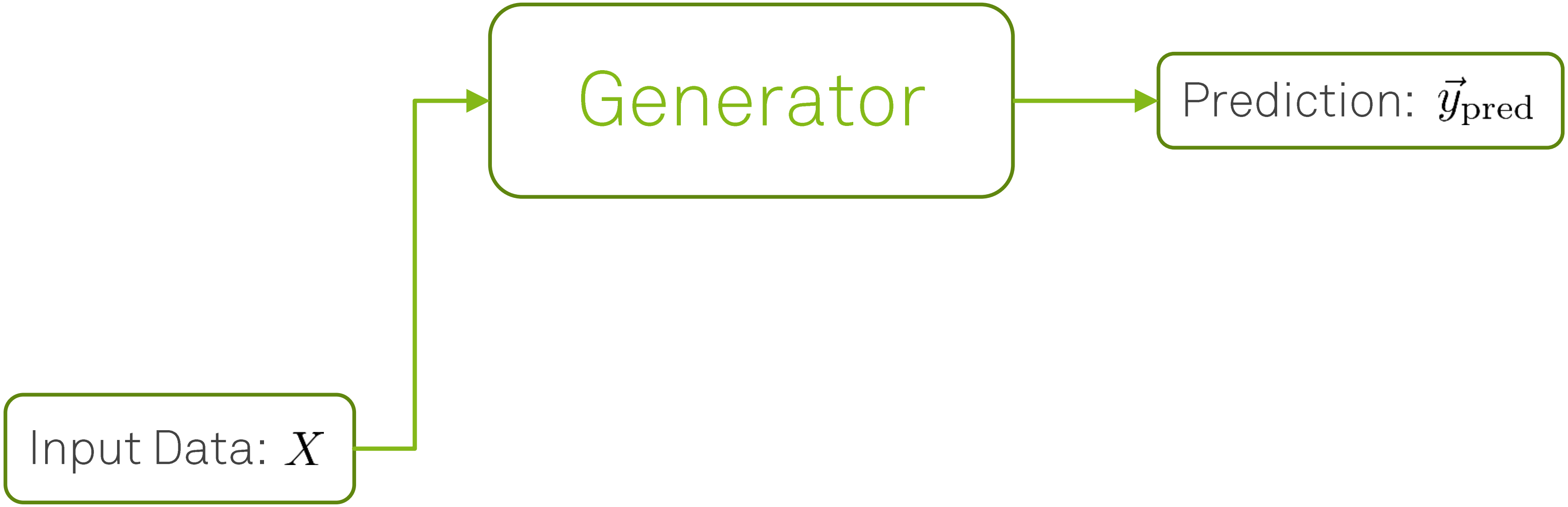


DOI: 10.1109/ICCV.2015.324

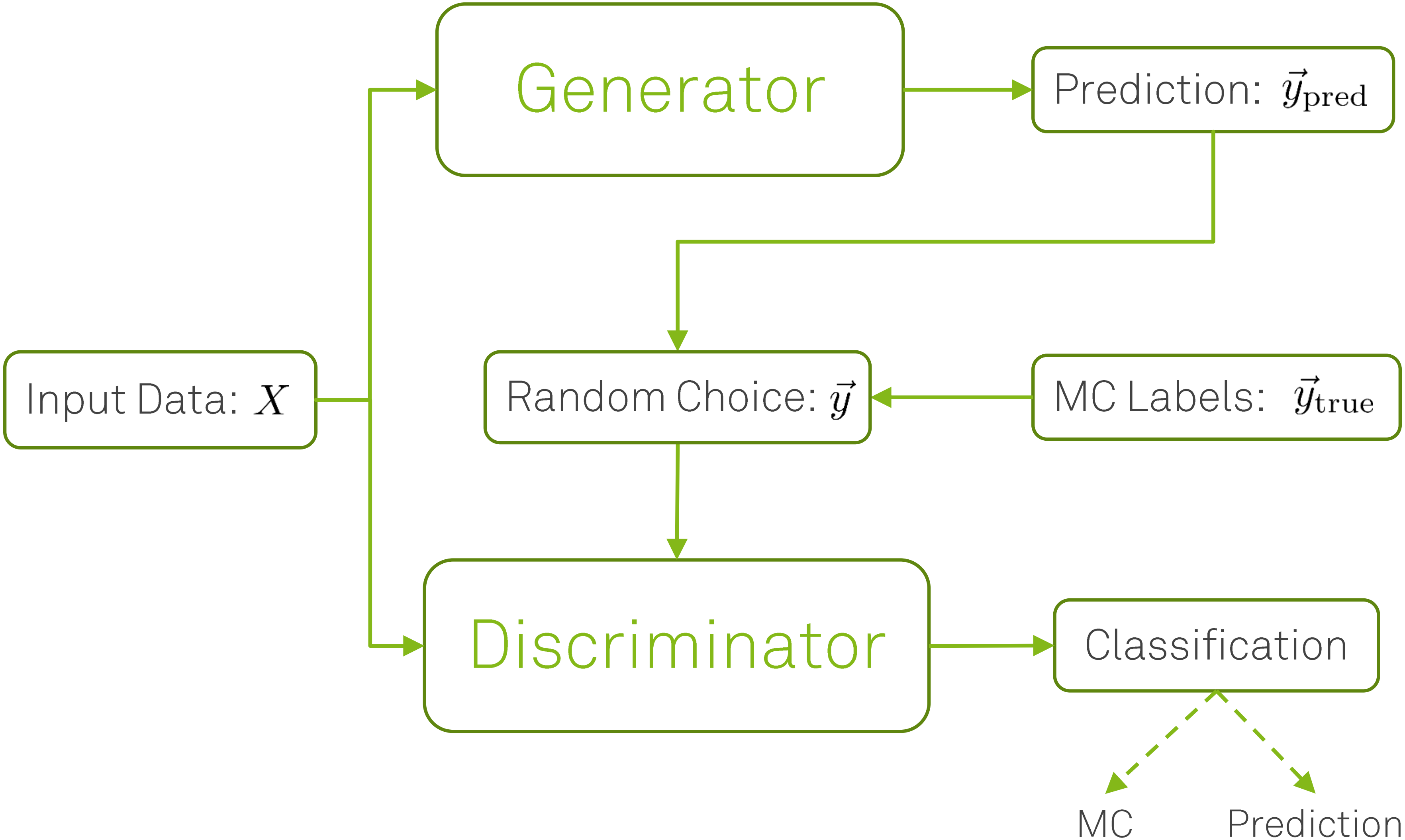
Adversarial Training



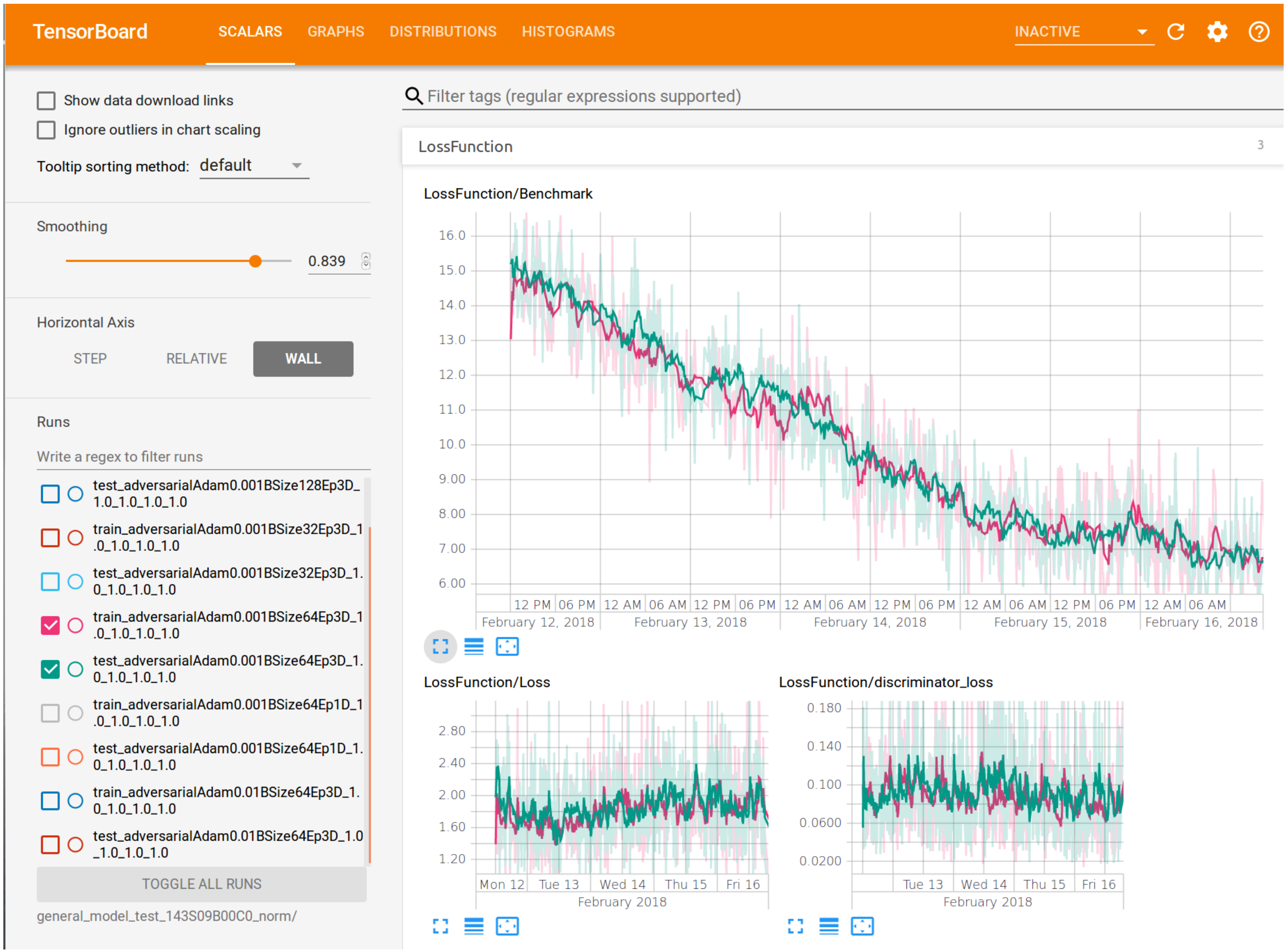
Adversarial Training



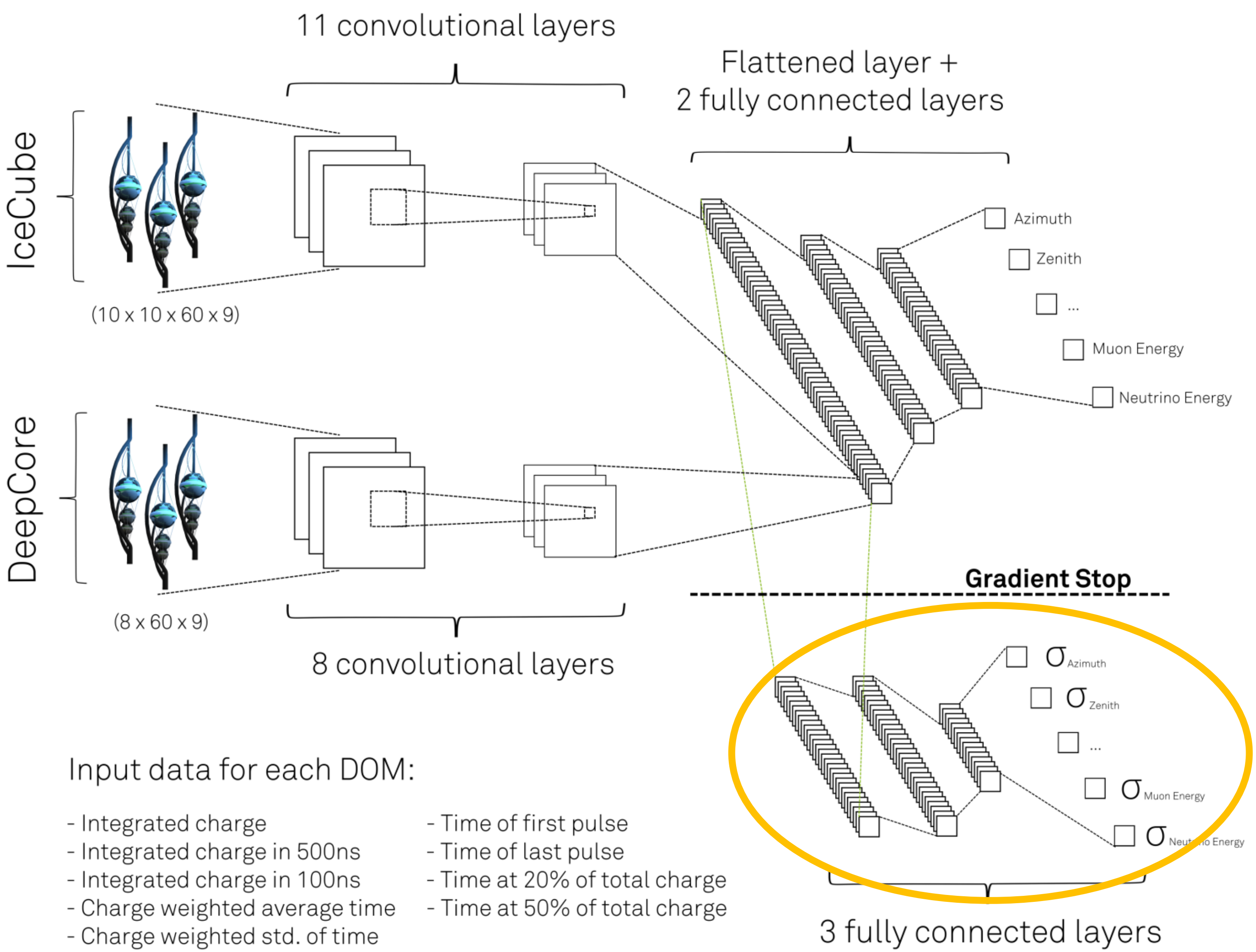
Adversarial Training



Adversarial Training



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:

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

  - Network is trained to correctly estimate absolute residual on average: $\sigma_{\text{pred}} \approx \langle y_{\text{abs}} \rangle$

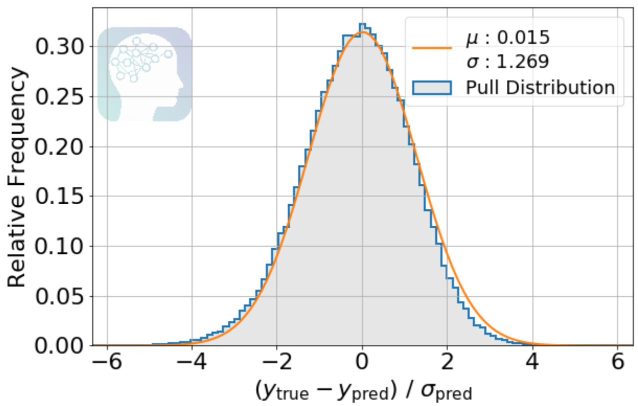
  - If residuals are Gaussian-distributed, the mean absolute residual can be converted to the standard deviation: $\sigma = \sqrt{\frac{\pi}{2}} \cdot \langle y_{\text{abs}} \rangle \approx 1.253 \cdot \langle y_{\text{abs}} \rangle$

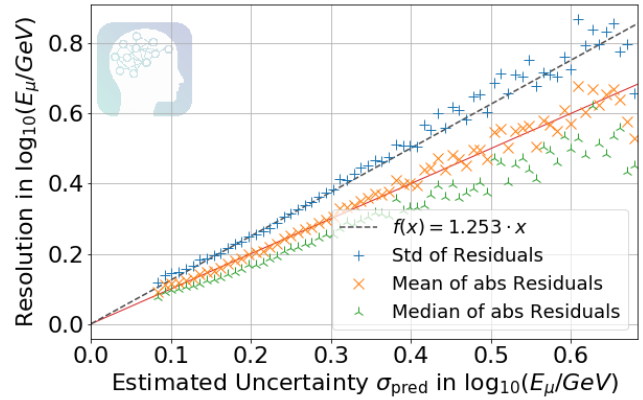
- Gaussian Likelihood Minimization:

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

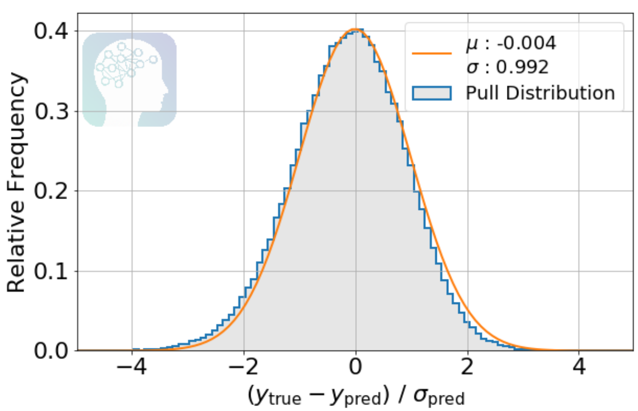
Uncertainty Estimation

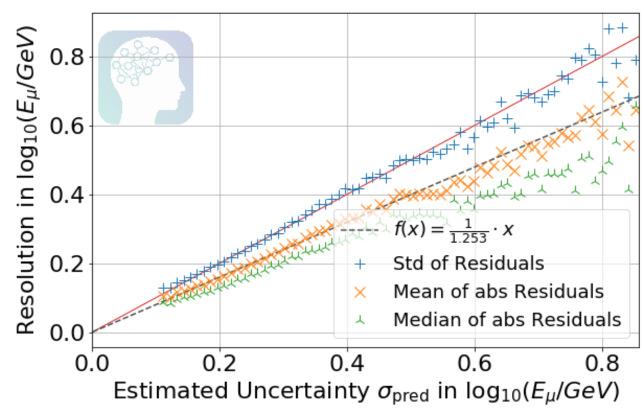
MSE of absolute Residuals



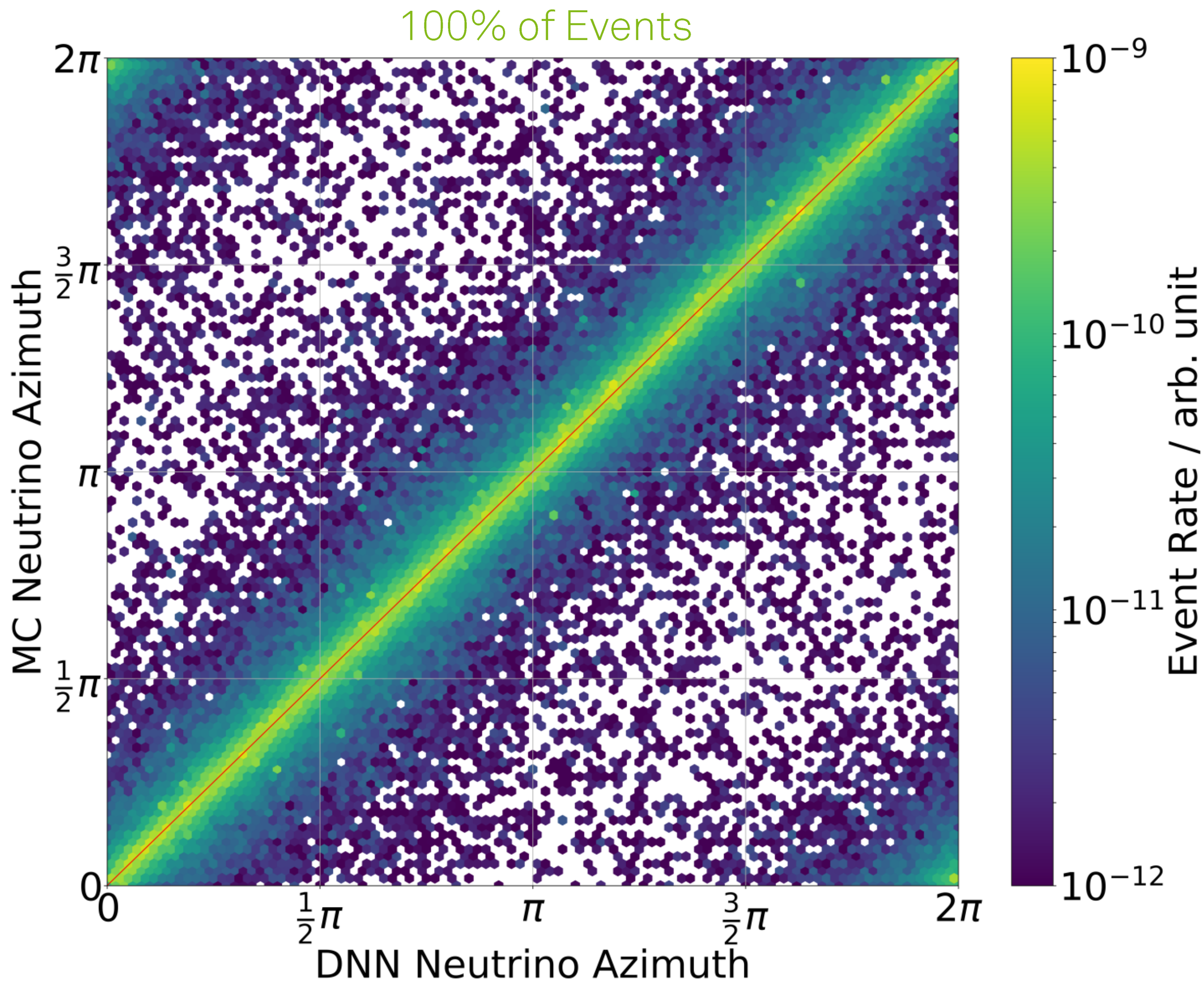


Gaussian Likelihood

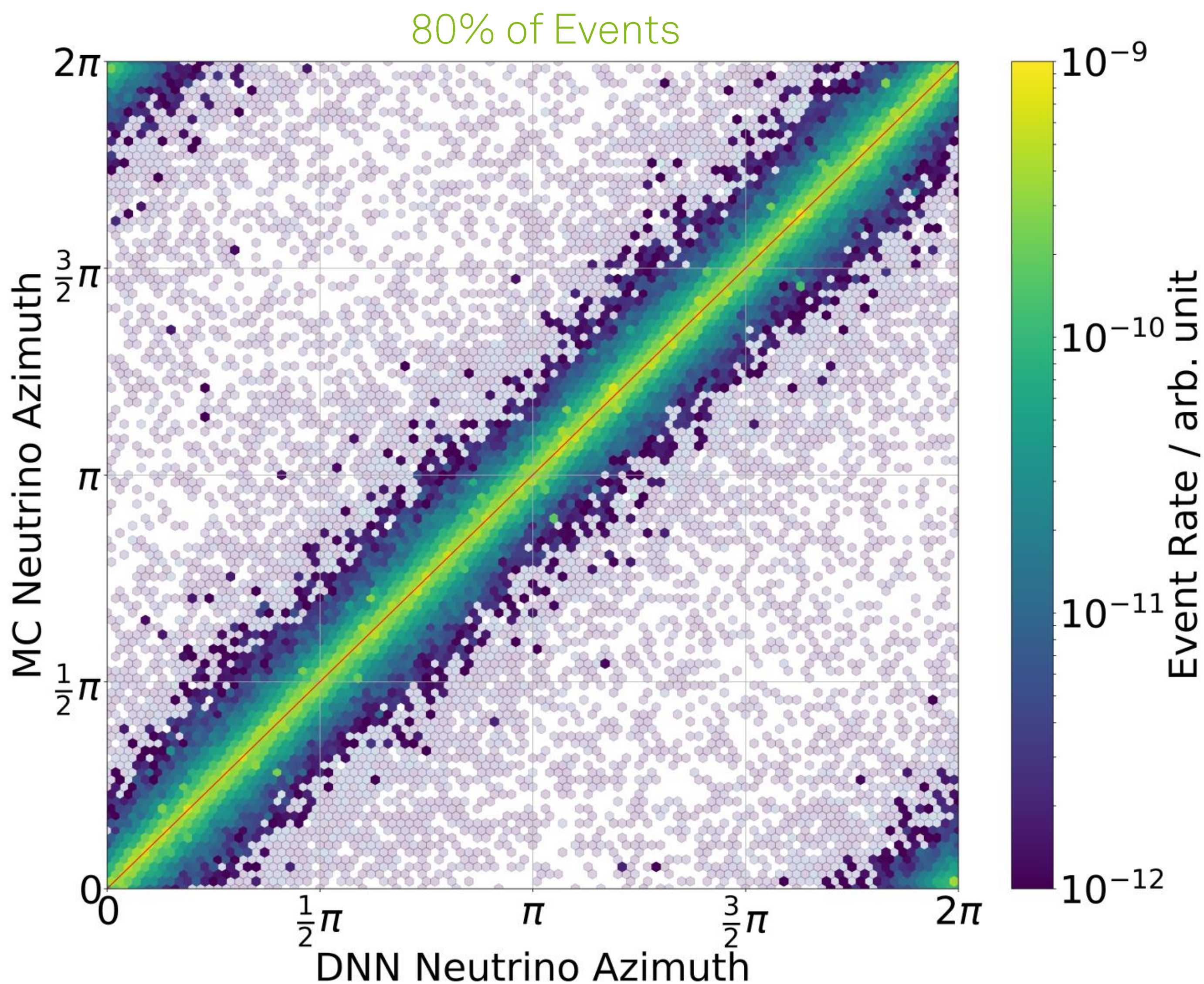




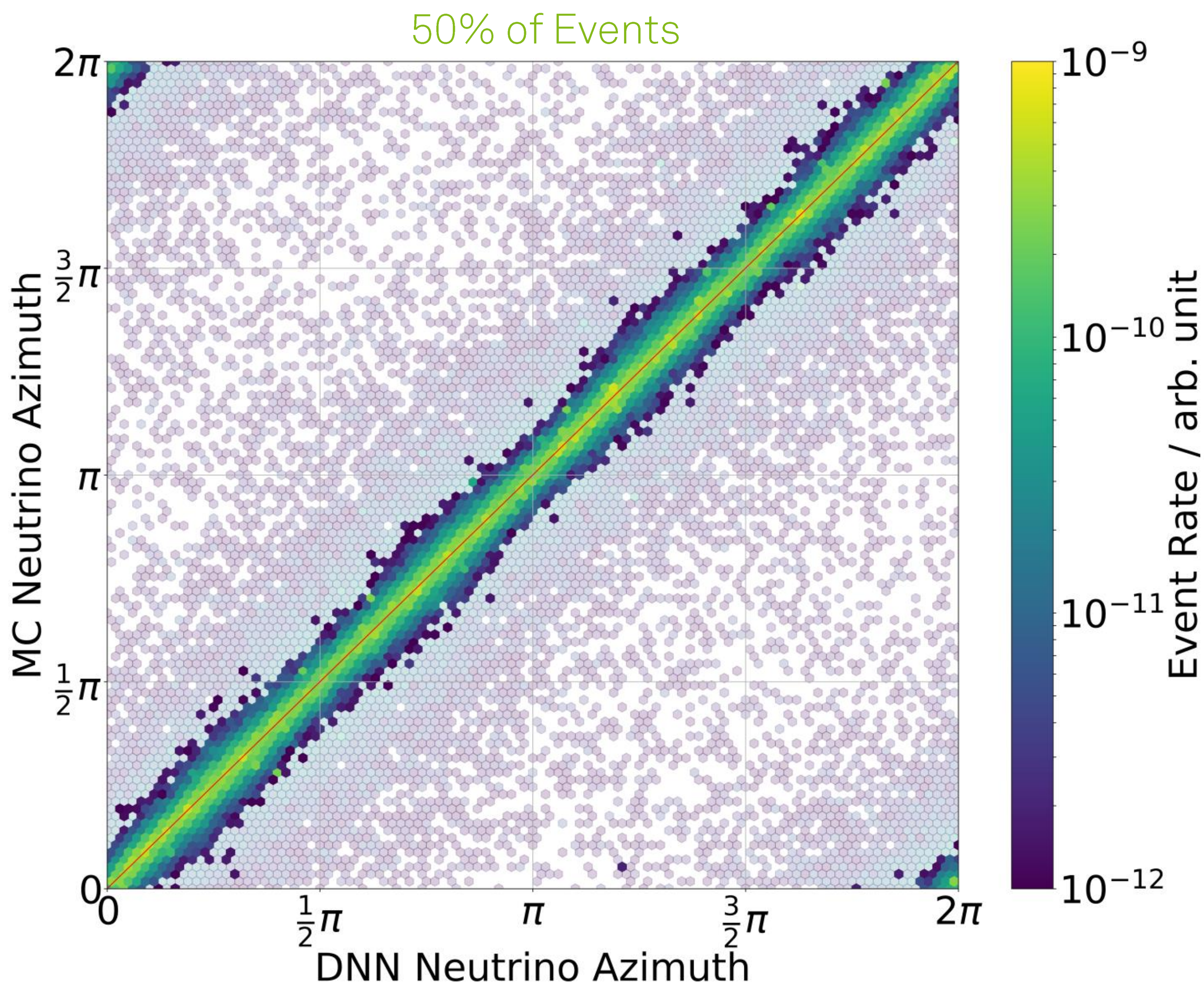
Uncertainty Estimation



Uncertainty Estimation

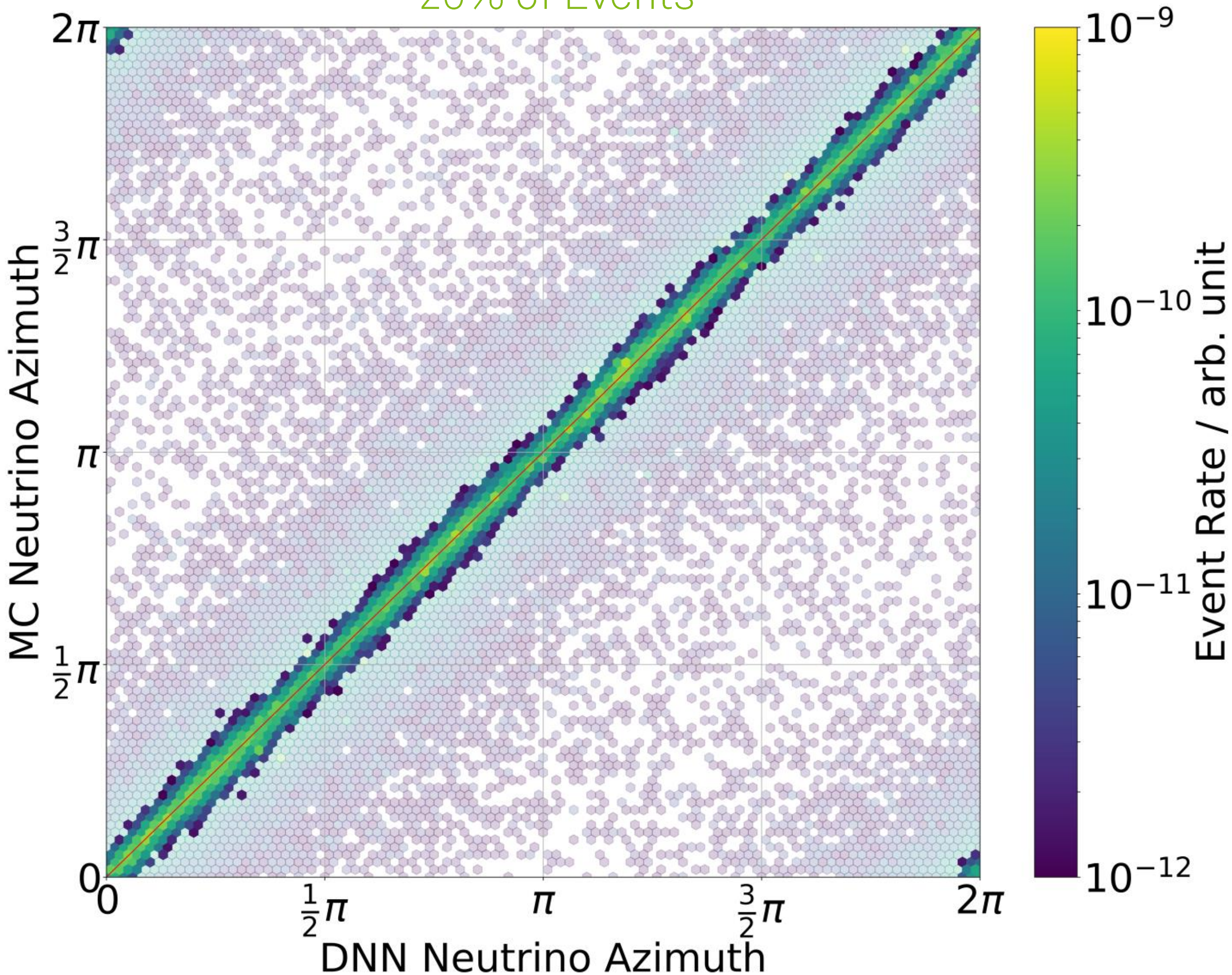


Uncertainty Estimation

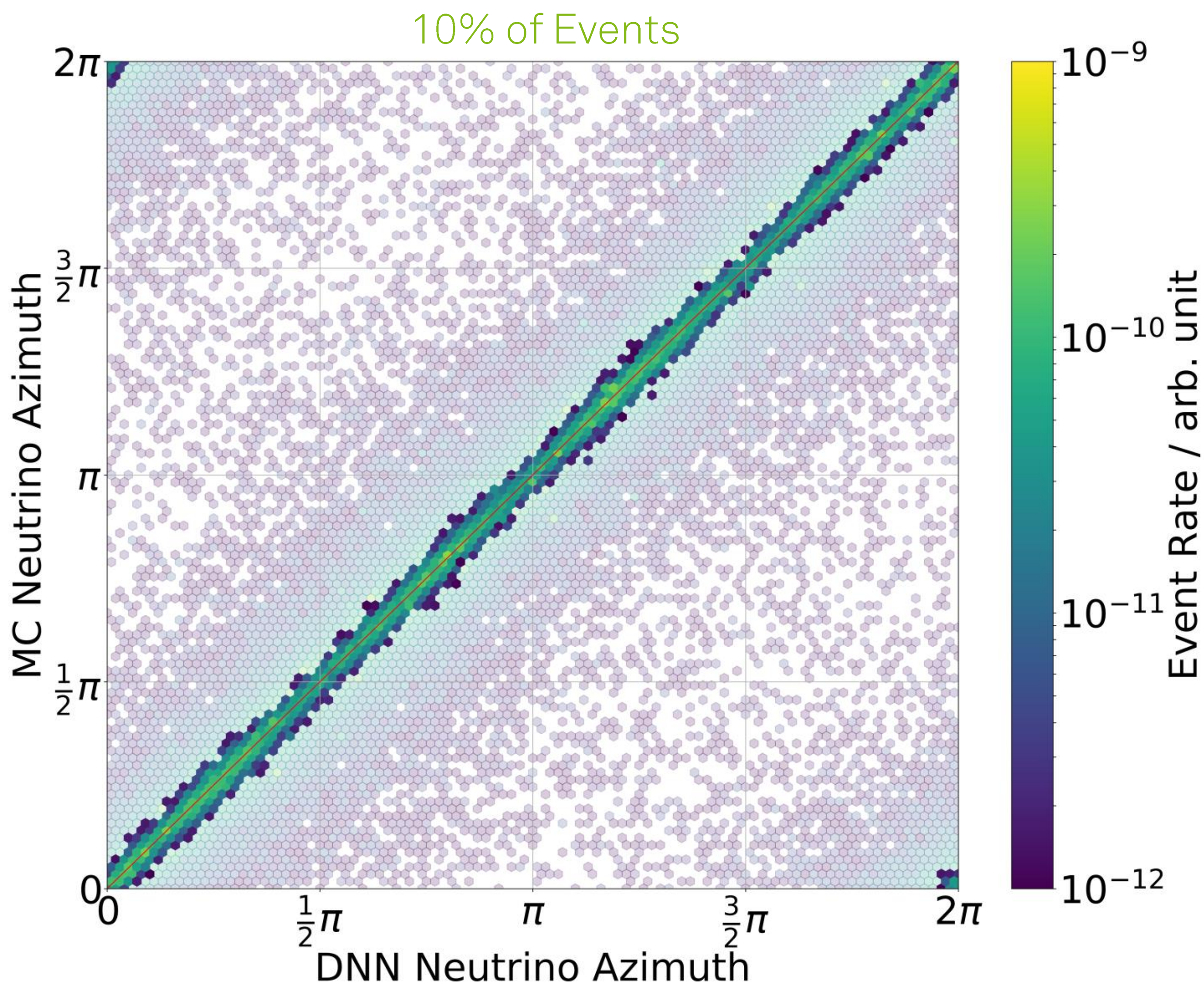


Uncertainty Estimation

20% of Events



Uncertainty Estimation



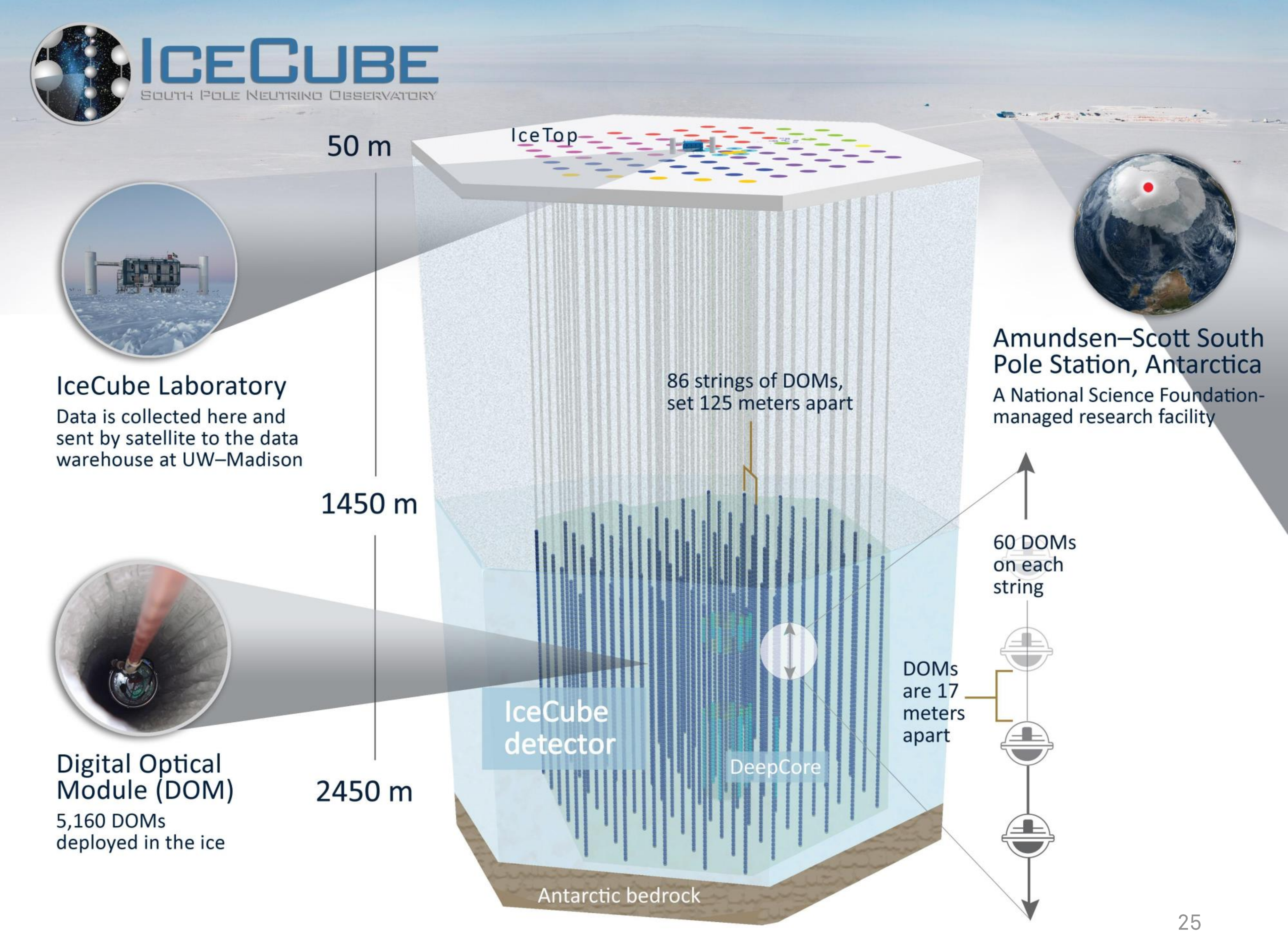
Summary & Outlook

- Application of deep learning-based techniques works well, however, just applying these methods is not enough!

- 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

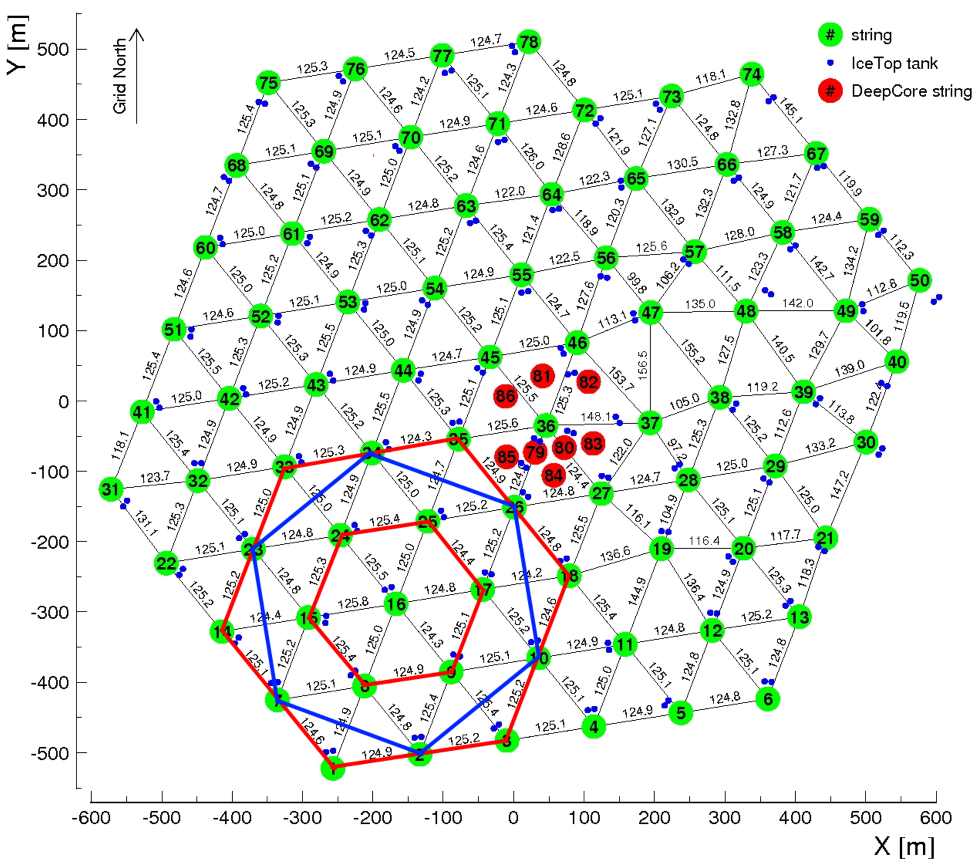
- Additional Topics:
  - Modelling detector
  - Fast and efficient simulation
  - Learning from the neural network
  - ...

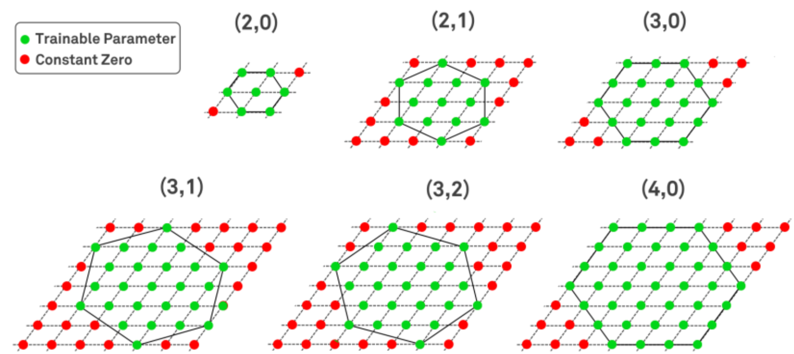
Appendix

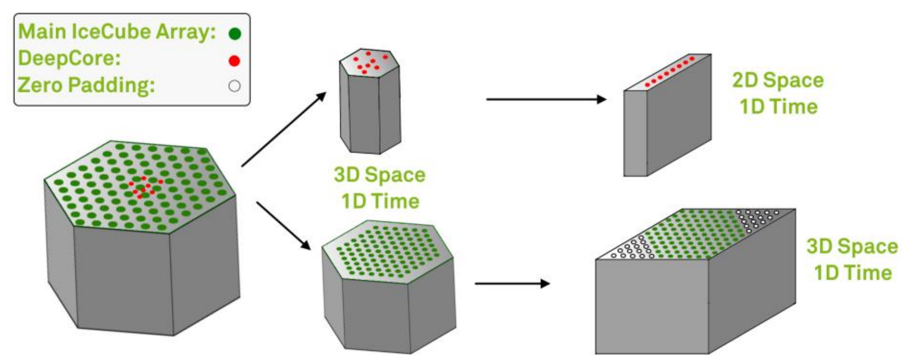


ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY
IceTop
50 m
1450 m
2450 m
IceCube Laboratory
Data is collected here and
sent by satellite to the data
warehouse at UW–Madison
86 strings of DOMs,
set 125 meters apart
Amundsen–Scott South
Pole Station, Antarctica
A National Science Foundation-
managed research facility
60 DOMs
on each
string
DOMs
are 17
meters
apart
IceCube
detector
DeepCore
Digital Optical
Module (DOM)
5,160 DOMs
deployed in the ice
Antarctic bedrock

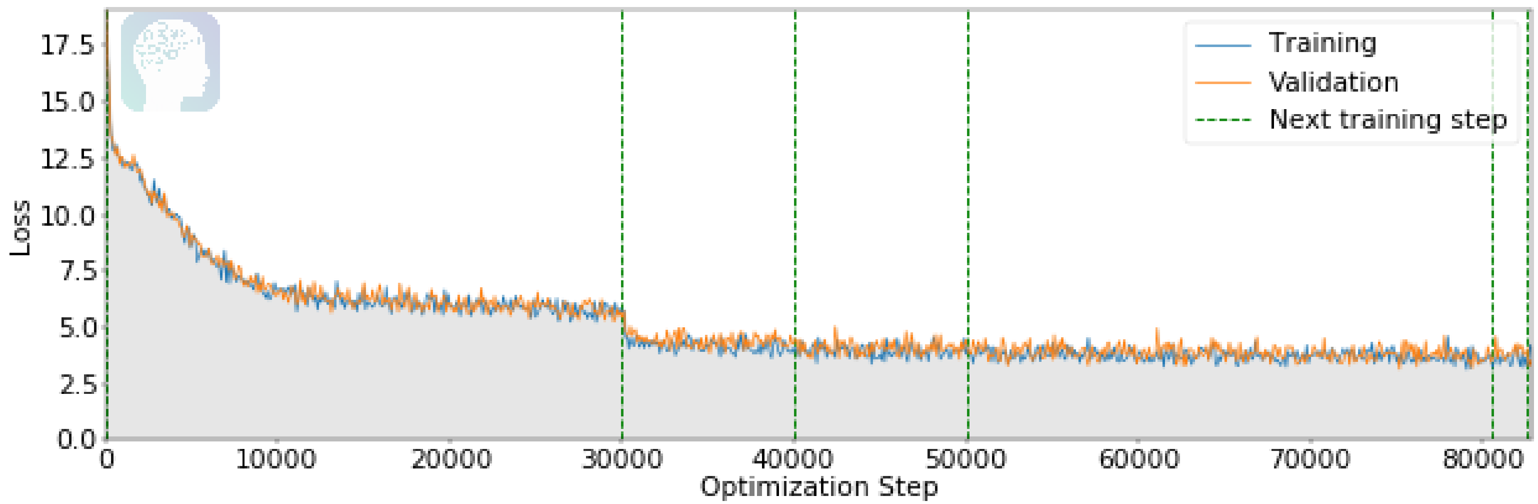
IceCube – Data Format

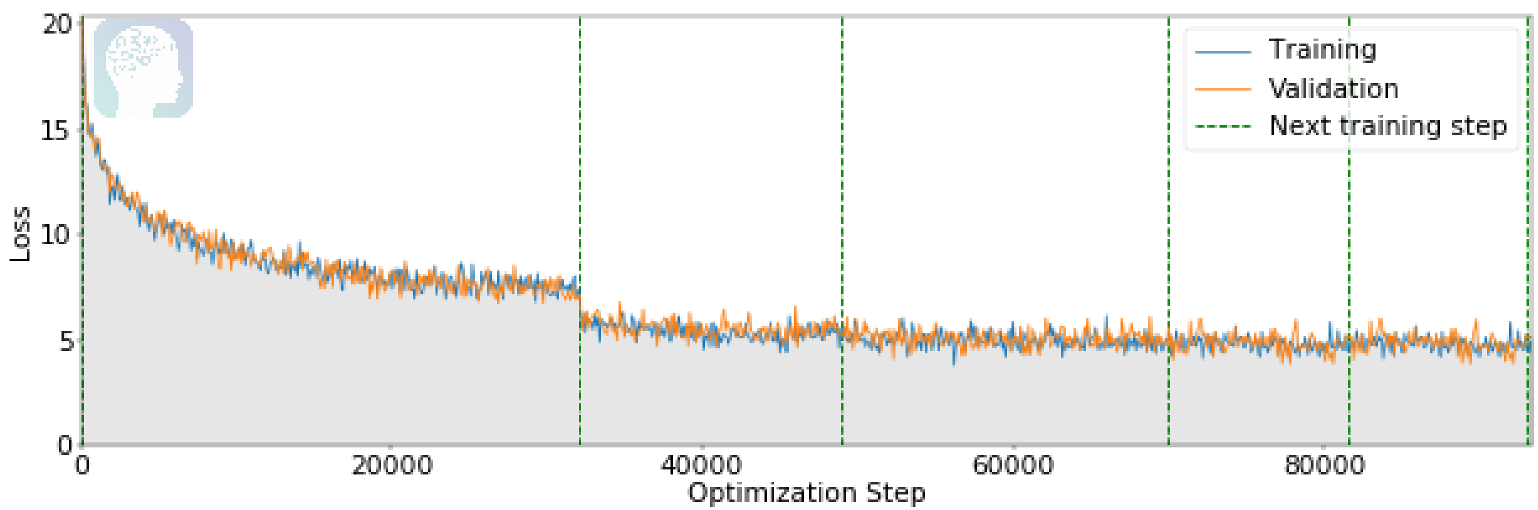






Deep Learning – Training

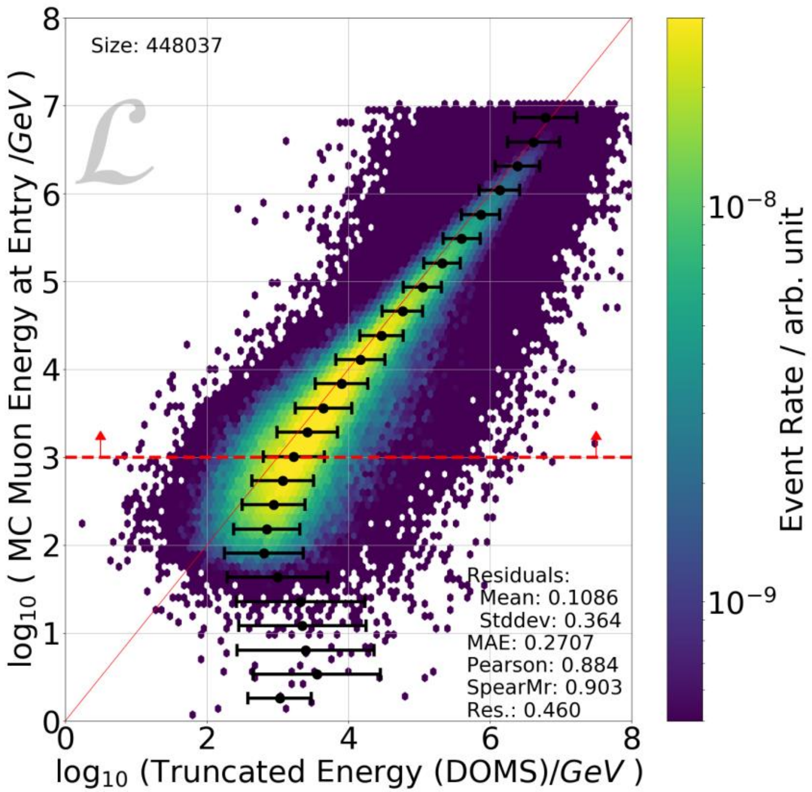




Energy Reconstruction – Muon Energy at Entry

OnlineL2 Muon Filter – CC events

Truncated Energy (DOMS)



Deep Learning