Simultaneous Track Finding and Track Fitting by the Deep Neural Network at BESIII

Yao ZHANG[1], Ye YUAN[1,2], Haiyong JIANG[2], Zhibin YANG[1],Wenniu ZHANG[2], Xiaorui LYU[2],Yangheng ZHENG[2],Jun XIAO[2]

1. Institute of High Energy Physics, Chinese Academy of Sciences, Beijing
2. University of Chinese Academy of Sciences

25 October 2022

ACAT 2022

中国科学院高能物理研究所
Institute of High Energy Physics, Chinese Academy of Sciences

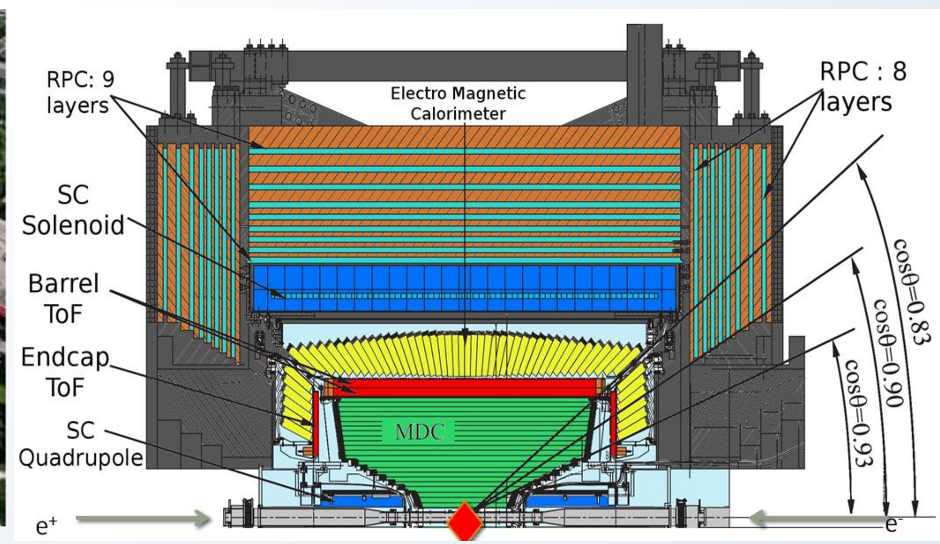
BESIII

中国科学院大学
University of Chinese Academy of Sciences

BEPCII and BESIII

- BEPCII is a double-ring e^+e^- collider running in the E_{cm}=2-5GeV in China
  - Highest luminosity : $10^{33}cm^{-2}s^{-1}$
- BESIII at the BEPCII is for the studies at hadron physics and τ-charm physics with the highest accuracy achieved until now
- World's largest J/ψ dataset : 10 billion

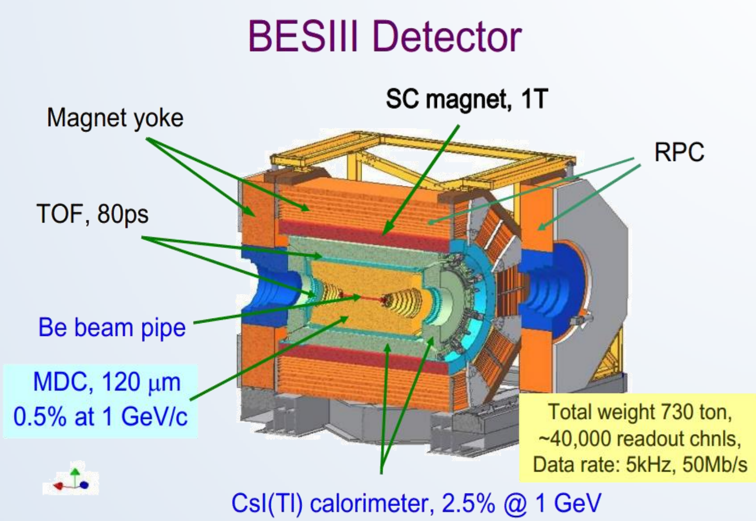
The Beijing Electron Positron Collider (BEPCII)

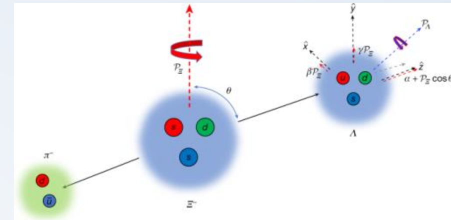


The BESIII detector

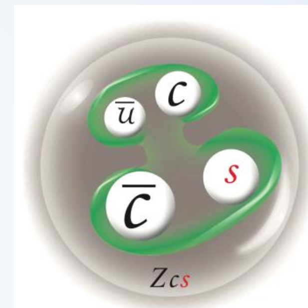
Detector and physics of BESIII

- Charged track reconstruction
  - A multi-layer drift chamber(MDC) for track momentum, position and secondary vertex
- Particle identification
  - A time-of-flight system (TOF)
  - An electromagnetic calorimeter (EMC)
- Muon-pion separation
  - Muon counter (MUC)

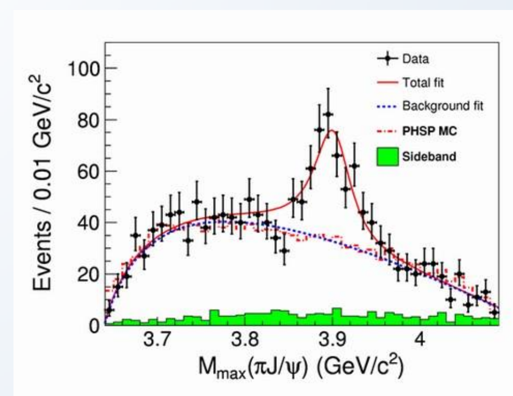




[Probing CP symmetry with Entangled Double-strange baryons]() @ nature



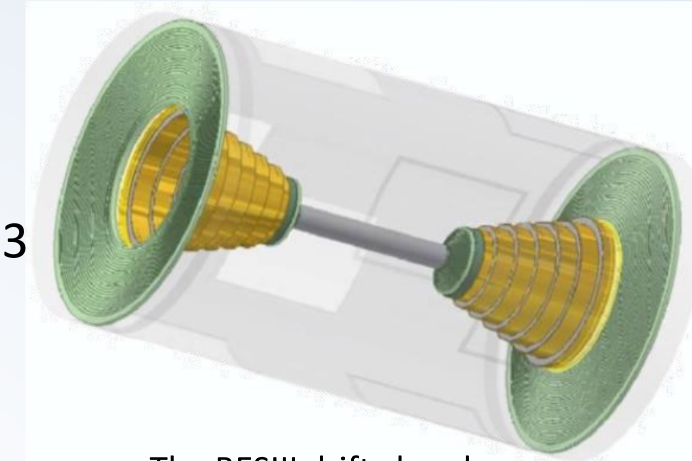
[Observation of the Zcs(3985) strange four-quark meson]()



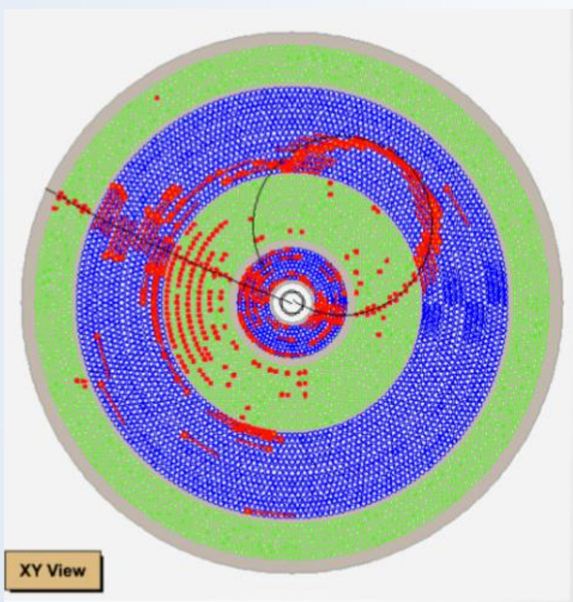
[Observation of a charged charmonium like structure at BESIII Zc(3900) is on the top of the "Highlights of the Year"]()

The BESIII Drift Chamber

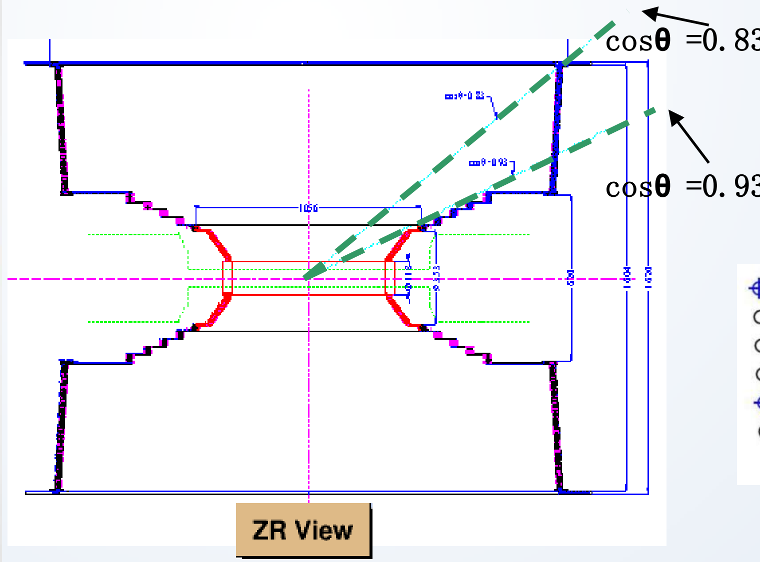
- Gaseous cylindrical drift chamber with |cosθ|<0.93
- 6796 wires arranged in 43 layers
- 3 or 4 layers are grouped to super-layers
- Axial and stereo super-layers alternatively



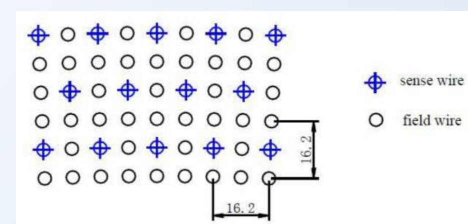
The BESIII drift chamber



Blue: stereo layers Green: axial layers



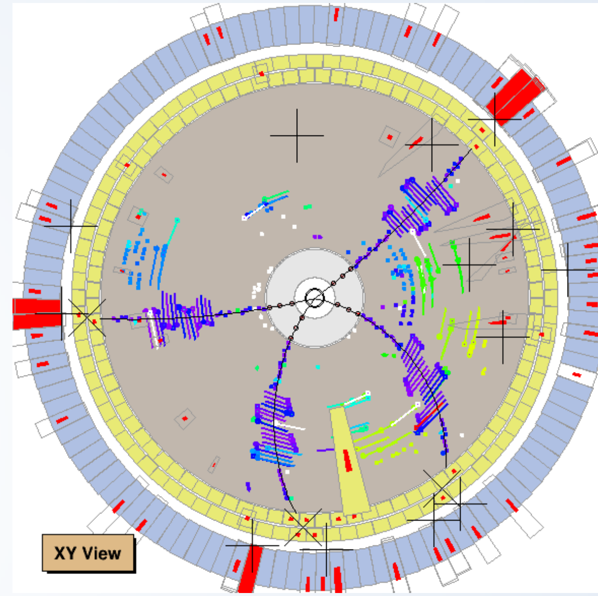
ZR view of drift chamber



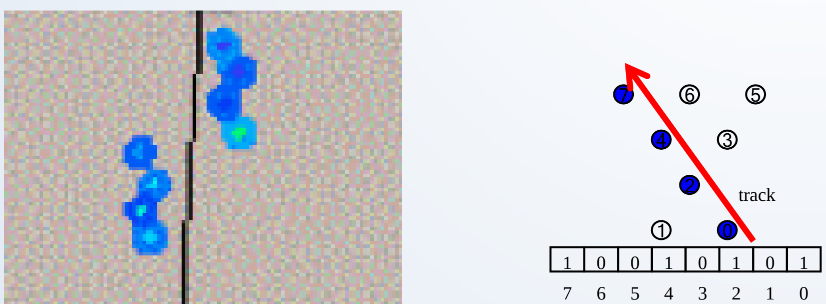
Cell structure

Traditional tracking of BESIII drift chamber

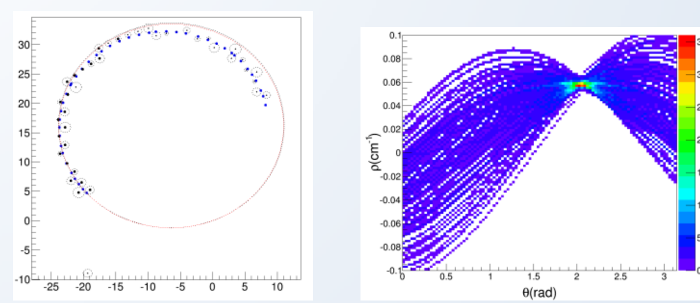
- Local method
  - Template matching for segment in super-layers
    Sensitive to wire inefficient, layer arrangement and momentum
  - Seeding and road following
    Affect by noise or background alone the track path
- Global method
  - Hough transform
    Affect by energy loss and overlapping track



An event display



Track segment finding in super-layers



Hough transform

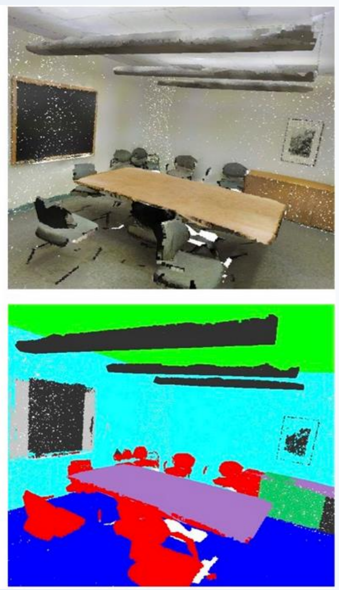
Motivation

- Increase the tracking efficiency and performance for special events
  - Low transverse momentum
  - Large dip angle
  - Secondary vertex
- Higher Background and noise with the upgrade of BEPCII
  - Noise hit resistance (another efforts please refer to Rostam's poster)
- The optimization of the tradition tracking algorithm could be risky and challenge

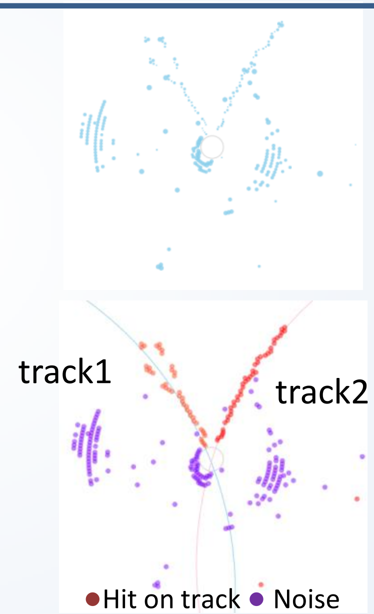
- Aim of our work
  - Explore the new tracking method with novel technics
  - Hit clustering with the auxiliary of the track parameter regression
    - Parallel "track finding and fitting" with neural network
  - Experiment independent tracking with 2-D measurement (drift chamber)

Neural network approach for track reconstruction

- Idea:
  1. Semantic segmentation of the point set –> track finding/hit clustering
  2. Extract local features –> track fitting/track parameter estimation



PointNet:Deep Learning on Point Sets for 3D Classification and Segmentation



Track reconstruction

Address track parameter estimation and hit clustering simultaneously in an end-to-end fashion

PointNet++:

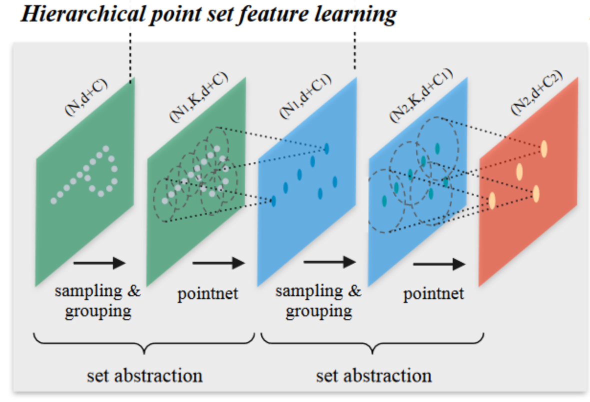
Point set segmentation for semantic scene labeling and feature extraction

Given an unordered point set $\{x_1, x_2, ..., x_n\}$ with $x_i \in \mathbb{R}^d$, one can define a set function $f : \mathcal{X} \rightarrow \mathbb{R}$ that maps a set of points to a vector:

$$f(x_1, x_2, ..., x_n) = \gamma \left(\underset{i=1,...,n}{\mathrm{MAX}} \{h(x_i)\} \right) \tag{1}$$

where γ and h are usually multi-layer perceptron (MLP) networks.

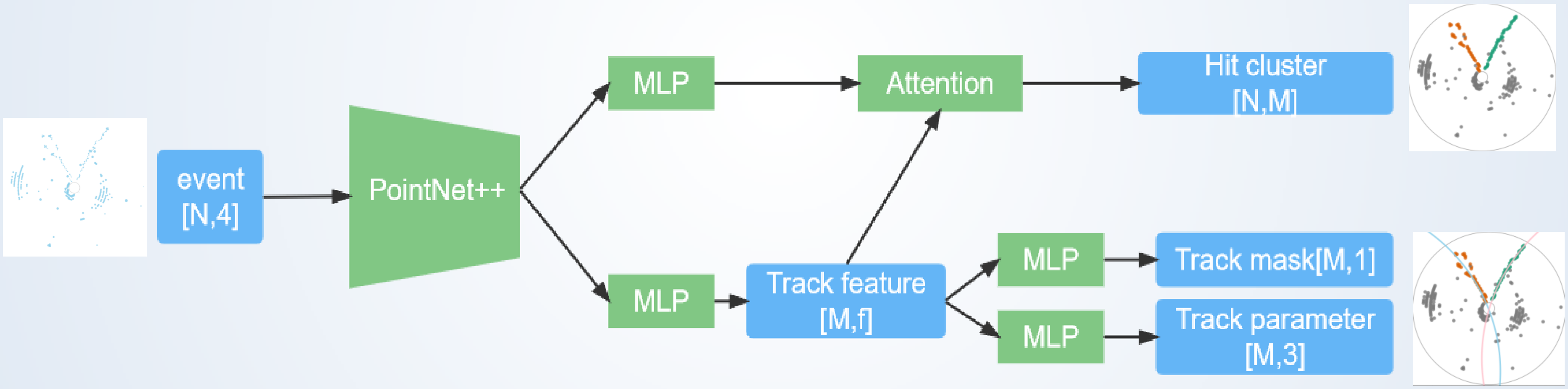
- The response of h can be interpreted as the spatial encoding of a point
- The set function f in Eq. 1 is invariant to input point permutations and can arbitrarily approximate any continuous set function



Track parameter estimated by learning the local feature of hit cluster

PointNet++ based tracking network

- End-to-end hit clustering with track parameter estimation
  1. Hit clustering for each hit
  2. track parameters for each predicted group of hits



PointNet++ based tracking network

Loss function

- Clustering loss
  - Cross entropy loss for hits clustering

$$\ell(x,y) = L = \{l_1,\dots,l_N\}^\top, \quad l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^{C} \exp(x_{n,c})} \cdot 1\{y_n \neq \text{ignore\_index}\}$$

where x is the input, y is the target, w is the weight, C is the number of classes, and N spans the minibatch

  - Binary loss of the classification for default clusters

$$\ell(x,y) = L = \{l_1,\dots,l_N\}^\top, \quad l_n = -w_n \left[y_n \cdot \log x_n + (1-y_n) \cdot \log(1-x_n)\right],$$

where N is the batch size. If `reduction` is not `'none'` (default `'mean'`), then

- Track parameters loss
  - Frobenius Norm Loss

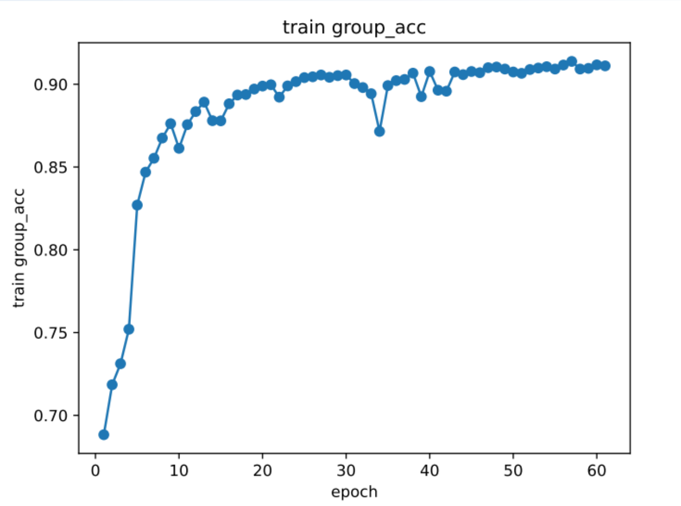
$$L_{Frob} = \| A_{pred} - A_{gnd} \|_F = \sqrt{\sum_i^m \sum_j^n |a_{ij}^{pred} - a_{ij}^{gnd}|^2}$$

Training

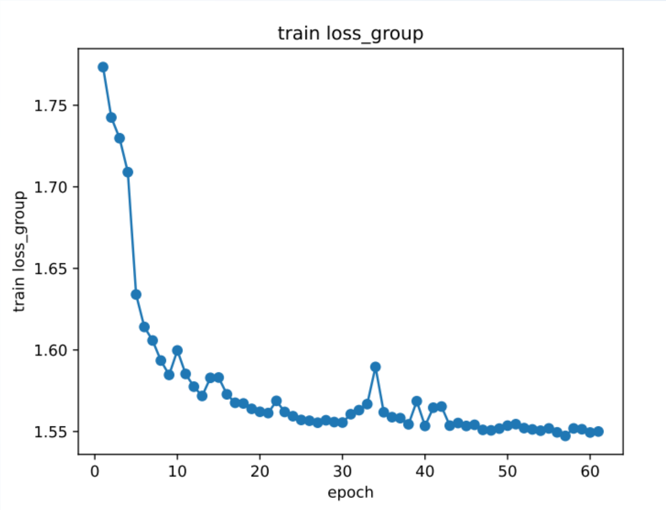
- Dataset
  - $J/\psi \to \rho\pi$ with one or two tracks as final state
  - Training/validation/test: 120k/15k/15k events
- Pre-processing
  1. Initial layer < 8
  2. Number of gap layers < 5
  3. Total number of hits >= 10
  4. Total number of layers >= 8
- GPU environment
  - Tesla V100

Training loss and accuracy

- Converged after ~30 epoch

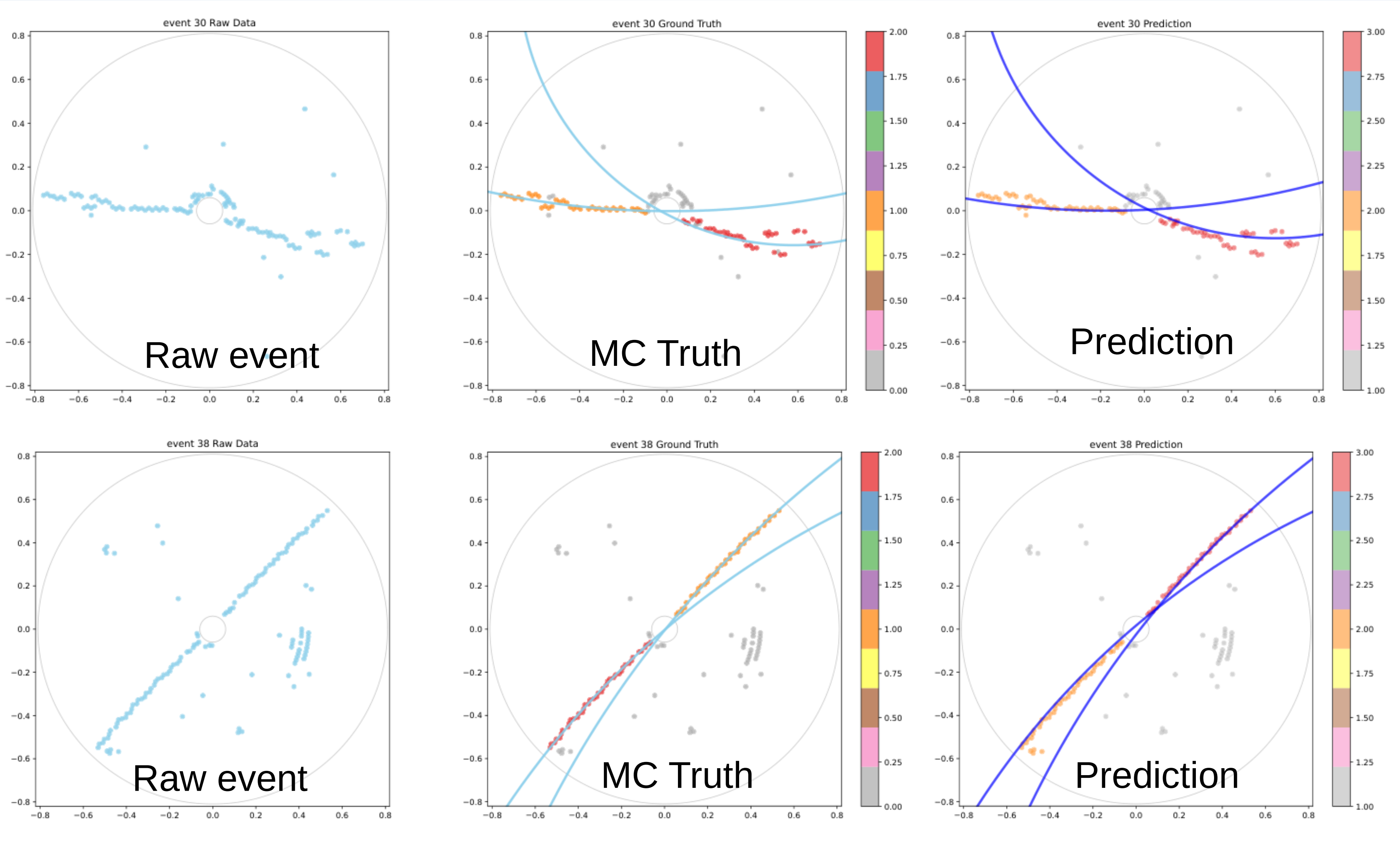


Training loss v.s. epoch



Training accuracy v.s. epoch

Prediction of the events



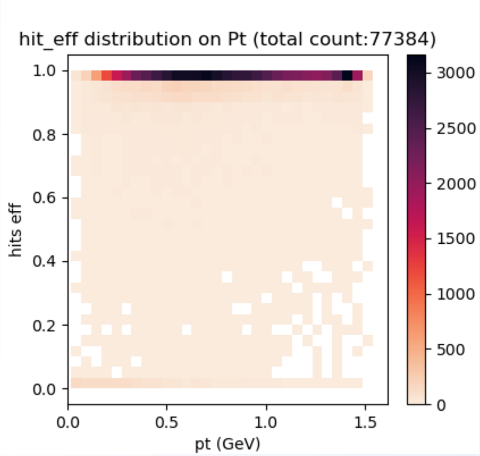
Hit clustering and track parameter performance

$$hit\ efficiency = \frac{N_{hit\ predicted\ correctly}}{N_{truth\ hit\ on\ track}}$$

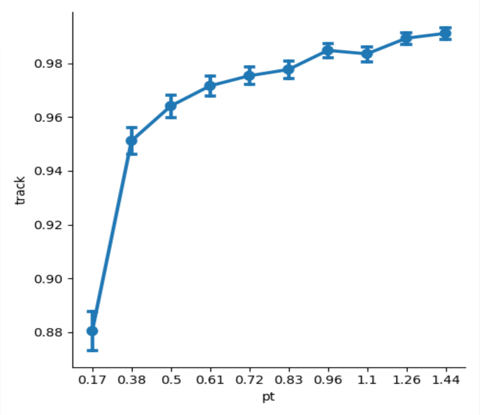
$$hit\ purity = \frac{N_{hit\ predicted\ correctly}}{N_{predicted\ hit\ on\ track}}$$

$$track\ efficiency = \frac{N_{predicted\ track}}{N_{truth\ track}}$$

- J/ψ–>ρπ
  - Hit efficiency: 93.2%
  - Hit purity: 91.9%
  - Track efficiency for high p_T tracks > 96%



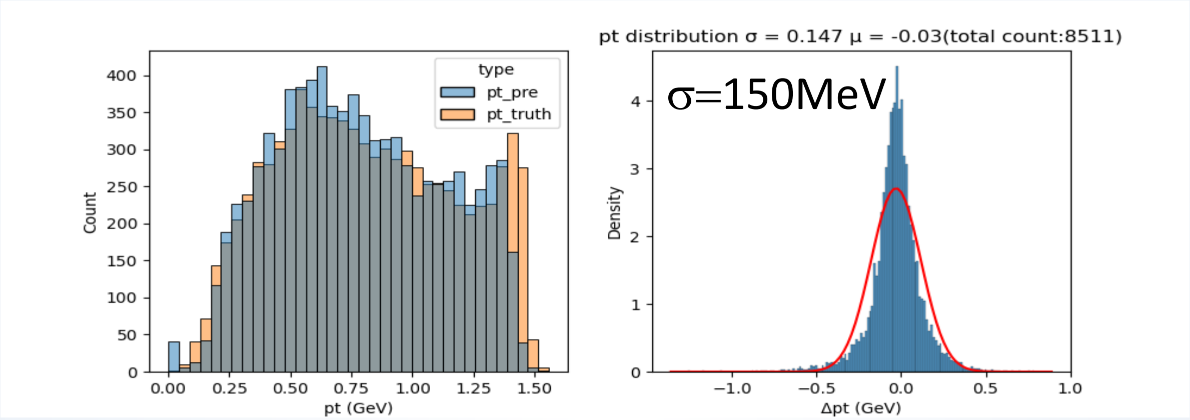
Hit efficiency vs p_T



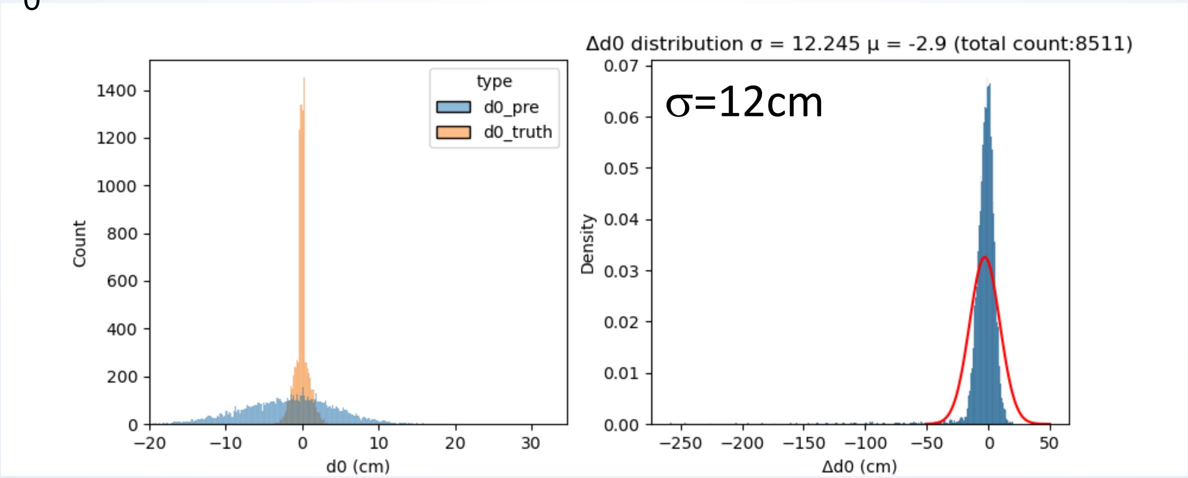
Track efficiency vs p_T

Performance of track parameter

- p_T distribution



- d_0 distribution



The performance of track parameter estimation is promising

Further approach

- Optimize current model considering physical mechanic
  - Use dice-loss to evaluate the hit efficiency
  - Distance between track and hits
  - Axial and stereo layers separately
  - Put kinematics and track model
  - Add an refinement of the track

- More efforts
  - DeepFit: 3D Surface Fitting via Neural Network Weighted Least Squares
  - Attention mechanism with transformers

DeepFit

Conclusions

- We propose a novel neural network approach for drift chamber tracking
  - An end-to-end multi-track tracking
  - Hit clustering and track estimation simultaneously
- Preliminary performance of this work is promising

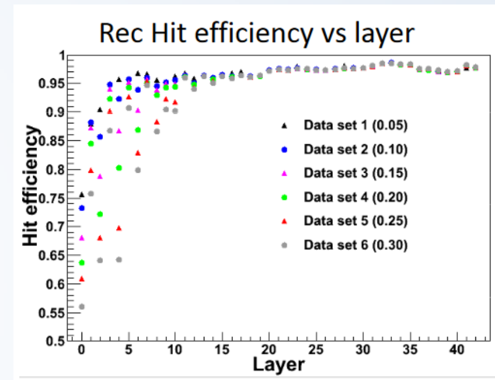
- More approach is under investigate such as
  - The combination of least-square fit
  - Attention mechanism with transformers

Thank you for your attention!

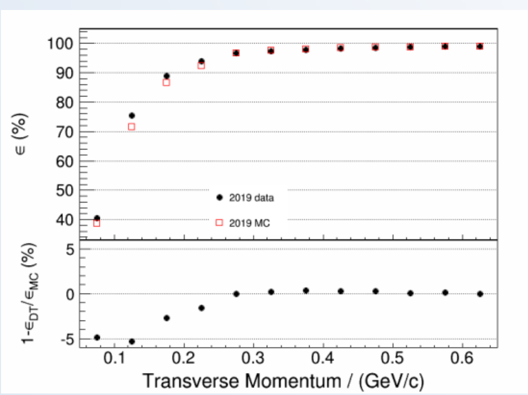
Backup

Challenge for BESIII tracking

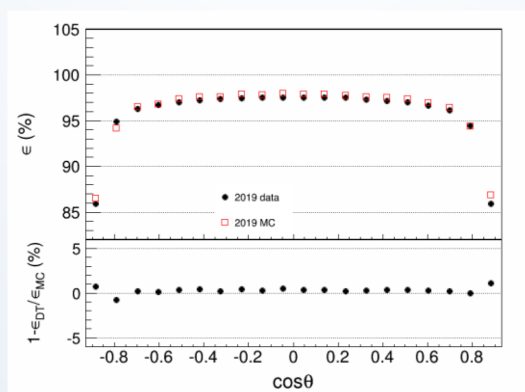
- Tracking quality need to improve for following situation
  - Noise resistance
  - Low momentum
  - Large dip angle
  - Secondary vertex



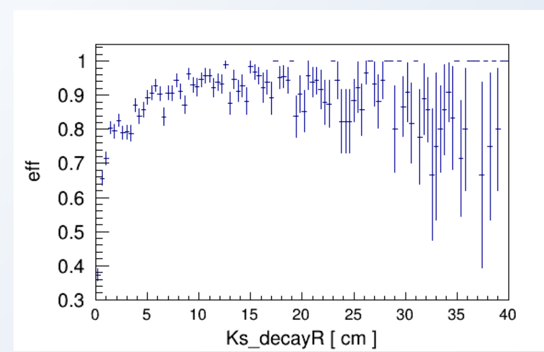
Hit efficiency vs noise level



Pion tracking efficiency vs p_T



Pion tracking efficiency vs $\cos\theta$



Ks tracking efficiency decay length