ID/Track ML meeting

02/20/2023

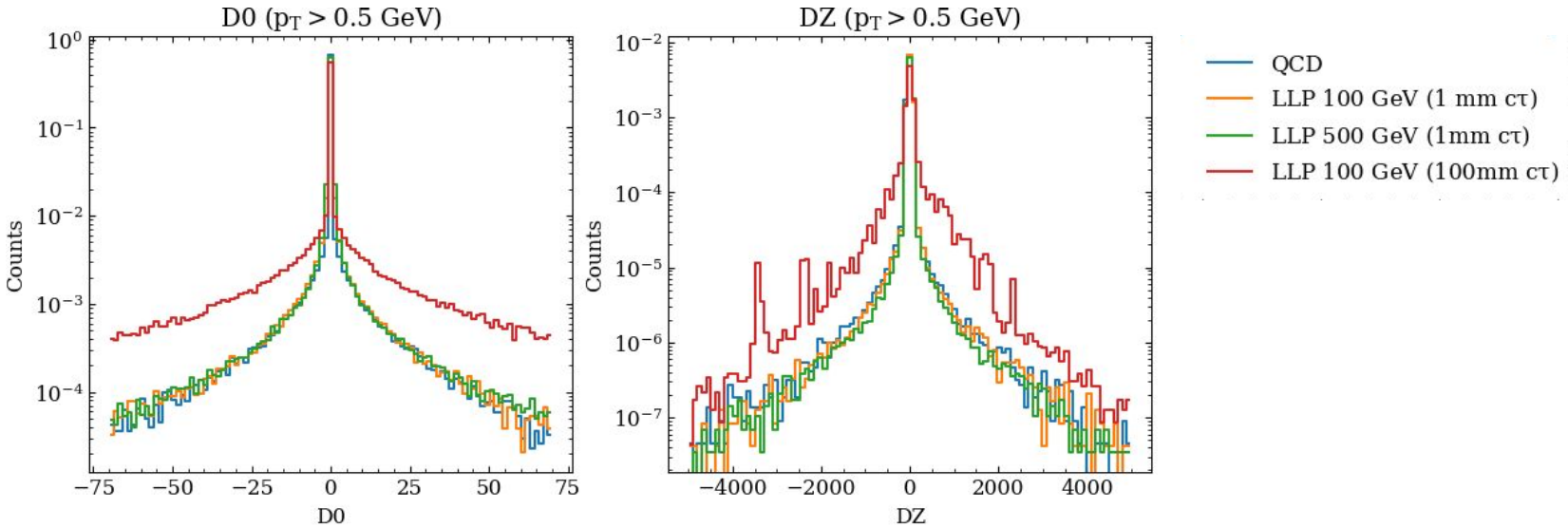
Sam Young

Stanford University

SLAC NATIONAL ACCELERATOR LABORATORY

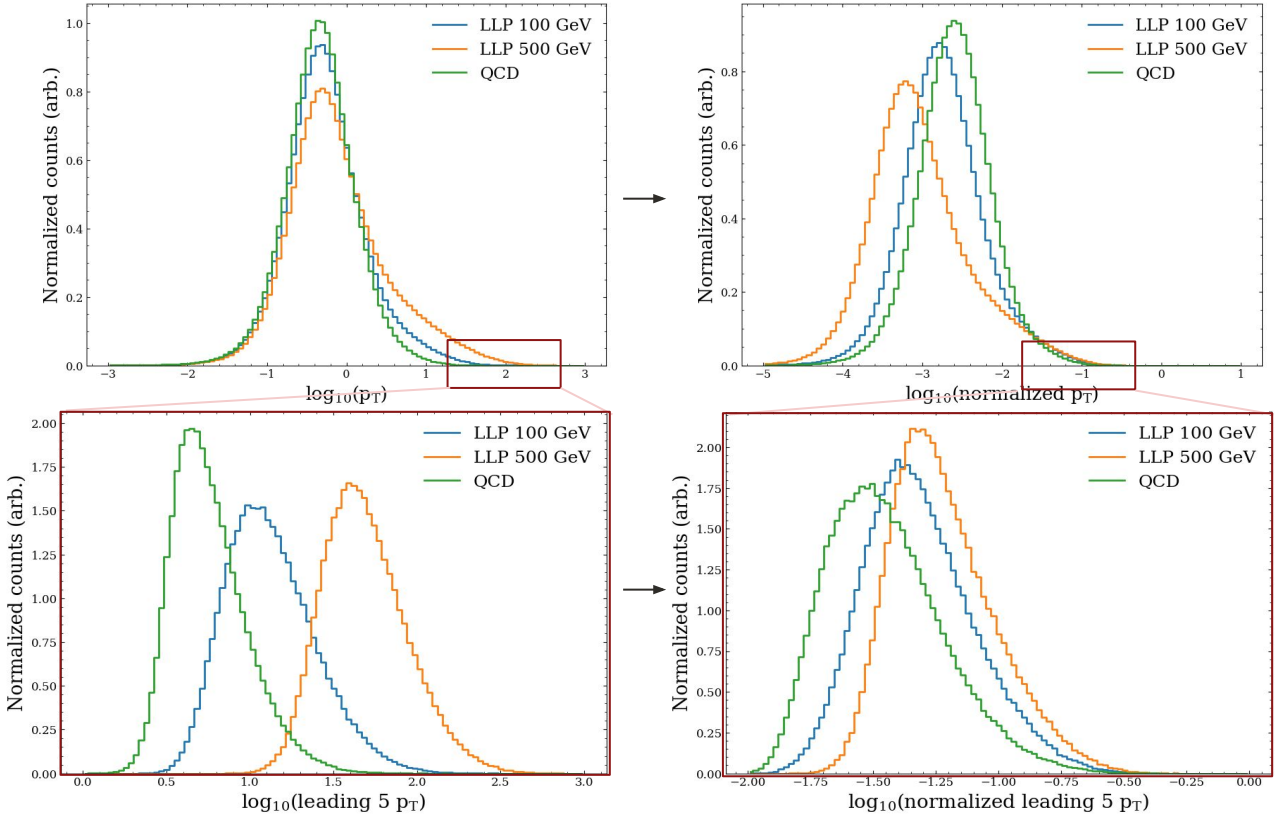
Checking that dataset contains LLP tracks

- Looking at distribution with an LLP with O(100 mm) proper decay length
- Just 1000 events so low statistics



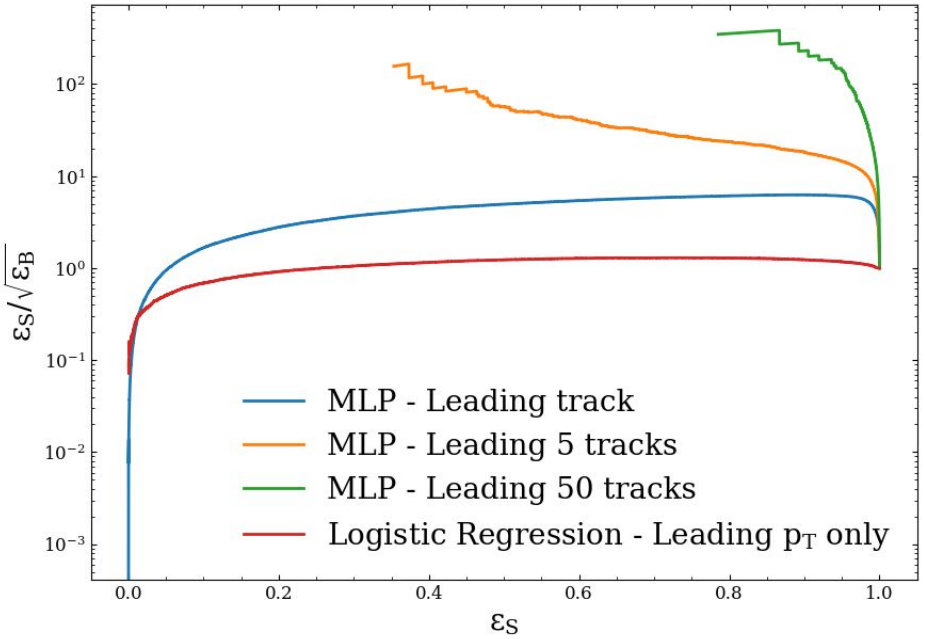
Normalizing p_T distribution

- Madgraph doesn't simulate p_T jet slices for QCD events, leading to small tail
- Solution? Normalize p_T distribution:
  - Emphasize **shape** of distribution rather than scale
  - Other parameters may be impacted (eta, phi)?



MLP Classification after p_T normalization

- 99% test acc on 50 highest p_T tracks / event
- 97% test acc on 5 highest p_T tracks / event
- 92% test acc on 1 highest p_T track / event
- 69% test acc on only highest p_T / event (LR)



Extras

Simple MLP

Architecture:

Input (N, 5) $(p_T, \eta, \phi, d_0, d_z)$

⟶ Shuffle tracks + batch normalization

⟶ Linear(5⟶32) * mask

⟶ GELU

⟶ 3 (Linear(32⟶32) + GELU)

⟶ Linear(32⟶2)

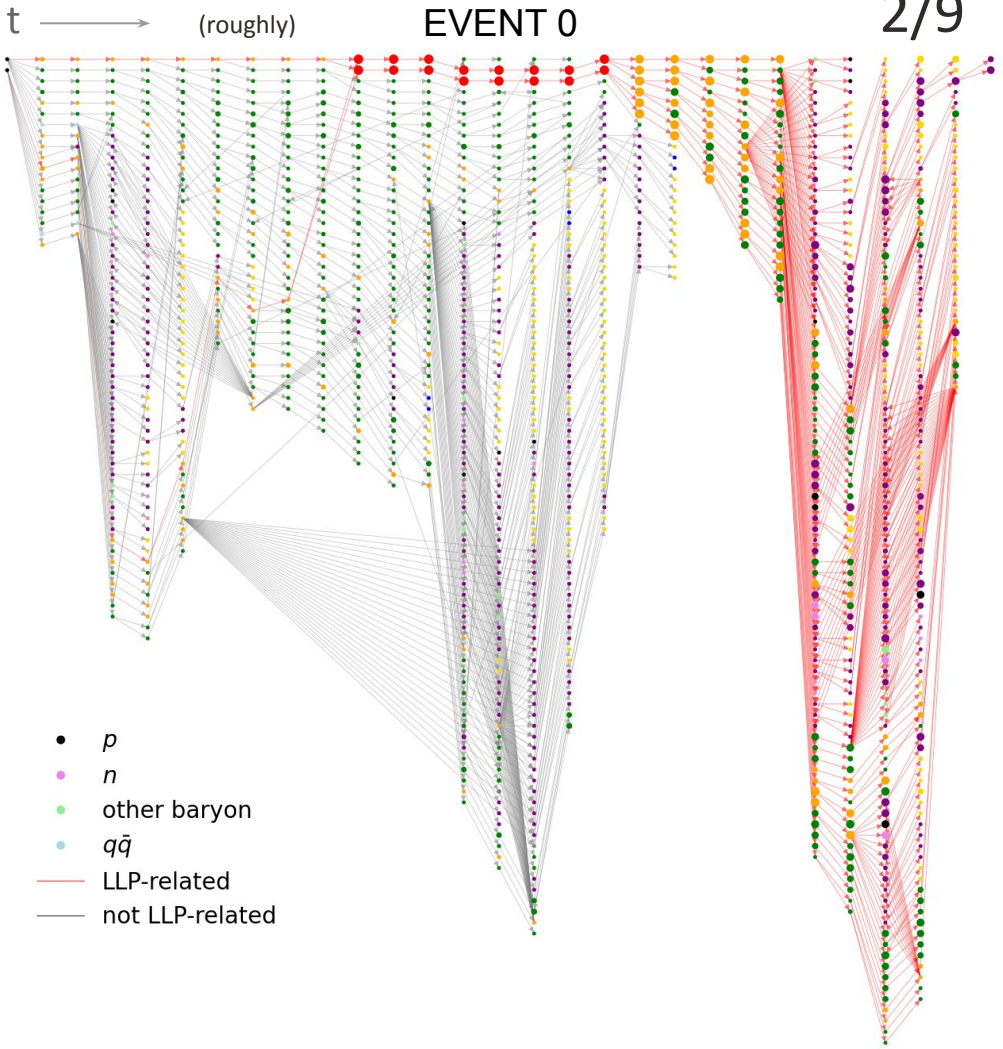
⟶ Softmax

⟶ Argmax

Particle Decay Graph

χ^0_3 mass = 500 GeV

pp $\rightarrow \chi^0_3 \chi^0_3 \rightarrow$ jjj jjj
+ ISR/FSR



PT > 0.5 GeV Cut

- Entire dataset constrained to low DZ

No PT cut:

% LLP-related: 22.2

% ISR/FSR: 77.8

PT > 500 MeV:

% LLP-related: 33.3

% ISR/FSR: 66.7