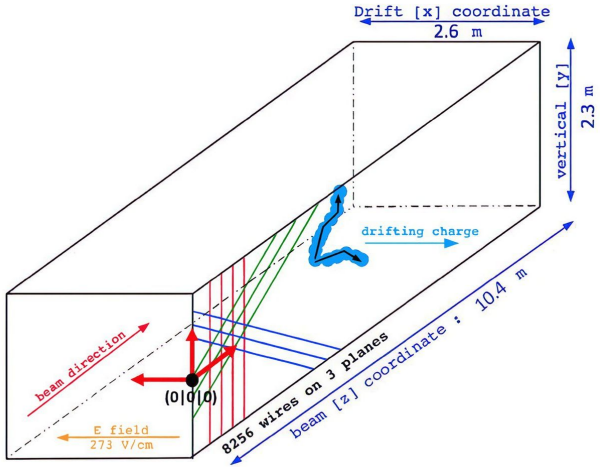
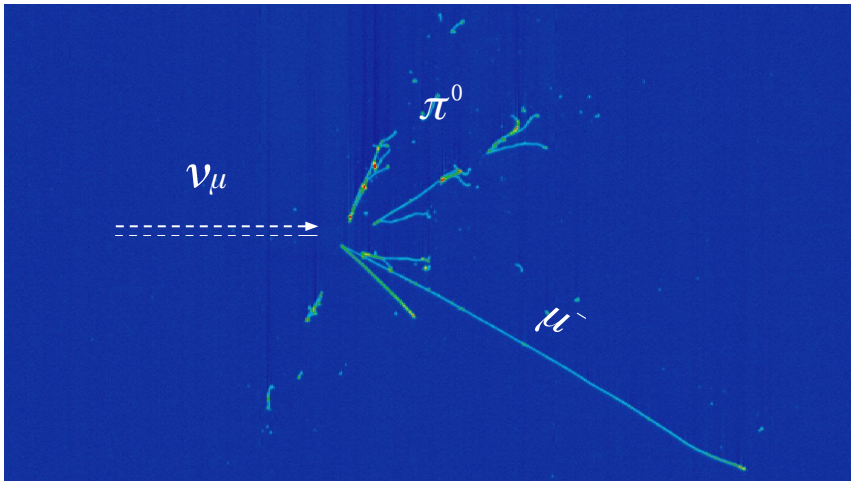
Optimal Transport for e/π^0 Particle Classification in LArTPC Neutrino Experiments

Chuyue “Michaelia” Fang
University of California, Santa Barbara
New Perspectives - 7/8/2024
FERMILAB-SLIDES-24-0157-V

LArTPC Neutrino Detectors and MicroBooNE



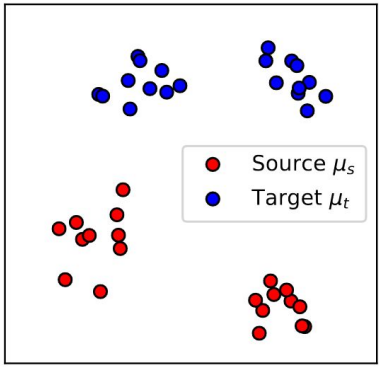
Operational Principle of MicroBooNE LArTPC



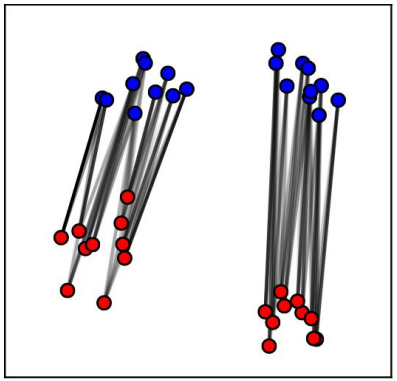
MicroBooNE Event Display of A Charged Current ν_μ Interaction

- π^0 is a crucial background to oscillation experiments and BSM searches
- both e and π^0 present as EM showers, making it a reconstruction challenge to separate them
- currently using MicroBooNE Public Datasets for samples input

What is Optimal Transport?



Distributions (Flamary 2019)



Transport plan visualized

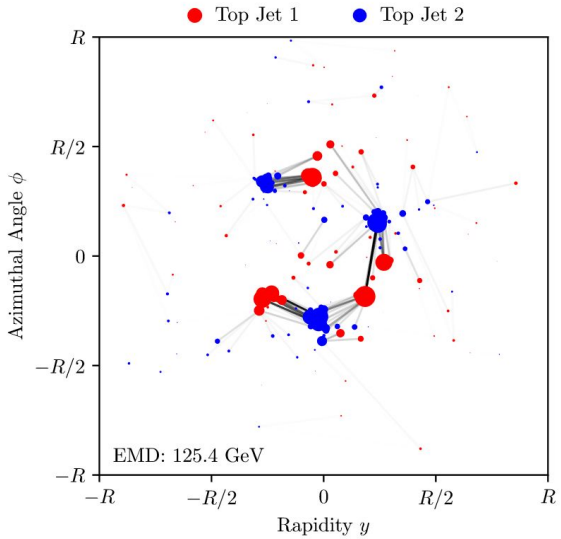
- “the general problem of moving one distribution of probability mass to another as efficiently as possible”
- provides a transport plan and an optimal transport distance, which is used to compare two probability distributions

Why Optimal Transport?

- advantages of optimal transport
  - optimal transport performs well with sparse dataset
  - more transparent in how it's achieving the results
  - can be used as pre-processing for further analysis (ex.kNN)
- optimal transport has different variants and metrics which each has their own benefits
  - currently using 2-Wasserstein distance

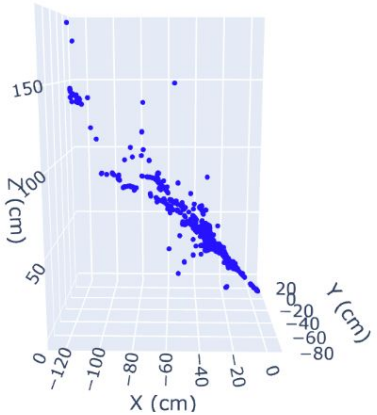
Optimal Transport in HEP

- optimal transport has been used for jet classification in LHC data by several groups, including N. Craig and J. Howard at UCSB who we're working with
- optimal transport outperforms traditional methods in jet classification; it's competitive with standard machine learning methods and it's also easy to interpret

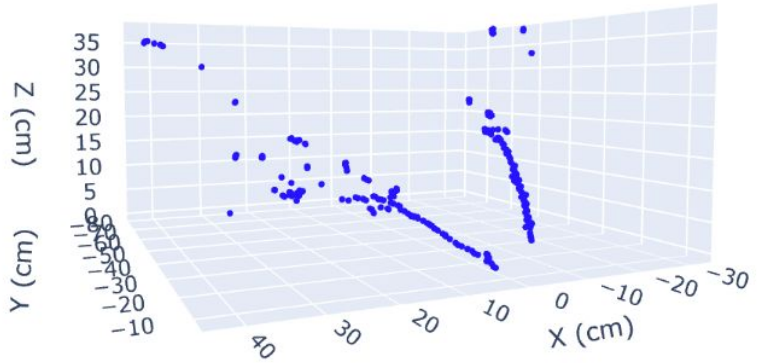


Optimal Transport for Jets (Komiske 2019)

e/π^0 Events in LArTPC



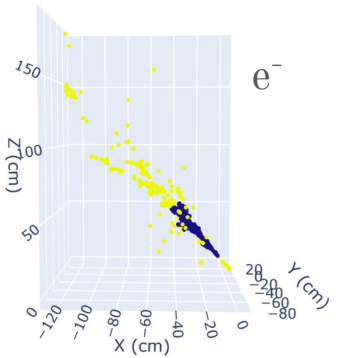
3D Reconstructed e Event



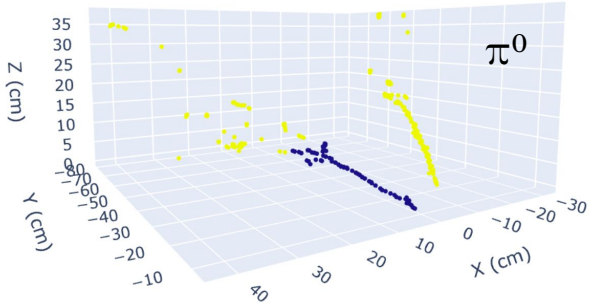
3D Reconstructed π^0 Event

- e produces one EM shower
- π^0 decays into two photons which produce two EM showers
- we aim to use OT for classification without directly reconstructing the EM showers separately

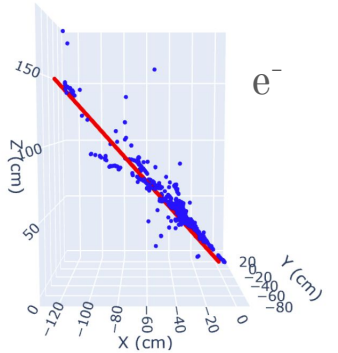
Identifying Principal Axis of a 3D Reconstructed Event



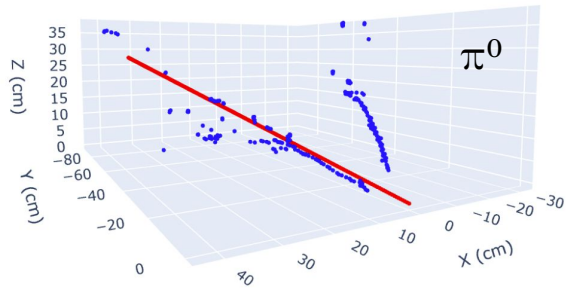
3D Reconstructed e^- event with identified largest cluster



3D Reconstructed π^0 event with identified largest cluster



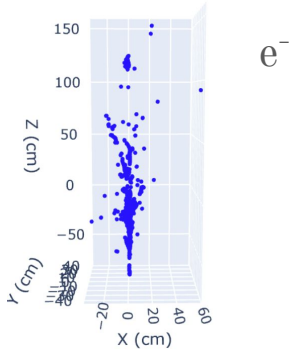
e^- event with principal axis



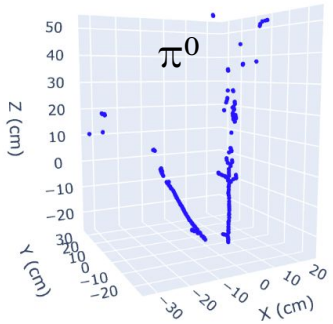
π^0 event with principal axis

- proximity clustering finds largest cluster
- Principal Component Analysis (PCA) on largest cluster to identify principal axis of the event

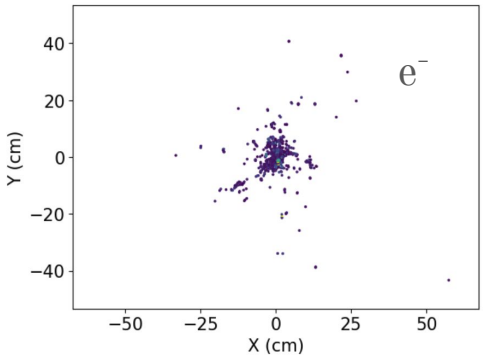
Taking Planar Projections of 3D Reconstructed Sample



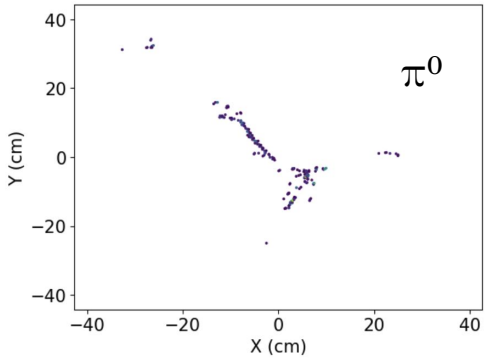
Rotated e^- event



Rotated π^0 event



Planar projection of e^- event



Planar projection of π^0 event

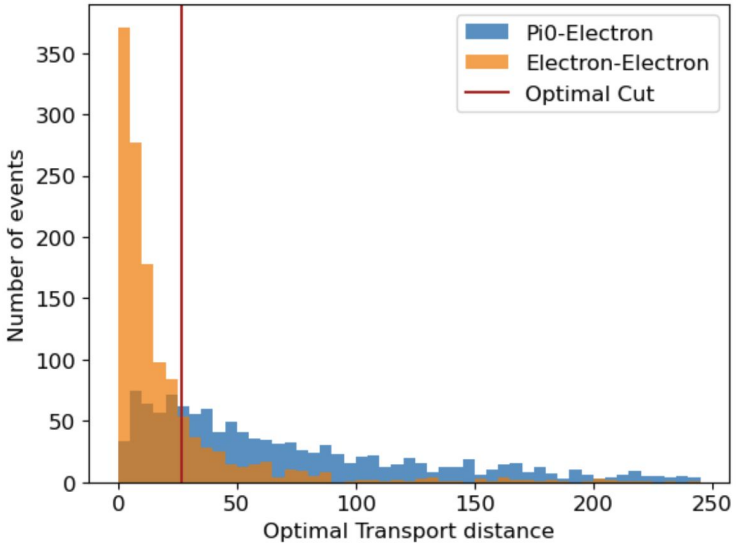
- rotate all the spacepoints so that principal axis aligns with Z-axis
- project all spacepoints onto XY-plane

Optimal Transport Computation

- e and π^0 samples are separated into 8 different energy bins
- optimal transport distances are computed between events in the same energy bin with equal numbers of e and π^0 events
  - planar projections of 3D reconstructed samples are used as input
- OT distances are used for classification
  - different machine learning methods could be used for classification with OT distances as input

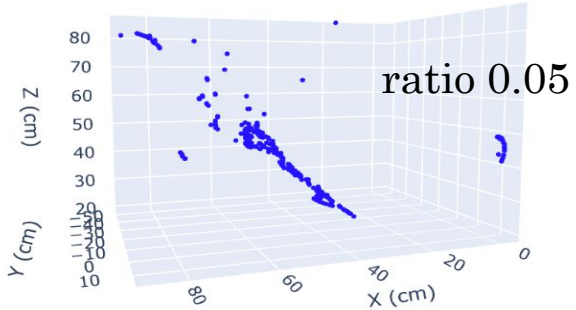
Results - Performance of Optimal Transport

- using a cut on OT distances
  - accuracy: 0.764
- using OT distances as input for machine learning methods
  - k-Nearest Neighbors (kNN)
    - accuracy: 0.786
  - Support Vector Machine (SVM)
    - accuracy: 0.809

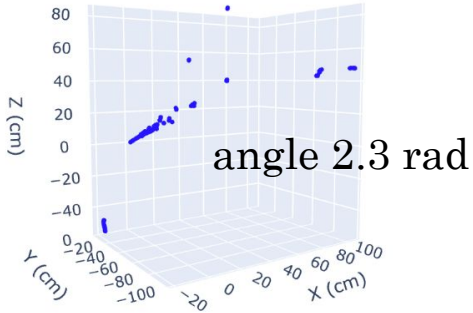


Optimal Transport Distance for π^0 and e Events Compared to Electron Events for First Energy Bin

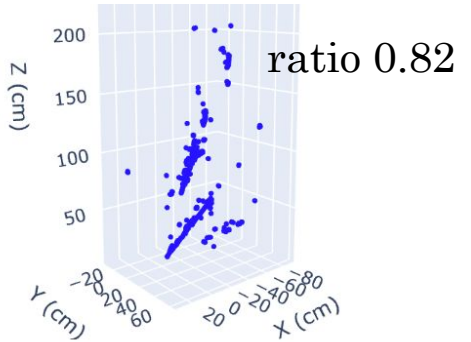
π^0 Kinematic Variables



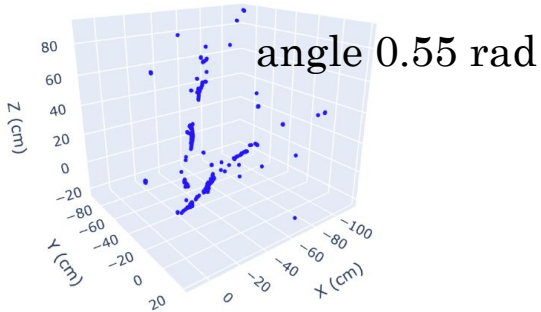
π^0 event with high shower asymmetry



π^0 event with large opening angle



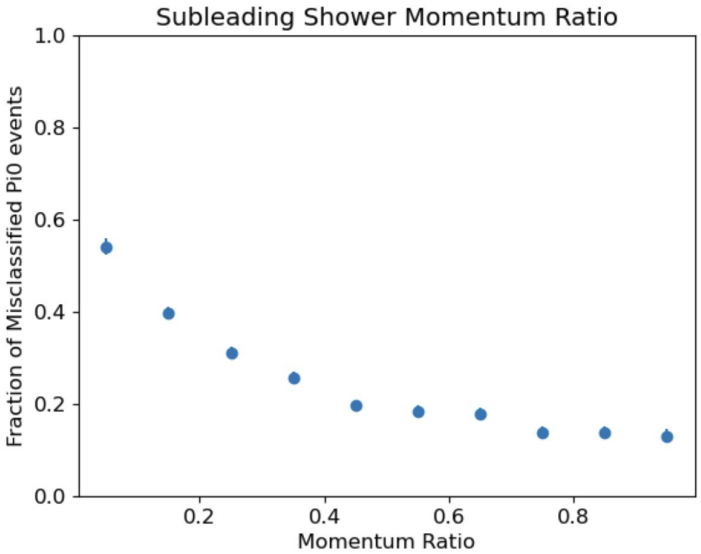
π^0 event with low shower asymmetry

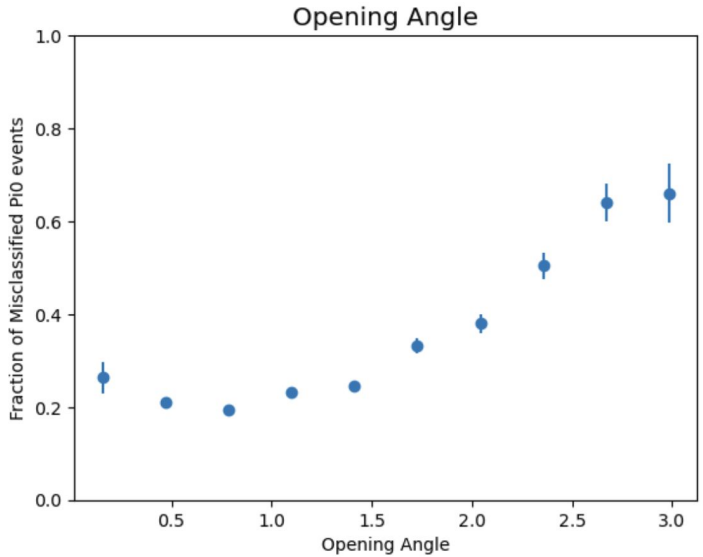


π^0 event with small opening angle

- shower asymmetry ($p_{subleading} : p_{primary}$ ratio)
- opening angle between two showers

Performance Compared with Kinematic Variables





- accuracy increases with less shower asymmetry as expected
- low accuracy at high end for opening angle

Summary

- application of optimal transport for LArTPC neutrino experiments
  - have implemented optimal transport on MicroBooNE public datasets
  - overall able to separate π^0 from e using OT distances
  - finalizing first implementation of optimal transport for neutrino event classification
  - possible future implementation in SBN and DUNE analyses

Backup slide - p-Wasserstein distance

$$W_p(\mathcal{E}, \tilde{\mathcal{E}}) = \min_{g_{ij} \in \Gamma(\mathcal{E}, \tilde{\mathcal{E}})} \left(\sum_{ij} g_{ij} \|x_i - \tilde{x}_j\|^p \right)^{1/p}$$