Deeply Learning from Neutrino Interactions with the KM3NeT neutrino telescope

Journées de Rencontre des Jeunes Chercheurs 2023

Santiago Peña Martínez

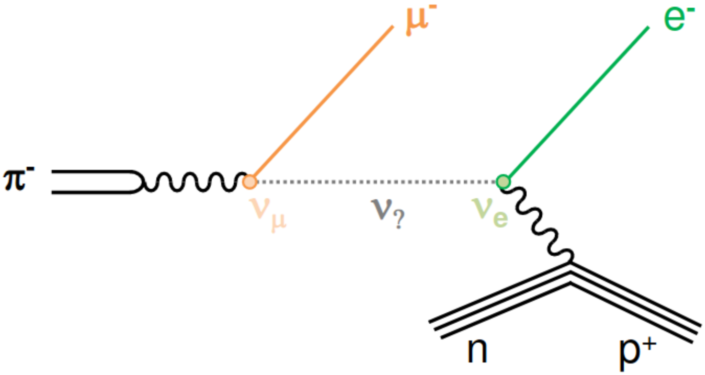
APC - KM3NeT

October 27, 2023

Overview

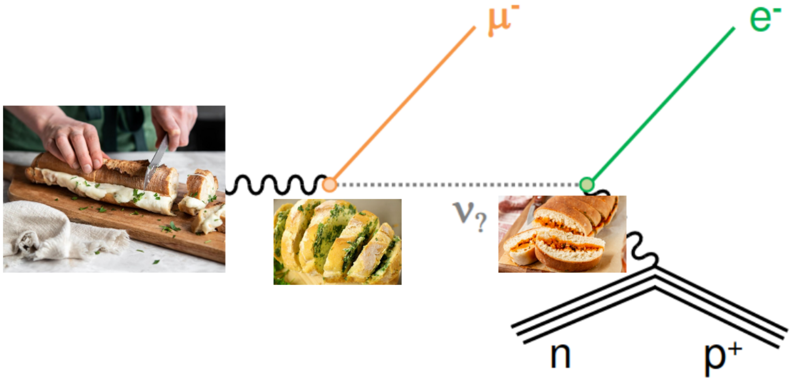
- Neutrino oscillations and Neutrino Mass Ordering
- About KM3NeT
- Motivation
- Neural Network for combined energy estimate
- Graph Neural Networks

Neutrino oscillations in a nutshell

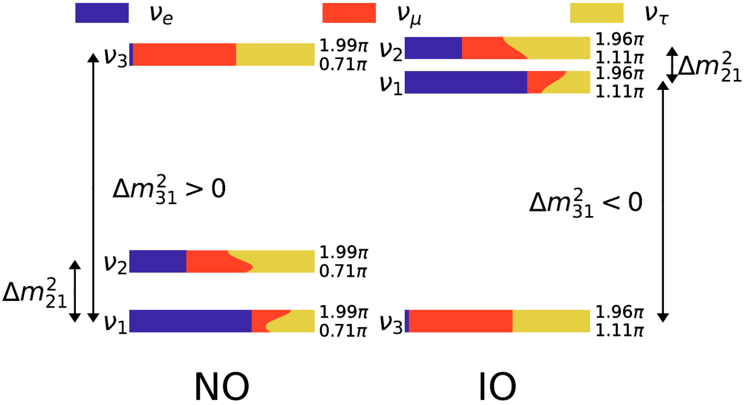


Credits J. Coelho

~~Neutrino~~ Préfou oscillations in a nutshell

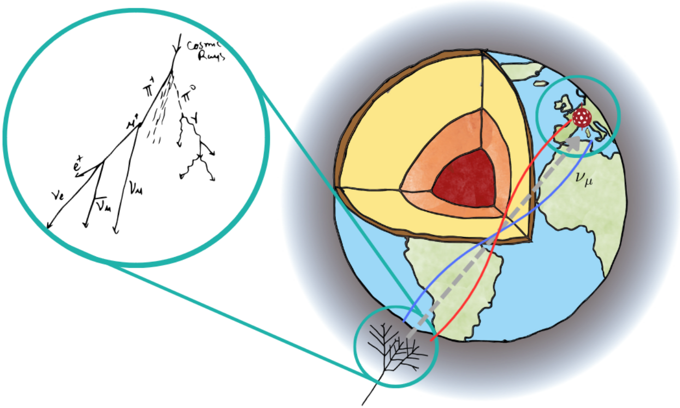


Neutrino mass ordering (NMO)

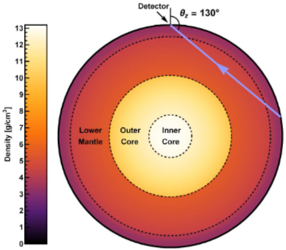


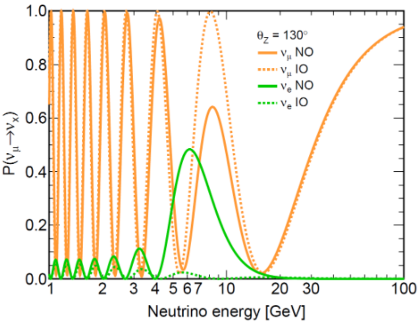
https://globalfit.astroparticles.es/2018/07/03/neutrino-mass-ordering/

Atmospheric neutrinos production



Measuring ν mass ordering using matter effects



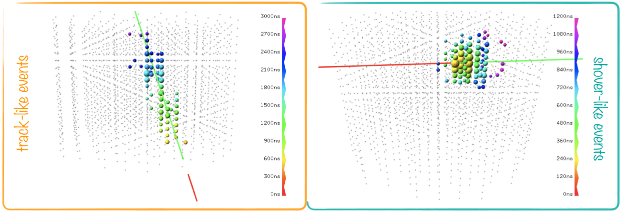


Credits J. Coelho

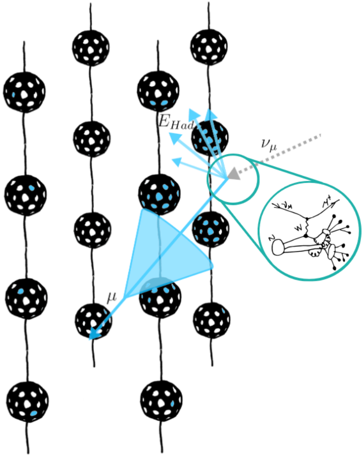
KM3NeT detector

Neutrino telescope using a large sea water volume as detection volume:

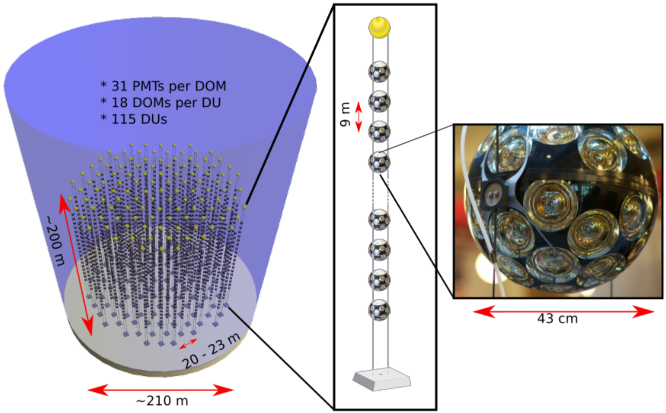
- Neutrino interaction in water produces charged secondary particles which induce Cherenkov radiation.
- Detector composed of large array of photosensors.
- Designed to have two construction sites with different physics goals but same technology.



J. Phys. G: Nucl. Part. Phys. 43 084001 (2016)



KM3NeT - ORCA



PMT: Photomultiplier Tube, DOM: Digital Optical Module, DU: Detection Unit (string of DOMs)

ORCA

Study ν properties, determine the Neutrino Mass Ordering (NMO) and measure oscillation parameters.

Status

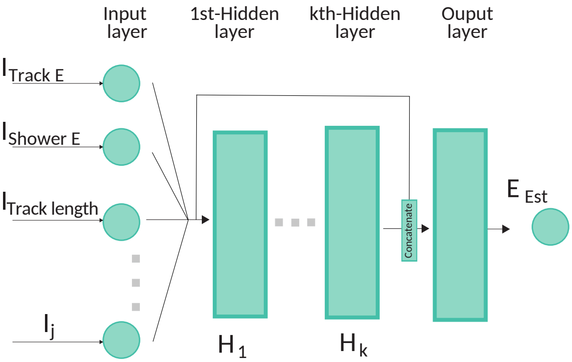
ORCA site currently has 18 working DUs, expected to have 28 by the end of the year.

The work presented here is done for the detector with 6 DUs $\longrightarrow$ ORCA6.

Using Deep Neural Networks (DNN) for a combined energy estimate

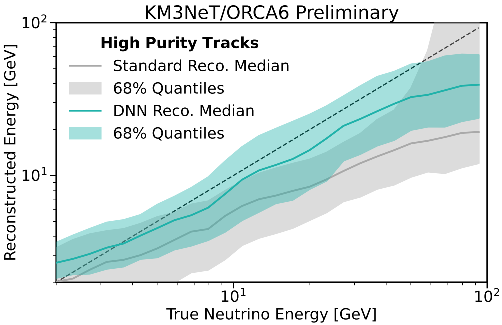
Problem and motivation

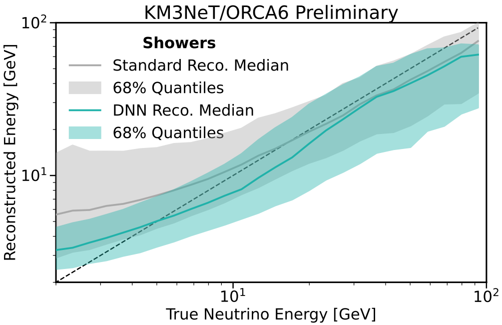
- Currently the collaboration uses physics based energy reconstructions.
- Information about triggered hits are shared in all reconstructions.
- Reconstructions provide auxiliary variables (Direction, vertex position, Cherenkov variables, etc...).
- A Neural Network can use all this information to estimate an optimal combination having the true energy as target.



DNN energy resolutions

- Energy reconstructions given by the DNN show less bias than the standard reconstructions for track and shower-like events.
- Energies around the oscillation ranges 5-20 GeV show improved resolution for the DNN. This energy range is relevant for oscillations analysis.

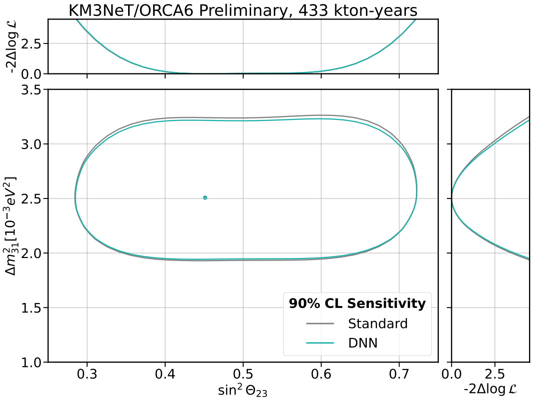


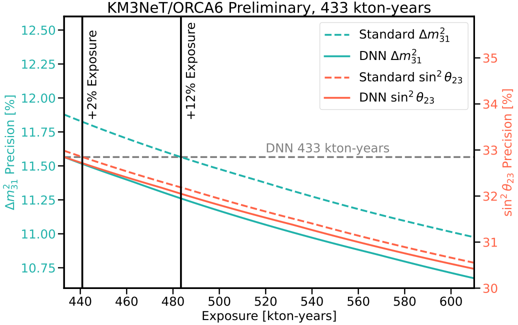


Sensitivity to oscillation parameters

- Energy reconstructions given by the DNN show less bias than the standard reconstructions for track and shower-like events.

- Energies around the oscillation ranges 5-20 GeV show improved resolution for the DNN. This energy range is relevant for oscillations analysis.

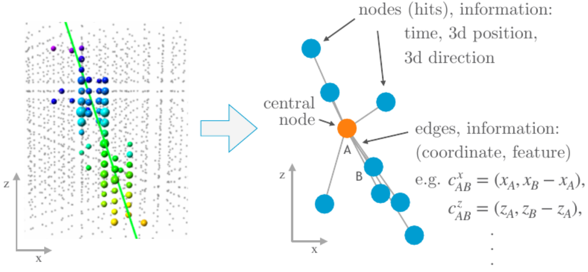




Graph Neural Networks (GNN) in KM3NeT

About Graph Neural Networks

- Event of neutrino going through the detector can be represented as a graph composed of nodes and edges.
- Graph given to the network as an input fully encodes the information of the event.
- Already used in KM3NeT for different tasks:
  - Signal vs noise classification
  - Track vs shower topology classification
  - Energy regression
  - Direction regression



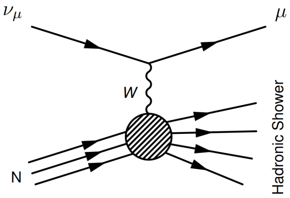
From: Development of detector calibration and graph neural network-based selection and reconstruction algorithms for the measurement of oscillation parameters with KM3NeT/ORCA, Daniel Guderian, PhD Thesis

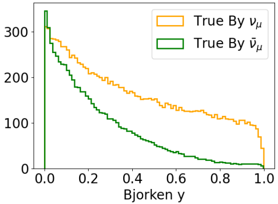
Inelasticity of a neutrino interaction

Inelasticity is given by the variable called Björken y, defined as the fraction of the lepton's energy transferred to the nucleon rest frame:

$$y = 1 - E'/E = E_{Sh}/E_{Tot}$$

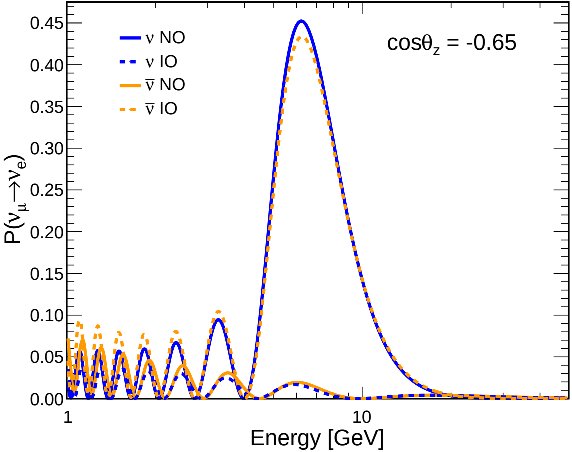
For neutrinos interacting with matter, the distribution of Björken y is different for ν and $\bar{\nu}$.





Problem and motivation

- NMO effects are visible when looking at ν and $\bar{\nu}$
- But the KM3NeT detector cannot distinguish ν from $\bar{\nu}$ so far.
- Björken y distributions can help make this distinction.
- Information of the inelasticity can be retrieved from the track and shower components.
- A Graph Neural Networks may have enough power to reconstruct the Björken y.



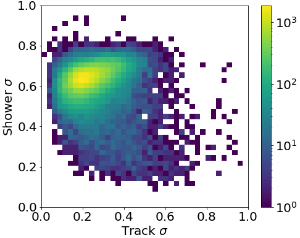
Total energy estimation

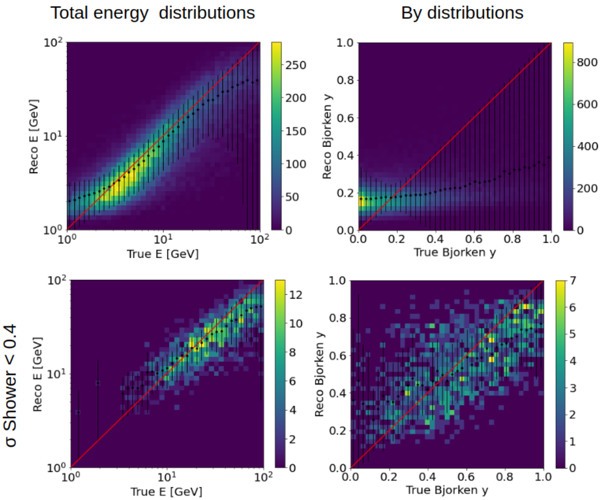
Upper row: Without quality cut

Total E reco of the event is slightly underestimated. Björken y is mostly reconstructed at a fixed value.

Lower row: With cut $\sigma_{Sh} < 0.4$

Total E reco continues be correlated. Björken y has a wide range of estimated values.





Conclusions

Additional information can be extracted when considering track and a shower component of events.

DNN for energy reconstruction

- Hits contain additional information.
- Shows to improve oscillation parameter estimation.
- Gain in sensitivity is limited by systematic uncertainties.

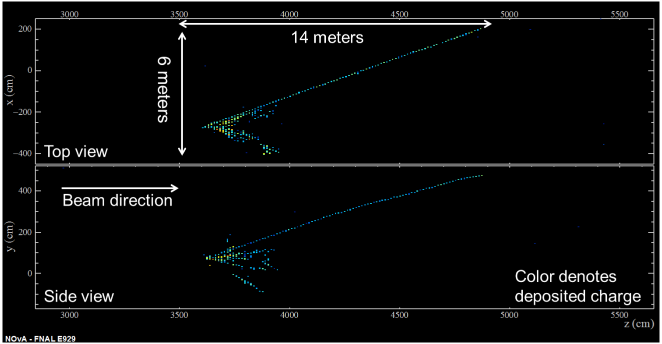
GNN for Björken y reconstruction

- GNN can reconstruct the energy of the full event.
- Björken y mostly reconstructed at a fixed value.
- Certain events are correctly reconstructed, this can be exploited.

Both tasks are expected to improve performance with bigger detector size.

Backup slides

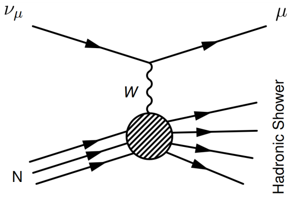
ν event is composed of track and shower



Event display from the NOνA experiment. Taken from: https://nusoft.fnal.gov/nova/public/neutrinos/zoomed-notes.png

Reconstruction methods used at the moment do not include information about both components.

But how do you predict the Björken y?



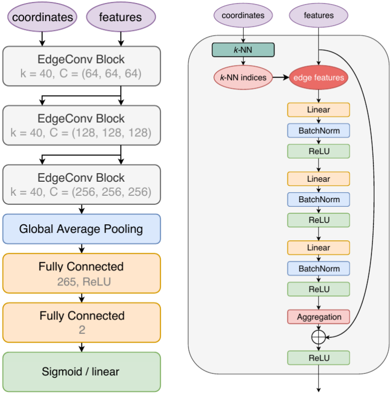
Björken y is a continuous value between 0 and 1 which is annoying from the point of view of a loss funcion. Regression loss functions work with unbounded continuous values. To try to solve this issue one can ask the network for two output neurons with their corresponding uncertainties:

- Energy of the track
- Energy of the shower

Train on ν_μ and $\bar{\nu}_\mu$ charged current (CC) events, this allows to have shower and track in one event. Training done for events in the $0 - 100 GeV$ range for simplicity to compensate the detector size.

Architecture of the network

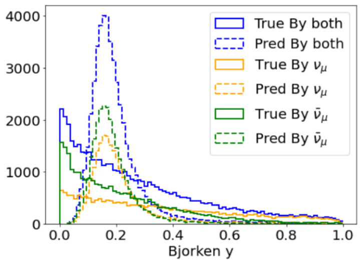
- Nodes are passed through three edge convolution blocks.
- Afterwards the output is passed through a fully connected network.
- The output is subject to an activation function depending on the task.



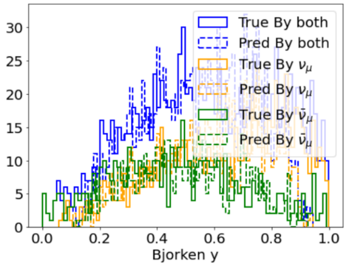
From: Development of detector calibration and graph neural network-based selection and reconstruction algorithms for the measurement of oscillation parameters with KM3NeT/ORCA, Daniel Guderian, PhD Thesis

Björken y distributions

Without any cut

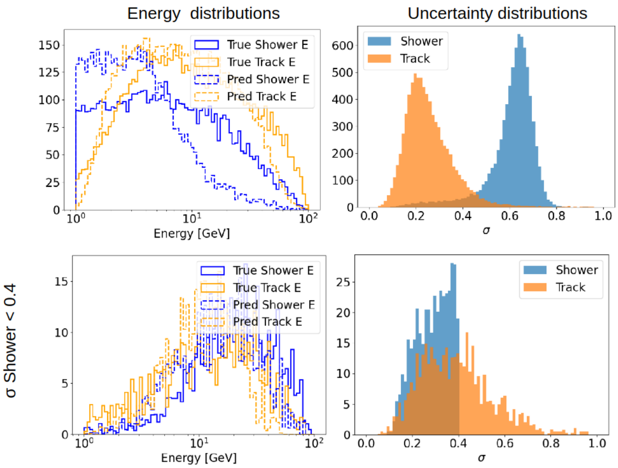


With σ Shower < 0.4 cut

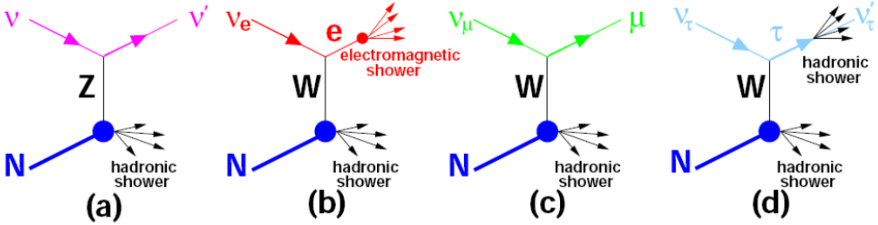


GNN reconstructs using the average of the Björken y distrubution. The events with improved shower reconstruction have a high Björken y.

GNN Reconstructed energy distributions



ν interaction topologies



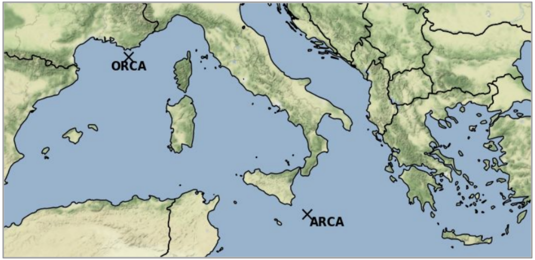
Summary of Deep Inelastic Scattering (DIS) neutrino event topologies in neutrino telescopes: (a) flavour-insensitive NC, (b) ν_e CC, (c) ν_μ CC, (d) ν_τ CC.

Taken from: A. Trovato. “Development of reconstruction algorithms for large volume neutrino telescopes and their application to the KM3NeT detector”. PhD thesis. Università degli Studi di Catania, Scuola Superiore di Catania, 2010

KM3NeT objectives

ARCA

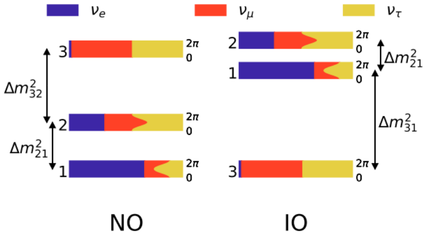
Discover/Observe high-energy neutrino sources in the universe.



From: https://indico.cern.ch/event/855372/contributions/4454016/

ORCA

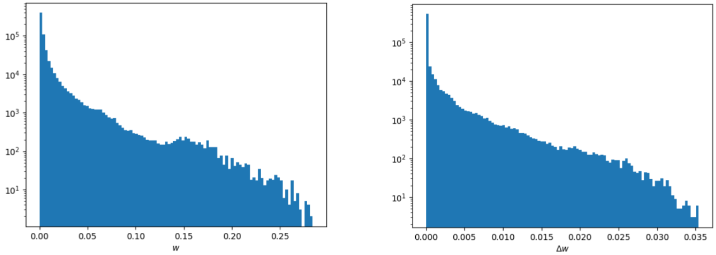
Study ν properties, determine the Neutrino Mass Ordering (NMO) and measure oscillation parameters.



From: https://doi.org/10.3389/fspas.2018.00036

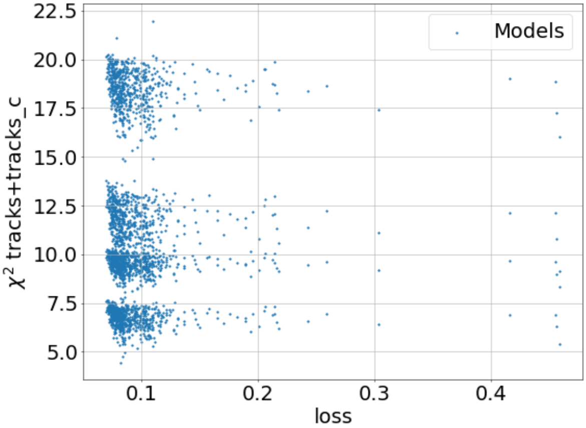
Using oscillation weights for training: Defining weight

- We want the network to be right specially for events sensitive to oscillation effects.
- Define the weights for training as $w = \Delta w * k + w(\Delta m_{31} = 2.5e-3, \sin^2(\theta_{23}) = 0.5)$, where $\Delta w = |w(\Delta m_{31} = 2.5e-3, \sin^2(\theta_{23}) = 0.5) - w(\Delta m_{31} = 2e-3, \sin^2(\theta_{23}) = 0.6)|$ and k is a hyperparameter which controls the importance of the difference in the oscillation weights.



HPO doing the right thing?

- Short answer, no.
- The apparent best model selected is not the one with the best sensitivity possible.
- So far it is still the best option.
- Phase space of hyperparameters has a very complicated shape.
- Longer warm-up for bayesian algorithm allows to better explore the phase space.



- Train many models.
- Compute log likelihood ratio for different oscillation parameter combinations.
- Select the best in terms of sensitivity.
- Go to contours and compare with JEnergy contours.