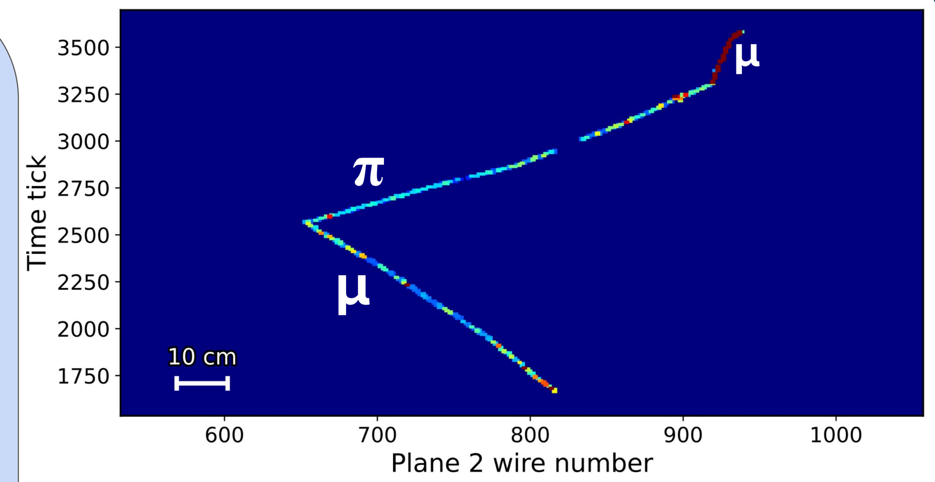
μ/π Separation in LArTPCs Using Optimal Transport-Based Machine Learning

Christopher Sauer and Xiao Luo at UC Santa Barbara

Contact Us:
christophersauer@ucsb.edu, xiaoluo@ucsb.edu

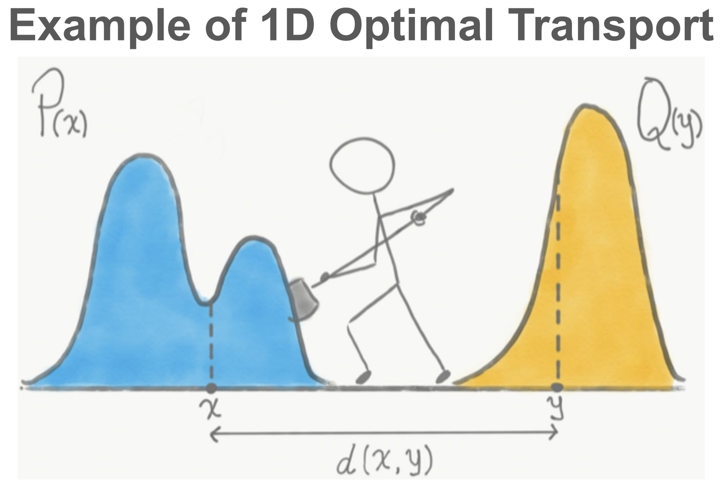
Physics Goals

- Many Long-lived particle models predict **dimuon signatures**
- CC1π events are an "irreducible" background for dimuon signals with current reconstruction methods
- Need high-purity mu/pi separation to reduce this background and improve LLP sensitivity



Standard LArTPC reconstruction methods lack the ability to effectively distinguish the MIP-like profiles of μ and π tracks, we introduce a topology based ML method to improve the separation efficiency for beyond standard model searches

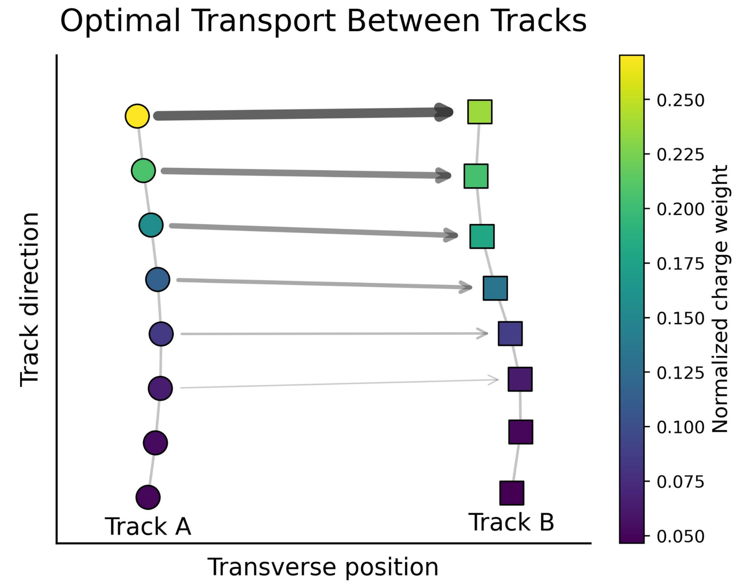
Example of 1D Optimal Transport



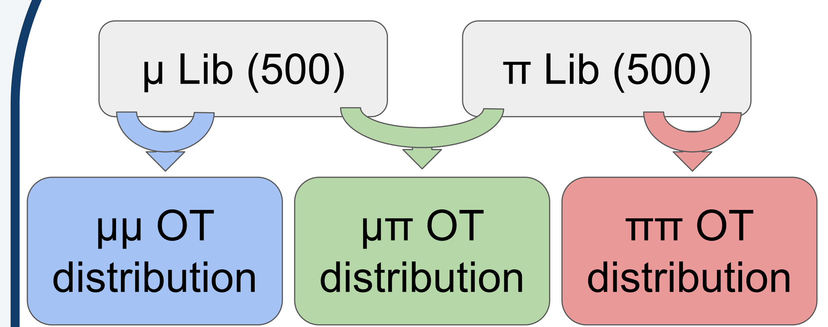
Optimal Transport (OT): Optimization method that determines the minimal "cost" to morph one distribution P(x) into another Q(y)

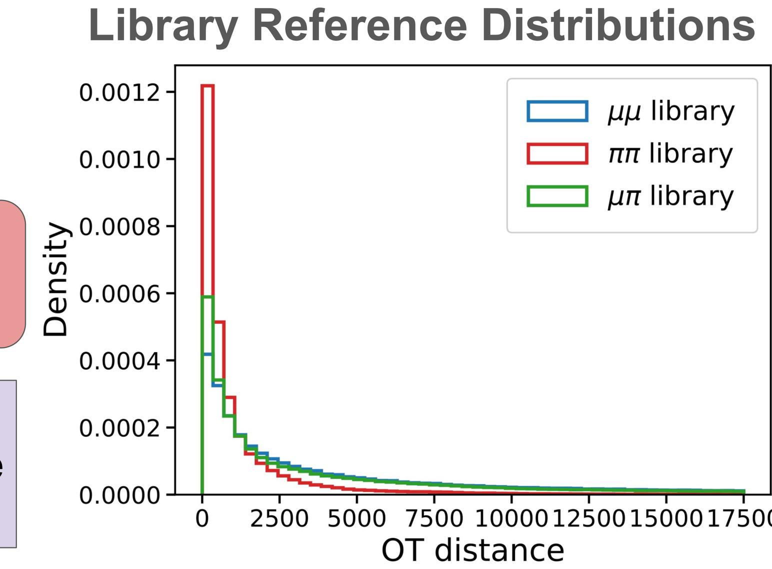
- Compute the distance based cost matrix for points in both tracks
- **Minimize the total cost** over all transport plans
- We use the 2-Wasserstein distance to compute the cost of transportation between points

$$W_2^2(\mu,\nu) = \min_{\gamma \in \Pi(a,b)} \sum_{ij} \gamma_{ij} \, \|x_i - y_j\|^2$$



Analysis Method

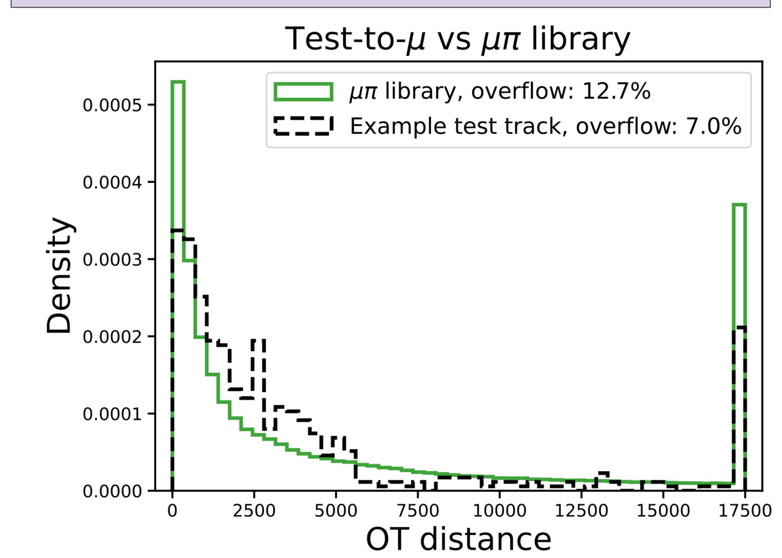


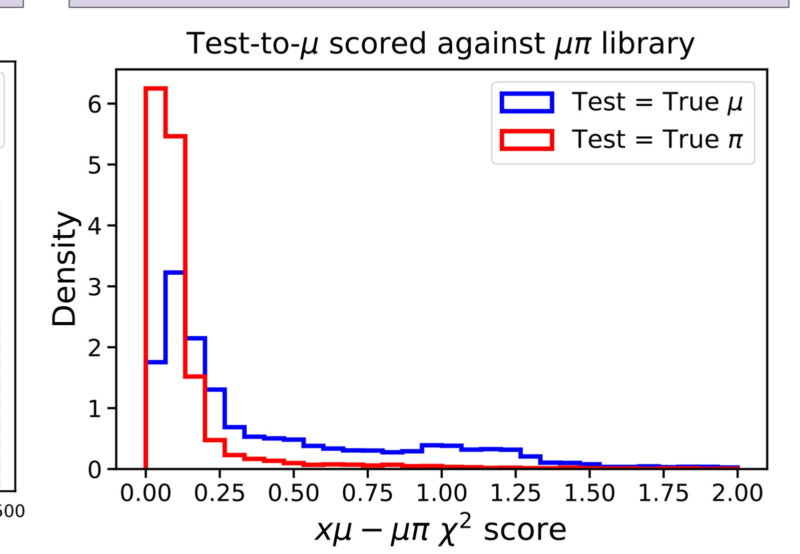


Step 1: Compare **unknown track** with all library tracks to create 2 OT distance distributions

Step 2: Compare test-to-library distributions to lib distributions using χ^2

Step 3: Repeat for all unknown tracks, for each 1 retrieve 4 χ^2 scores

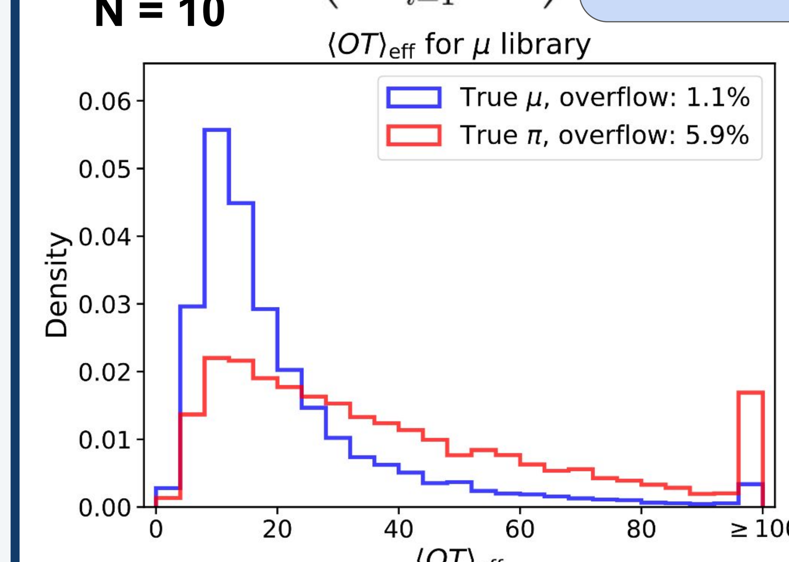


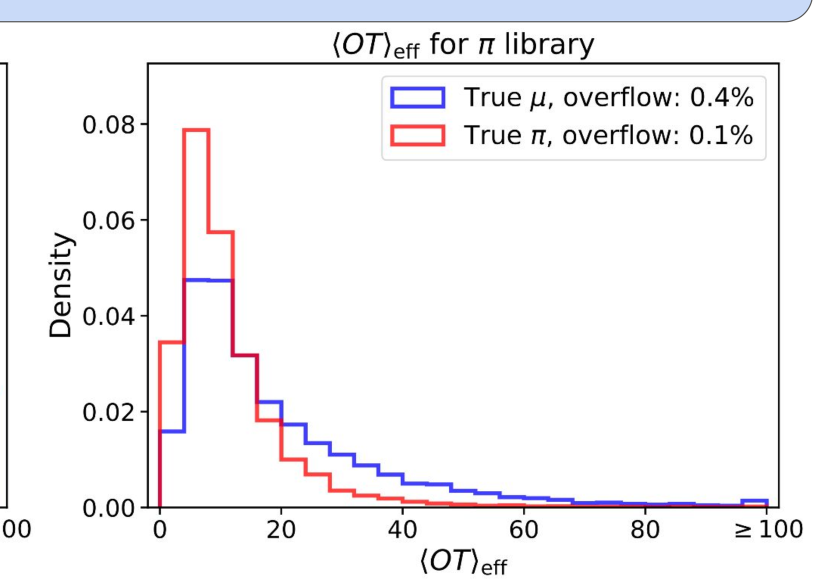


$$\langle \mathrm{OT} \rangle_{\mathrm{eff}} = \frac{1}{L} \left(\frac{1}{N} \sum_{i=1}^{N} d_{\mathrm{OT}}^{(i)} \right)$$

N = 10

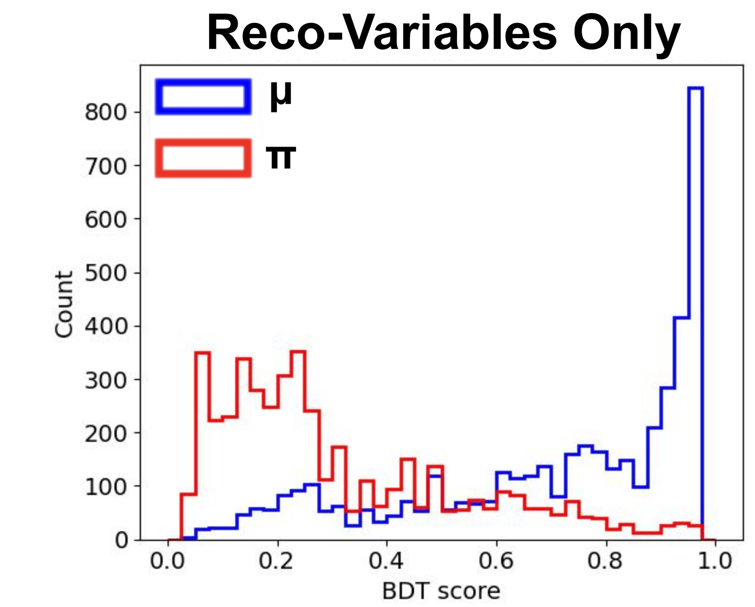
Nearest neighbor (SP count) study to produce **2 average OT distances** to be used in the BDT.

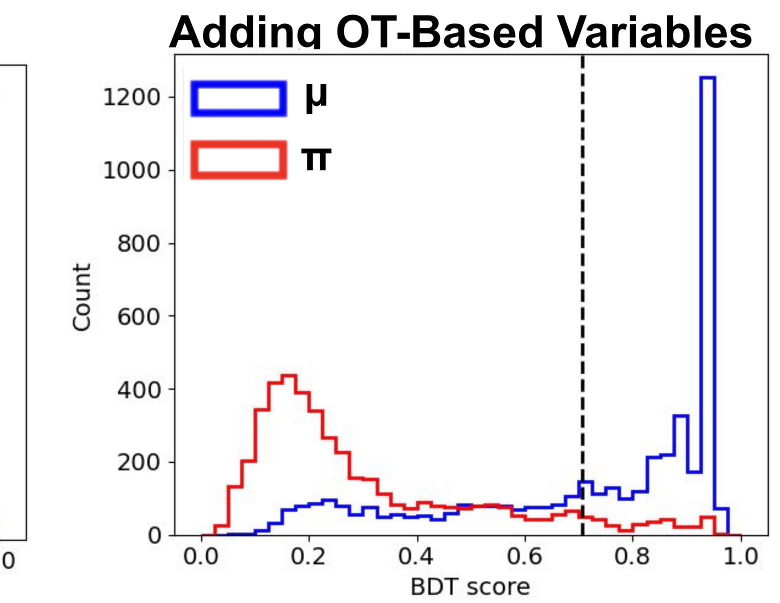




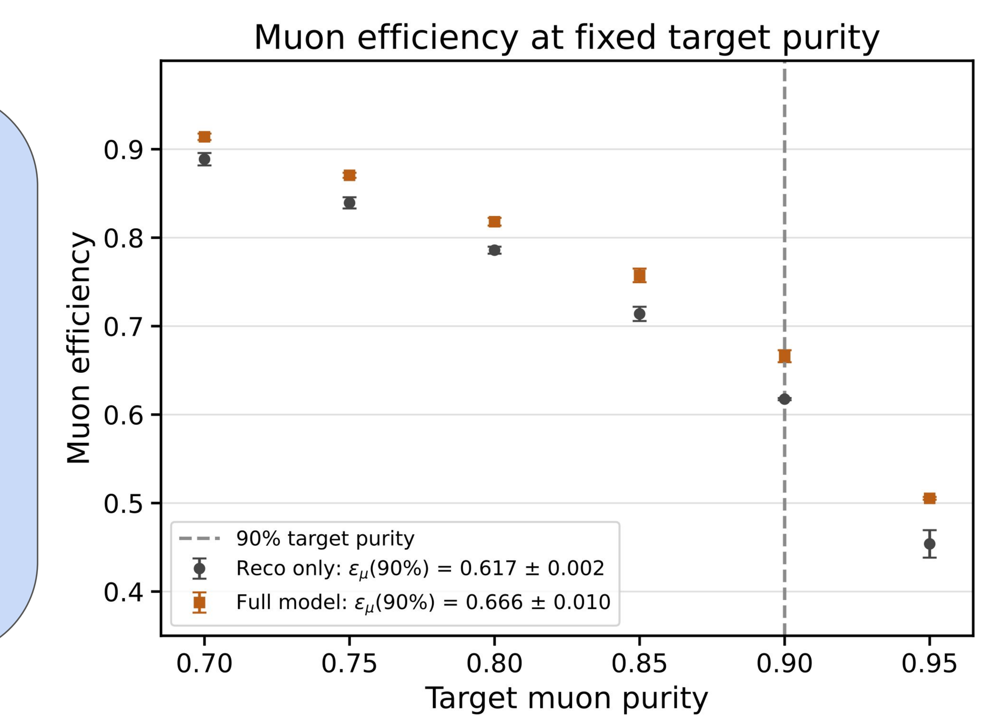
Results

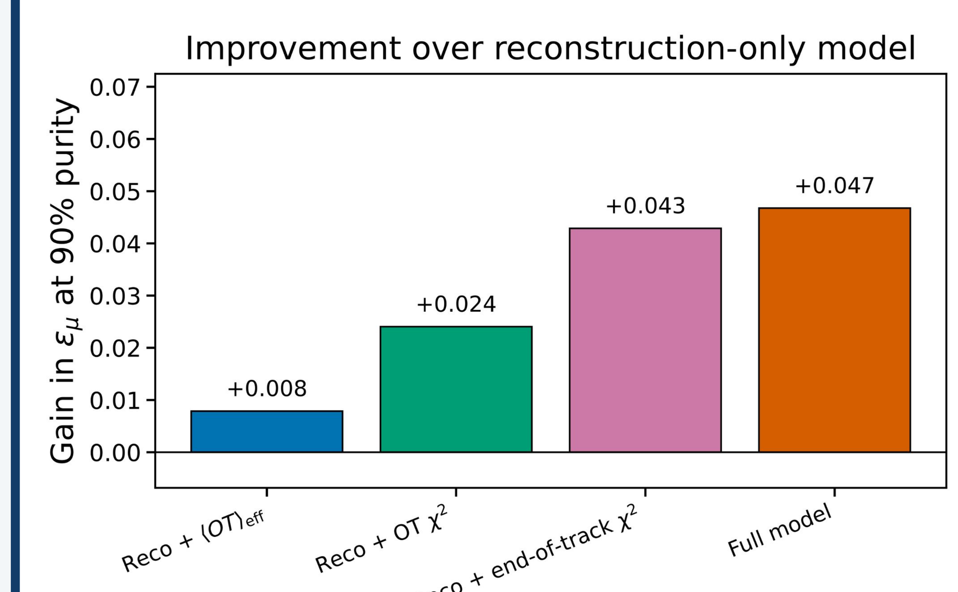
Goal: Improve sensitivity for dimuon event searches by reducing the overwhelming background of CC1pi events. I**mproving the achieved muon efficiency at a fixed high muon purity of 90%** minimizes the pion misidentification rate that contributes to the primary background for dimuon searches





- **67% muon efficiency** at 90% purity
- **5% average efficiency increase** from inclusion of OT features





Major improvements from the end of track features: topology based analysis leverages OT for improved separation

Outlook

Estimate dimuon sensitivity (Z) as the signal rate divided by the square root of the background

$$Z \propto \frac{\epsilon_\mu^2}{\sqrt{\epsilon_\mu f_{\pi\to\mu}}} \propto \epsilon_\mu^{3/2}$$

12% increase in sensitivity to dimuon events achieved by OT above standard reconstruction variables

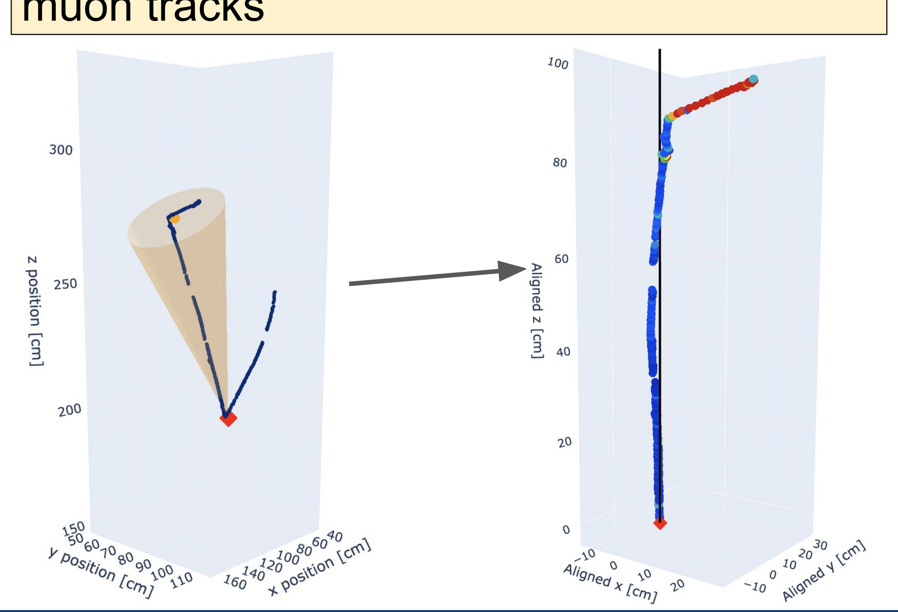
OT scaling:
1e21 POT → ~200k tracks. Using a library size of 500 this workflow takes **~2000 CPU hours.**

Pre-Processing

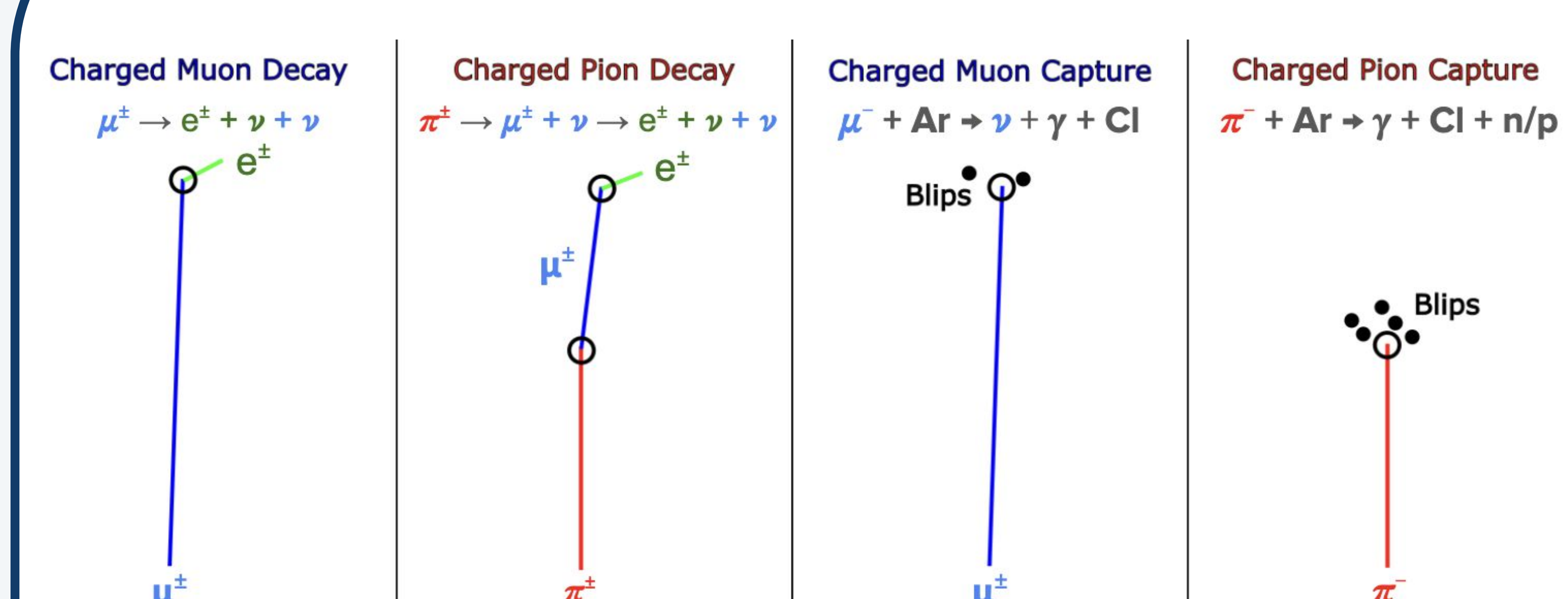
Quality Cuts:
- Fully contained tracks
- no large wire gaps
- > 50 reconstructed space points
- > 30 cm length

Pre-Processing Stages:
1. Spacepoint selection
2. Origin alignment
3. Rotation to beam direction
4. Charge weighting

MicroBooNE Public Dataset – BNB Inclusive Overlay of **~753k simulated neutrino events with ~14k** well reconstructed pion tracks, abundance of muon tracks



End of Track Analysis



- Leverage end of track topology differences by isolating the last 30 cm
- Produces **4 more Chi-Squared** scores to include in the BDT