dN/dx Reconstruction with Machine Learning for Drift Chamber

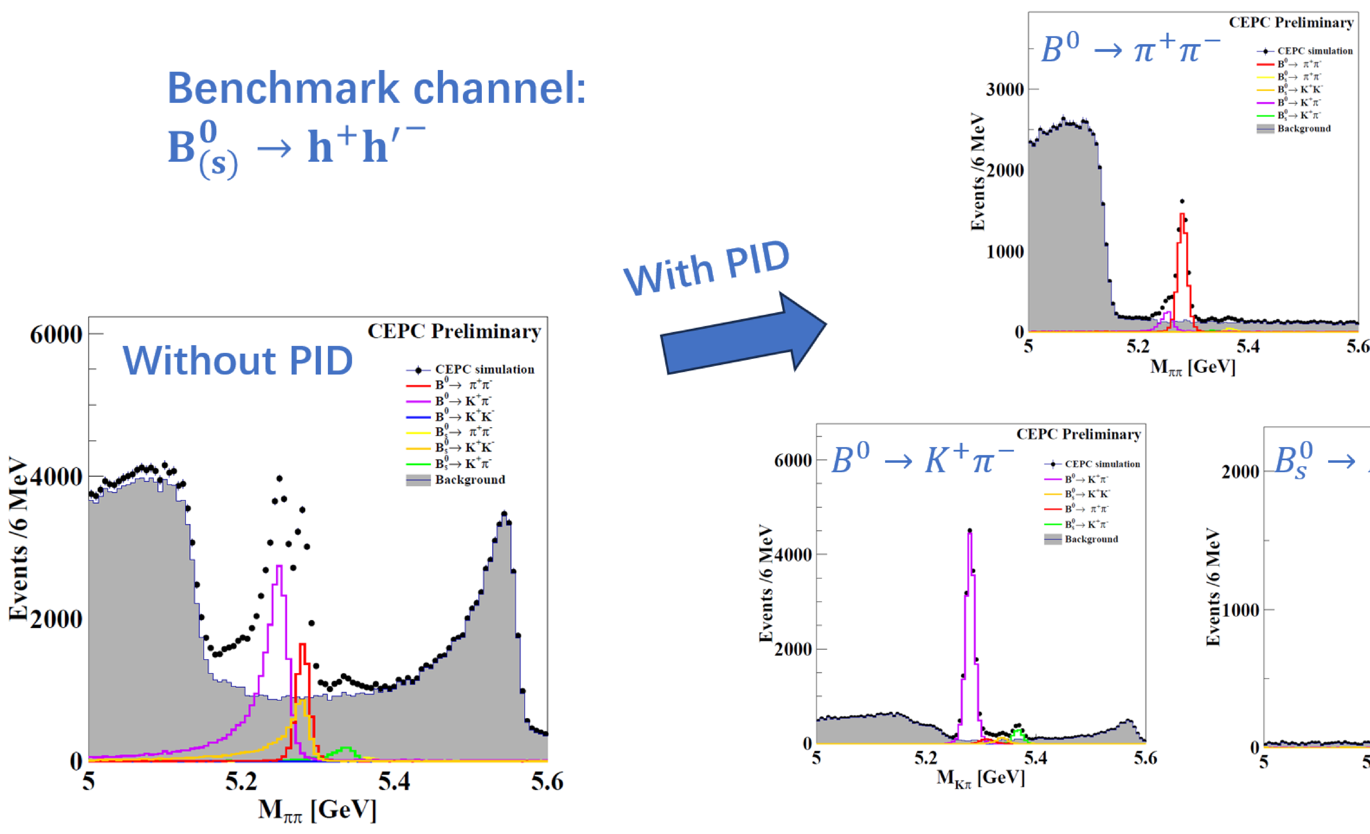
Guang Zhao[1], Zhefei Tian[2], Zhenyu Zhang[2], Linghui Wu[1], Mingyi Dong[1], Gang Li[1], Shengsen Sun[1], Francesco Grancagnolo[3], Nicola De Filippis[3]

[1] Institute of High Energy Physics, China
[2] Wuhan University, China
[3] Istituto Nazionale de Fisica Nucleare, Italy

1. Introduction

1.1 Cluster counting (dN/dx) as a breakthrough in PID techniques

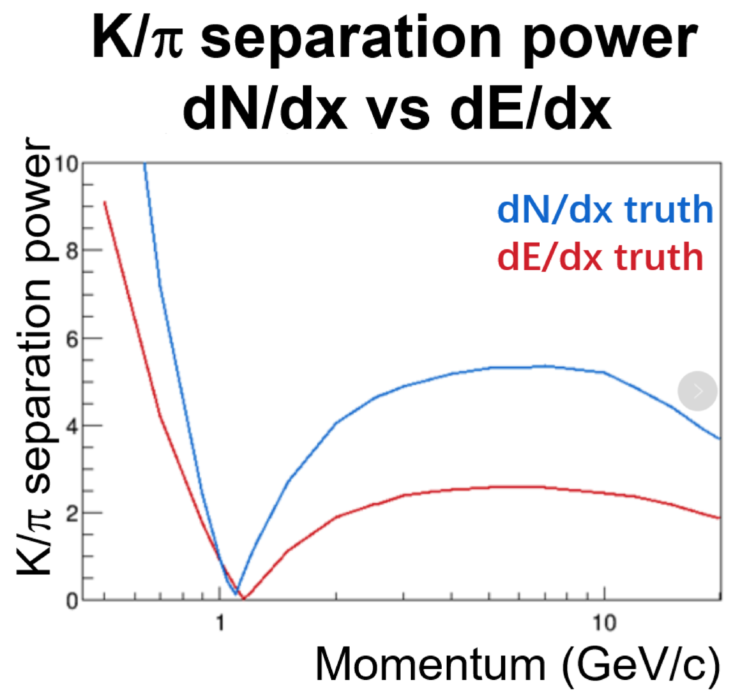
- PID is essential for flavor physics in future large collider experiments
  - Suppressing combinatorics
  - Distinguishing between same topology final-states
  - Adding valuable additional information

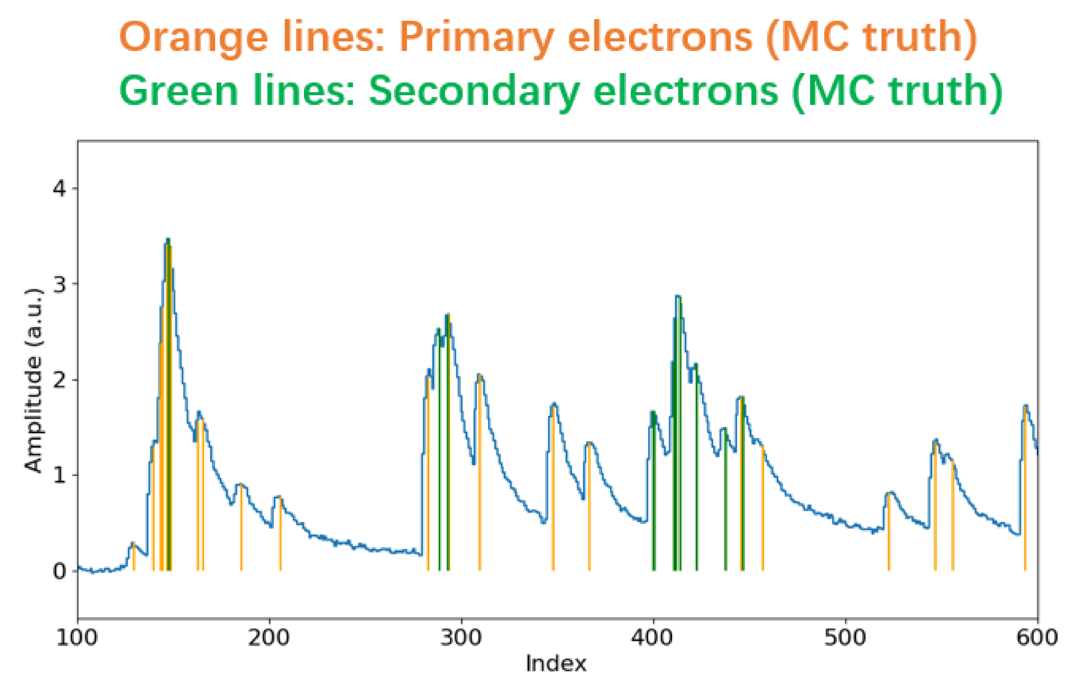


- Drift chamber with cluster counting (dN/dx) could provide powerful PID
  - Poisson distributed → No tails
  - Small fluctuation → Potentially improve the resolution (from dE/dx) by a factor of 2

1.2 dN/dx reconstruction

- Determine the # of **primary electrons** in the waveform





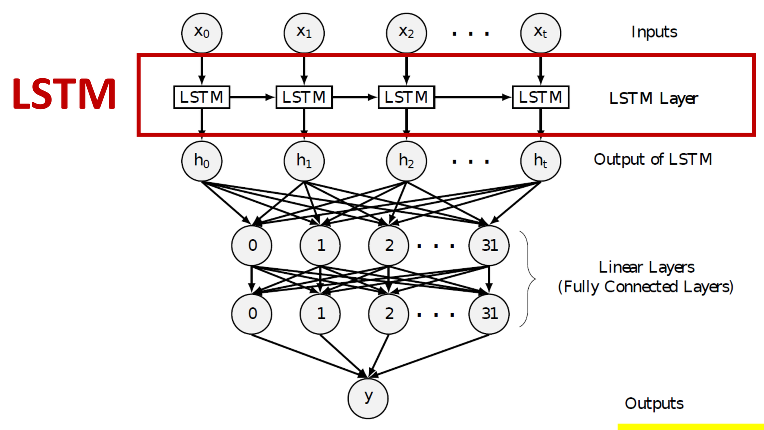
2. Supervised model

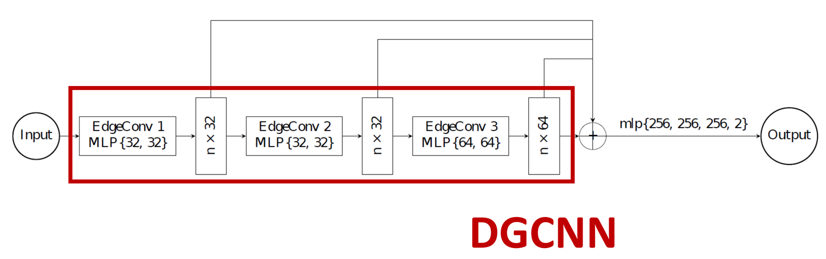
2.1 LSTM-based peak finding

- **LSTM:** A specified recurrent neural network (RNN) that deals with the vanishing gradient problem; Can handle long sequences efficiently
- **Peak-finding:** Waveform as sliding windows; Binary classification of signals and noises

2.2 DGCNN-based clusterization

- **DGCNN:** A specified graph neural network (GNN) that incorporates local information and stacked to learn global properties, which is very suited for clusterization
- **Clusterization:** Peak timing as the node feature. Edges are initially connected by timing similarity; Binary classification of primary and secondary electrons

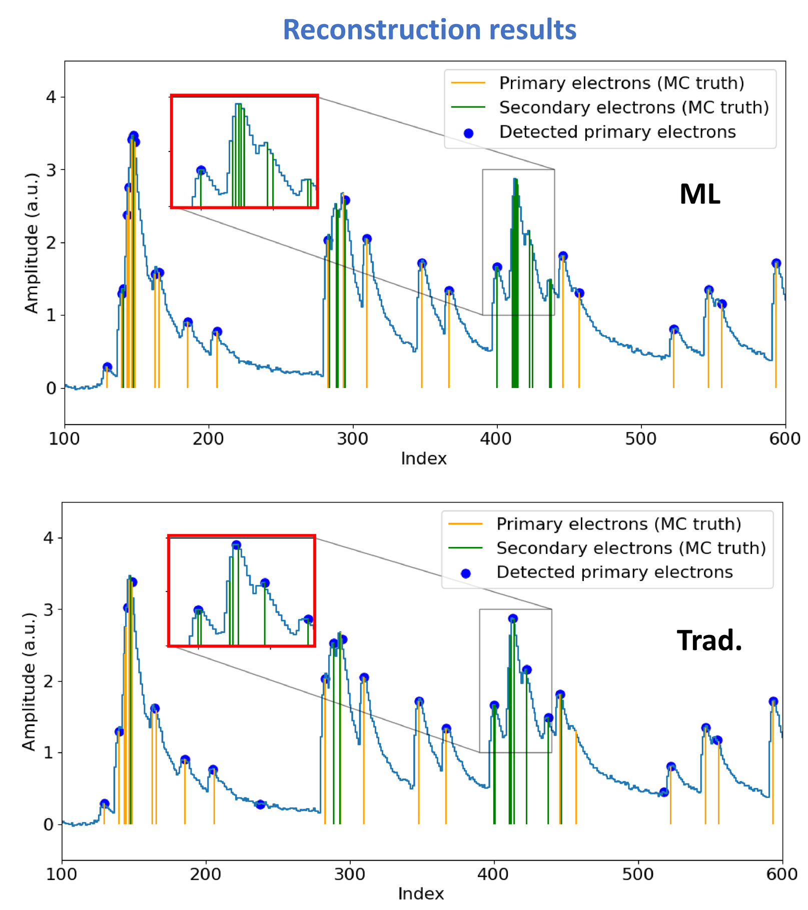




arXiv: 2402.16493

2.3 PID performance

- Supervised model is more efficient for pile-up detection and secondary electrons removal
- Supervised model has a **10% improvement** for K/pi separation
- **dN/dx resolution is less than 3%** for high momentum K/pi



3. Semi-supervised domain adaptation

3.1 Challenges for applying supervised model on real data

- Imperfect simulation
- Incomplete labels in real data

3.2 Solution: Domain adaptation

- Transfer knowledge between simulation and real data domain via optimal transport

$$\min_{f,g}\left[\sum_{i=1}^{m} L_s\left(y_i^s, f\left(g(x_i^s)\right)\right) + \frac{1}{m_l}\sum_{i=1}^{m_l} L_t\left(y_i^{t,l}, f\left(g(x_i^{t,l})\right)\right)\right] + \min_{\gamma\in\Delta}\sum_{i,j}\gamma_{ij}\left(\alpha\left\|g(x_i^s)-g(x_j^t)\right\|^2 + \lambda_t L_t\left(y_i^s, f\left(g(x_j^t)\right)\right)\right)$$

Loss for labeled samples in source domain

Loss for labeled samples in target domain (THIS WORK)

Cost of feature differences between source and target

Cost of 'label' differences between source and target

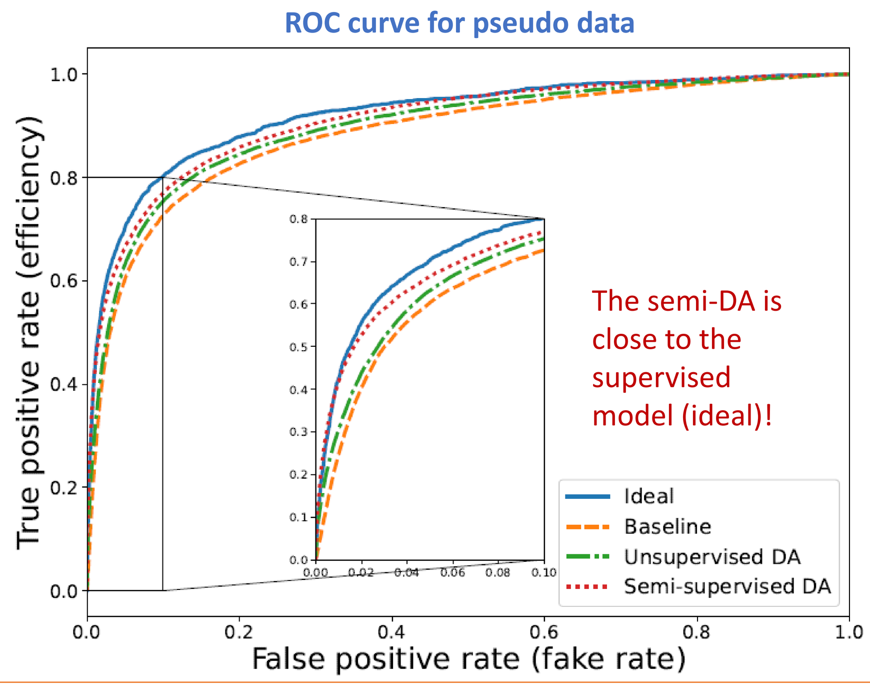
Cost of joint feature-label distribution for OT

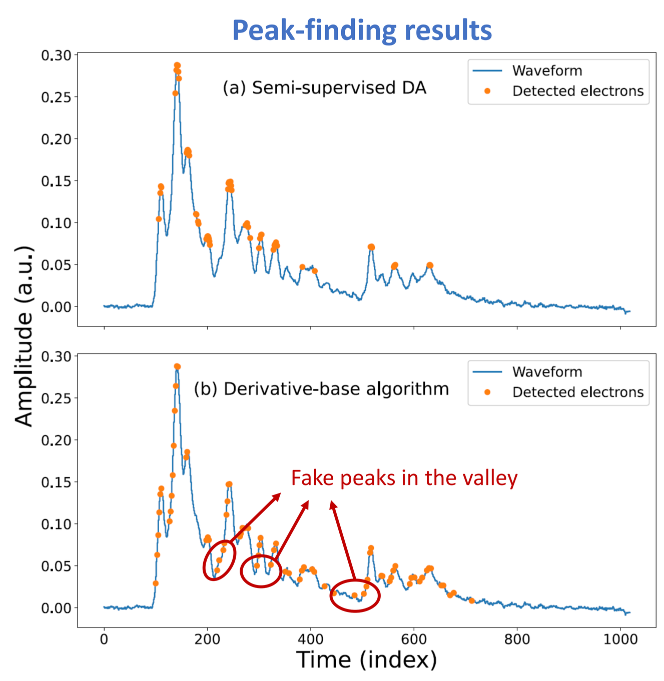
Semi-supervised domain adaptation

Computer Physics Communications 300, 109208 (2024)

3.3 Results

- Validated by pseudo data (Performance close to supervised model)
- Applying to the **CERN testbeam data**, the DA model is more powerful than the traditional algorithm





4. Conclusion

- For simulated samples, the supervised model has **10% improvement** on K/pi separation w.r.t. traditional algorithm.
- For testbeam data, the semi-supervised domain adaptation model **successfully transfers information from simulation** and achieves stable performances.

Contact

Guang Zhao (Institute of High Energy Physics)

Email: zhaog@ihep.ac.cn

Institute of High Energy Physics
Chinese Academy of Sciences

42nd INTERNATIONAL CONFERENCE ON HIGH ENERGY PHYSICS 18-24 July 2024