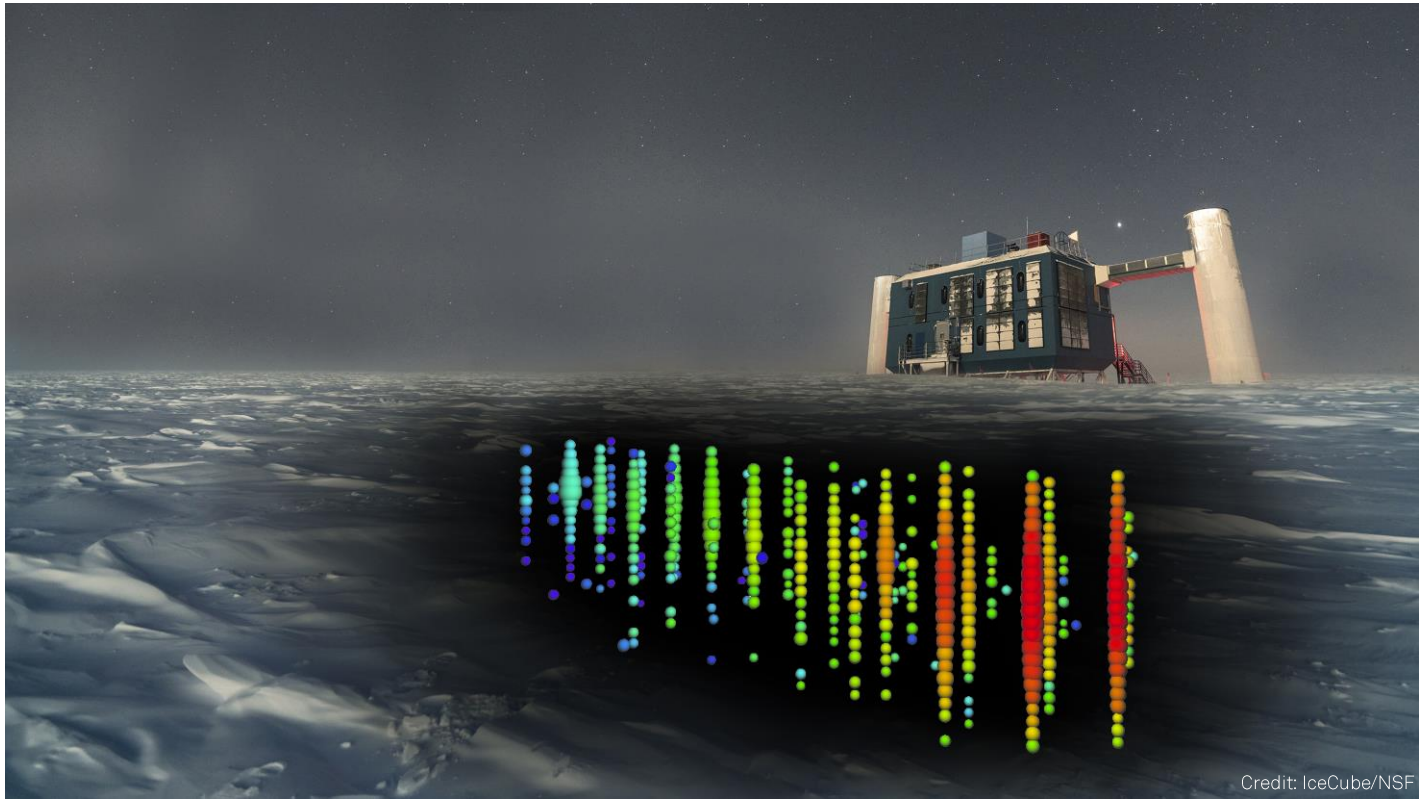
technische universität dortmund

ICECUBE SOUTH POLE NEUTRINO OBSERVATORY

SFB 876 Providing Information by Resource-Constrained Data Analysis

lehrstuhl physik e5



Reconstruction Techniques in IceCube using Convolutional and Generative Neural Networks

Mirco Huennefeld for the IceCube Collaboration*

mirco.huennefeld@tu-dortmund.de

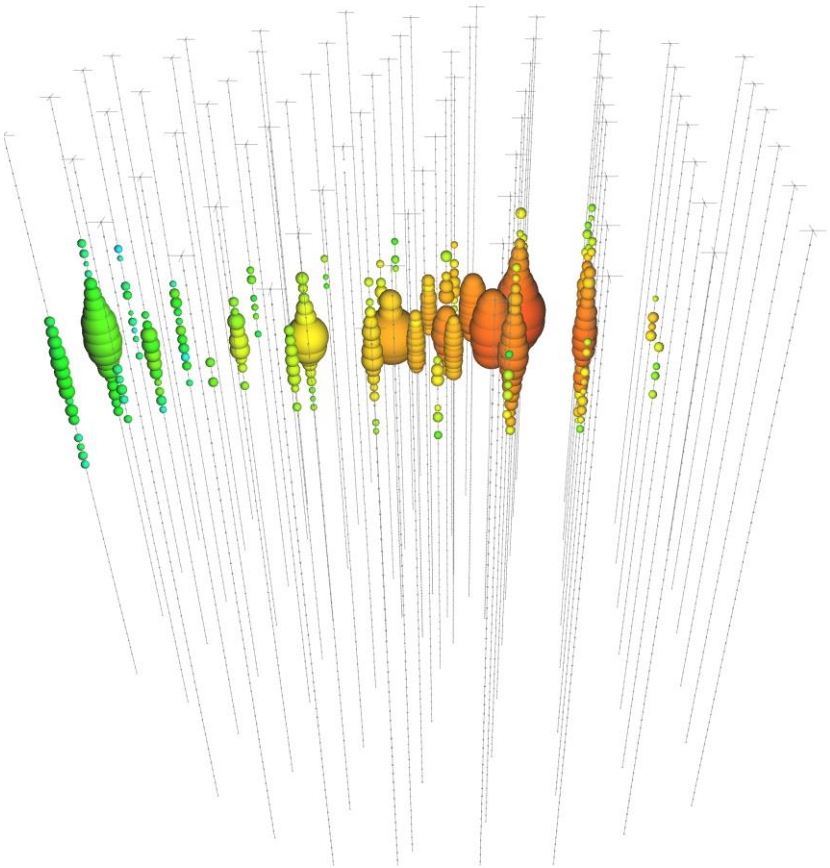
VLVnT 2018

Dubna – October 4th, 2018

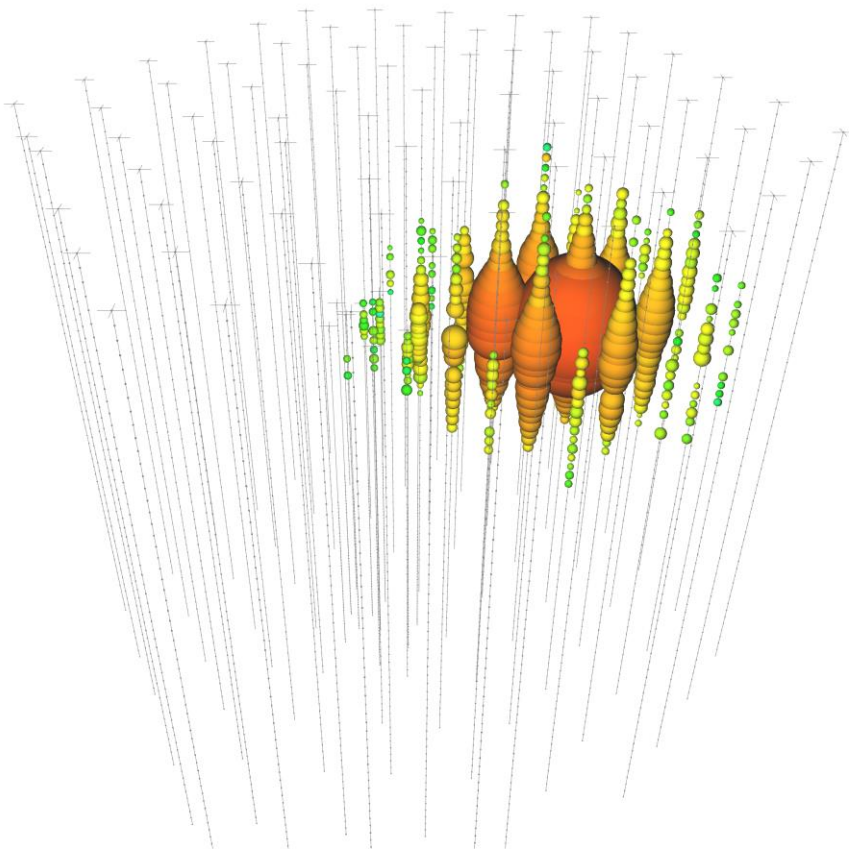
*http://icecube.wisc.edu

IceCube Data – Starting Point of Event Reconstruction

Tracks



Cascades



Can we further improve our event reconstruction?

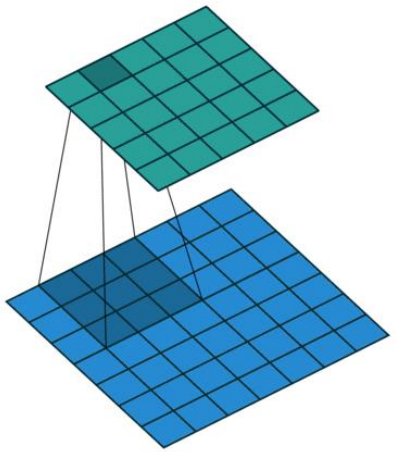
Why Deep Learning?

- Powerful and fast reconstruction methods desired for real-time analyses and follow-up programs
  - Limited resources
  - High data rate
  - Low level data
- Maximum-Likelihood methods have their limitations:
  - Often simplifications and approximations necessary
  - Can be extremely expensive to compute

Deep learning is an excellent candidate for these requirements!

Outline

- Brief introduction to neural networks
- Current Status:
  - Convolutional Neural Network (CNN)
  - Comparison to standard reconstruction methods
- Limitations of CNNs:
  - Need for better suited architectures
  - Exploitation of a priori knowledge
- Next Generation:
  - Generative Neural Network
  - First results



Deep Learning – Deep Neural Networks

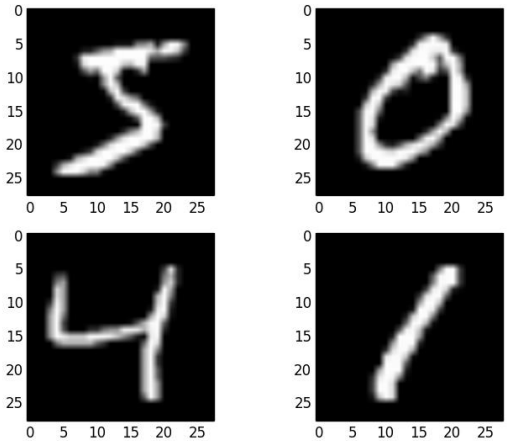
- Neural network defines a function:

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

θ: Free parameters defined by model architecture

I: Input data
Greyscale values of pixels (image recognition)
Pulse information of DOMs (IceCube)

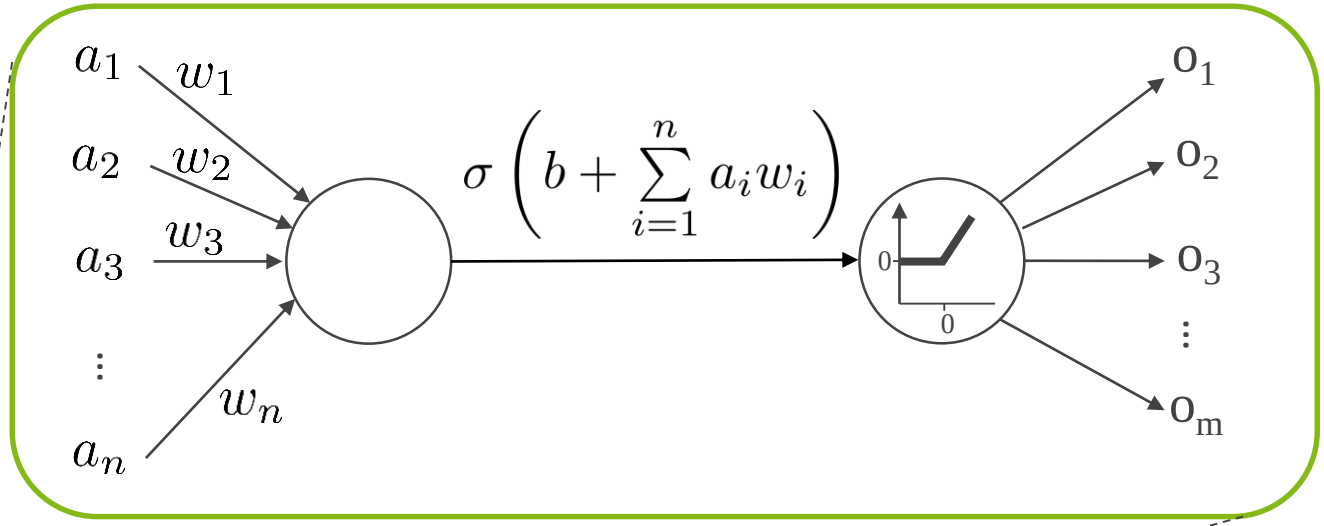
O: Output
Digit (image recognition)
Energy/direction of particle (IceCube)

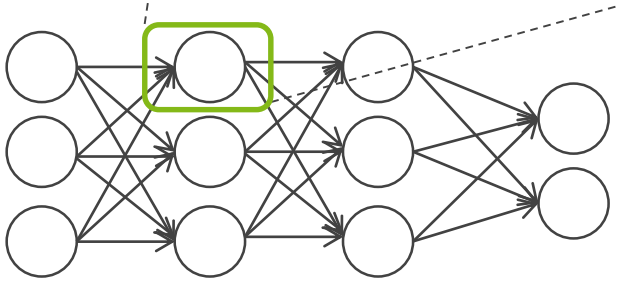


MNIST Dataset

Deep Learning – Deep Neural Networks

Artificial Neuron and fully connected layer





Weights and bias

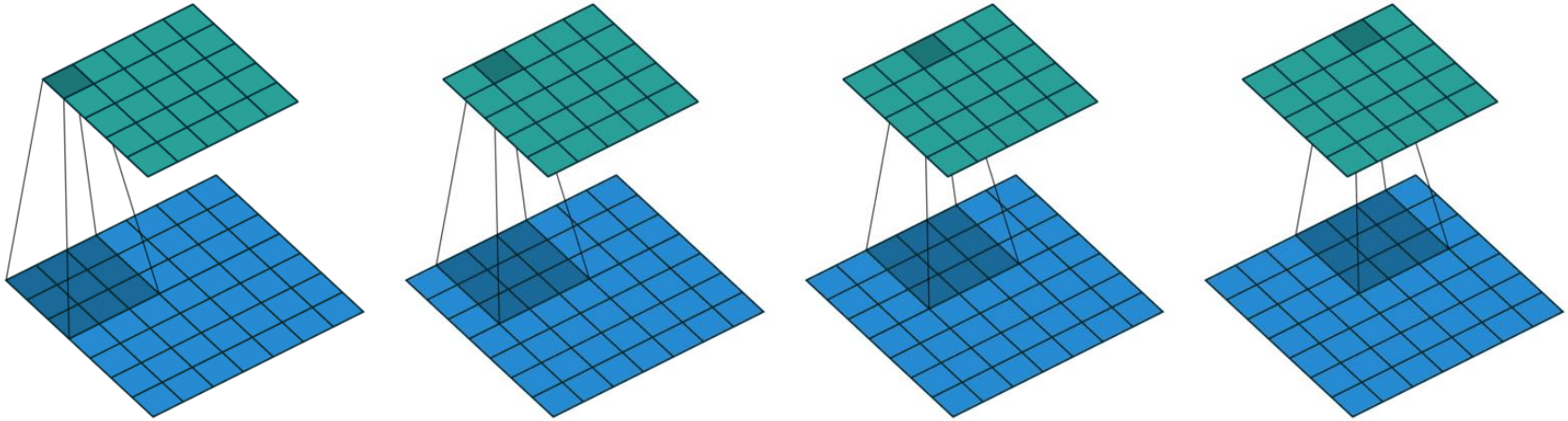
$n + 1$ free parameters per neuron

Nonlinear activation function e.g. ReLU

0 up to a fixed threshold

Deep Learning – Convolutional Neural Nets

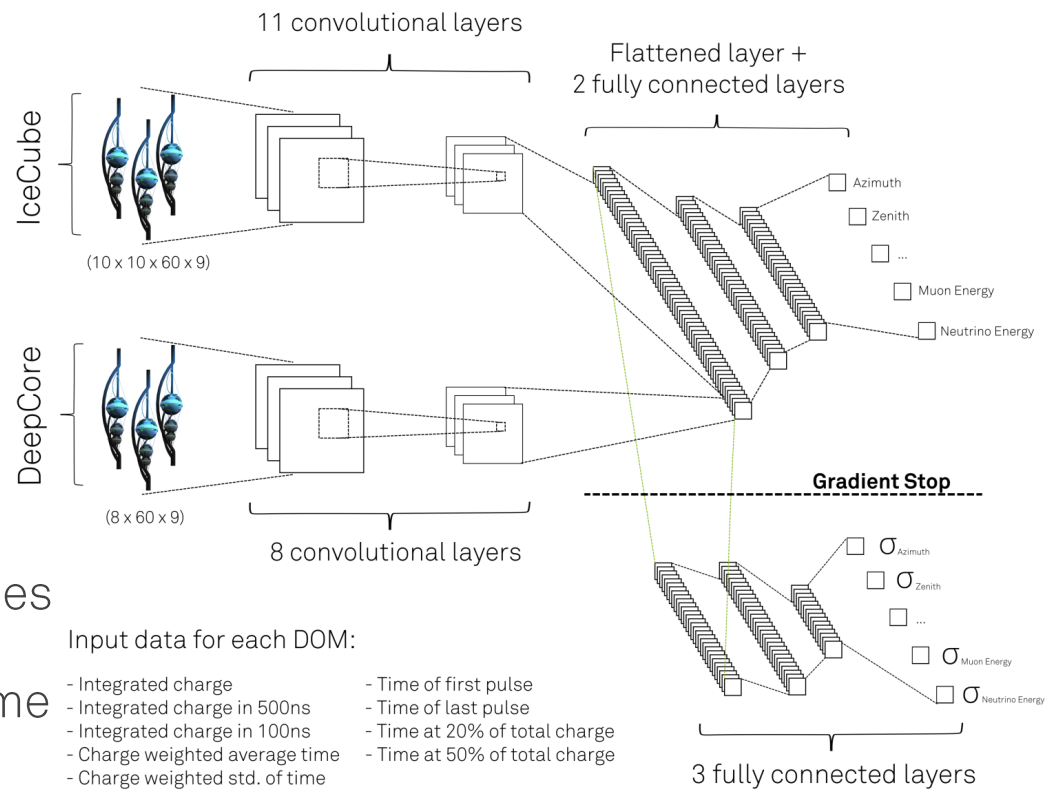
Convolutional Layer



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

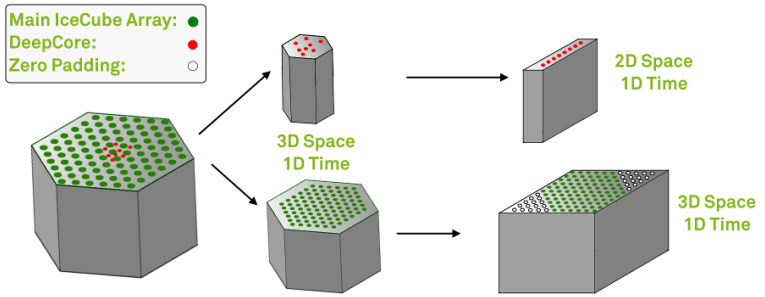
Current Status – CNNs

- Deep convolutional neural network (CNN)
- Energy and directional reconstruction for any event topology
- Easily extendable to other desired reconstruction quantities
- Fast and nearly constant runtime (data preprocessing adds energy dependence)
- Per event uncertainty estimate on each reconstruction quantity



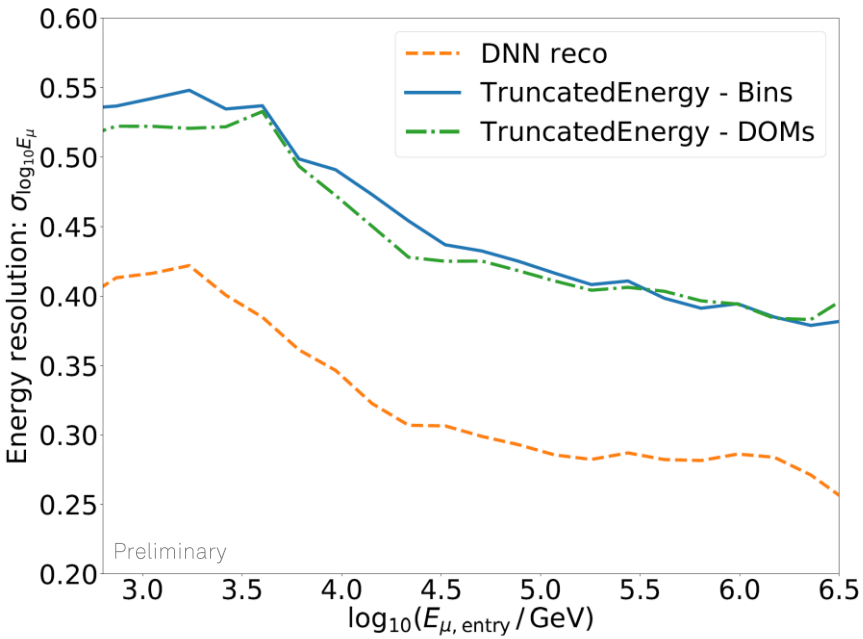
Input data for each DOM:

- 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

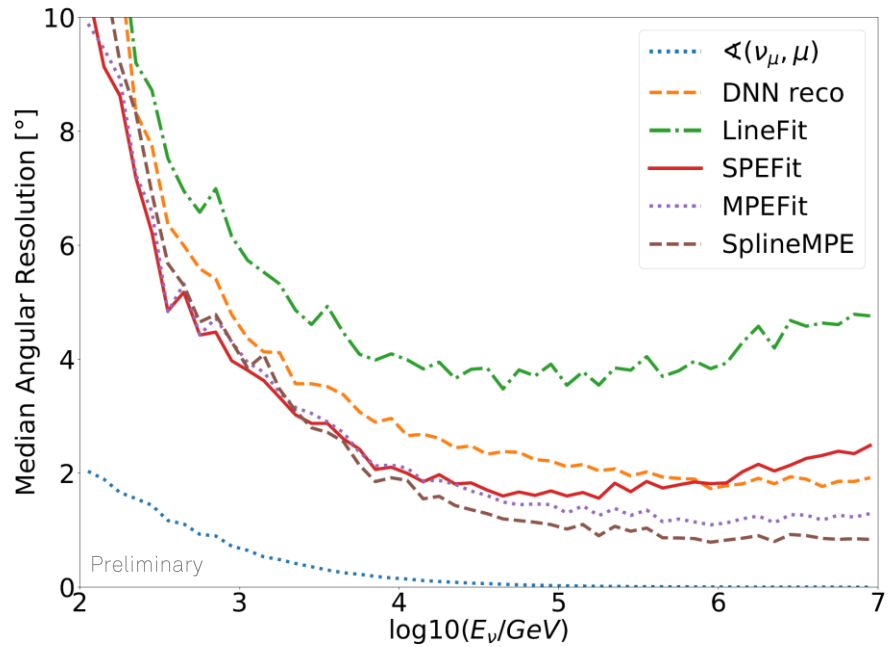


Current Status – Track Performance

Muon Energy at Detector Entry



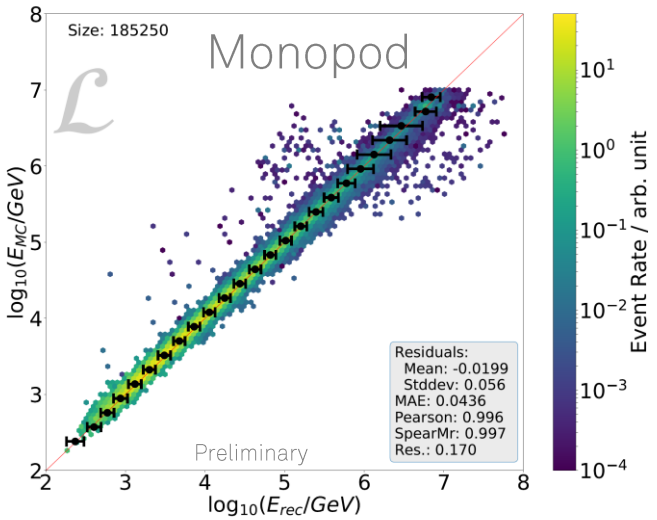
Median Angular Resolution

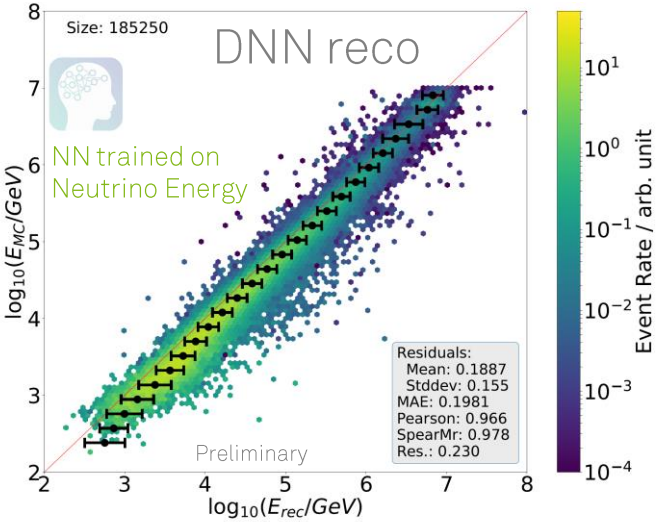


- Level 2 data: Final samples apply additional quality cuts
- Systematic uncertainties not included

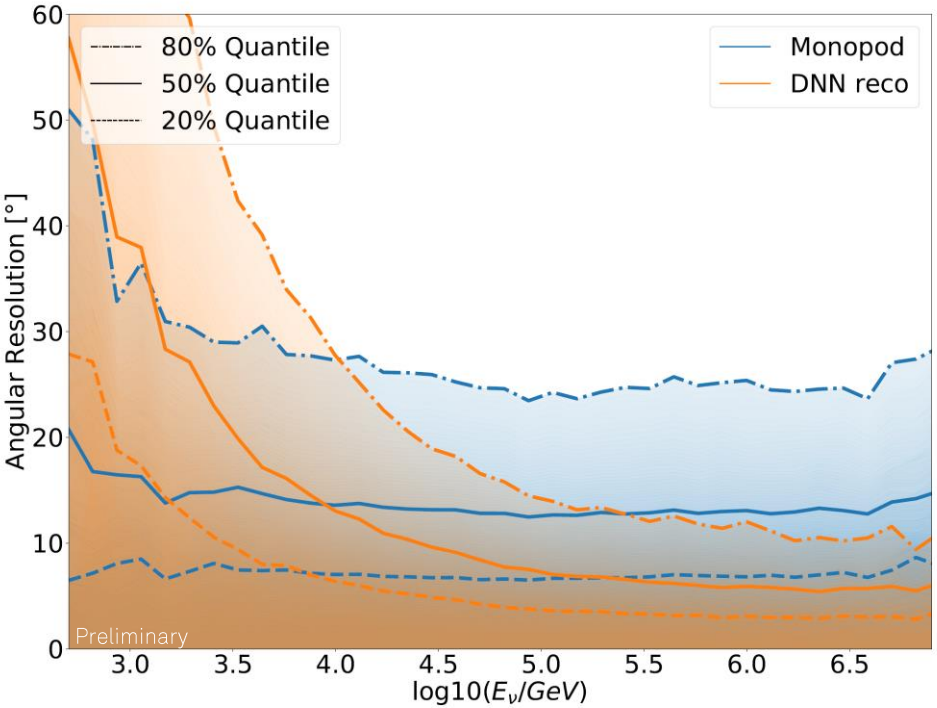
Current Status – Cascade Performance

Deposited Energy





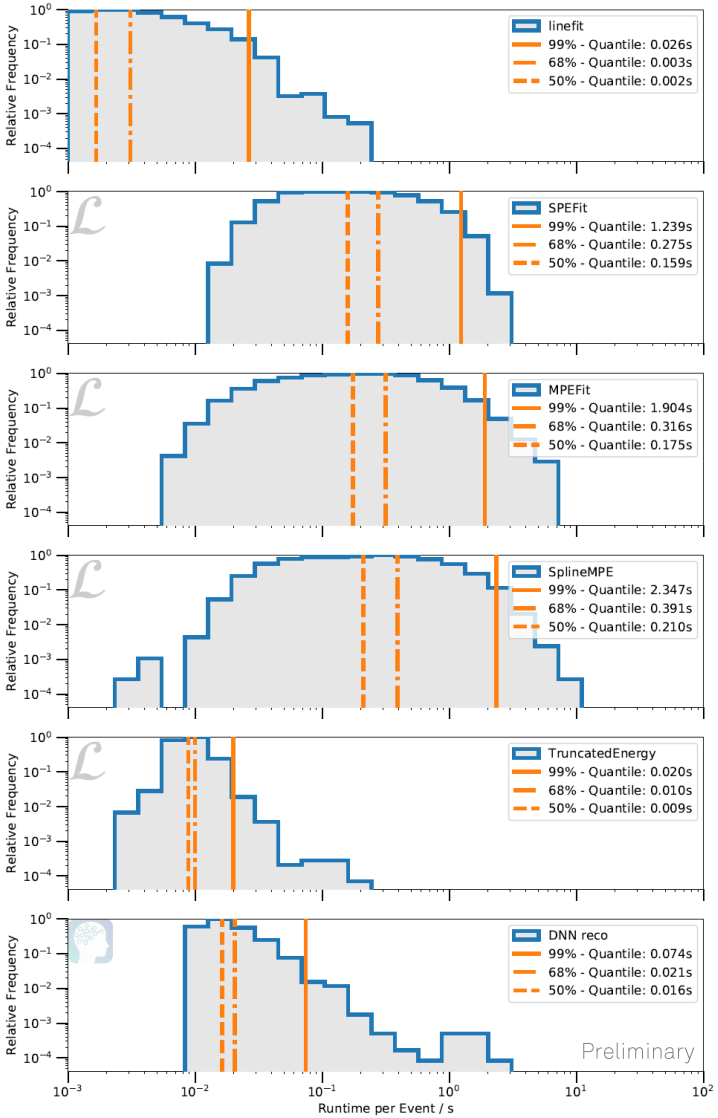
Cascade Angular Resolution



- Systematic uncertainties not included
- Final samples may apply additional quality cuts

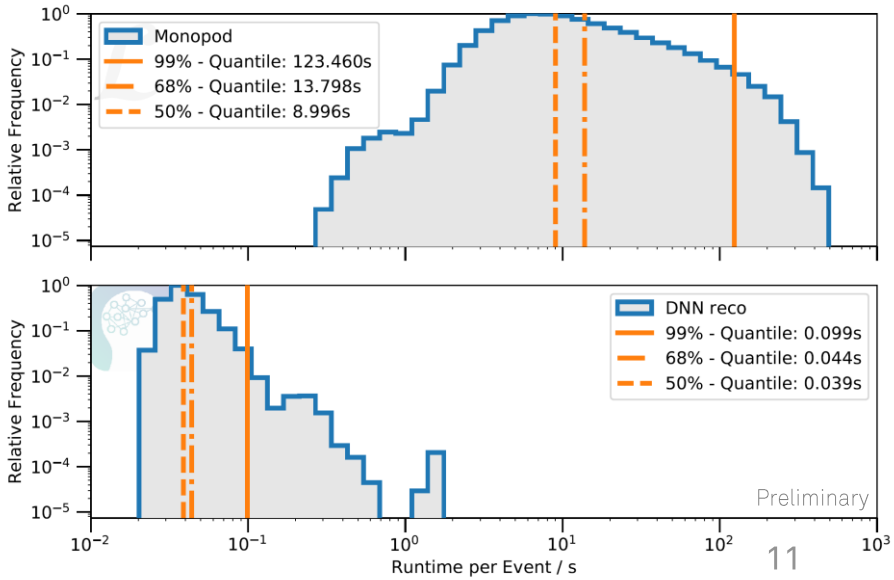
Current Status – Runtime

Track Reconstructions



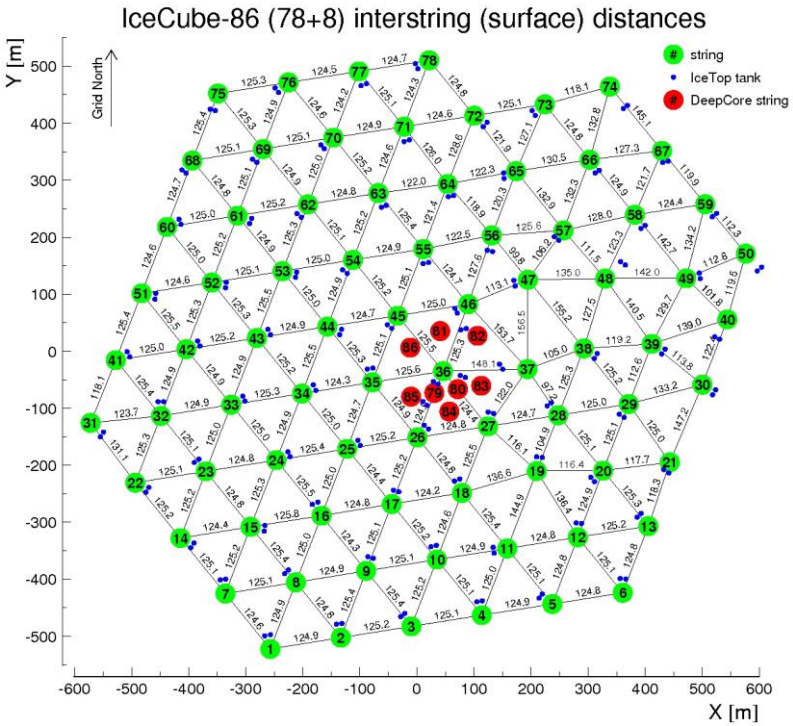
- Per event runtime for reconstruction
- More sophisticated reconstructions excluded in plots
- Smaller neural network architecture used for track reconstruction

Cascade Reconstructions



Limitations of Convolutional Neural Networks (CNNs)

- Only translational invariance and locality is used – More information and symmetries available
- Assumptions imposed by CNNs are only approximately met in IceCube
  - Irregularities in detector grid
  - No real translational invariance in observable space
- CNNs "wash out" data:
  - Great for robustness
  - Bad if exact information is needed

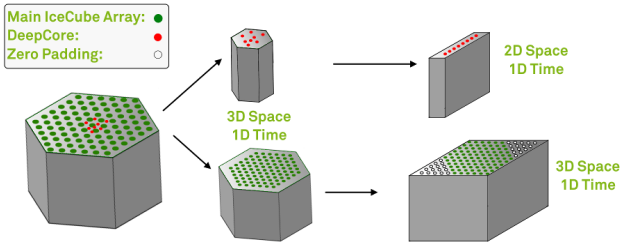


How can we improve?

Requirements for Next Generation

- Need for better suited network architectures:
  - Graph Neural Networks (GNNs) can help with irregularities in detector grid and uniformly handle IceCube + DeepCore
  - Combination of CNNs and recurrent neural networks can better handle time information
- Need to exploit more a priori knowledge:
  - Data generation process is known very well, but not exploited
  - Likelihood methods use this information

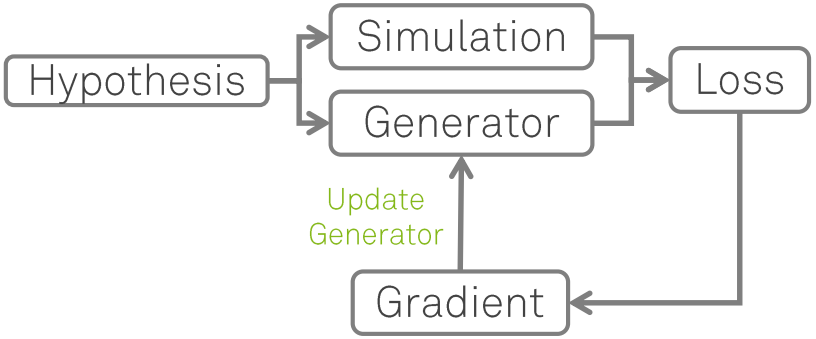
Can these approaches be combined?



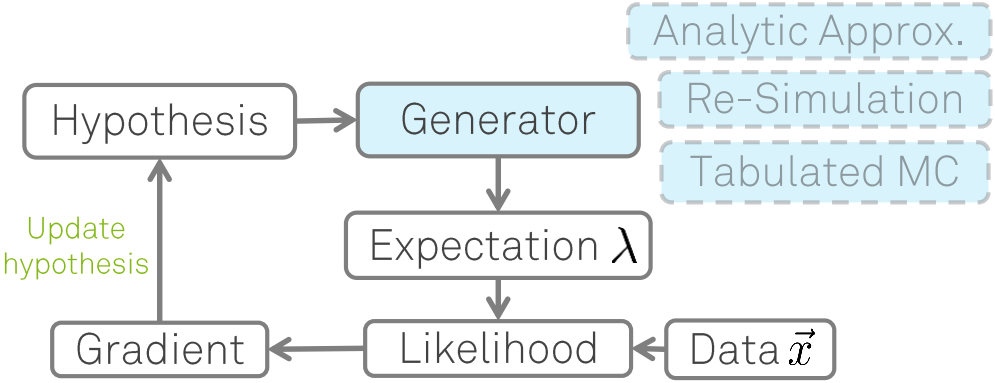
Cascade Generator – General Idea

- Combine strengths of neural networks and maximum-likelihood methods
- Generative network to obtain fast approximation of simulation
- Possible in between of spline-based reconstruction and direct re-simulation
- Once generator is trained, it can be used in reverse mode for reconstruction

Train Generator

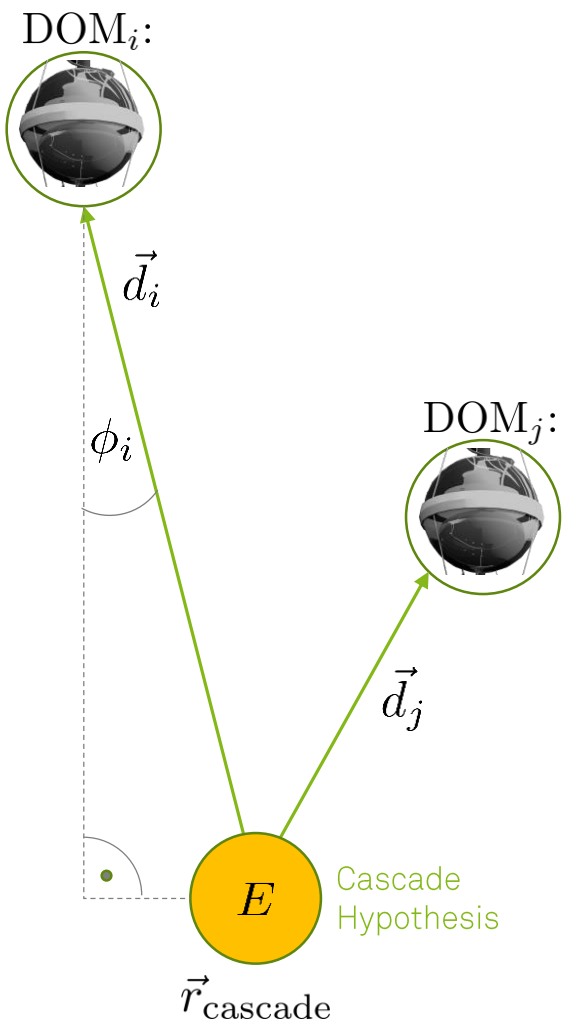


Reconstruct Events



Cascade Generator – Network Architecture

- No adversarial or randomness needed as usually required for Generative Adversarial Networks
- Knowledge of exact detector geometry is included
- Translational invariance in latent variable space is exploited as opposed to observable space
- Currently only trained to generate charge information



Input per DOM:

$\vec{d}, \phi, \ldots$ } Include geometry

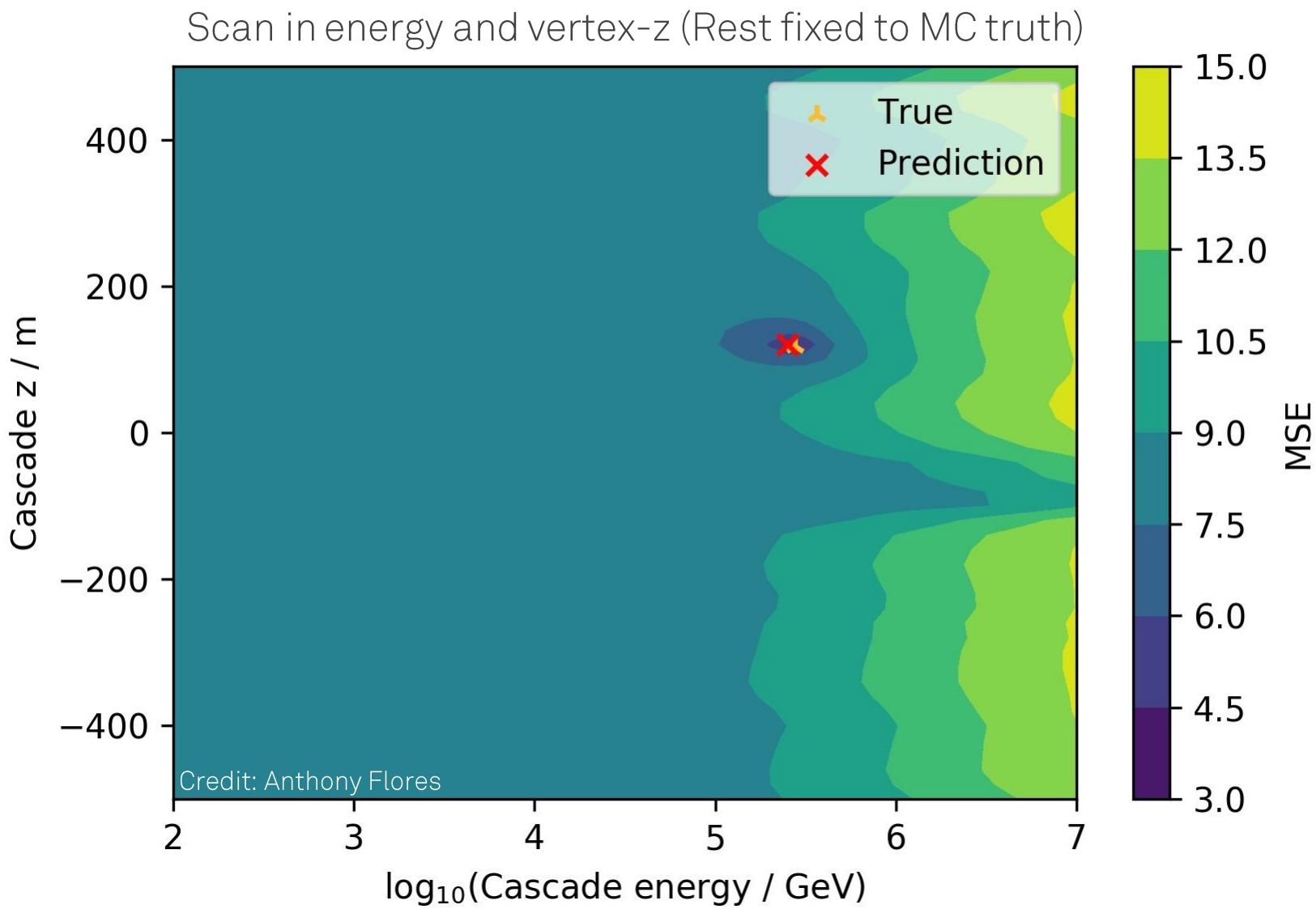
$\varphi_{\mathrm{cascade}}$
$\theta_{\mathrm{cascade}}$
$\vec{r}_{\mathrm{cascade}}$

Weight sharing over DOMs:
Exploit invariances in latent variable space

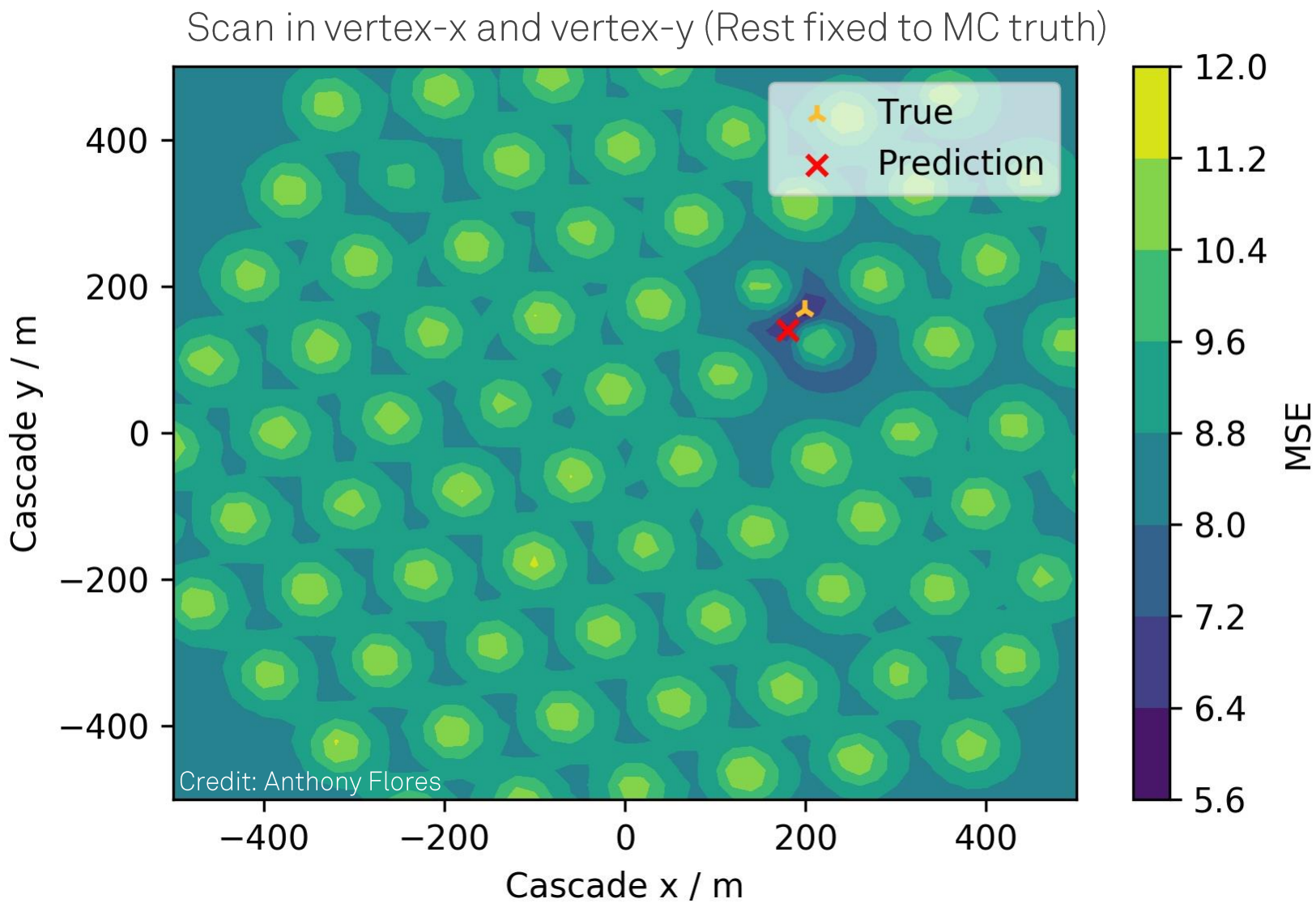
$E, \ldots$

Additional independent weights per DOM:
Account for asymmetries and inhomogeneities

Cascade Generator – Scans



Cascade Generator – Scans



Summary & Outlook

- Conventional neural network architectures provide good results and are well suited for online applications, but have limitations
- Need to adjust methods to fit to our data, not the other way around
- First Cascade Generator results look promising
- Next: create dedicated MC for training, include time information

THE ICECUBE COLLABORATION

AUSTRALIA
University of Adelaide

BELGIUM
Université libre de Bruxelles
Universiteit Gent
Vrije Universiteit Brussel

CANADA
SNOLAB
University of Alberta–Edmonton

DENMARK
University of Copenhagen

GERMANY
Deutsches Elektronen-Synchrotron
ECAP, Universität Erlangen-Nürnberg
Humboldt–Universität zu Berlin
Ruhr-Universität Bochum
RWTH Aachen University
Technische Universität Dortmund
Technische Universität München
Universität Mainz
Universität Wuppertal
Westfälische Wilhelms-Universität Münster

JAPAN
Chiba University

NEW ZEALAND
University of Canterbury

REPUBLIC OF KOREA
Sungkyunkwan University

SWEDEN
Stockholms Universitet
Uppsala Universitet

SWITZERLAND
Université de Genève

UNITED KINGDOM
University of Oxford

UNITED STATES
Clark Atlanta University
Drexel University
Georgia Institute of Technology
Lawrence Berkeley National Lab
Marquette University
Massachusetts Institute of Technology
Michigan State University
Ohio State University
Pennsylvania State University
South Dakota School of Mines and Technology
Southern University and A&M College
Stony Brook University
University of Alabama
University of Alaska Anchorage
University of California, Berkeley
University of California, Irvine
University of Delaware
University of Kansas
University of Maryland
University of Rochester
University of Texas at Arlington
University of Wisconsin–Madison
University of Wisconsin–River Falls
Yale University

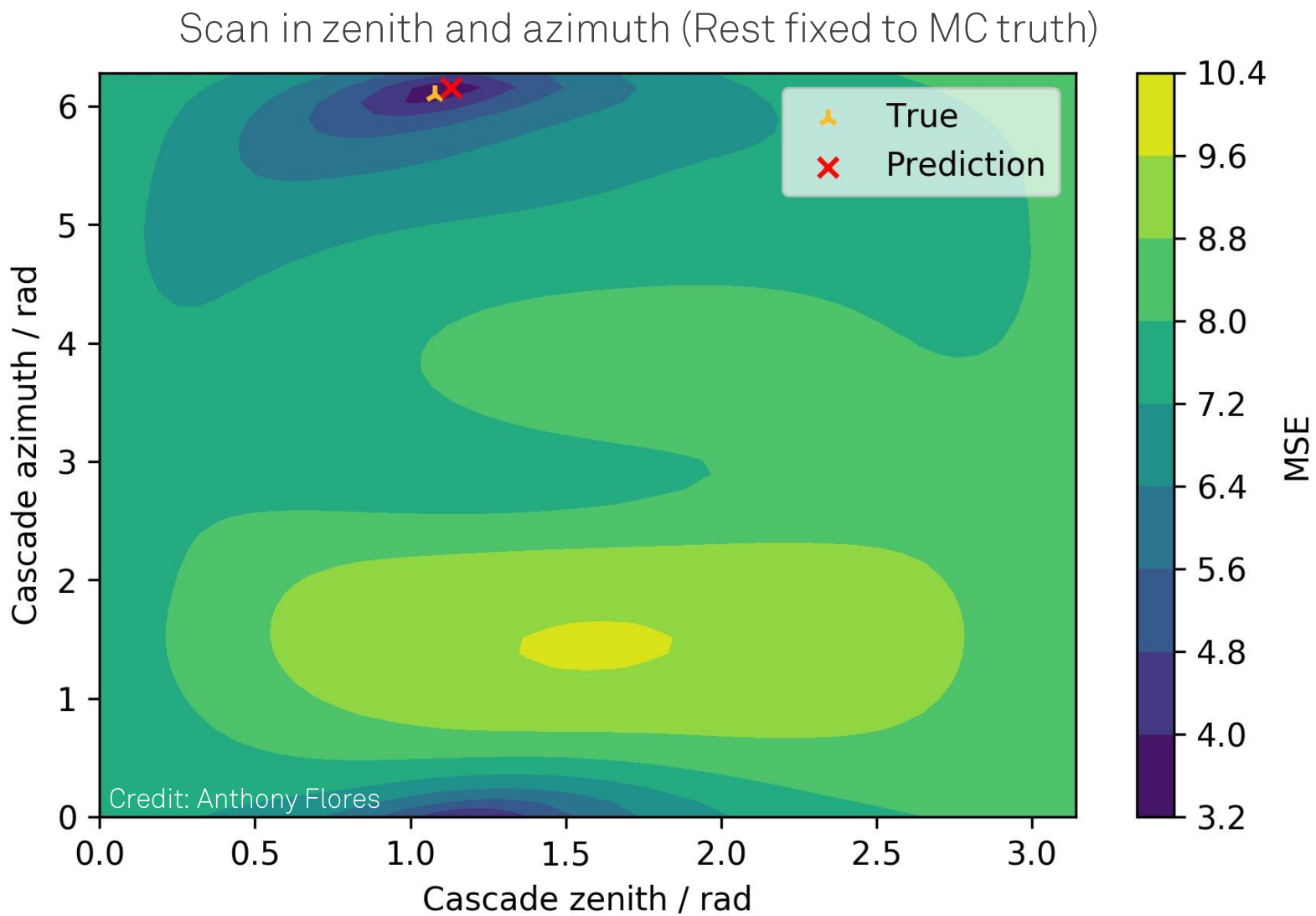
FUNDING AGENCIES

Fonds de la Recherche Scientifique (FRS-FNRS)
Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO-Vlaanderen)
Federal Ministry of Education and Research (BMBF)
German Research Foundation (DFG)
Deutsches Elektronen-Synchrotron (DESY)
Japan Society for the Promotion of Science (JSPS)
Knut and Alice Wallenberg Foundation
Swedish Polar Research Secretariat
The Swedish Research Council (VR)
University of Wisconsin Alumni Research Foundation (WARF)
US National Science Foundation (NSF)

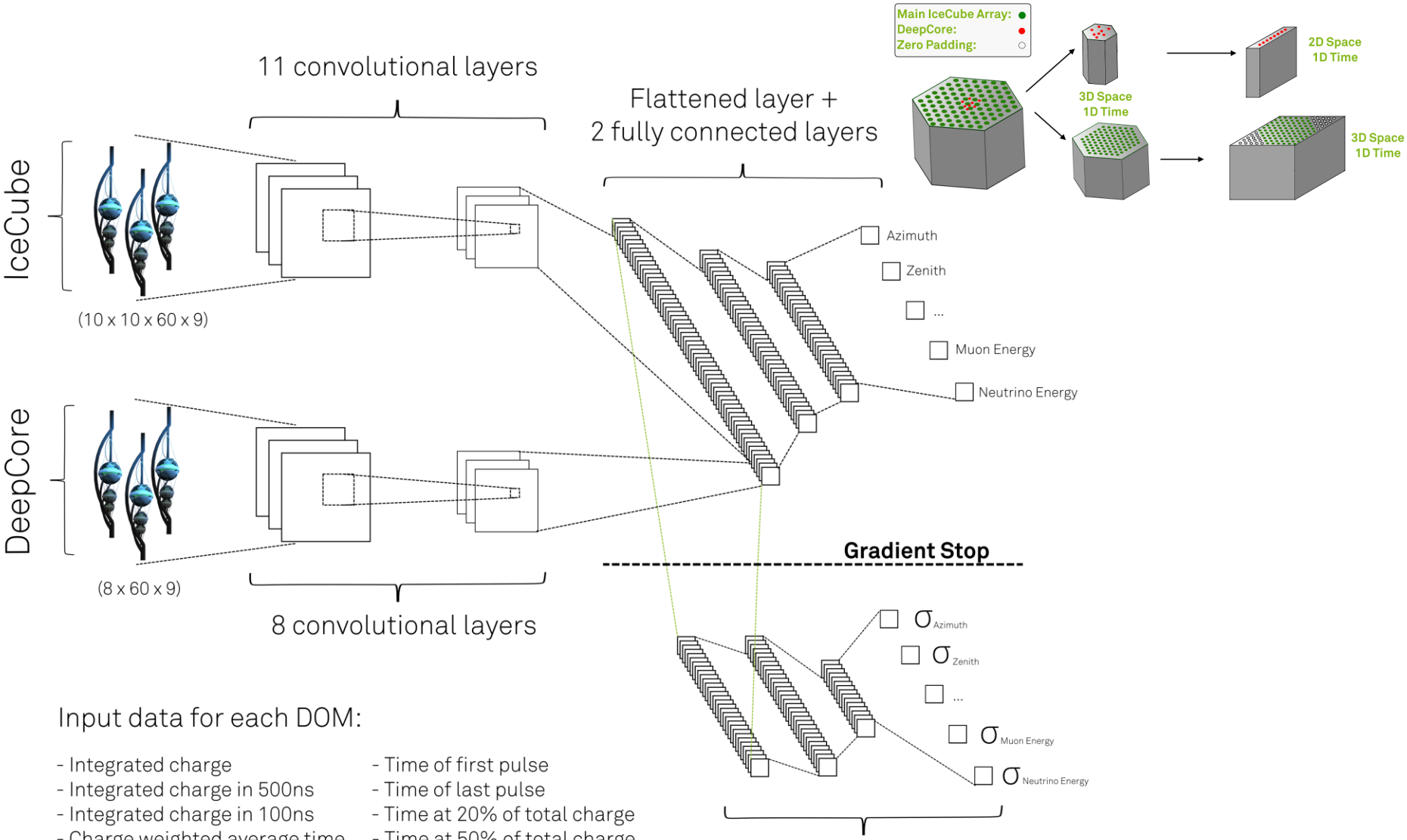
ICECUBE
SOUTH POLE NEUTRINO OBSERVATORY
icecube.wisc.edu

Backup

Cascade Generator – Scans



Network Architecture



Input data for each DOM:

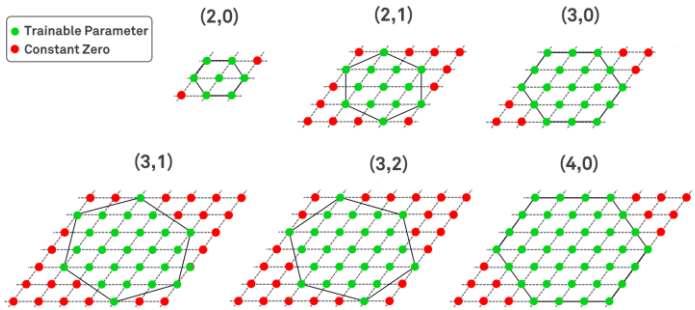
- Integrated charge
- Integrated charge in 500ns
- Integrated charge in 100ns
- Charge weighted average time
- Charge weighted std. of time
- Time of first pulse
- Time of last pulse
- Time at 20% of total charge
- Time at 50% of total charge

Network Architecture

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

  DOI: 10.1109/CVPR.2016.90

- Hexagonally shaped convolution kernels

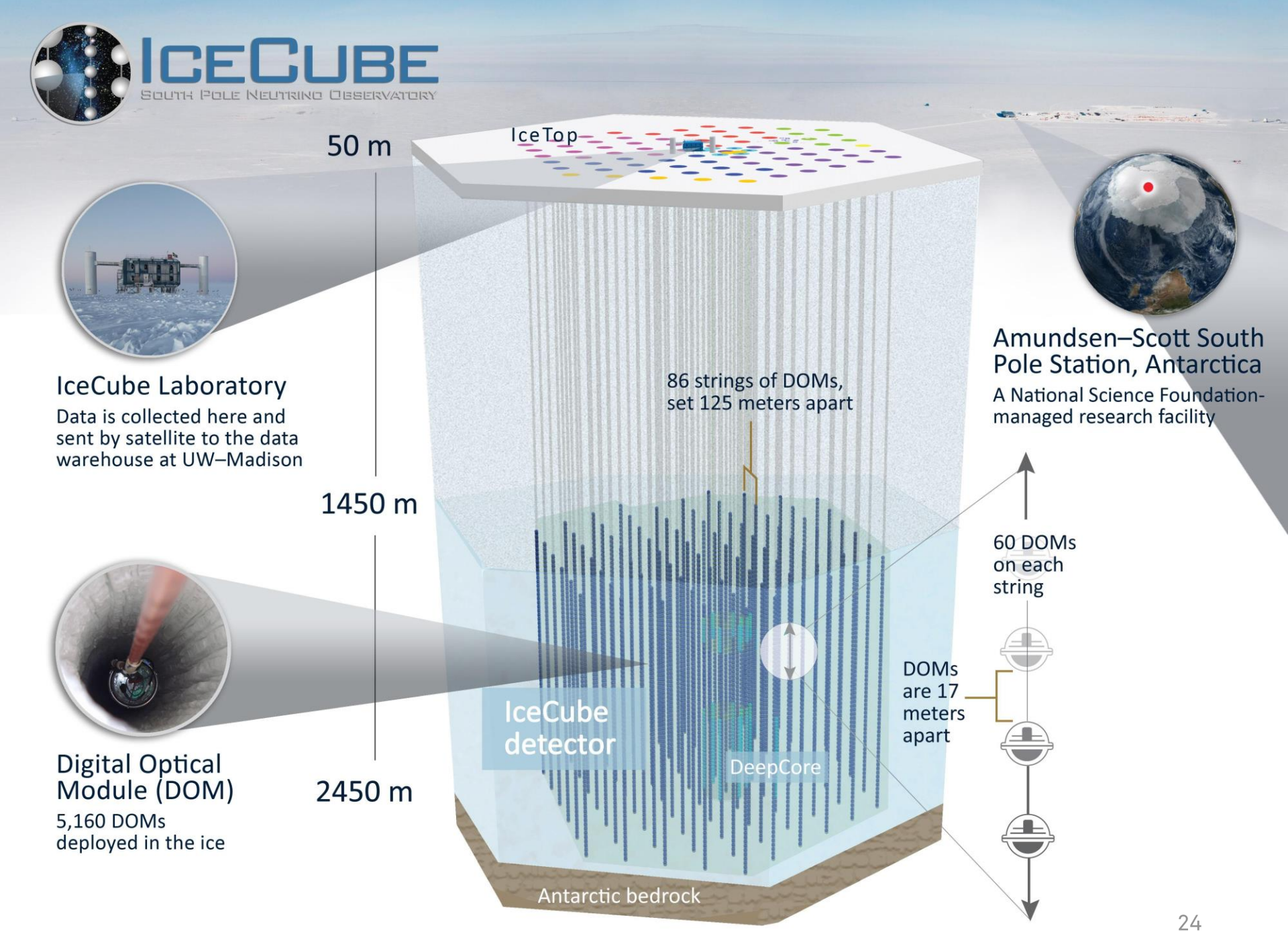


- Normalization of input and labels to mean 0 and variance 1

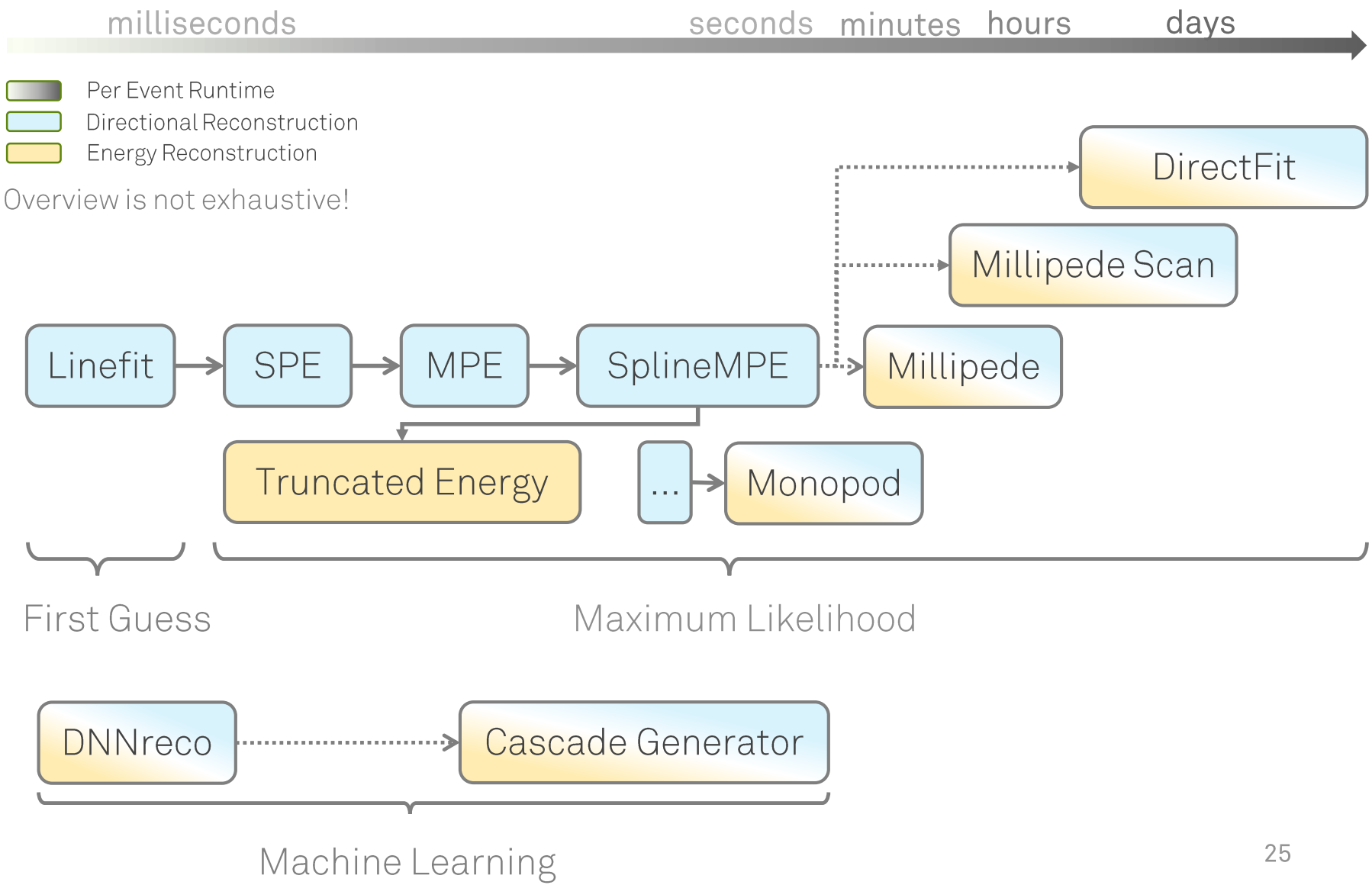
- Variance control in layers
  - Assuming input into a layer is normalized:

    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


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

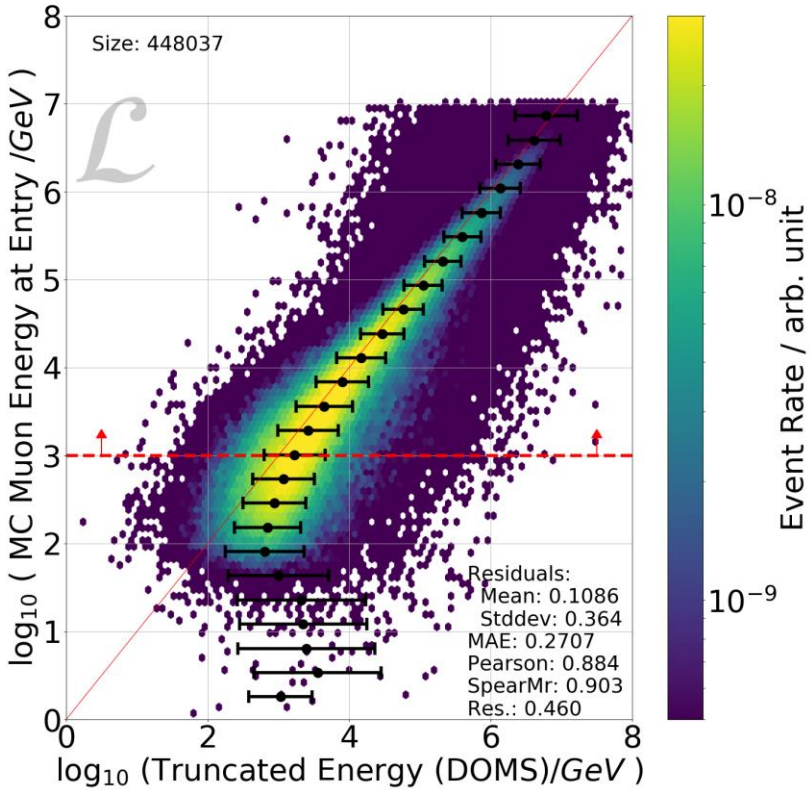
Overview of Reconstruction Methods



Energy Reconstruction – Muon Energy at Entry

OnlineL2 Muon Filter – CC events

Truncated Energy (DOMS)



Deep Learning