DUNE-PRISM

Efficiency correction and detector response matching

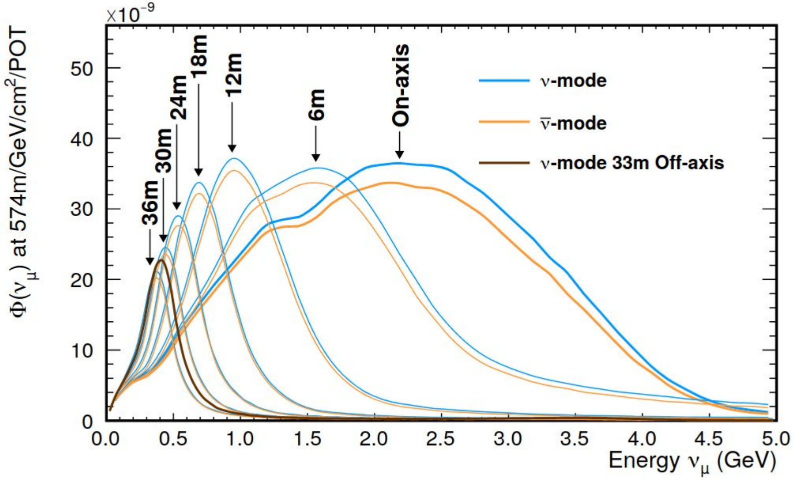
December 3 2020

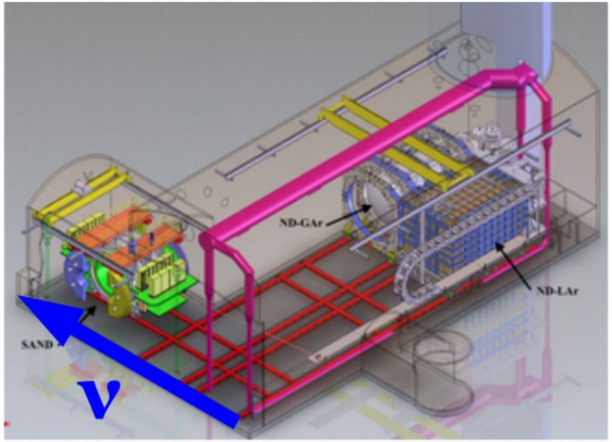
EP-NU Meeting

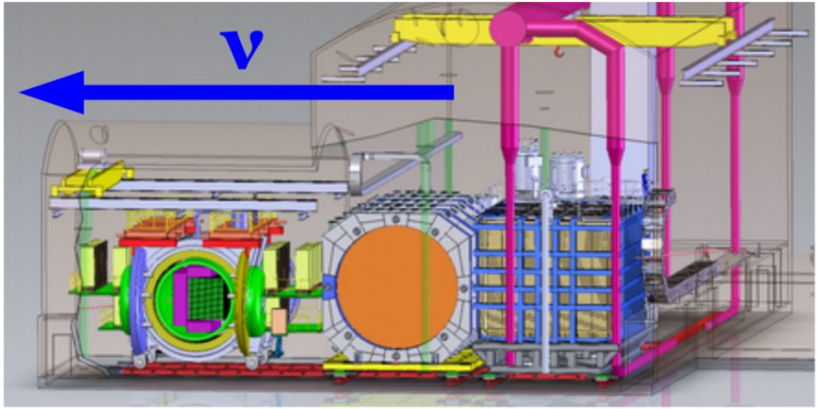
Cristóvão Vilela

The DUNE Near Detector

- **LArgon**: similar response to FD.
- **GArgon**: "zoomed-in" interactions on Argon + magnetic field.
- **Carbon**: on-axis monitoring + neutron detection in carbon interactions.
- **Off-axis** movement: resolve degeneracies in interaction model + data-driven oscillation analysis.







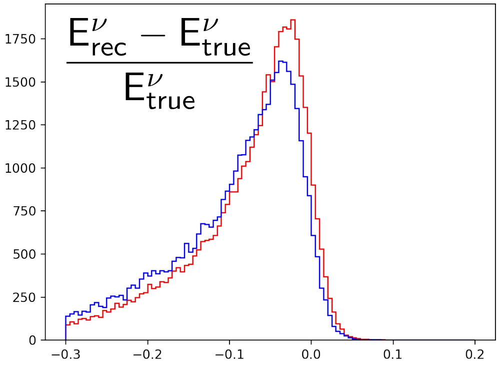
SAND ND-GAr ND-LAr

Neutrino measurement degeneracies

- A threefold conspiracy makes it very hard to make precise neutrino measurements:
  - Lack of neutrino **initial state** knowledge: neutrino energy spectrum wider than oscillation features we want to resolve, by construction.
  - Incomplete **final state** knowledge: even in LArTPCs we miss parts of the final state, like neutrons.
  - Imprecise **nuclear models:** nuclei are messy environments… (ask Stephen about this!)

Neutrino measurement degeneracies

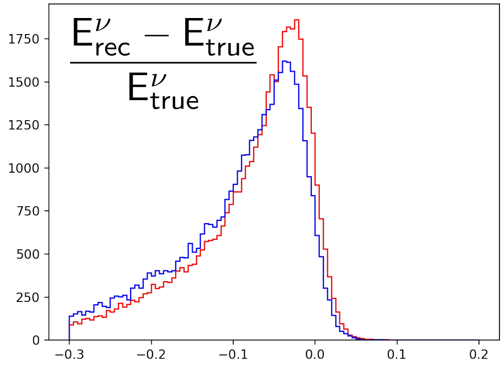
- A threefold conspiracy makes it very hard to make precise neutrino measurements:
  - Lack of neutrino **initial state** knowledge: neutrino energy spectrum wider than oscillation features we want to resolve, by construction.
  - Incomplete **final state** knowledge: even in LArTPCs we miss parts of the final state, like neutrons.
  - Imprecise **nuclear models:** nuclei are messy environments... (ask Stephen about this!)
- Illustrate the problem with a mock data study:

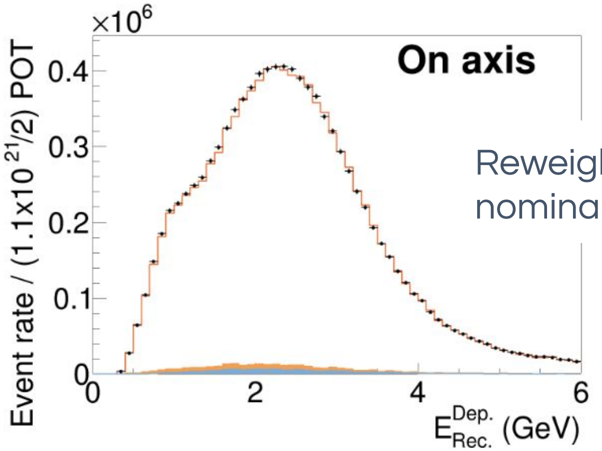


Introduce a bias in neutrino energy reconstruction by moving 20% of the energy carried by final state protons to neutrons.

Neutrino measurement degeneracies

- A threefold conspiracy makes it very hard to make precise neutrino measurements:
  - Lack of neutrino **initial state** knowledge: neutrino energy spectrum wider than oscillation features we want to resolve, by construction.
  - Incomplete **final state** knowledge: even in LArTPCs we miss parts of the final state, like neutrons.
  - Imprecise **nuclear models:** nuclei are messy environments… (ask Stephen about this!)
- Illustrate the problem with a mock data study:

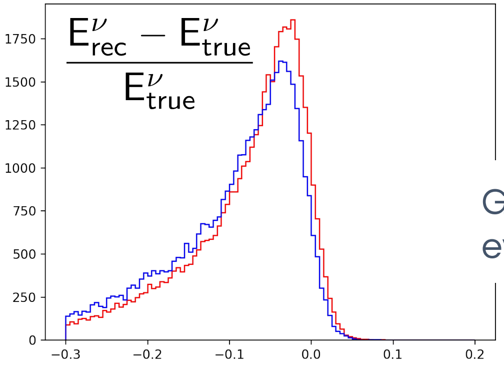


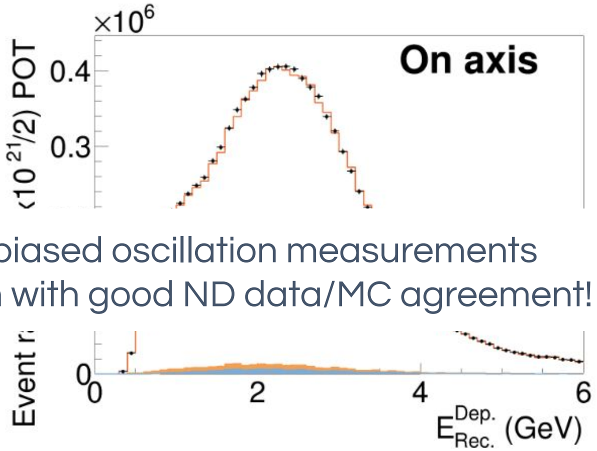


Reweight sample to match near detector nominal observation (i.e., fit the ND data)

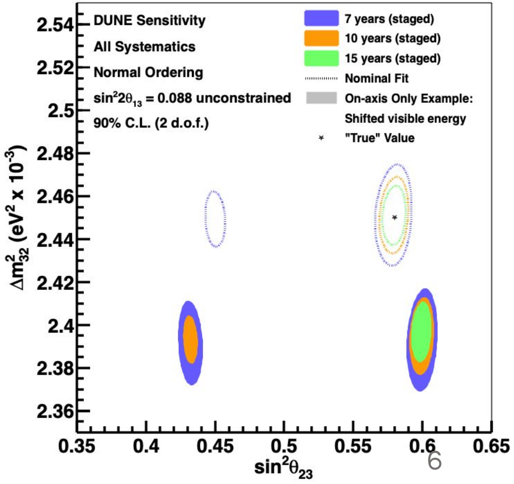
Neutrino measurement degeneracies

- A threefold conspiracy makes it very hard to make precise neutrino measurements:
    - Lack of neutrino **initial state** knowledge: neutrino energy spectrum wider than oscillation features we want to resolve, by construction.
    - Incomplete **final state** knowledge: even in LArTPCs we miss parts of the final state, like neutrons.
    - Imprecise **nuclear models:** nuclei are messy environments… (ask Stephen about this!)
- Illustrate the problem with a mock data study:



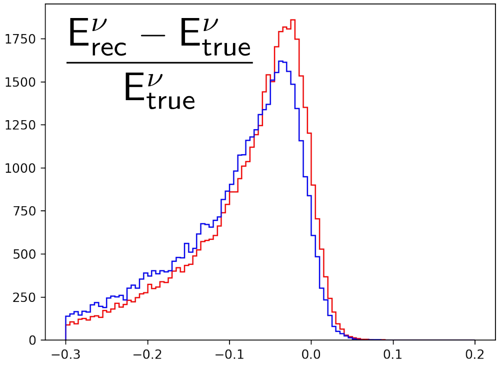


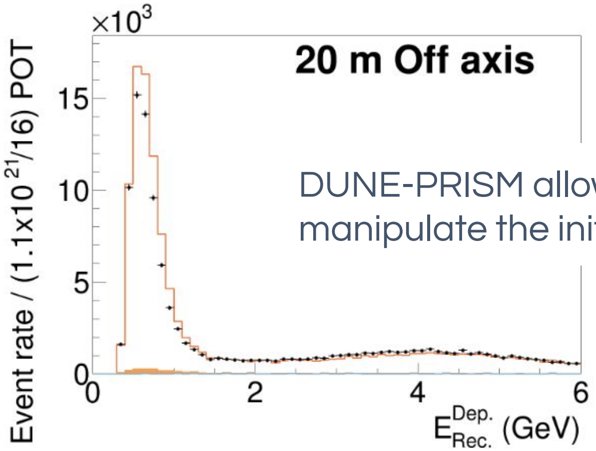
Get biased oscillation measurements even with good ND data/MC agreement!

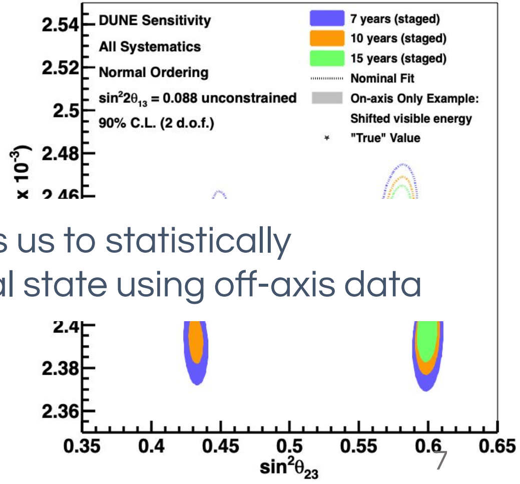


Neutrino measurement degeneracies

- A threefold conspiracy makes it very hard to make precise neutrino measurements:
  - Lack of neutrino **initial state** knowledge: neutrino energy spectrum wider than oscillation features we want to resolve, by construction.
  - Incomplete **final state** knowledge: even in LArTPCs we miss parts of the final state, like neutrons.
  - Imprecise **nuclear models:** nuclei are messy environments… (ask Stephen about this!)
- Illustrate the problem with a mock data study:



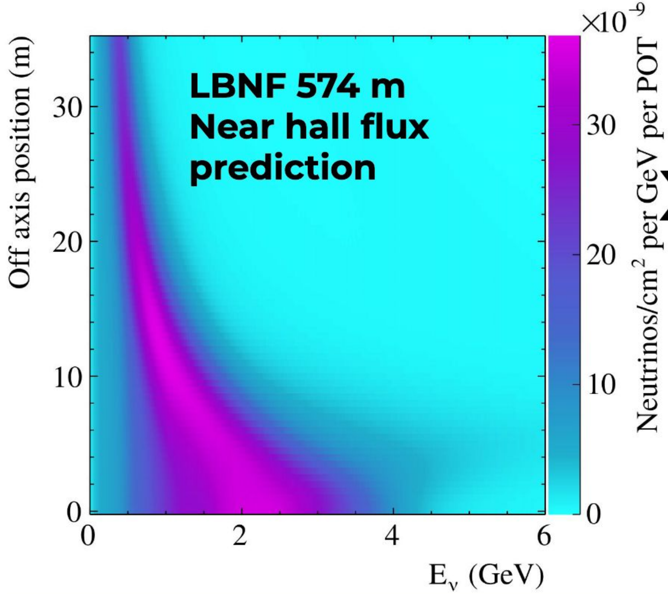




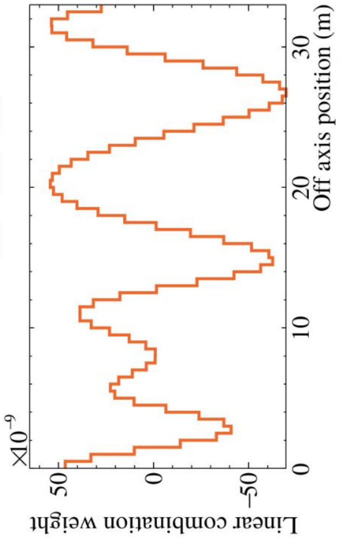
DUNE-PRISM allows us to statistically manipulate the initial state using off-axis data

DUNE-PRISM data-driven analysis

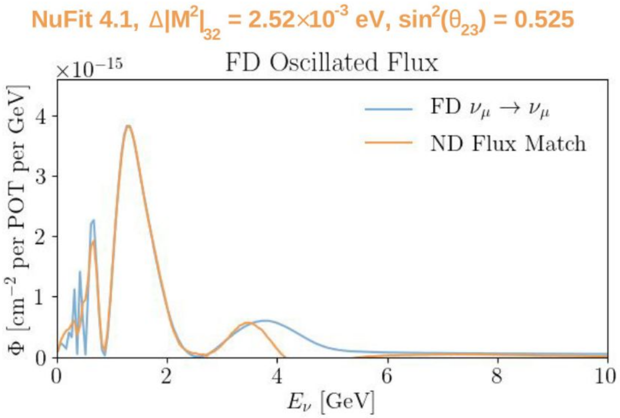
- Use **flux model** to solve linear algebra problem: which linear combination of ND fluxes matches the FD **oscillated** spectrum?



x



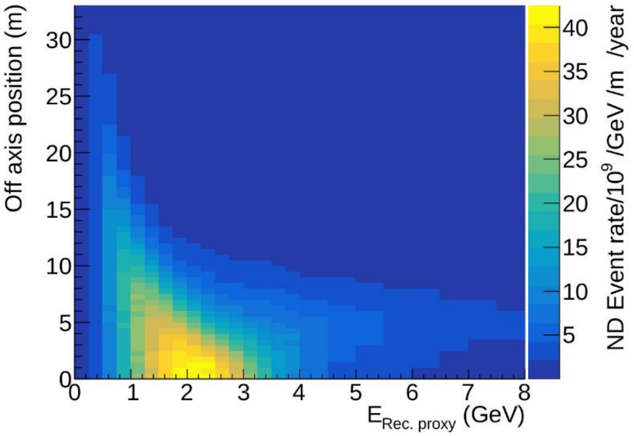
=



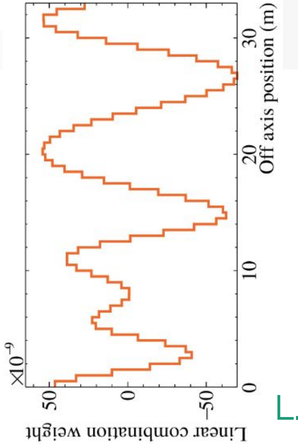
L. Pickering

DUNE-PRISM data-driven analysis

- Apply same coefficients to ND **data** to get FD prediction.
- Didn't use interaction model! (to first order...)

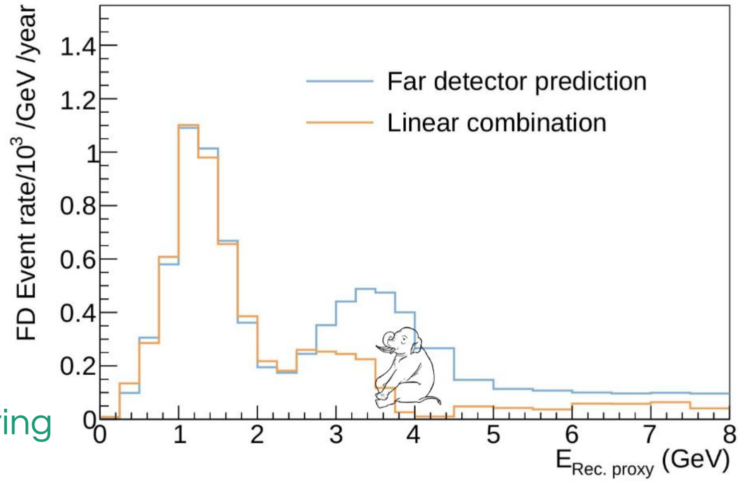


X



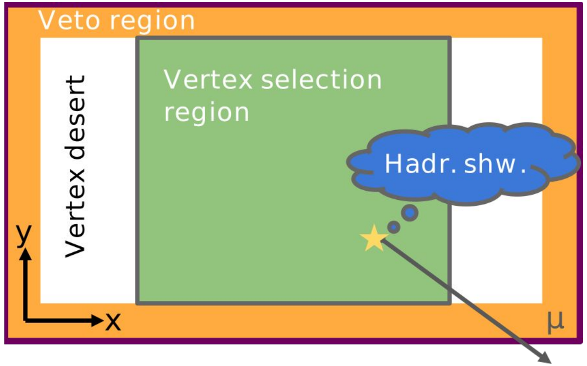
=

L. Pickering



Near detector efficiency

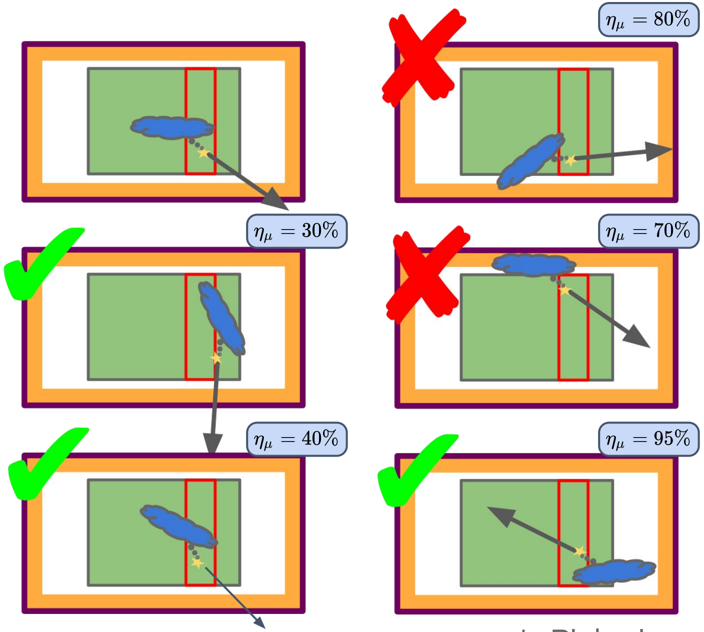
- For this to work, we need to understand differences in efficiency and response between the ND and the FD.
- Most obvious difference is the detector size:
  - ND will not contain very large hadronic systems.
  - ND does not contain high-ish energy muons.
    - But measures them in downstream tracker.
- Would like to know ND efficiency for a given event without relying on interaction model.



L. Pickering

Data-driven efficiency

- Use **symmetries** of neutrino interactions in ArgonCube:
  - **Translations** in LAr volume and **rotations** around beam axis.
- Algorithm:
  - For a **selected** ND event, rotate and translate 3D **hadronic** energy deposits and **muon position** and **momentum** vectors **N** times.
  - For the **hadronic** side:
    - Count how many of the trials would have passed the hadronic containment cut.
    - Take the ratio to the total number of trials get the "geometric" **hadronic containment** efficiency for that event.
  - For the **muon** side:
    - Use a **neural network** trained on particle gun MC to estimate the muon selection efficiency for a given translation/rotation.
  - **Combine** both to get event-level efficiency.

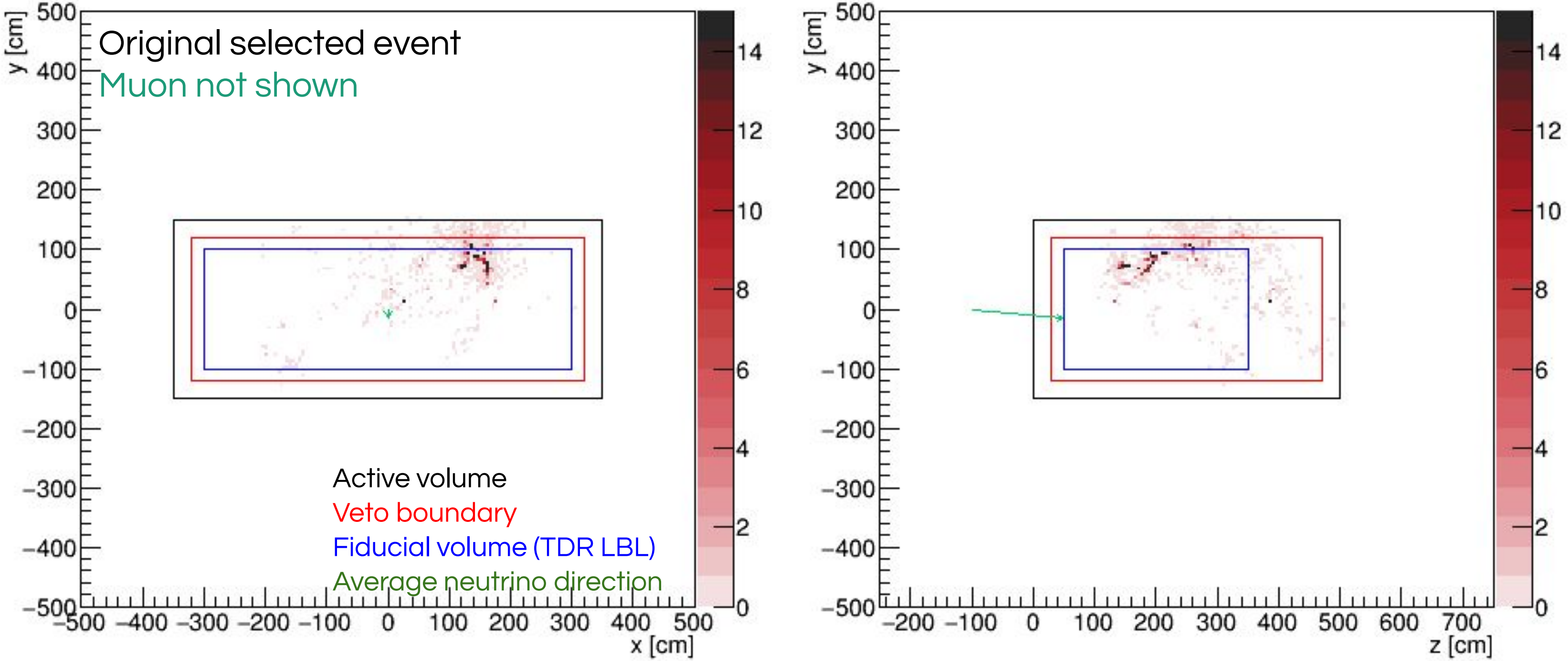


L. Pickering

$$\eta = \frac{0\times0.8+1\times0.3+0\times0.70+1\times0.4+1\times0.95}{5} = 33\%$$

Translations and rotations

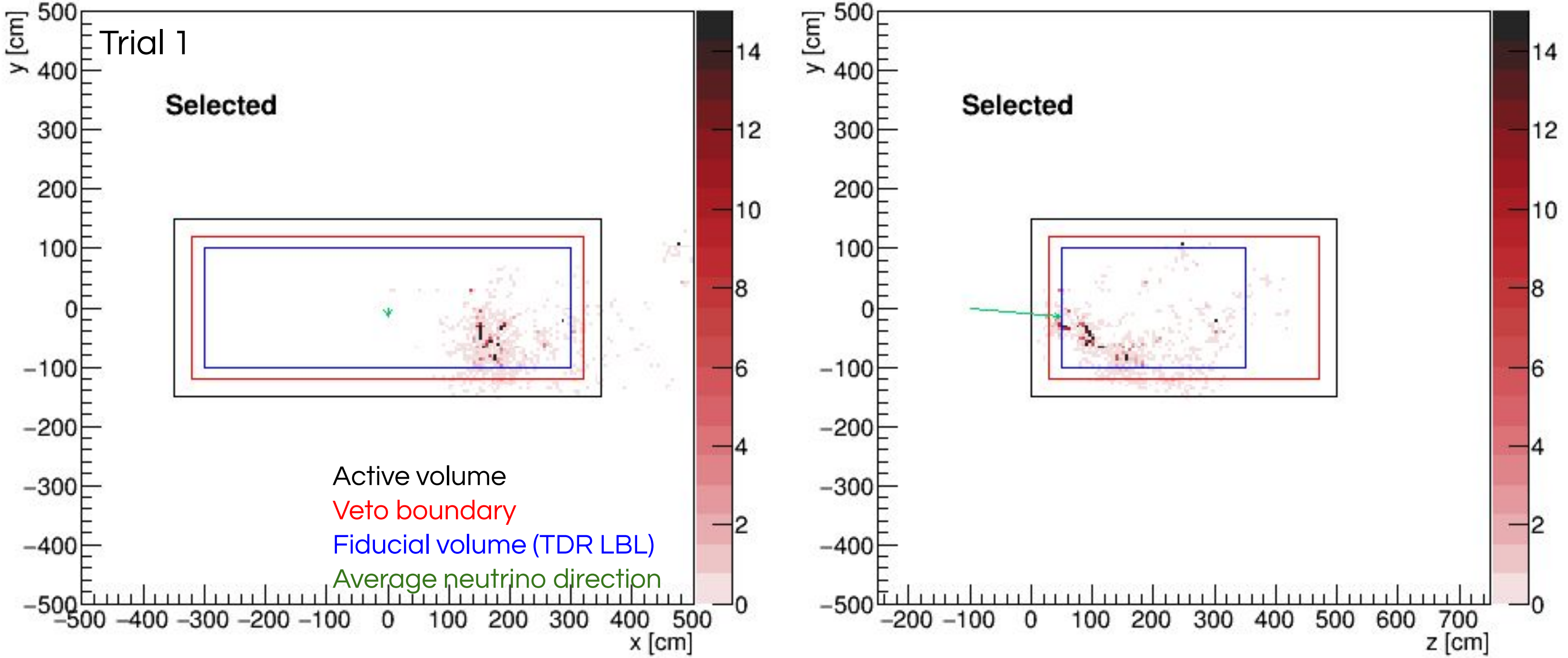
Hadronic system



Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

Translations and rotations

Hadronic system

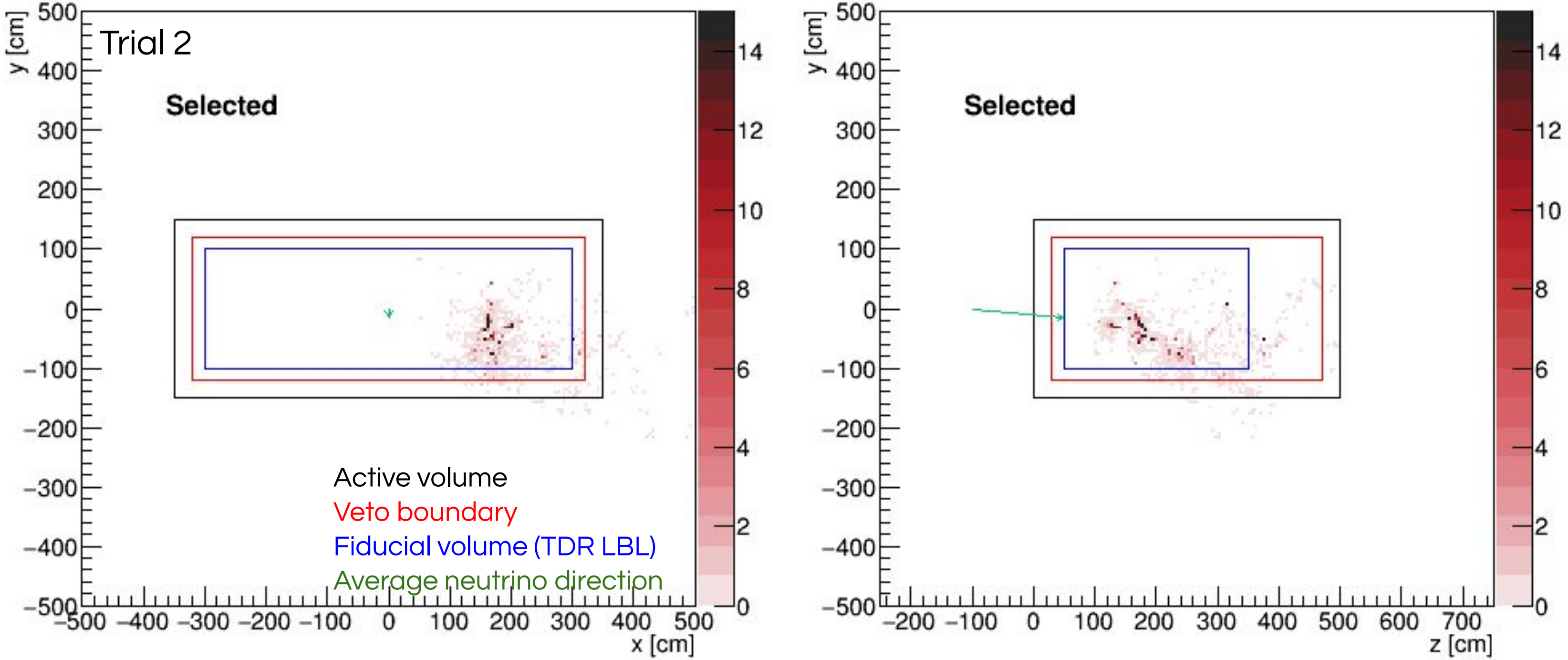


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{1}{1} = 100\%$$

Translations and rotations

Hadronic system

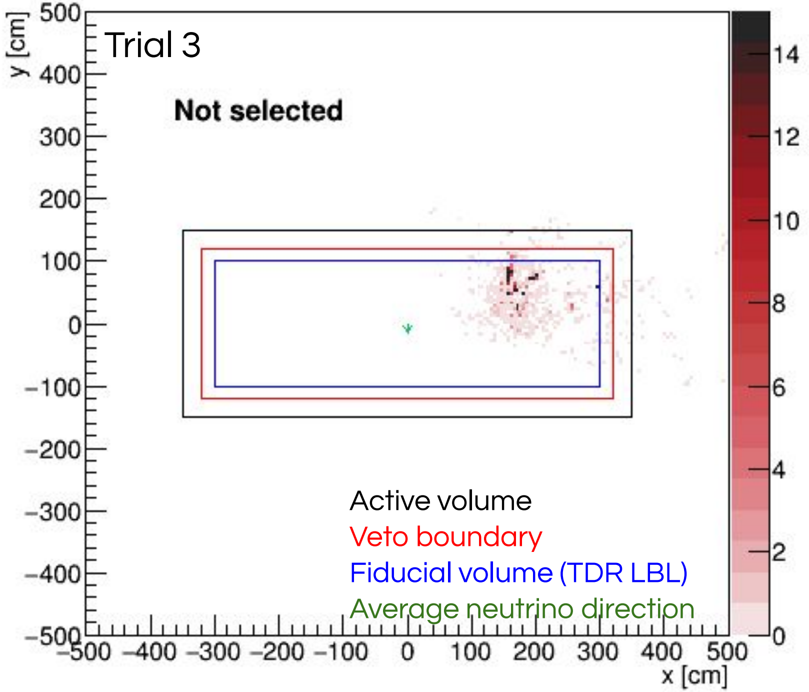


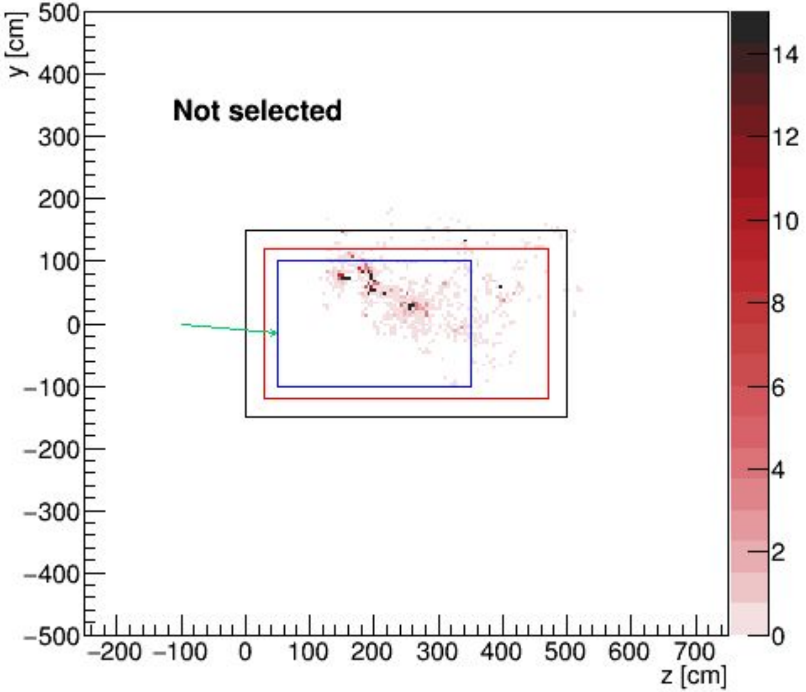
Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{2}{2} = 100\%$$

Translations and rotations

Hadronic system



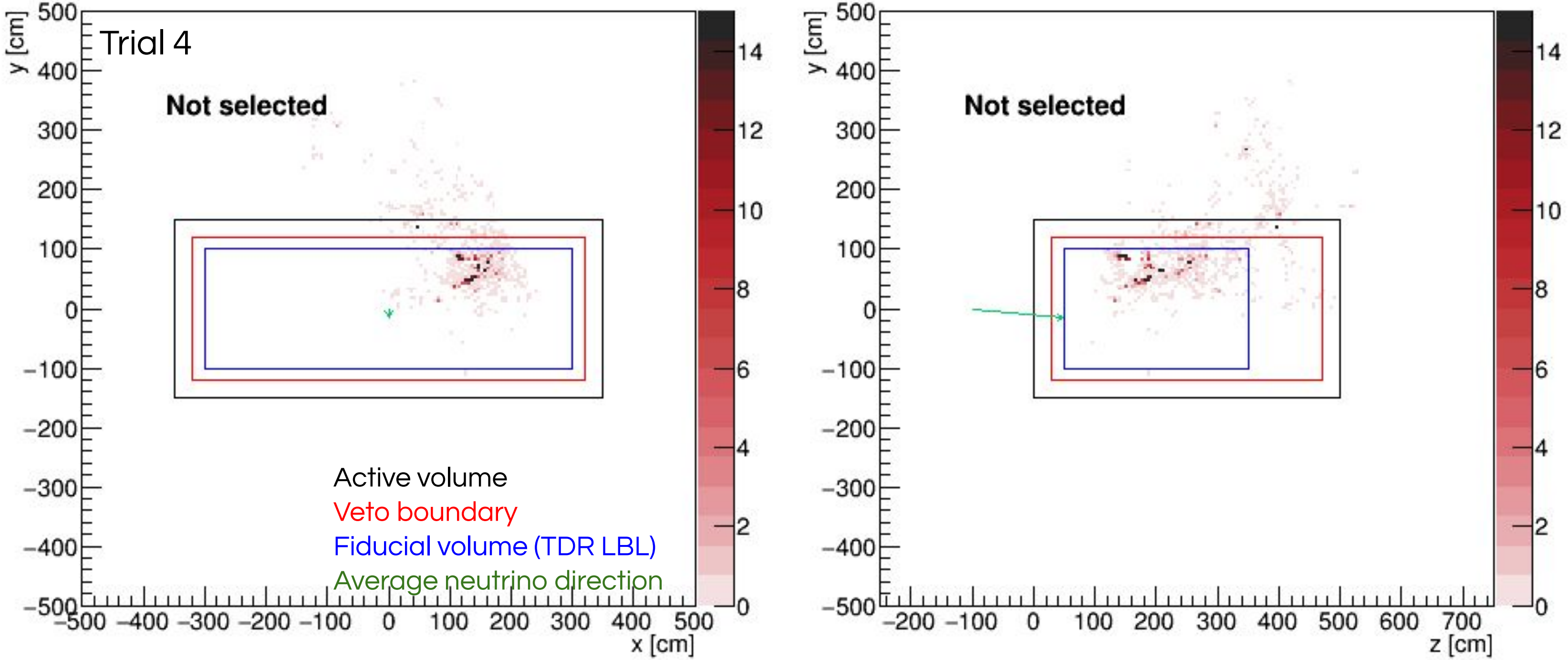


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{2}{3} = 66.7\%$$

Translations and rotations

Hadronic system

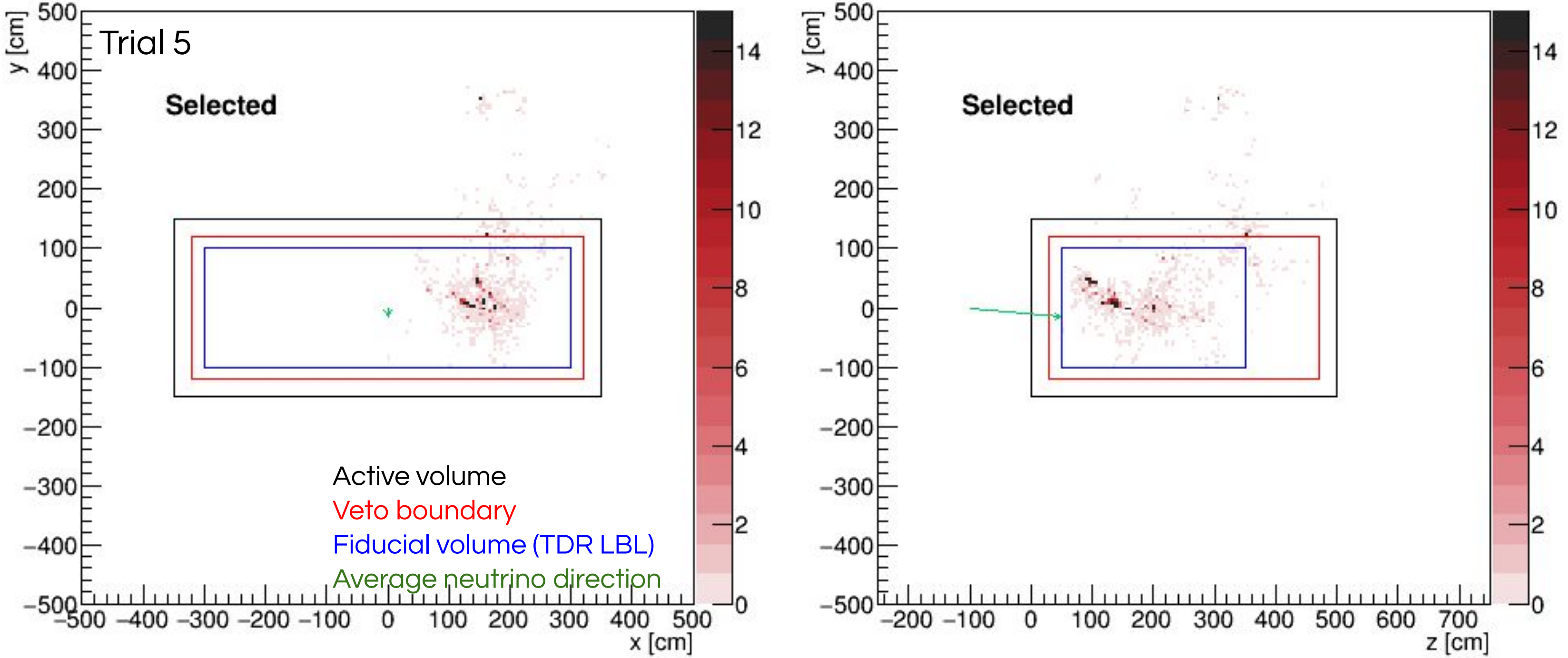


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{2}{4} = 50\%$$

Translations and rotations

Hadronic system

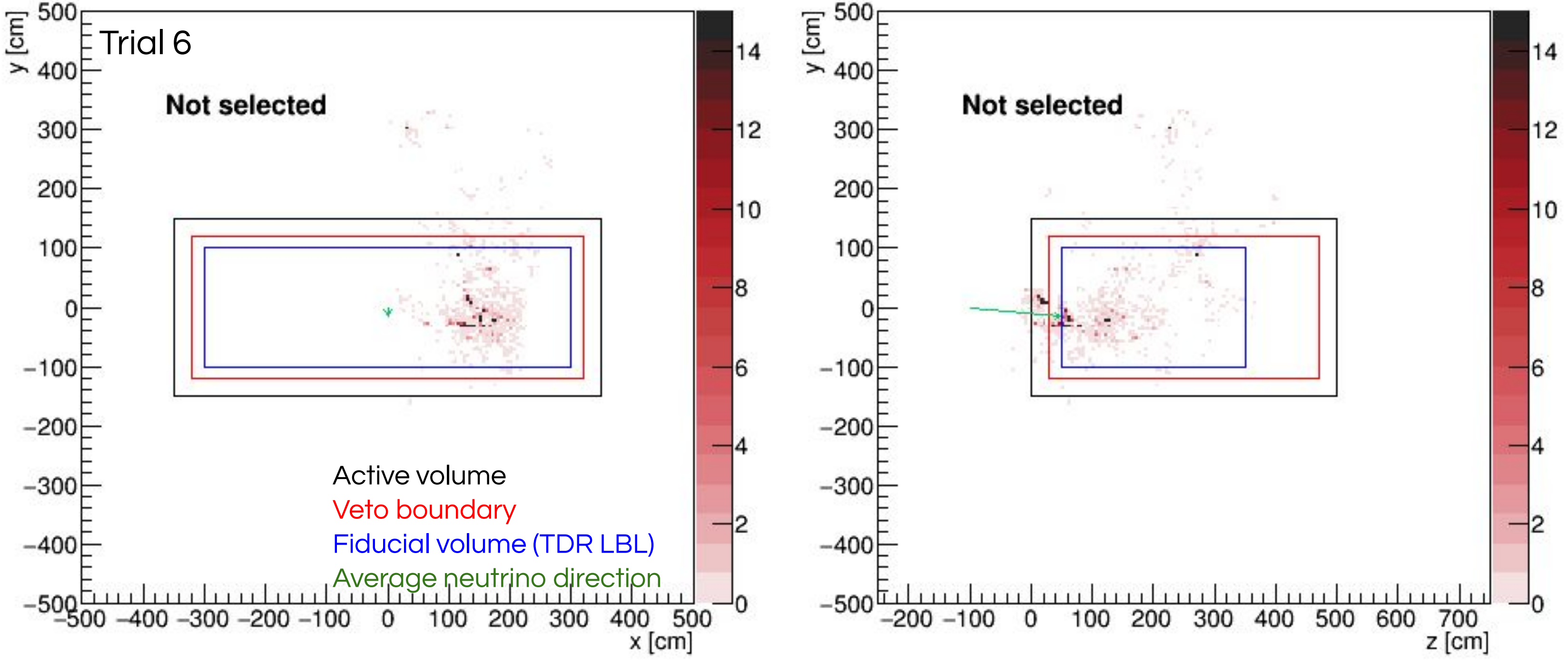


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{3}{5} = 60\%$$

Translations and rotations

Hadronic system

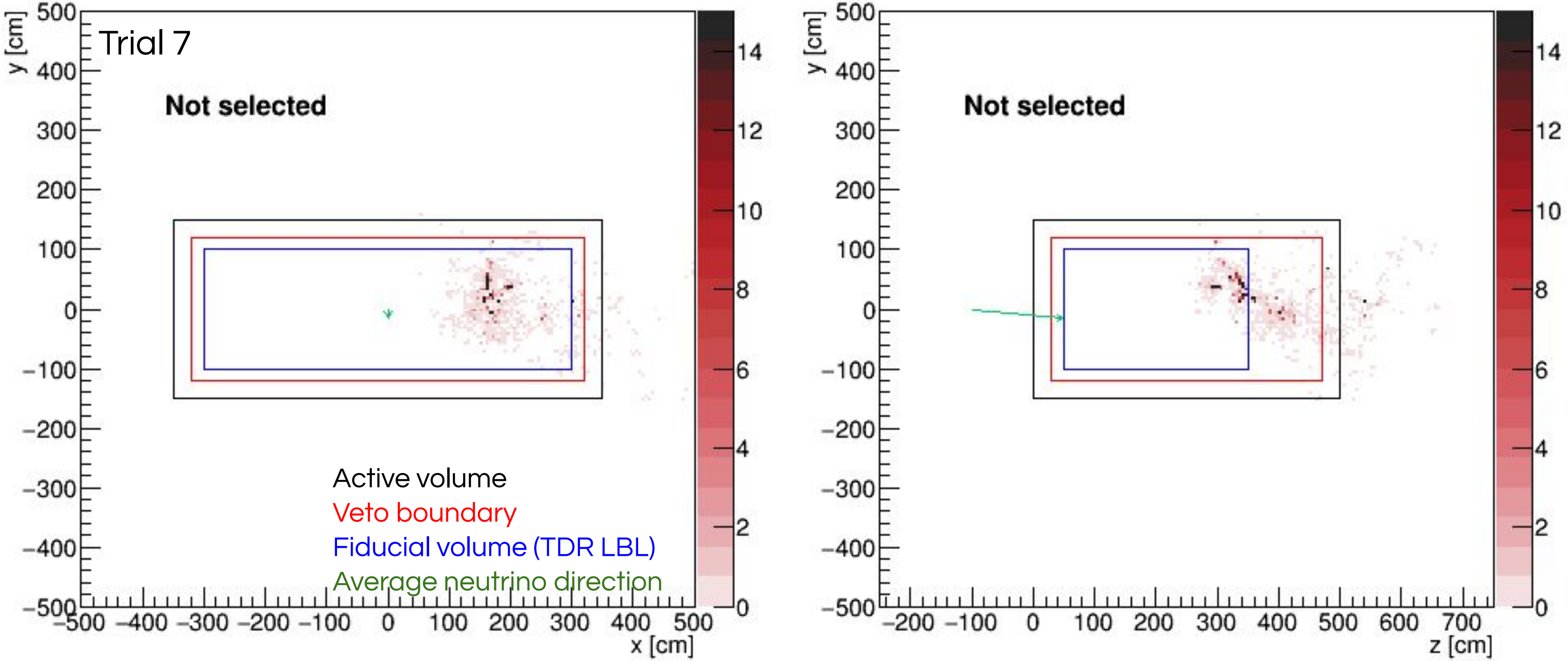


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{3}{6} = 50\%$$

Translations and rotations

Hadronic system

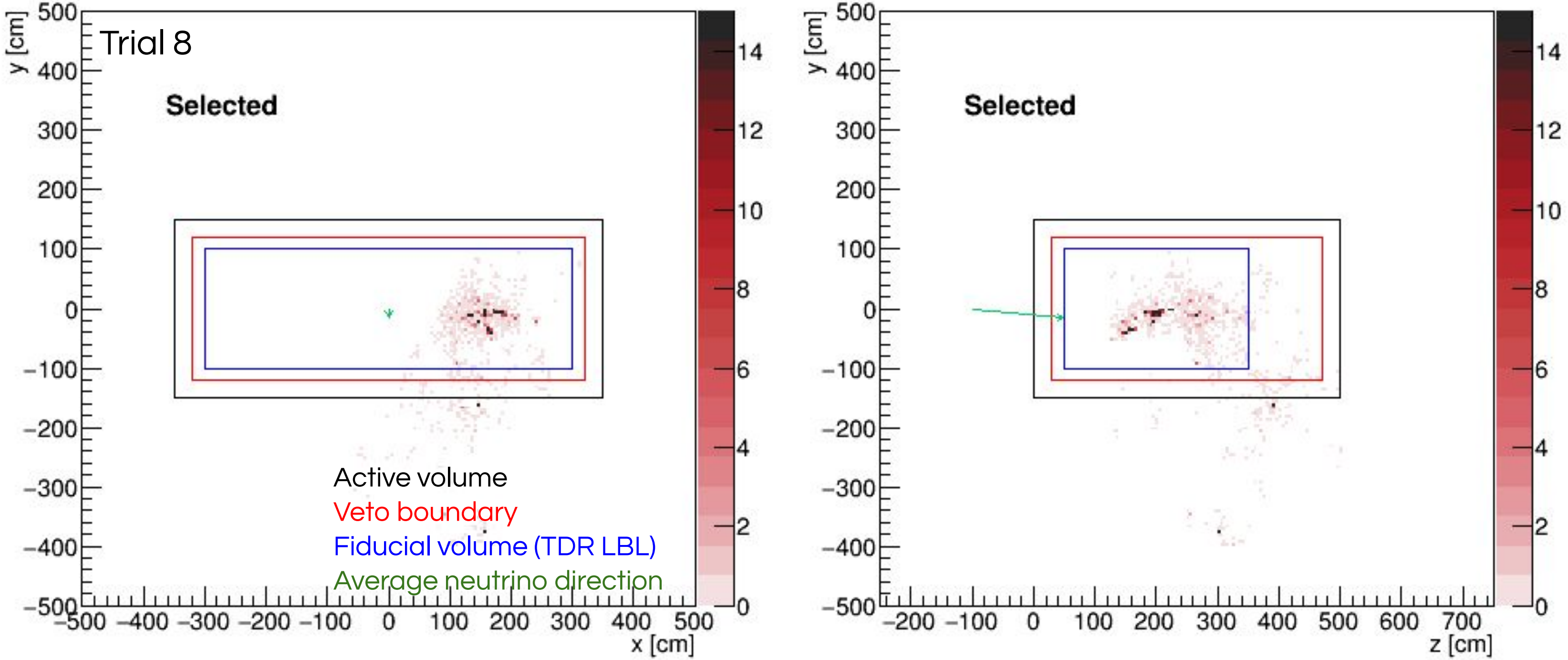


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{3}{7} = 42.8\%$$

Translations and rotations

Hadronic system

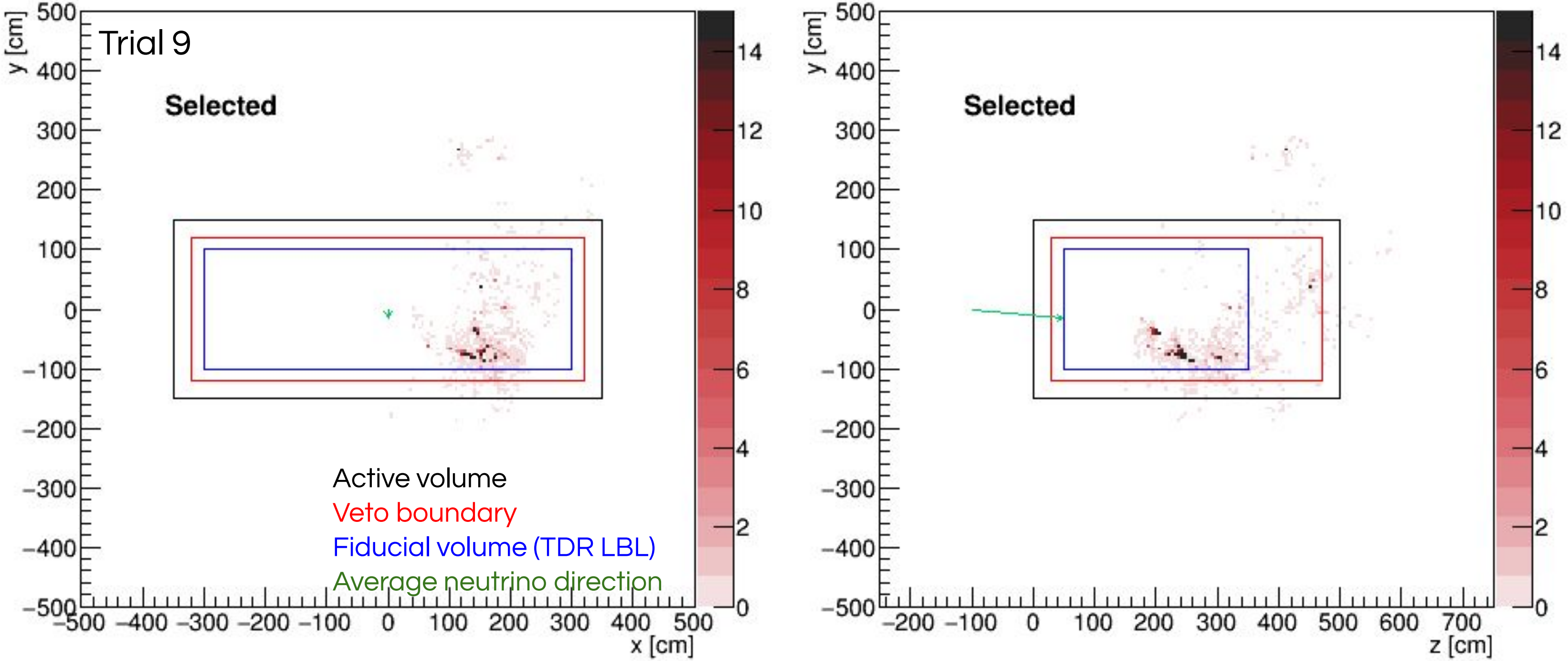


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{4}{8} = 50\%$$

Translations and rotations

Hadronic system

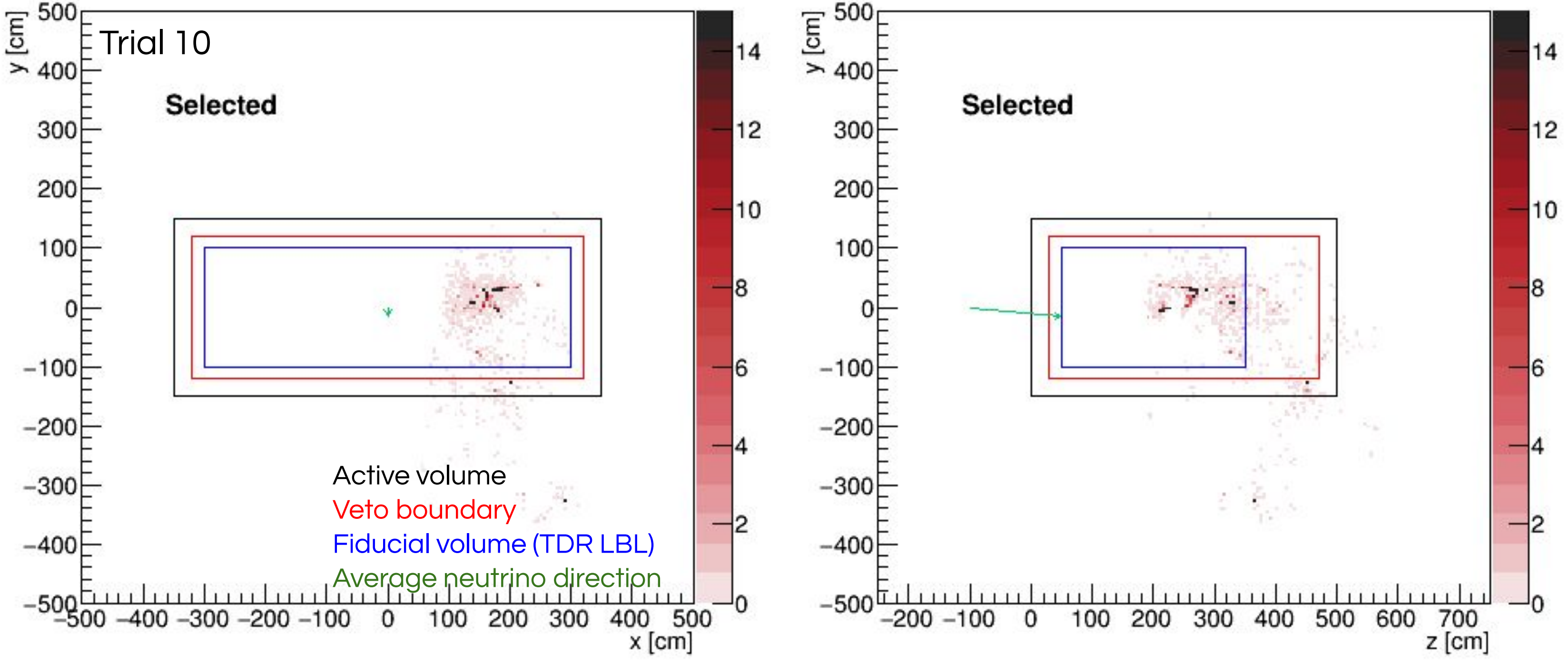


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

$$\eta \approx \frac{5}{9} = 55.6\%$$

Translations and rotations

Hadronic system

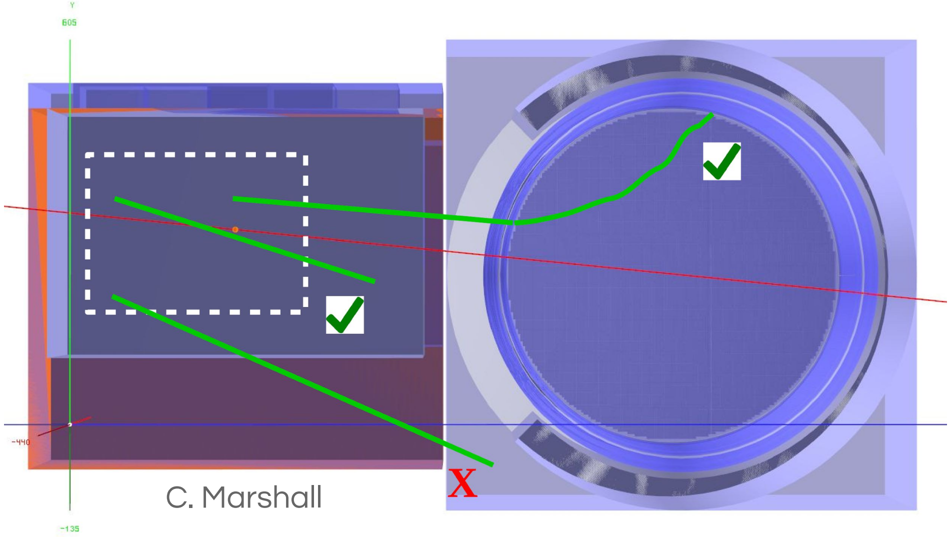


Event is selected if < 30 MeV is deposited in the veto region, 30 cm from active volume edge.

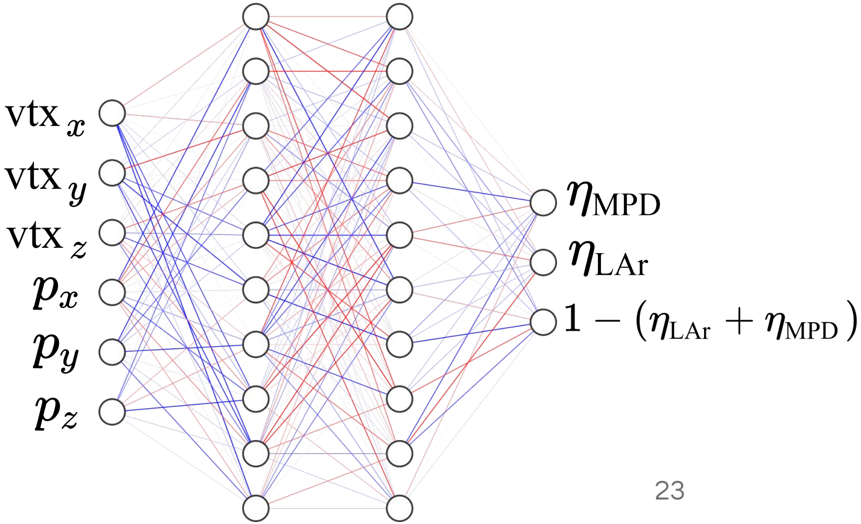
$$\eta \approx \frac{6}{10} = 60\%$$

Muon efficiency neural network

- Train neural network to predict fate of muon as a function of its position and momentum.
  - Output is the probability for the muon to be sampled in the **tracker**, be **contained** in the liquid argon, or **not** be **selected**.
- Start with simple neural network with 2 hidden layers with 64 nodes each and ReLU activation.
  - Implemented in PyTorch: https://github.com/cvilelahep/MuonEffNN

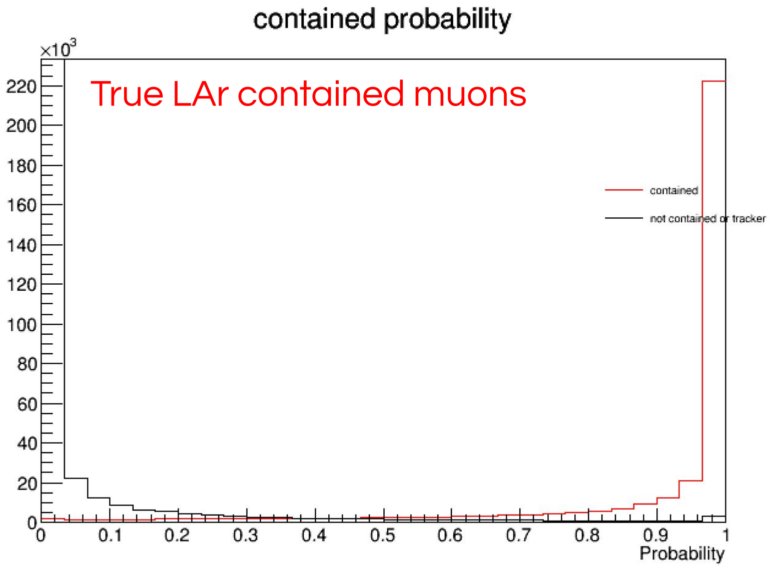


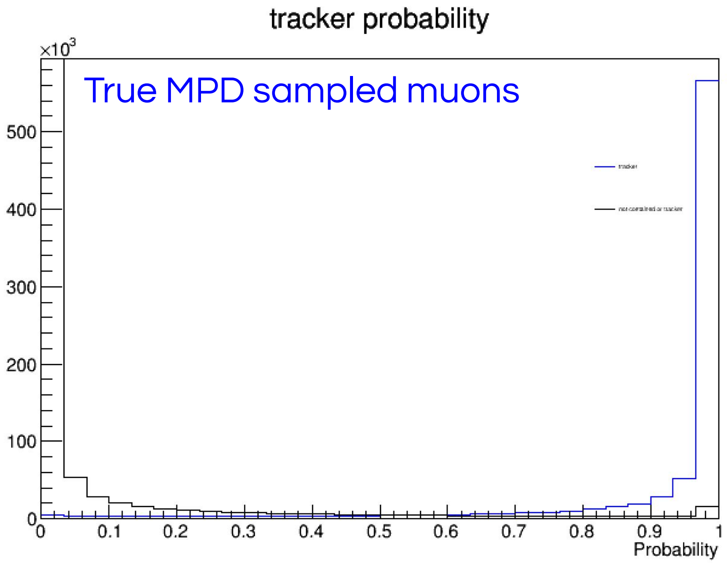
C. Marshall



Muon efficiency neural network output

- Neural network accurately predicts fate of muons based on initial state.
  - Encapsulates ND geometry and muon propagation physics.
- Can be trained on particle gun MC: no interaction model dependence.

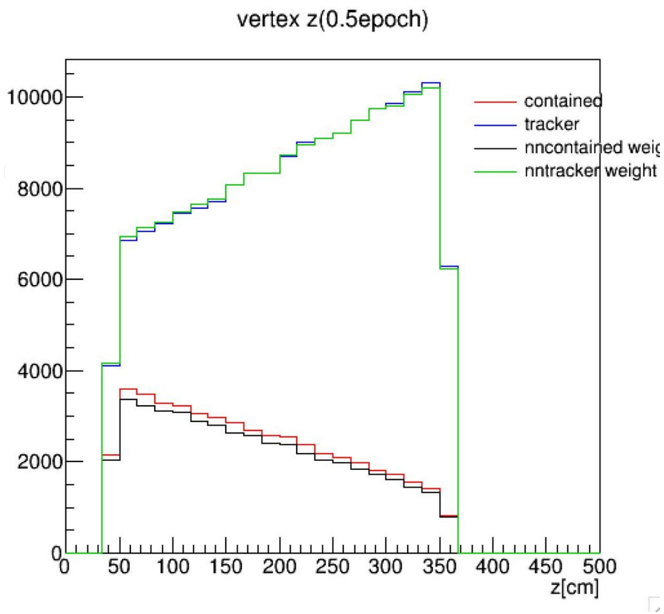


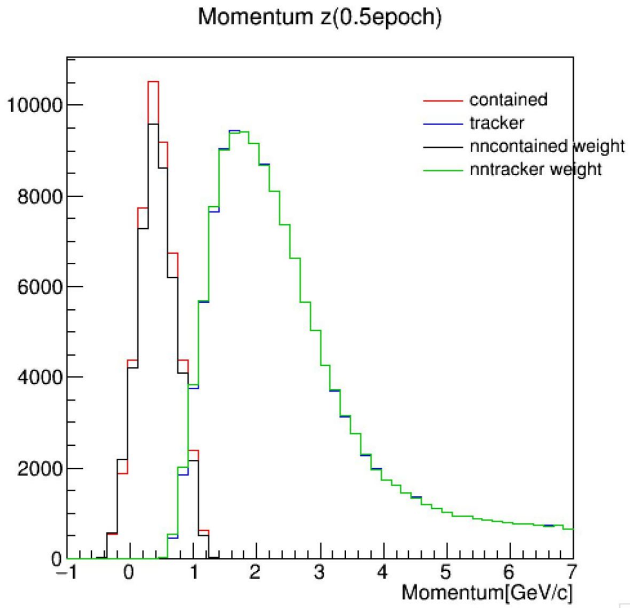


Z. Chen

Compare NN to simulation

- Reweight **all** MC events by the neural network output (“tracker” and “contained” probabilities) and compare to distribution of true contained and tracker muons.
- Neural network reproduces features in momentum and vertex distributions.



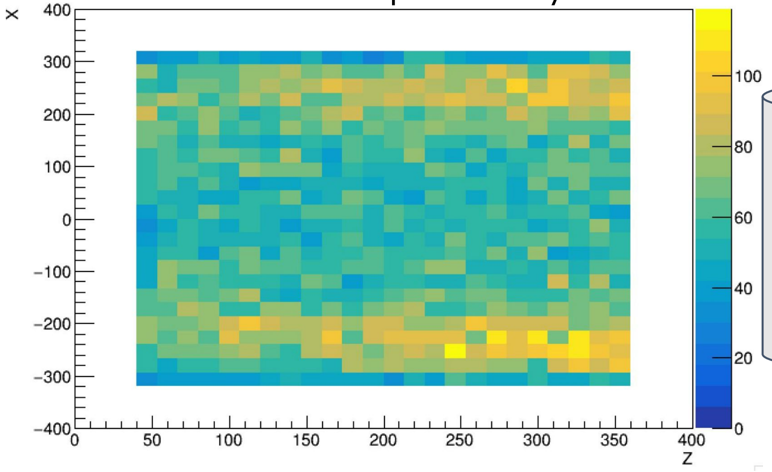


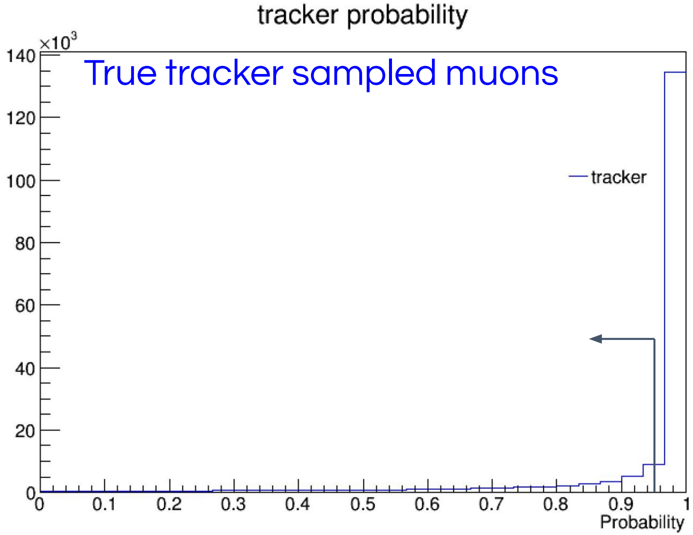
Z. Chen

Muon efficiency neural network output

- True tracker events in the low neural network score tail tend to be at the edges of the detector.
- Harder to predict whether these events will make it into the tracker just with initial position and momentum.

Vertices of true tracker sampled muons with NN probability < 95%

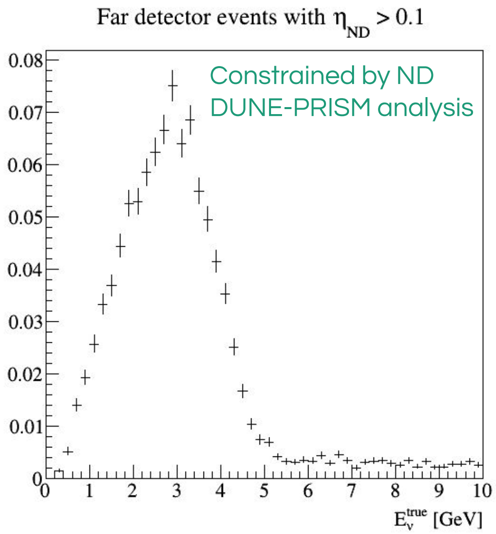


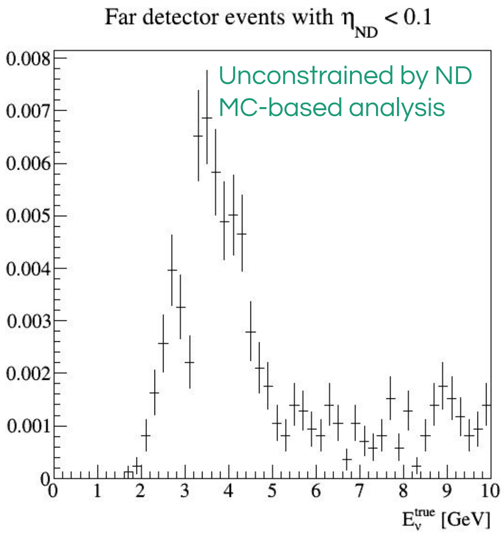


Z. Chen

ND/FD acceptance differences

- Data-driven efficiency estimation works for events that are **selected** in the ND.
- But there will be events at the FD that would **never** be contained in the ND.
- Obtain an ND efficiency for each FD event using the same algorithm.
- FD events with very low ND efficiency are not used in data-driven analysis.
  - Can still be used in traditional MC-based analysis, where the prediction is extrapolated from the ND data.
  - There is no direct ND constraint on these events!



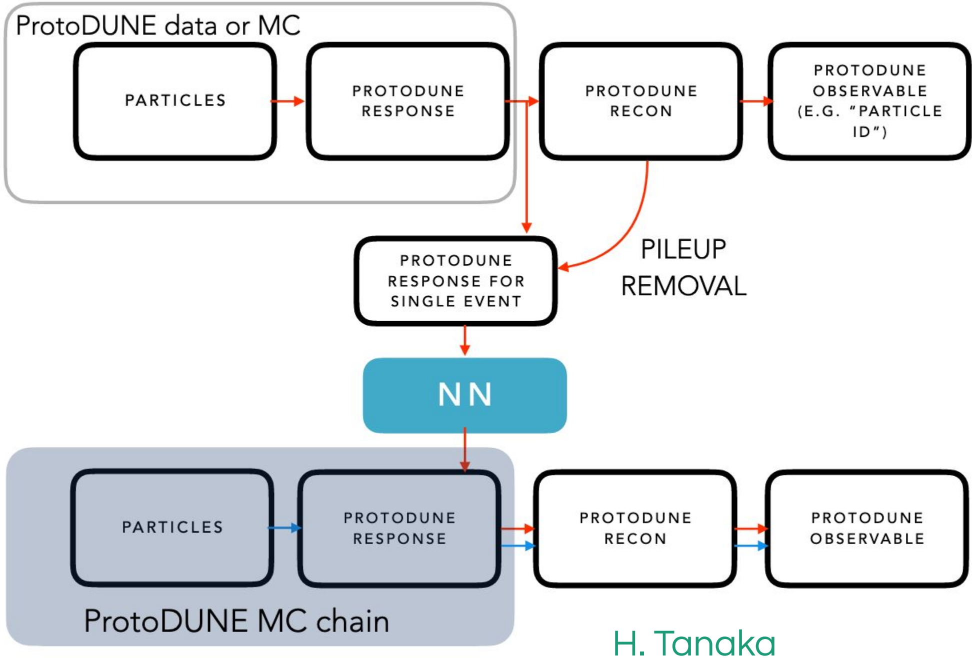


ND to FD response translation

- The missing piece is translating ND events to the FD:
  - Does a given ND event pass the FD NN event selection criteria?
  - How do ND observables map to the FD (Erec, etc)?
- Plan recently proposed by Hiro Tanaka:
  - Use machine learning to "unfold" ND events back to a level that is common with the FD. E.g., energy depositions in the LAr.
  - Propagate those energy depositions through the FD simulation and reconstruction chain.
  - Get an "FD-equivalent" for each ND event.

ND to FD response translation

- I'm interested in collaborating with the SLAC group on this.
- In particular, I would like to test this approach using ProtoDUNE data.
- Feedback welcome! ;)



Summary

- DUNE-PRISM will allow for a largely data-driven oscillation analysis.
- For this to be successful we need to be able to match ND events to FD events. This is challenging!
  - Developed a method to correct for first-order efficiency and acceptance differences.
  - Promising proposal to translate events between the two detectors.
- Taking a fundamentally different approach to oscillation analysis has led to the development of exciting new ideas!