Event reconstruction for KM3NeT/ORCA using convolutional neural networks

Stefan Reck,
GDR meeting, 24.11.2020

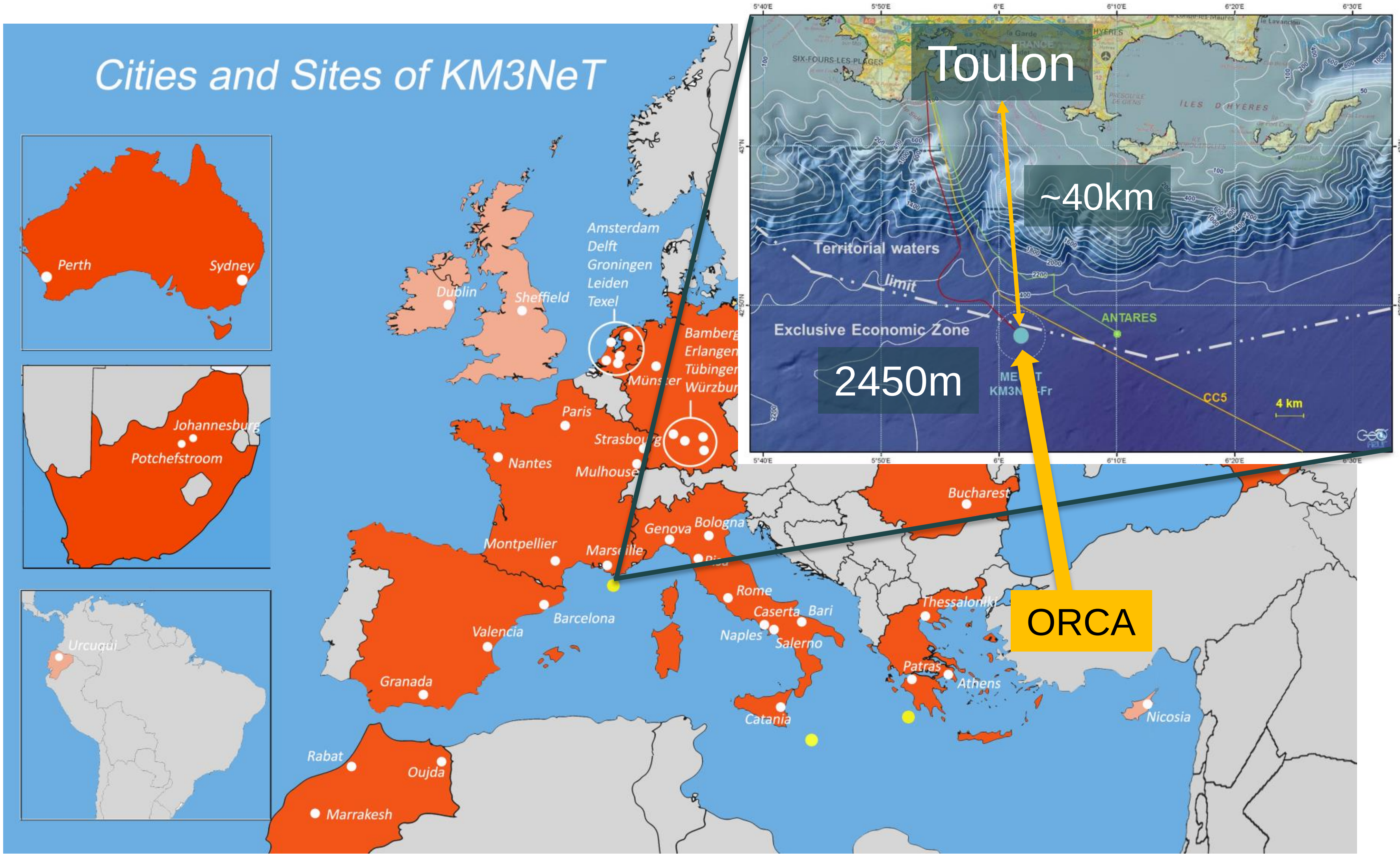
- This talk is based on the paper published in JINST:

  https://arxiv.org/abs/2004.08254

  by primary authors **Michael Moser** and **Thomas Eberl**

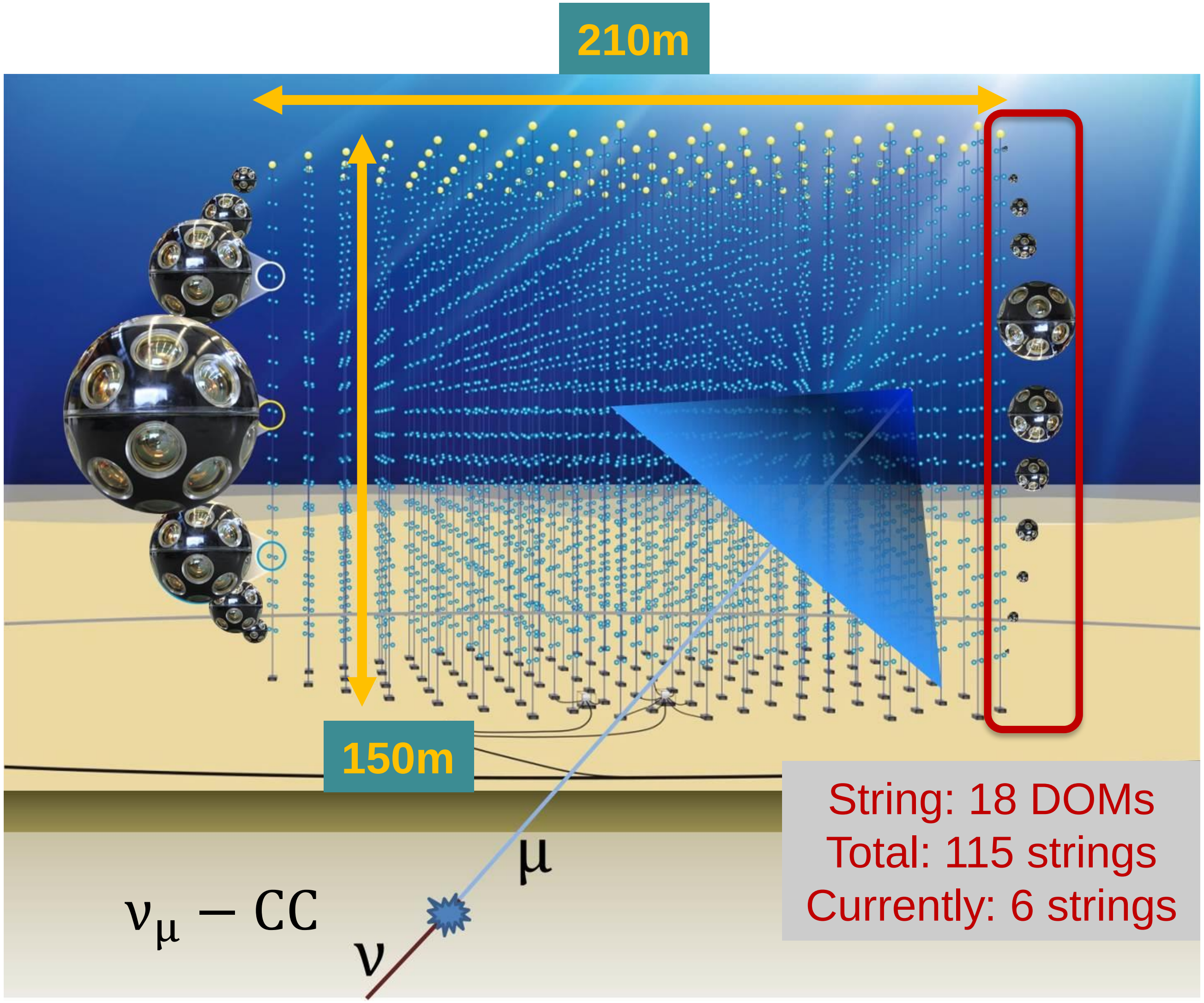
→ First application of a deep convolutional network in water-Cherenkov detector

KM3NeT/ORCA: a neutrino detector



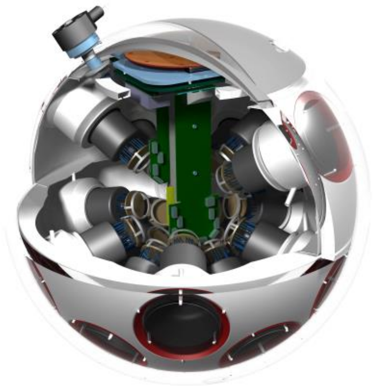
ORCA: a deep sea neutrino detector

Aim: measure atmospheric neutrinos with energies from 1-100 GeV



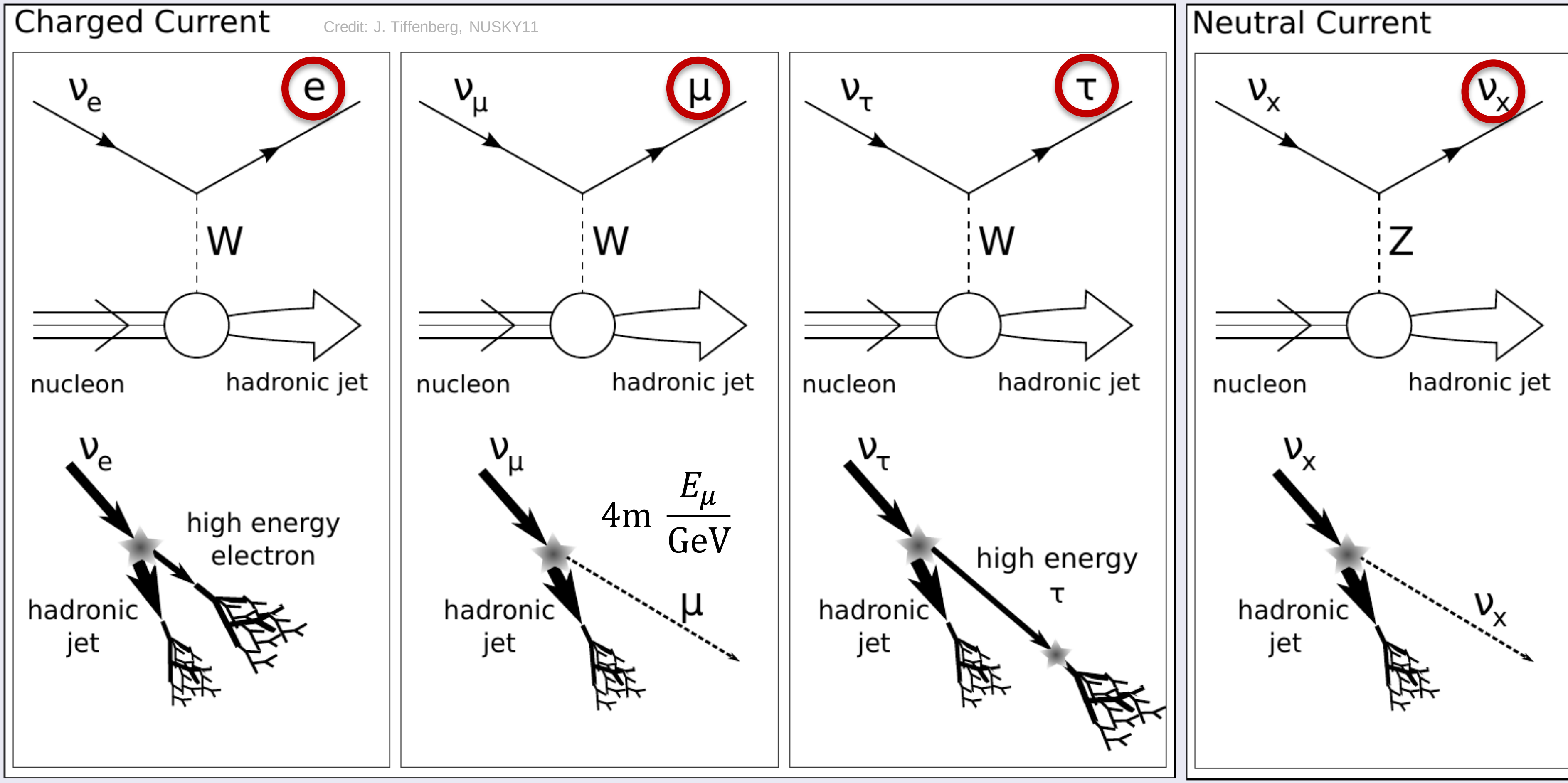
ORCA DOM

43 cm



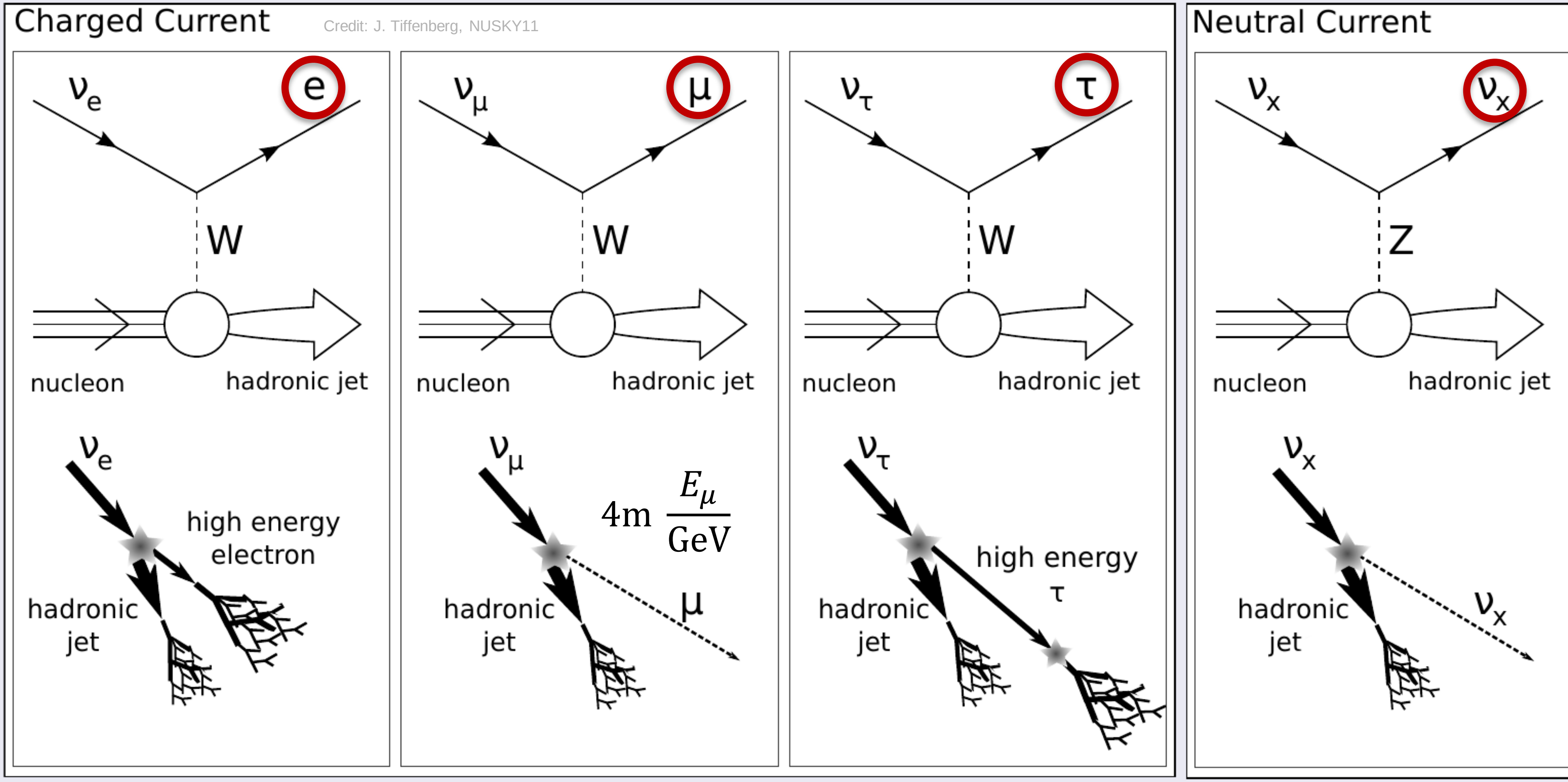
31 PMTs

Neutrino interactions



$\nu_e - \mathrm{CC}$ $\nu_\mu - \mathrm{CC}$ $\nu_\tau - \mathrm{CC}$ $\nu - \mathrm{NC}$

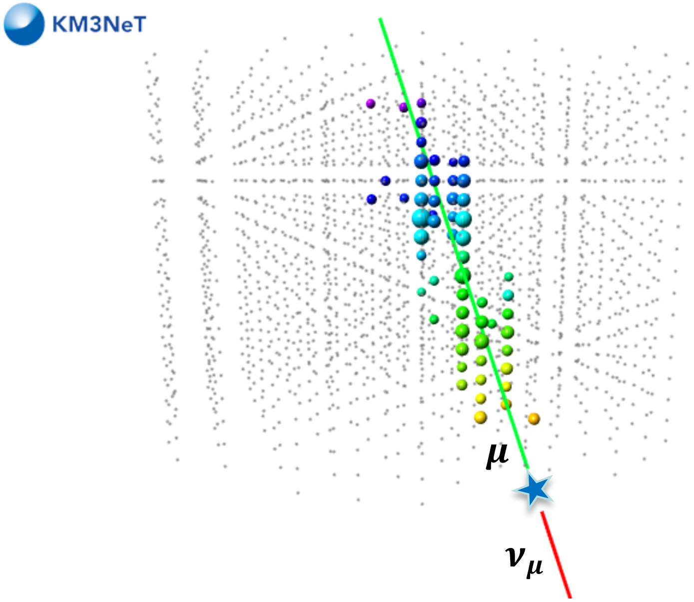
Neutrino interactions



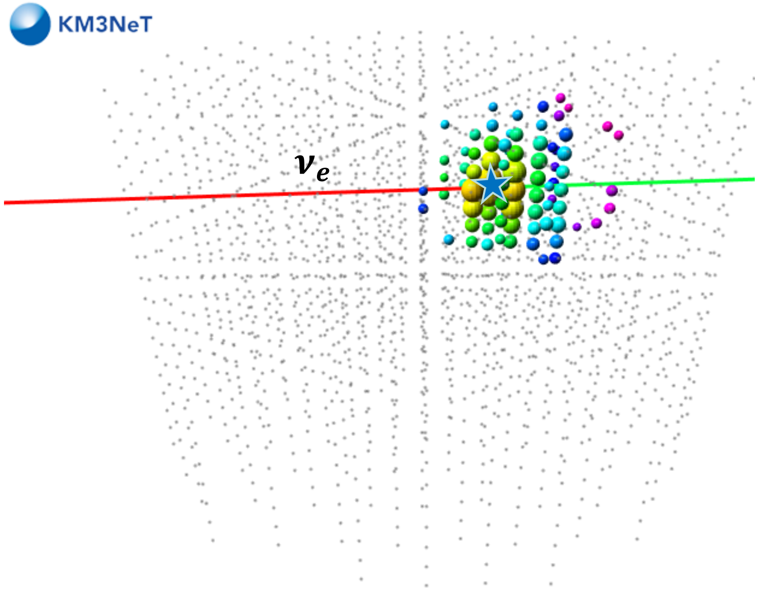
shower-like	track-like	shower-like (83%) track-like (17%)	shower-like

Event topologies

Up-going ν_μ – CC **track**-like event



ν_e – CC **shower**-like event



ORCA event reconstruction

Reconstruction of neutrinos from low-level detector data

1. Discriminate ***background*** from neutrinos
2. Classify neutrino events into ***track***- and ***shower***-like events
3. Reconstruct neutrino properties like ***energy*** and ***direction***

This work:

Perform this based on so-called ***deep learning*** techniques.

ORCA event reconstruction

Reconstruction of neutrinos from low-level detector data

1. Discriminate ***background*** from neutrinos
2. Classify neutrino events into ***track***- and ***shower***-like events
3. Reconstruct neutrino properties like ***energy*** and ***direction***

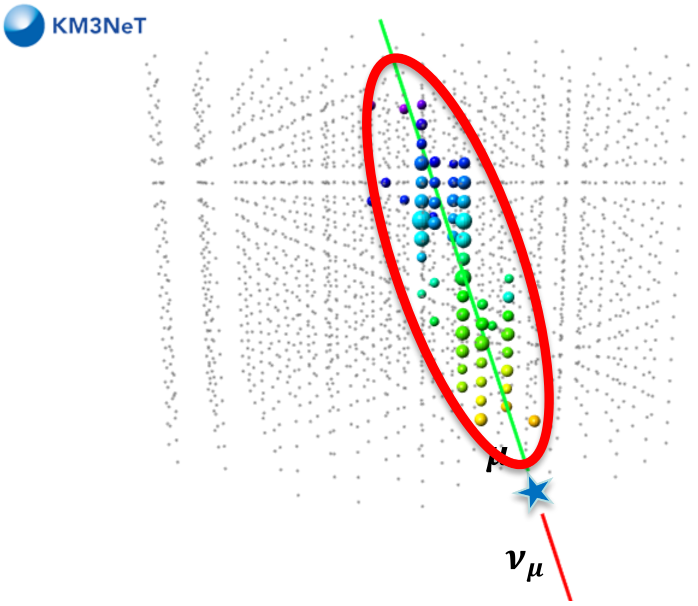
This work:

Perform this based on so-called ***deep learning*** techniques.

Event topology classification

- How to distinguish tracks and showers?
  → Classical way: come up with characteristic features, e.g. **Sphericity**

Up-going ν_μ – CC **track**-like event



ν_e – CC **shower**-like event

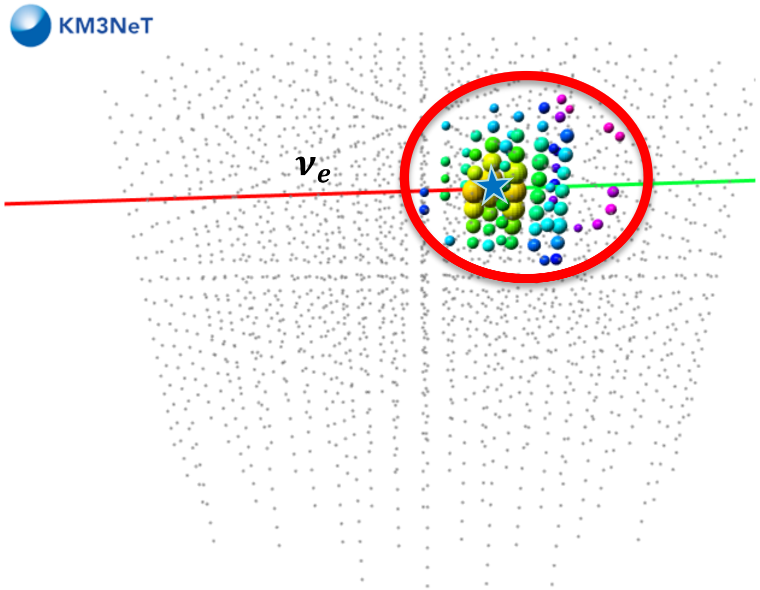


Image taken from [1]

Event topology classification

- Hand features to a machine learning based classifier
- Standard ORCA background classifier uses a Random Forest (RF) to combine the features

- Machine learning algorithm is „trained“ with simulations
- Problem: **feature design is not easy** and maybe we missed some good features?

Deep Learning in ORCA

- Solution: let an algorithm learn the features by itself based on simulations of low-level detector data

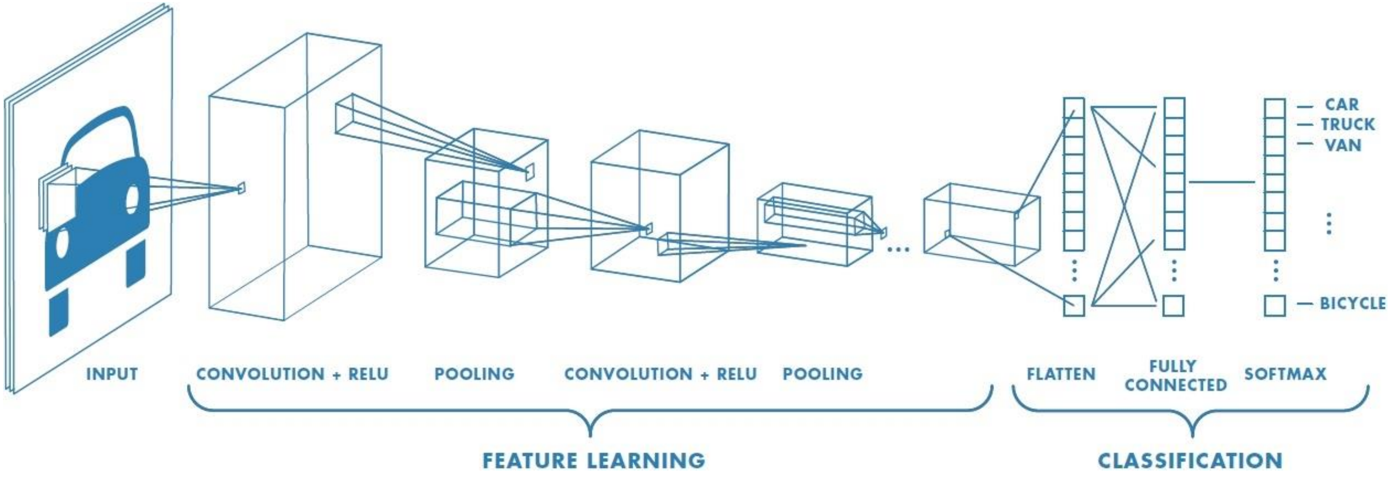
- Possible with recently emerging machine learning algorithms like deep neural networks, also called ***deep learning*** techniques

Deep Learning in ORCA

- How can we apply deep learning methods to ORCA data?

Successful model architecture in image recognition:

Convolutional neural networks (CNNs)



Simplified working principle of a CNN

Source: https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network/_jcr_content/mainParsys/band_copy_copy_14735_1026954091/mainParsys/columns_1606542234_c/2/image.adapt.full.high.jpg/1575485682772.jpg

Deep Learning in ORCA

- ORCA data can be interpreted to be 5D (XYZ, T, 31 PMT channels)

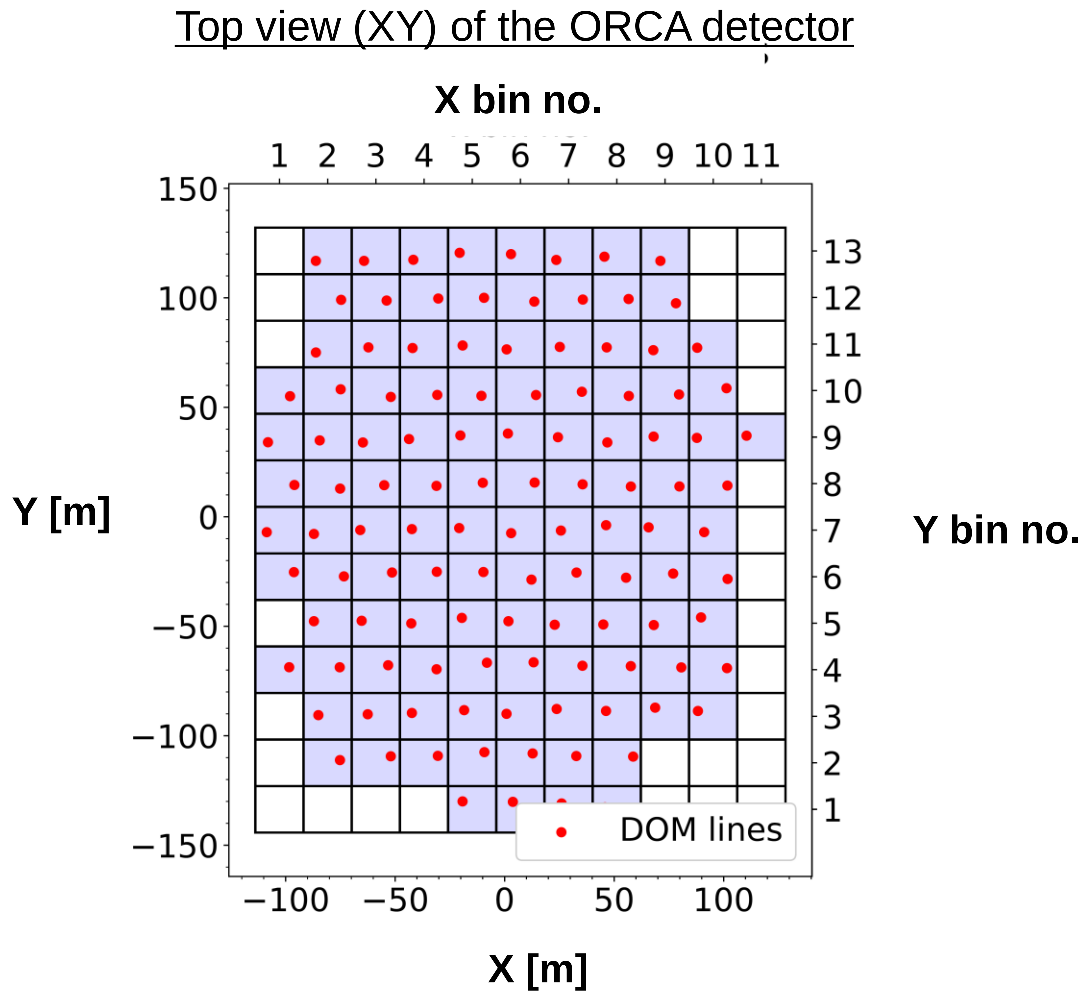
Most Deep Learning frameworks, like Google‘s Tensorflow, only support 4D input (colored videos)

- CNN expects image-like, pixelated input

➢ Bin XYZT dimensions to get pixelated **event images**

ORCA event images

- Spatial binning (XYZ): 1 DOM / 3D pixel

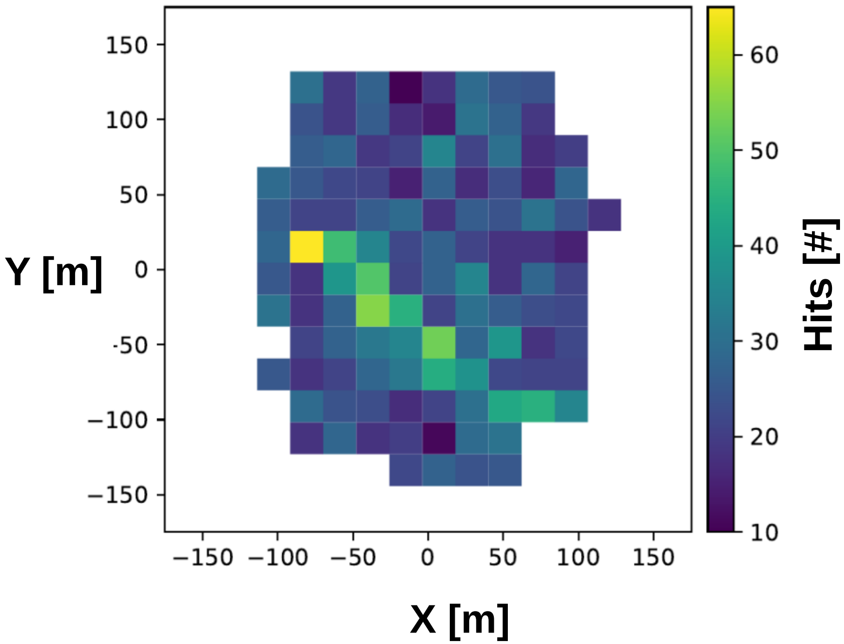


ORCA event images

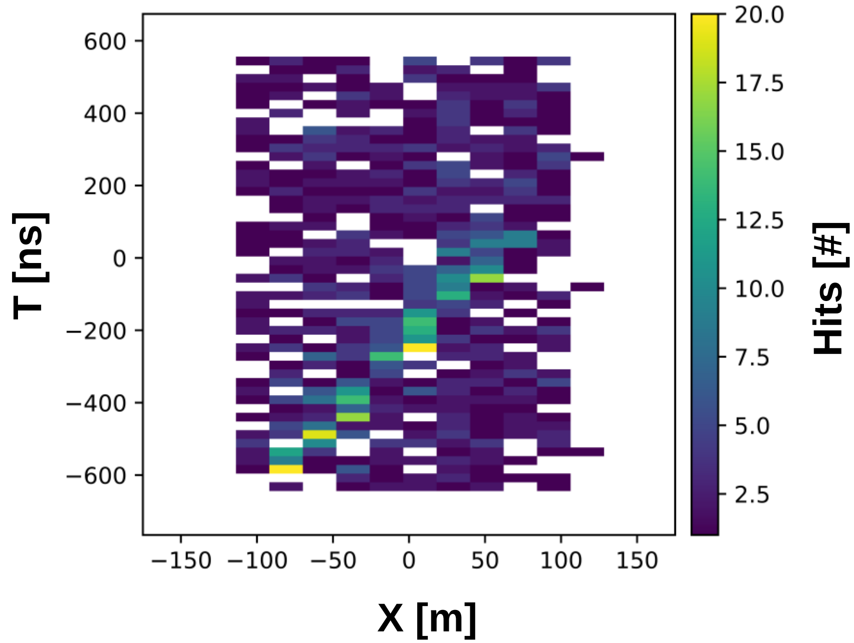
- Time binning:
  - 100 bins → ~ 9.5ns / bin
  - Number of bins limited by computational cost

XY & XT projection of the 4D XYZT „image" for an 80 GeV ν_μ - CC event

XY projection



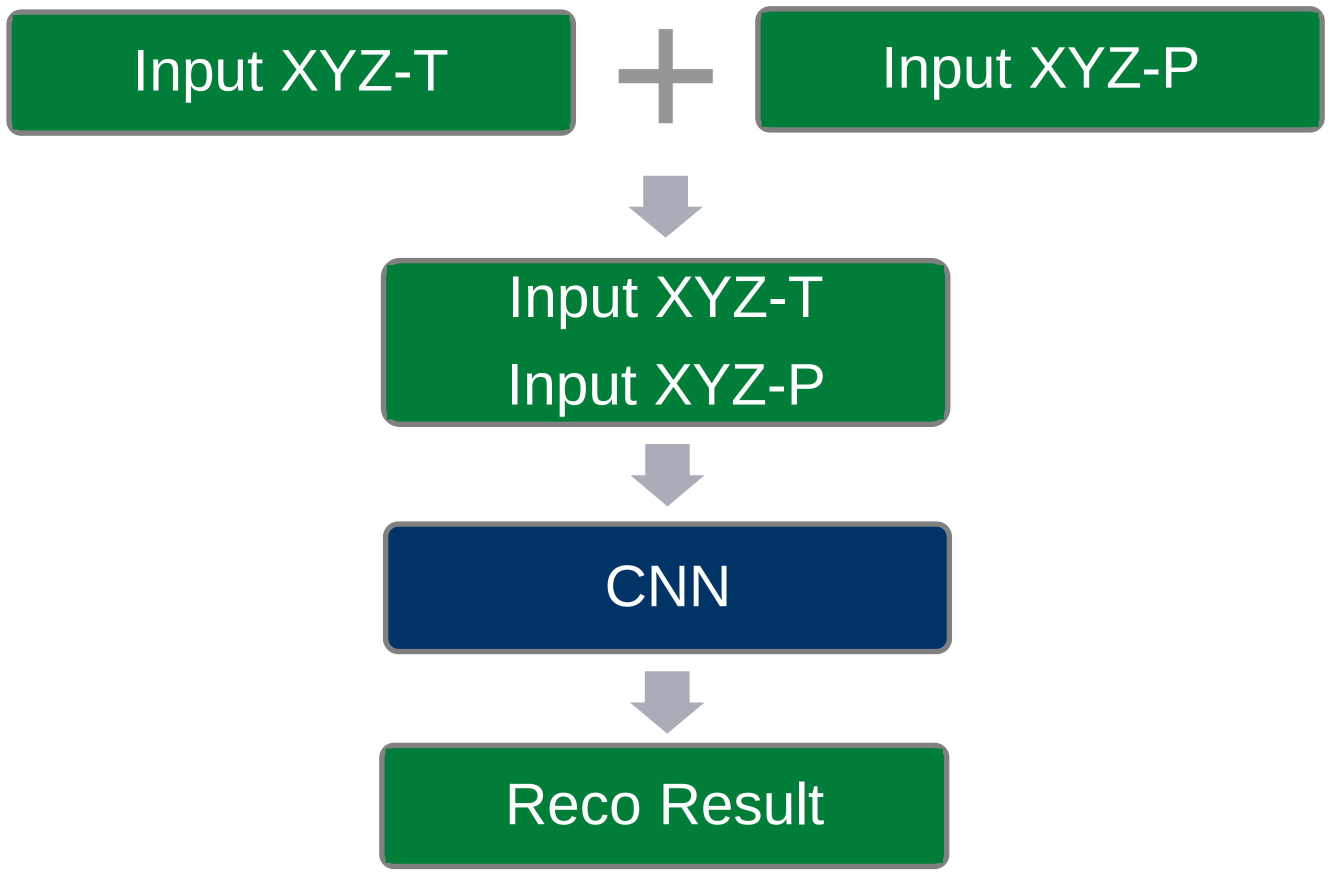
XT projection



ORCA event images

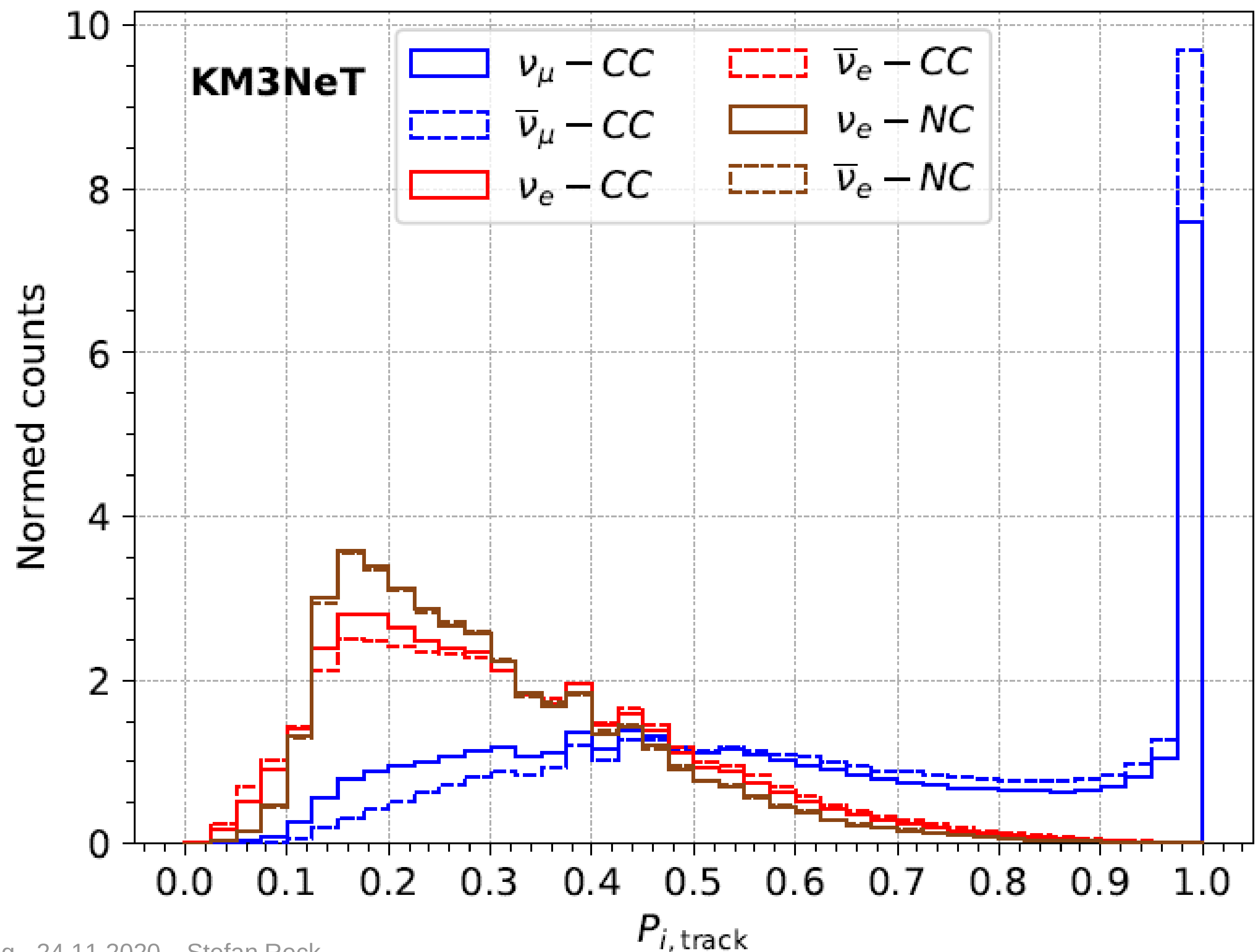
- How to use convolutions on our 5D XYZT-P input?

  → Supply network with two **4D projections**



Event topology classification

- Network output: Probability of event being a track



Event topology classification

- Separability between track and shower

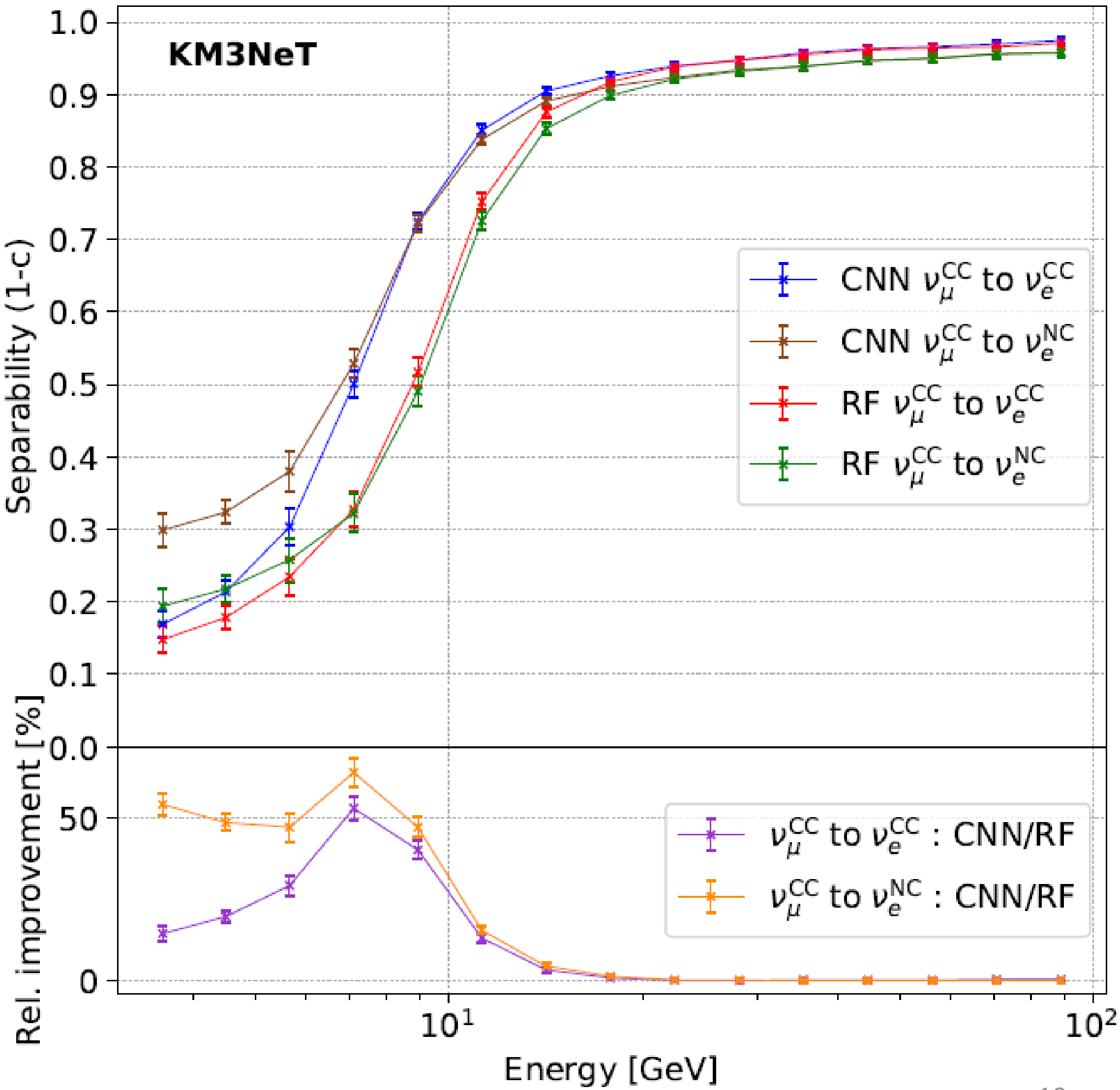
Neural network

(this work)

vs

random forest

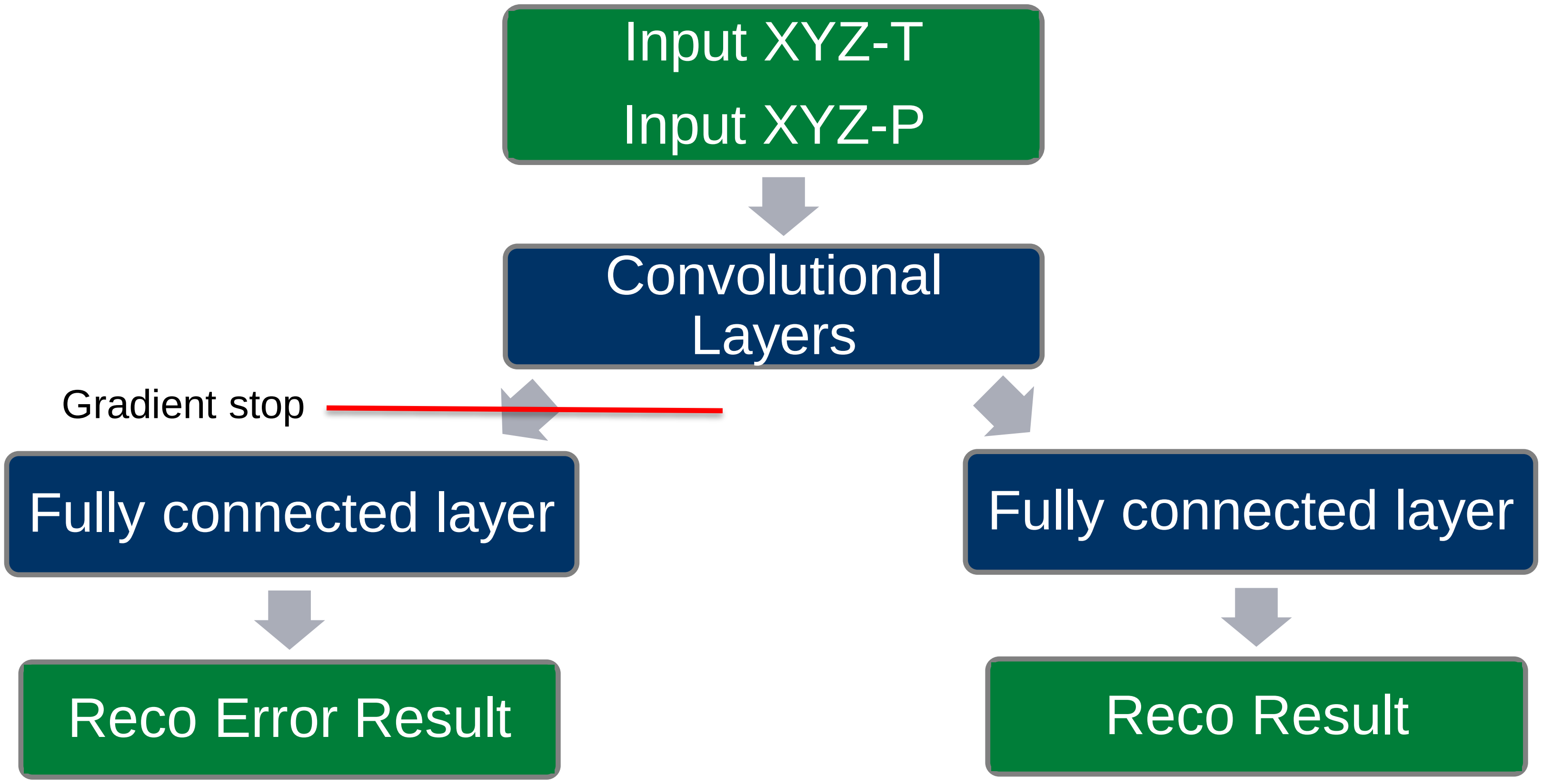
(standard method)



direction reconstruction

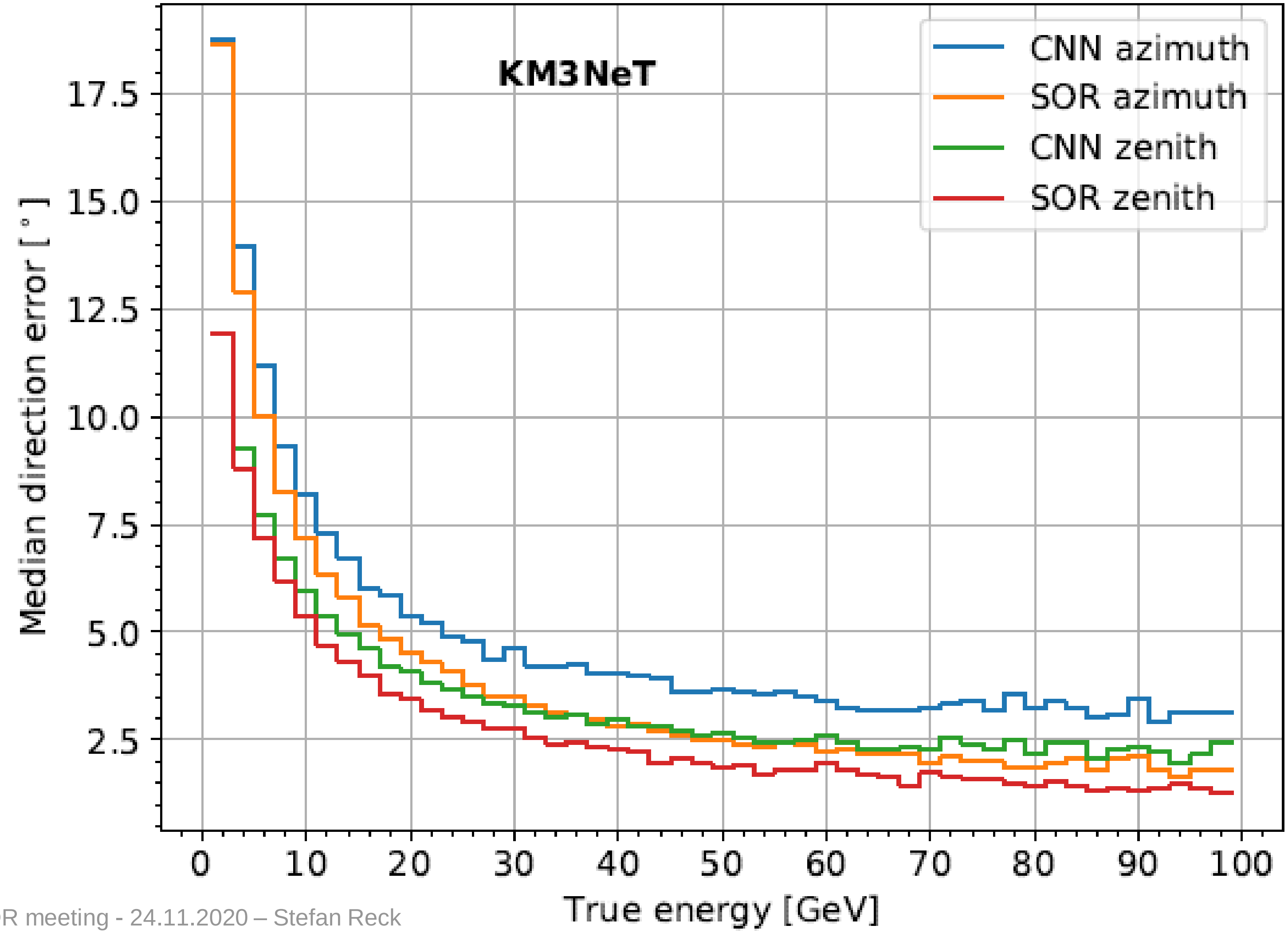
Regression results – direction

- Also predicts the standard deviation of any regression variable
- Done by adding a second dense network at the end



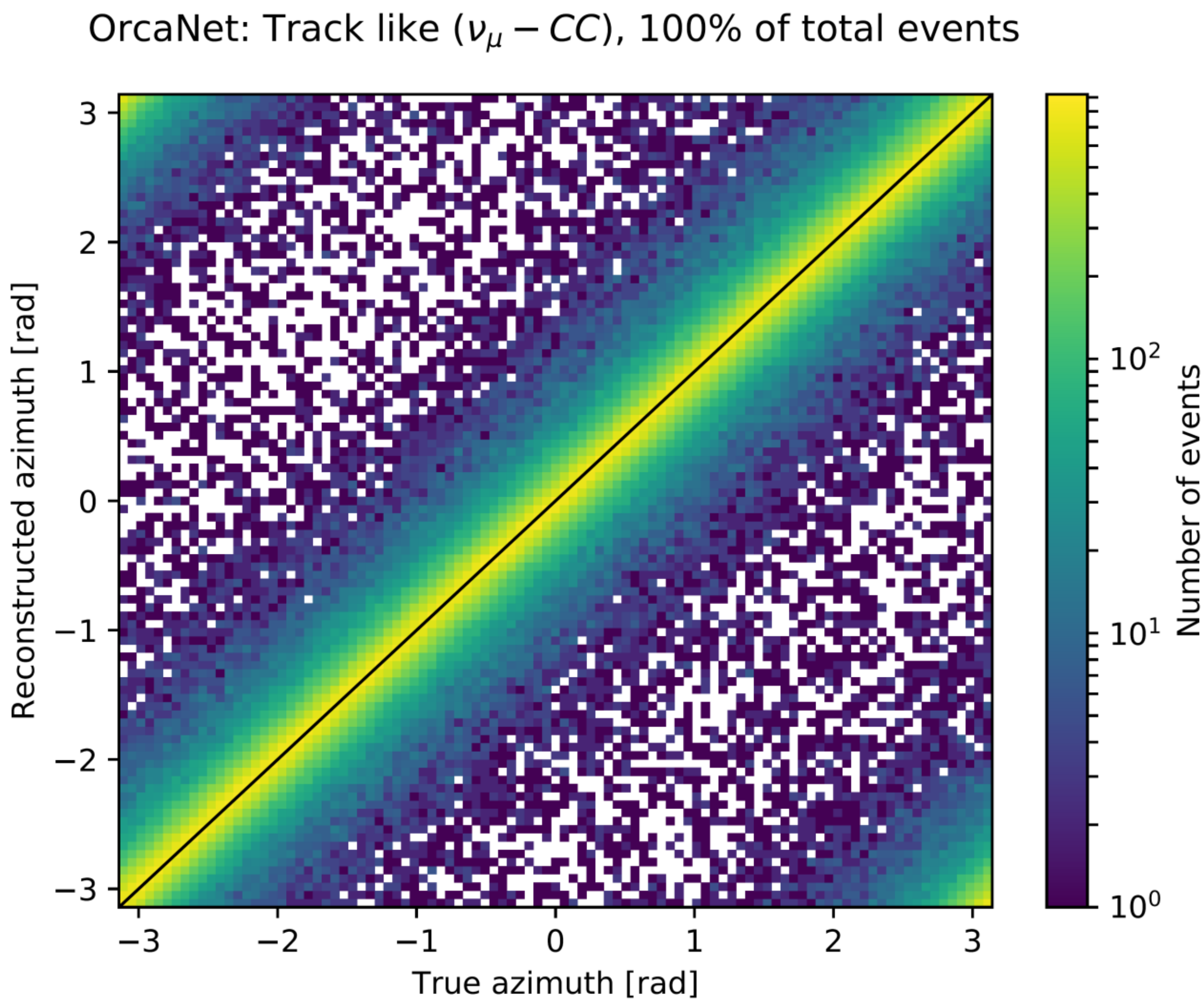
Regression results – direction

- CNN reco is comparable to classical likelihood based reco

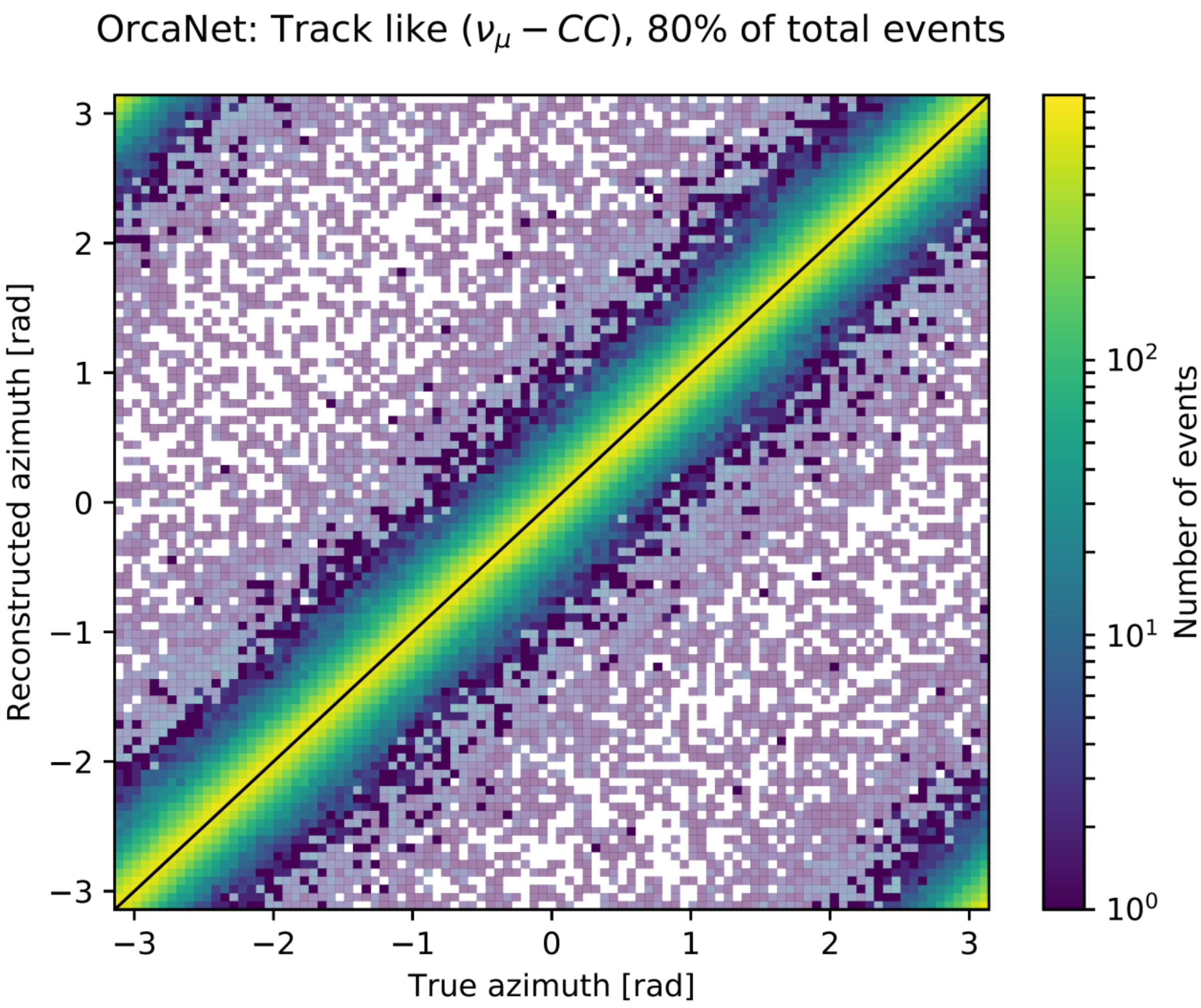


Regression results – direction

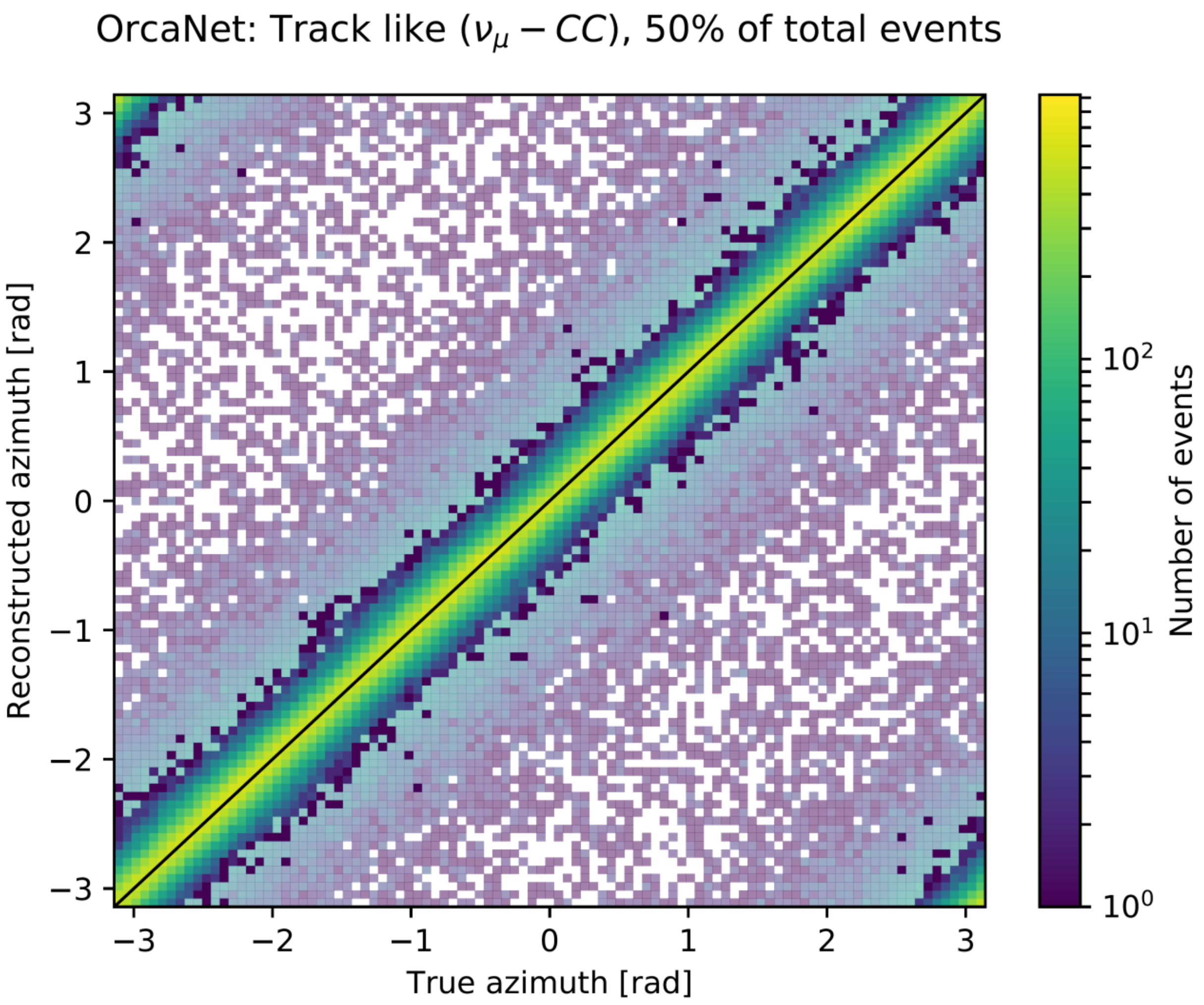
- can do cuts using network's error estimation



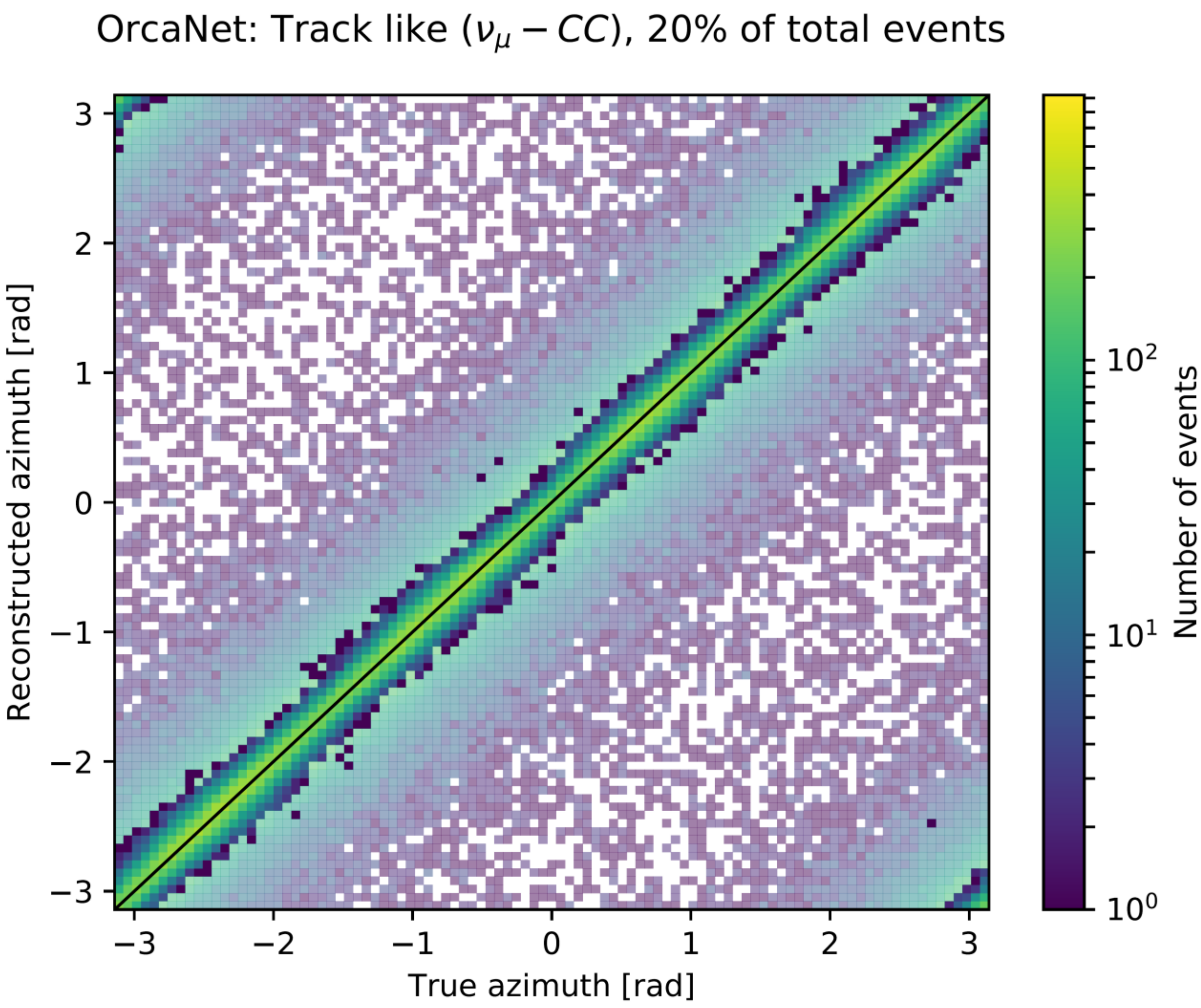
Regression results – direction



Regression results – direction



Regression results – direction



Summary

- **Background classification**
  significant improvement on rejection of atmospheric muons, same performance for random noise

- **Track-shower classification**
  significant improvement on the classification accuracy

- **Regression of energy, direction and neutrino interaction point**
  - Significant improvement for the energy reconstruction of track-like events
  - Competitive performance for other variables
  - error estimation for each reconstruction

backup

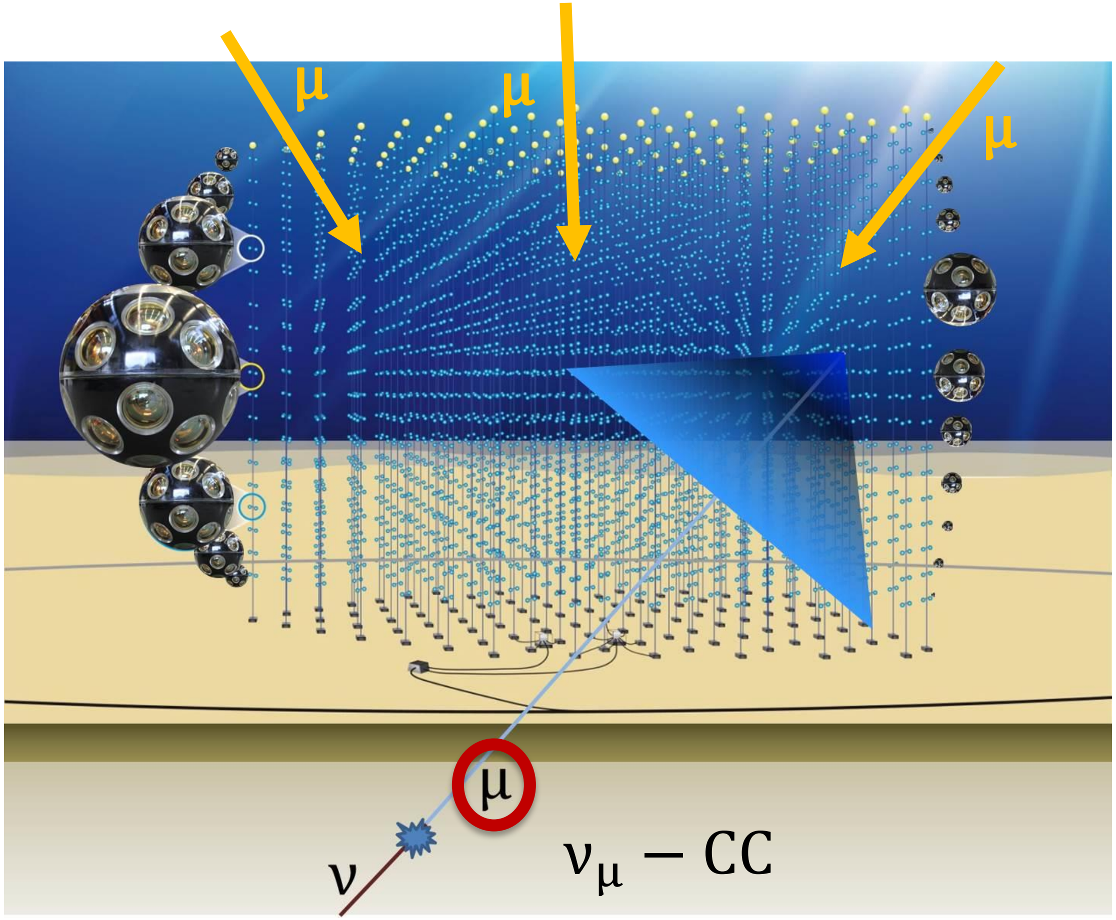
Network facts

- Network needs to be trained in order to distinguish the background
  - 43 Mio. simulated events used for the training
  - Consists of neutrinos, atmospheric muons, random noise events
  - CNN architecture with 10 layers
  - 3 million free parameters
  - Trained for about 1.5 weeks

background rejection

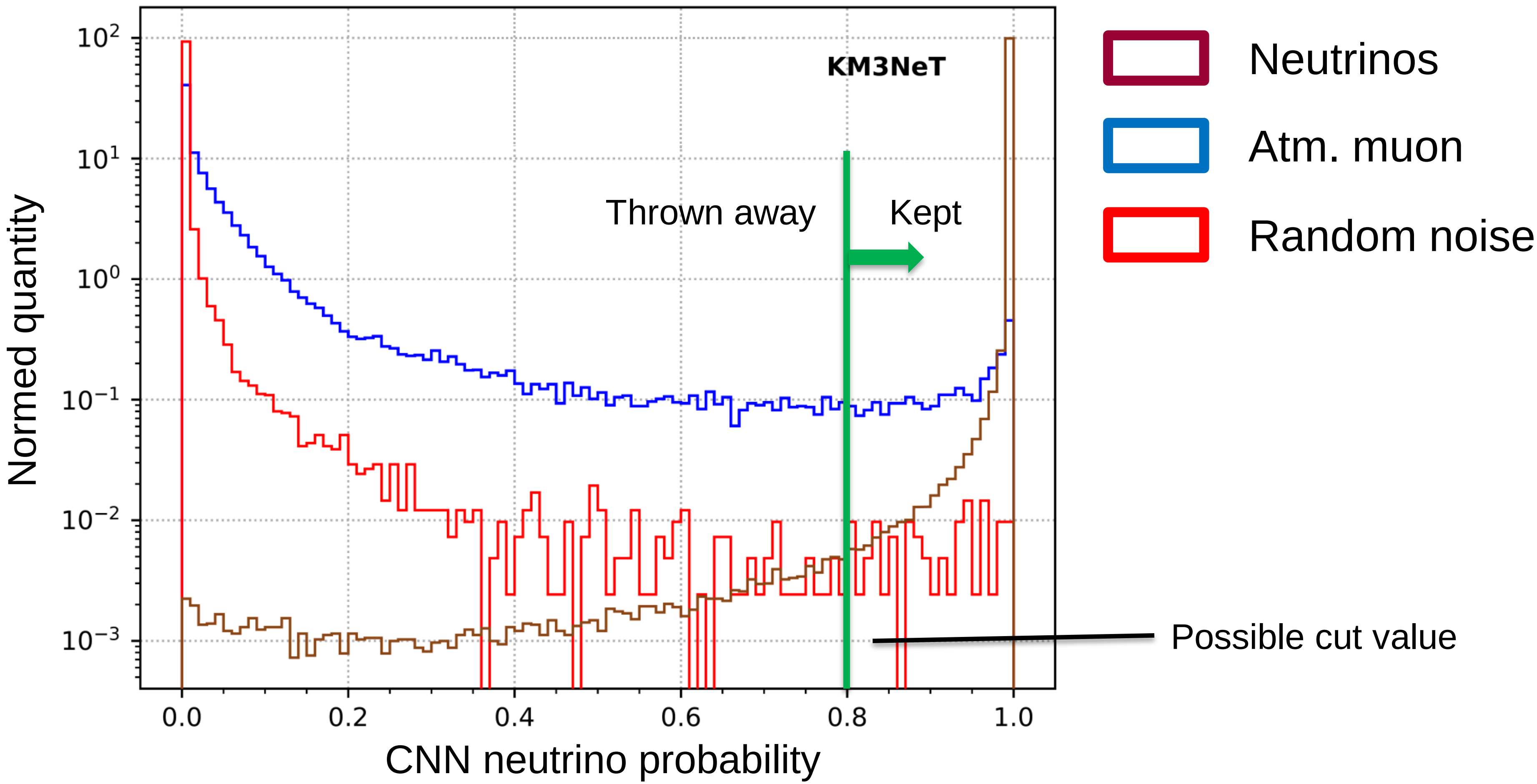
ORCA backgrounds

- Two types of backgrounds producing photons in the deep sea:
  1. **Atmospheric muons** passing the detector from above
  2. **Random noise**, by K-40 beta decays and bioluminescent animals



Background separation performance

- Output of the CNN on never-before-seen test data
- Reject backgrounds by cutting on the predicted neutrino probability



Background separation performance

- Assuming an oscillated neutrino and atmospheric muon flux

CNN (this work)

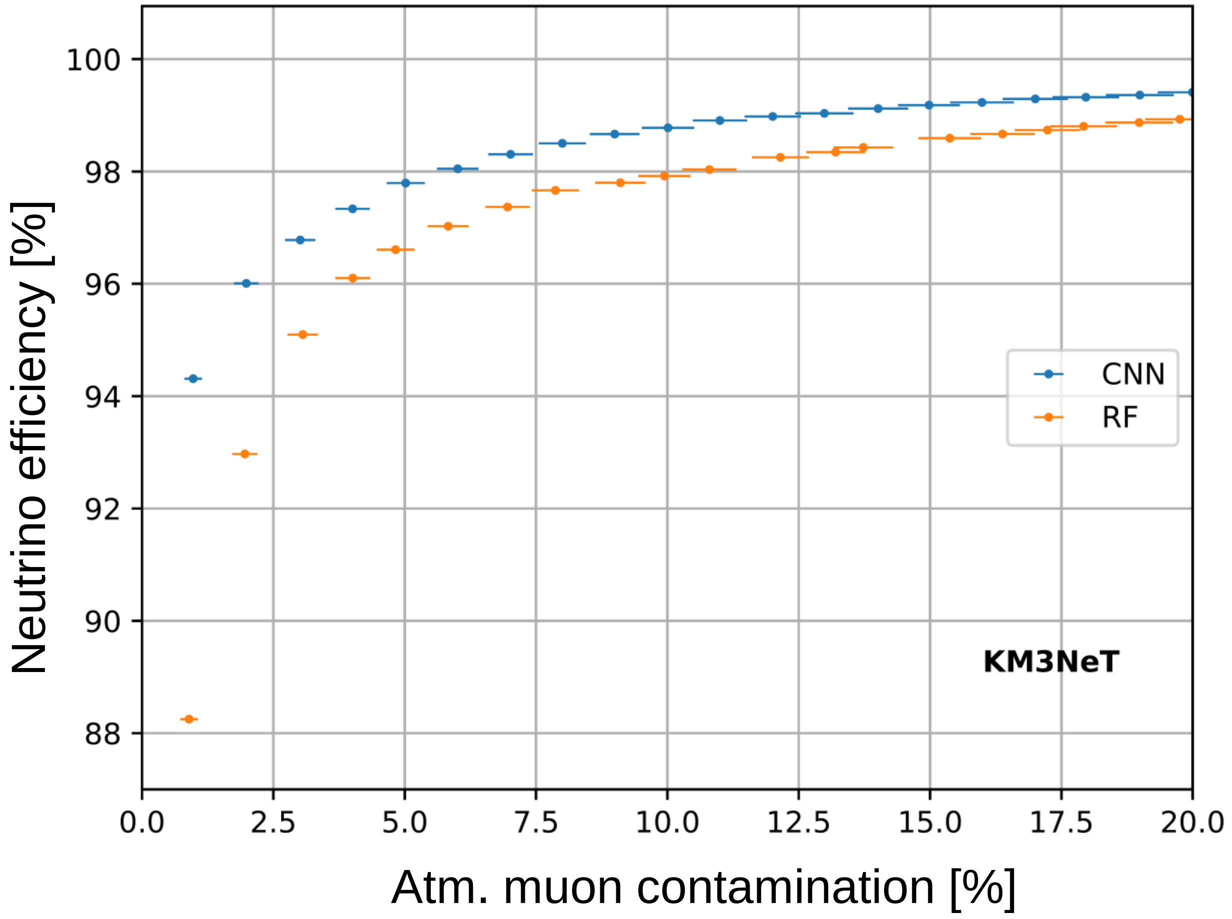
RF (standard method)

Muon contamination:

$$C_\mu = \frac{N_\mu(p)}{N_{total}(p)}$$

Neutrino efficiency:

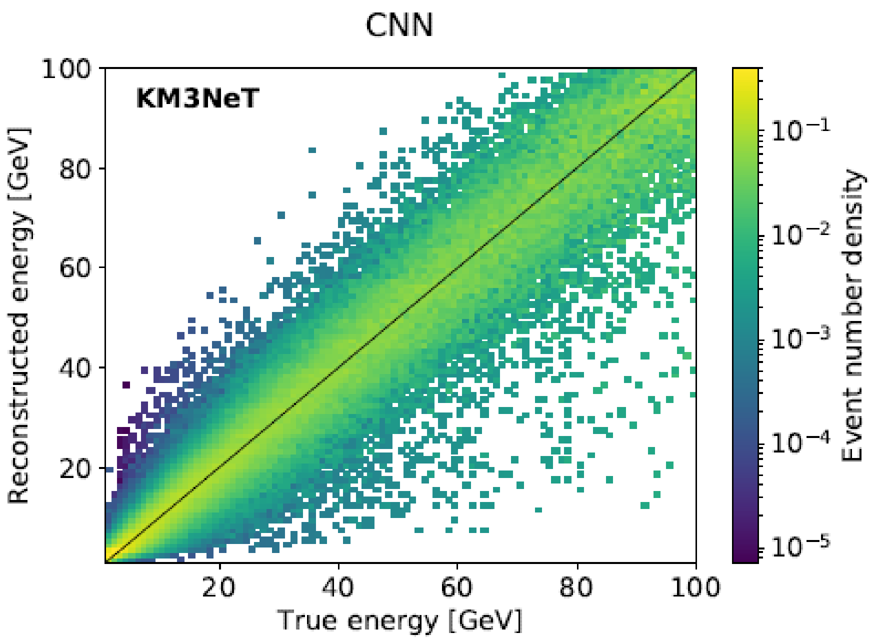
$$\nu_{eff} = \frac{N_\nu(p)}{N_{\nu,total}}$$



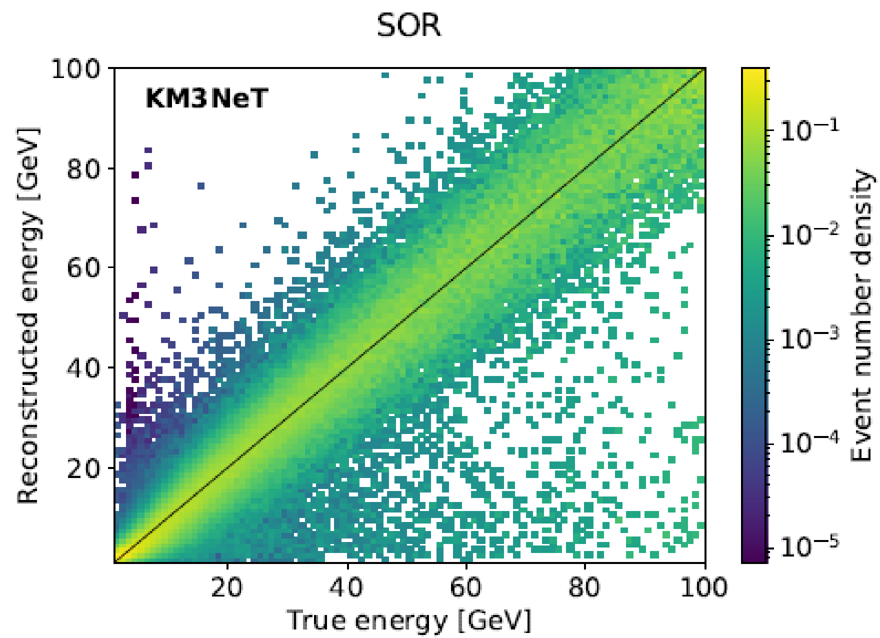
energy reconstruction

Regression results – shower energy

- OrcaNet shower energy reco is comparable to max. lkl. based reco



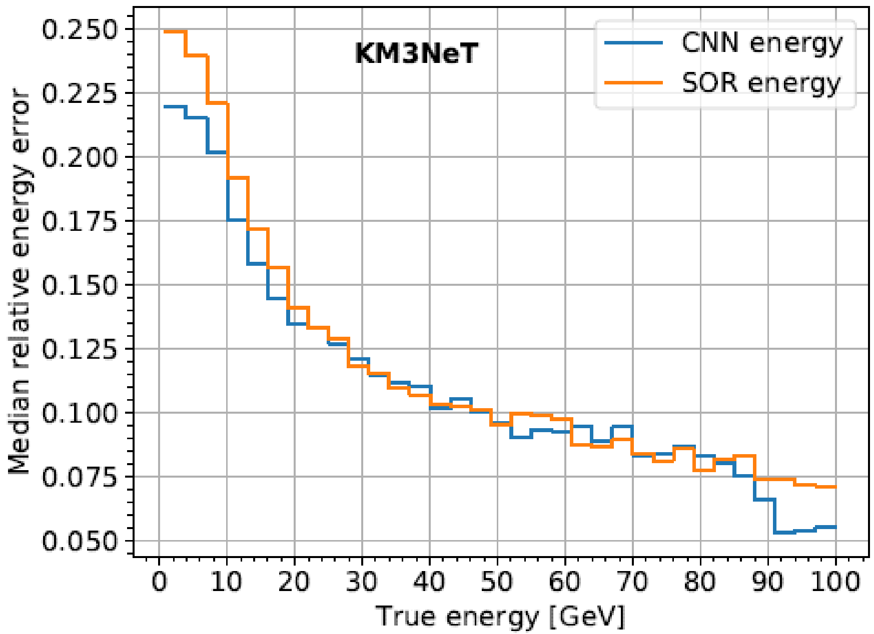
Deep Learning

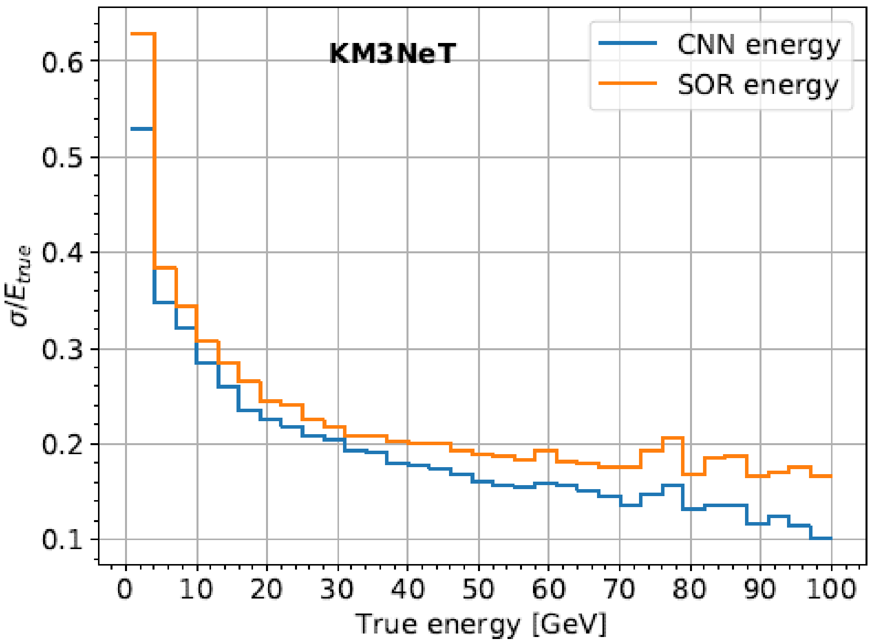


standard reco

Regression results – shower energy

- CNN energy reco is comparable to standard reco





Median relative error

$$\text{Median}\left(\left| \frac{E_{reco} - E_{true}}{E_{true}} \right| \right)$$

Relative standard deviation

$$\frac{\sigma}{E_{true}}$$