Machine Learning for Particle Track Reconstruction

13th Terascale Detector workshop

06-08 April 2021

W. Esmail, T. Stockmanns and J. Ritman

On behalf of PANDA Collaboration

Institut für Kernphysik (IKP), Forschungszentrum Jülich

Outline:

1. Introduction
2. PANDA Detector
3. Forward Tracking System FTS
4. DL Local Approach
5. DL Global Approach
6. Conclusion

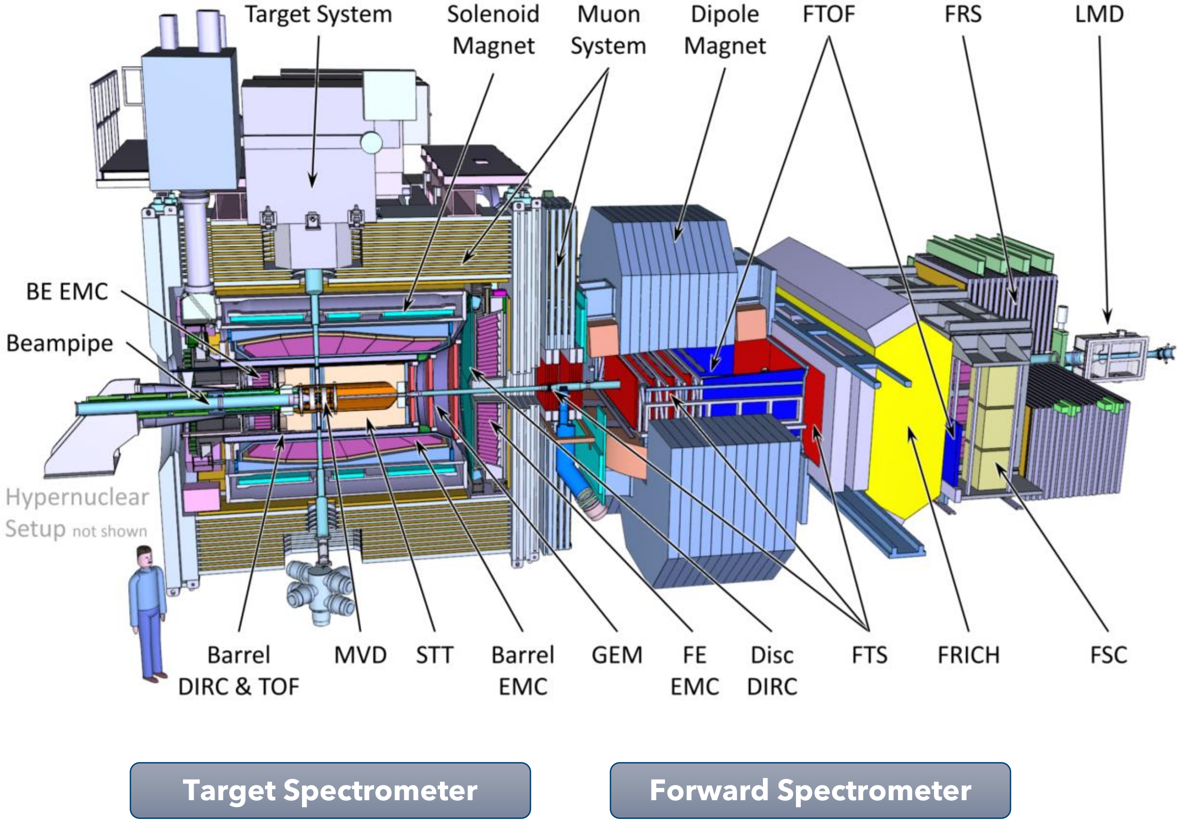
Introduction:

- Track reconstruction is a pattern recognition task
- Two main steps: $\boxed{\text{Track Finding}}$ and Track Fitting
- Track Finding: assign position measurements (**hits**) to track candidates (particle paths)
- Track Fitting: determine track parameters and covariance matrix for each track
- ML and DL address different aspects of the problem
- There are still open questions about how to incorporate such techniques

- Good tracking algorithm should be
  a) high efficiency
  b) high purity
  c) low fake/ghost rate
  d) fast algorithm

antiProton ANnihilation at DArmstadt P̄ANDA

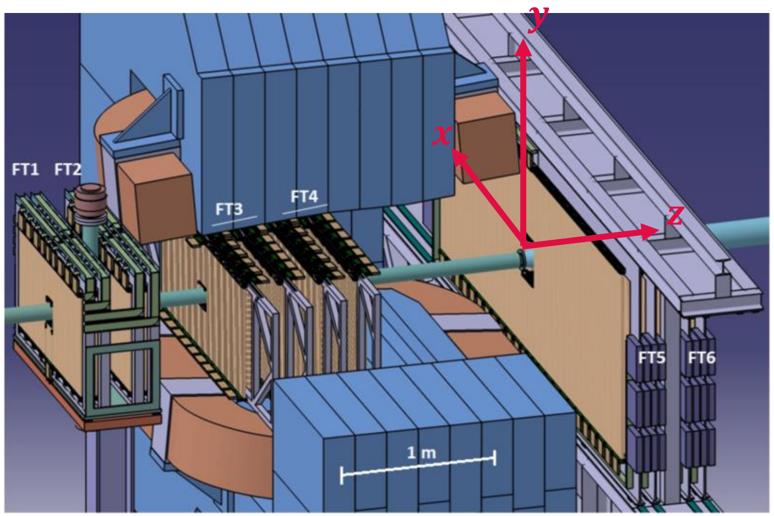
- A future fixed target experiment at the **F**acility for **A**ntiproton and **I**on **R**esearch (FAIR)



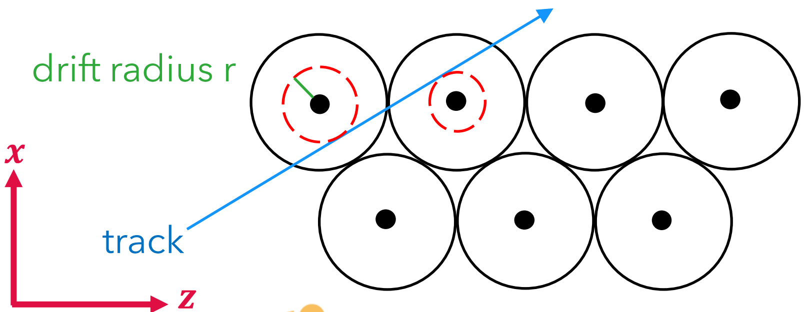
- Beam Momentum
  1.5 to 15 GeV/c
- Tracking & PID sub-systems
- Addresses questions of non pQCD:
  1) Hadron Spectroscopy
  2) Hadron Structure
  3) Hypernuclei

Forward Tracking System FTS

- FTS is dedicated to measuring tracks of particles boosted in the forward direction
- Planes of straw tubes arranged in chambers
- 6 chambers:
  - 4 double layers
  - Outer double layer are vertical
  - Inner layers are tilted by ±5°



- FT3 & FT4 are implemented inside a dipole magnet 2 T.m.
- Tracks are defined by distance of closest approach to the wire
- Hit position is the wire position



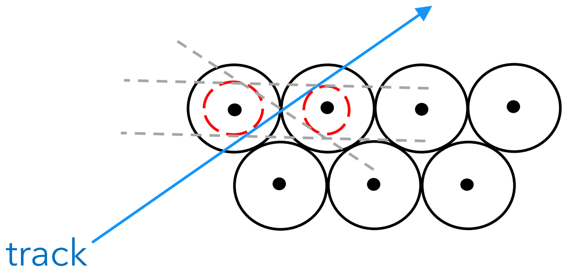
DL Local Approach 1

- A local approach for track finding at the FTS
- Identify the track in *x-z projection*
- Two main steps:

  1) Build tracklets before, inside and after the magnetic field (feed forward NN
  2) Connect tracklets (Recurrent NN)

1) Build Tracklets:

- Make all possible combinations of hit-pairs in adjacent (vertical) layers
- Train a NN to predict if the pair belong to the same track (classification)
- NN accuracy 96-98%
- Input observables are hit coordinates and drift radii
- Output is a probability
- **Output is used to connect the hits**

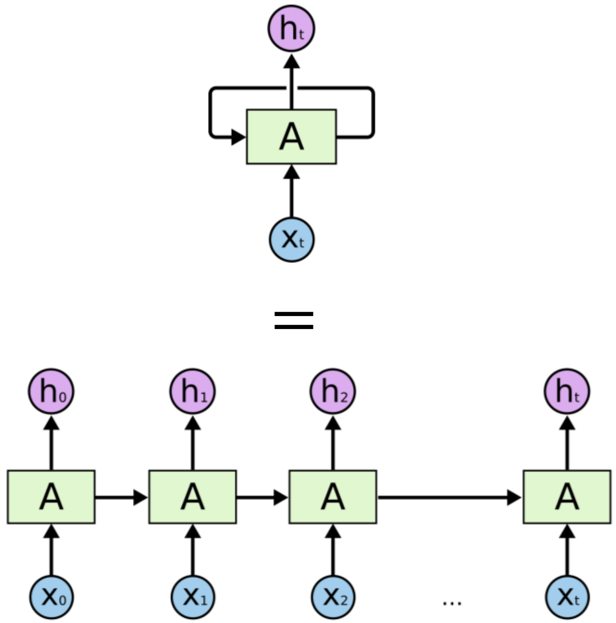


Machine Learning for Track Finding at PANDA 2019 arXiv:1910.07191

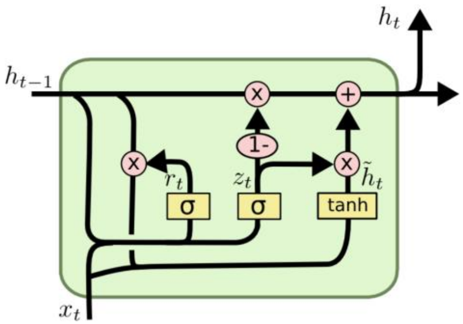
DL Local Approach 2

2) Connect Tracklets:

- Make all possible combinations of found tracklets
- Train a RNN to predict if the combination is a true track (classification)
- A **Recurrent Neural Network RNN** processes **sequences** by iterating through the sequence elements and maintaining a state containing information relative to what it has seen so far
- An implementation of **Long Short-Term Memory LSTM** is used



$$\mathbf{y}_{(t)} = \phi\left(\mathbf{W}_x{}^{\mathsf{T}}\mathbf{x}_{(t)} + \mathbf{W}_y{}^{\mathsf{T}}\mathbf{y}_{(t-1)} + \mathbf{b}\right)$$



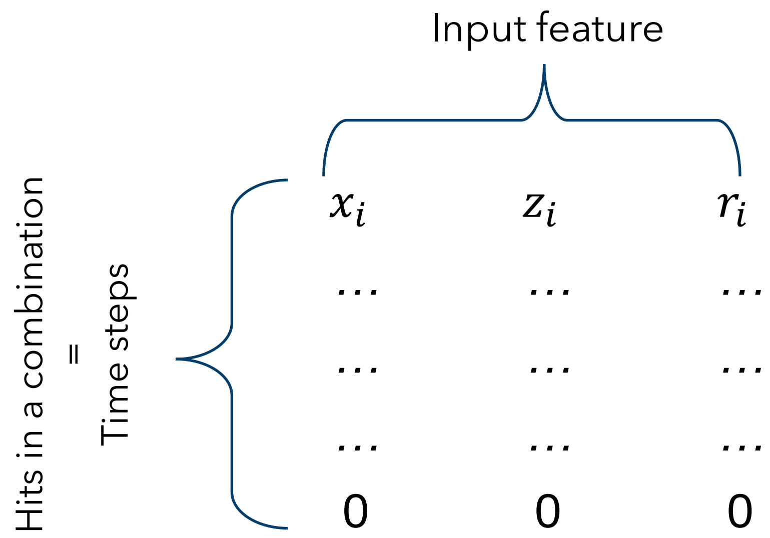
$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$
$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$
$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

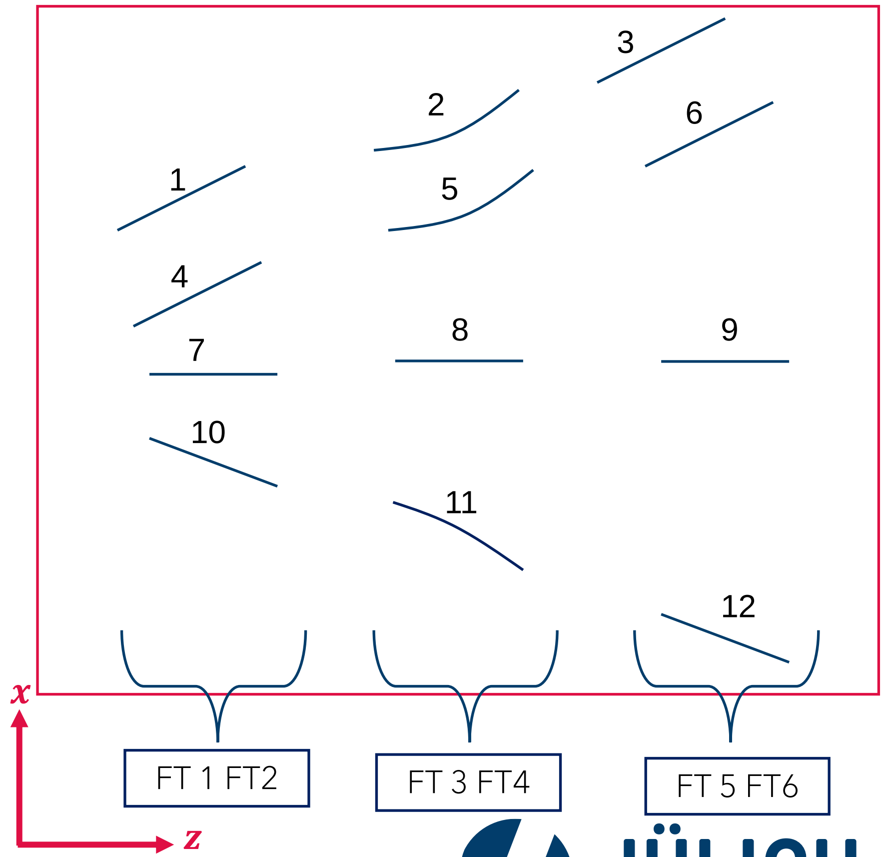
colah.github.io/Understanding LSTM Networks

DL Local Approach 3

2) Connect Tracklets:

- Make all possible combinations of found tracklets [1,2,3], [1,5,3], …
- Train a RNN to predict if the combination is a true track (classification)
- Accuracy 98%
- A sequence input can vary in length



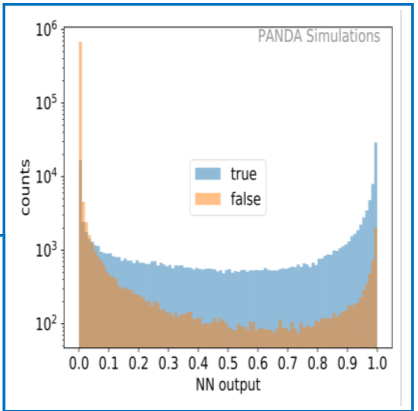


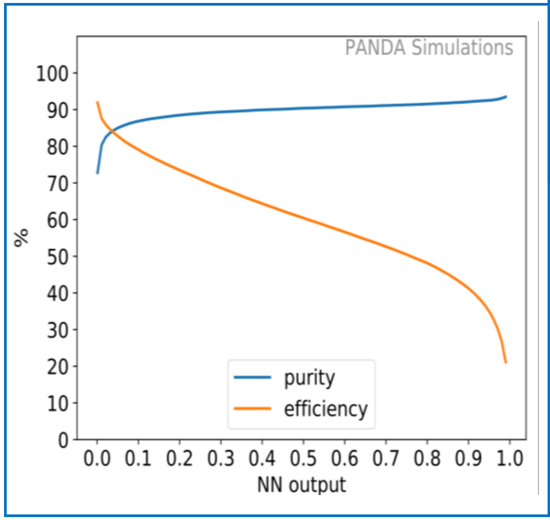
colah.github.io/Understanding LSTM Networks

DL Local Approach 4

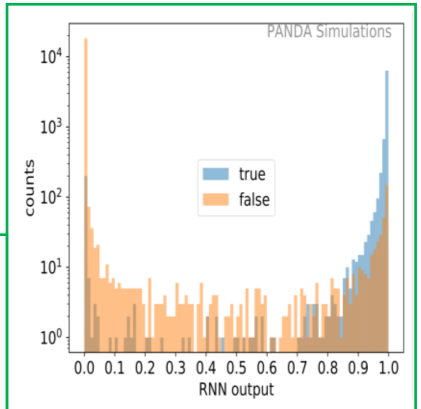
- NN and RNN performance parameters
  - **Purity = true that pass the cut / all that pass the cut**
  - **efficiency = true that pass the cut / all true**

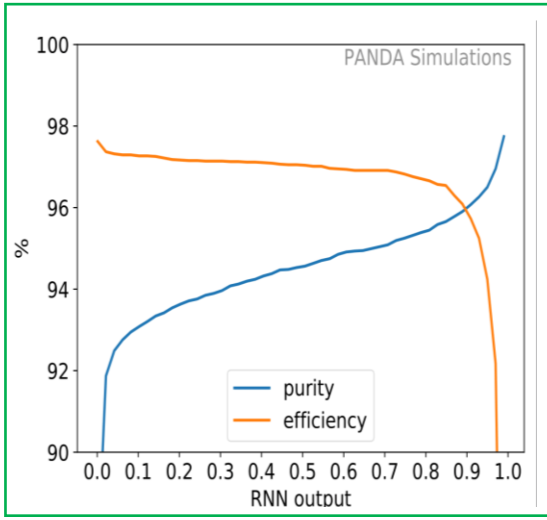
NN





RNN

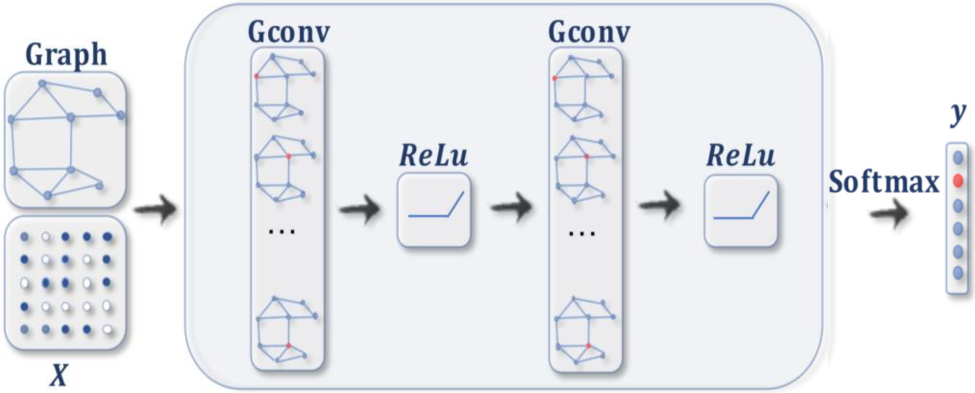




DL Global Approach 1

Graph Neural Networks GNN

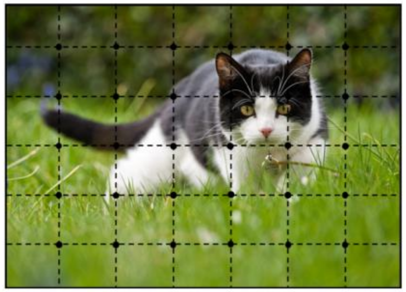
- Motivated by CNN and graph embeddings
- **RecGNNs, ConvGNNs, GAEs, …**



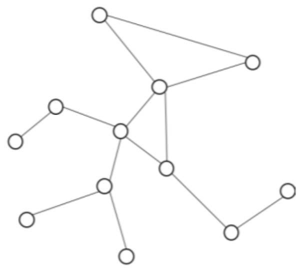
The target of GNN is to learn a state embedding (neighborhood relations)

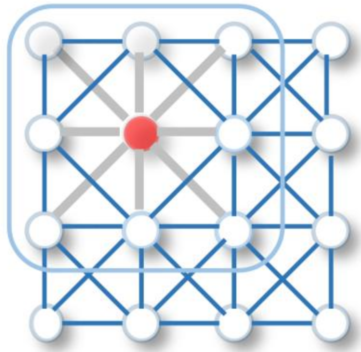
$$H^{t+1} = F(X, H^{t})$$

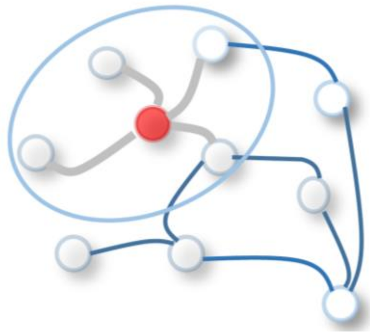
Euclidean



Non-Euclidean







Tasks: Node-level, **Edge-level**, Graph-level.

1. Graph Neural Networks: A Review of Methods and Applications Jie Zhou 2019
2. A Comprehensive Survey on Graph Neural Networks Zonghan Wu 2019

DL Global Approach 2

Graph Neural Networks

- GNN is used as a binary classifier (**hit-pairs classification or edge classification**)
- Input is a graph (FTS hits of one event).
- Two main components: **edge network** and **node network**
- Edge network uses the node features to compute edge weights
- Node network aggregates node features with the edge weights and updates node features

- With each graph iteration, the model propagates information through the graph, strengthens important connections, and weakens useless ones

    - node features = [x, z, r]
    - graph iterations = 2

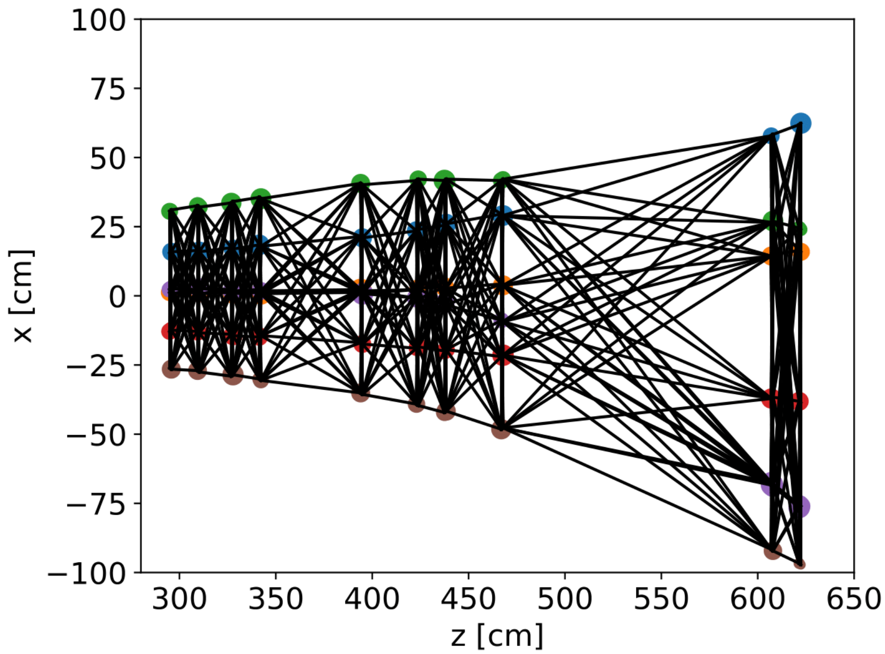
- GNN in a form of **Message Passing Neural Network MPNN** adapted from **HEP.TrkX**

1. Novel deep learning methods for track reconstruction Steve Farrell, CTD/WIT 2018

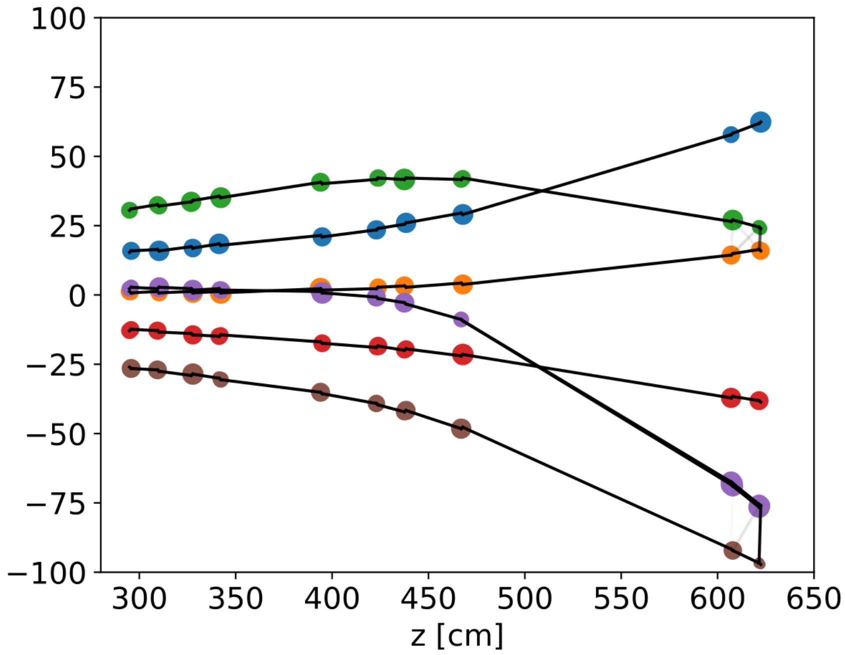
DL Global Approach 3

Graph Neural Networks

An Input Graph



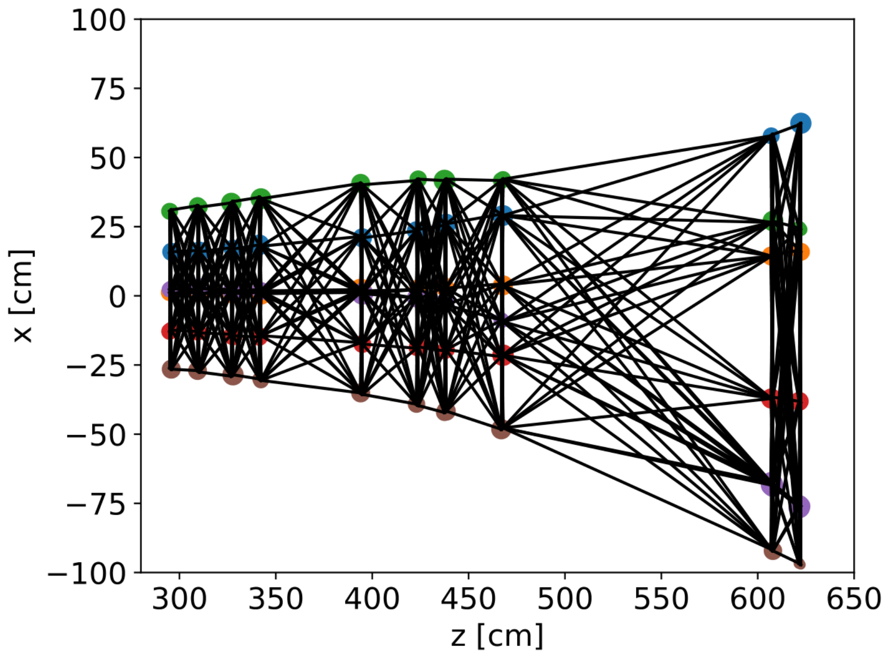
An Output Graph



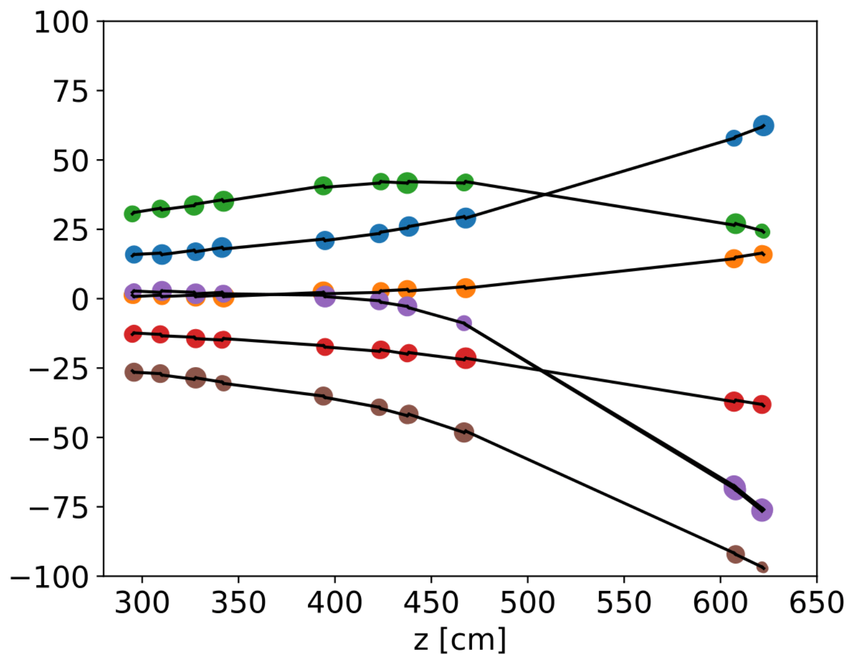
DL Global Approach 4

Graph Neural Networks

An Input Graph

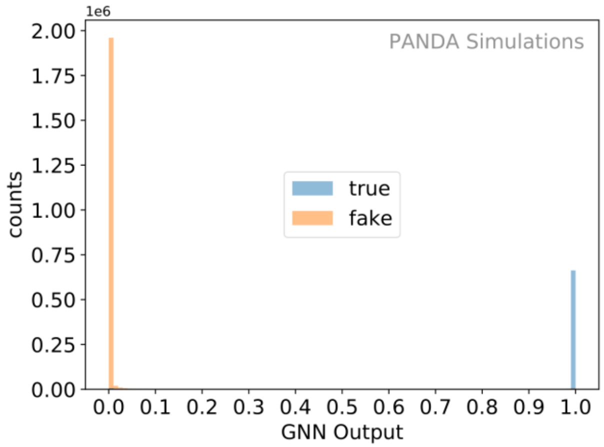


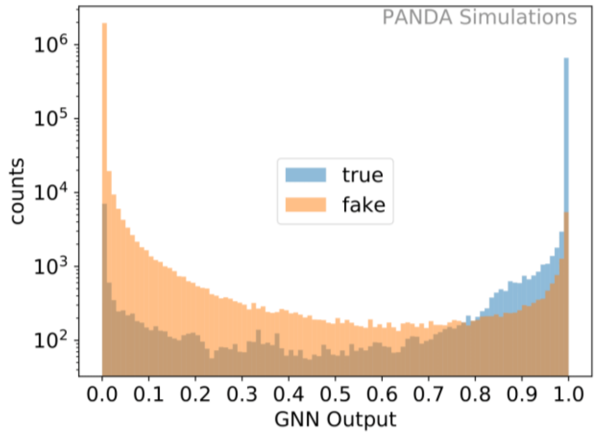
GNN output > 0.9

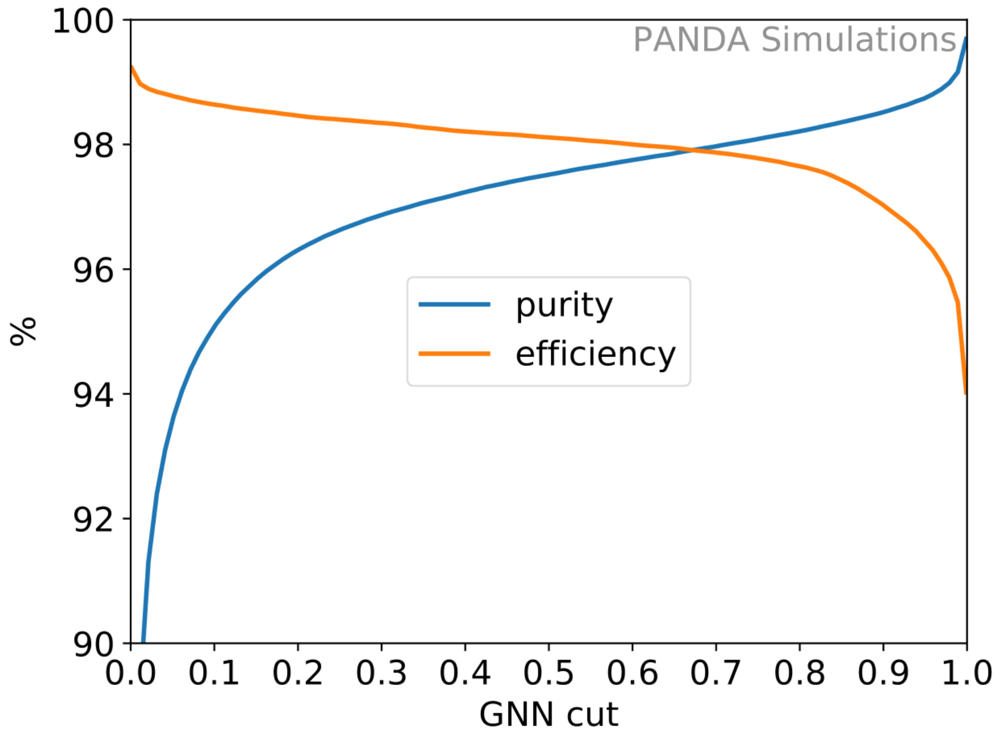


DL Global Approach 5

- GNN performance parameters
  - **Purity = true that pass the cut / all that pass the cut**
  - **efficiency = true that pass the cut / all true**







DL Global Approach 6

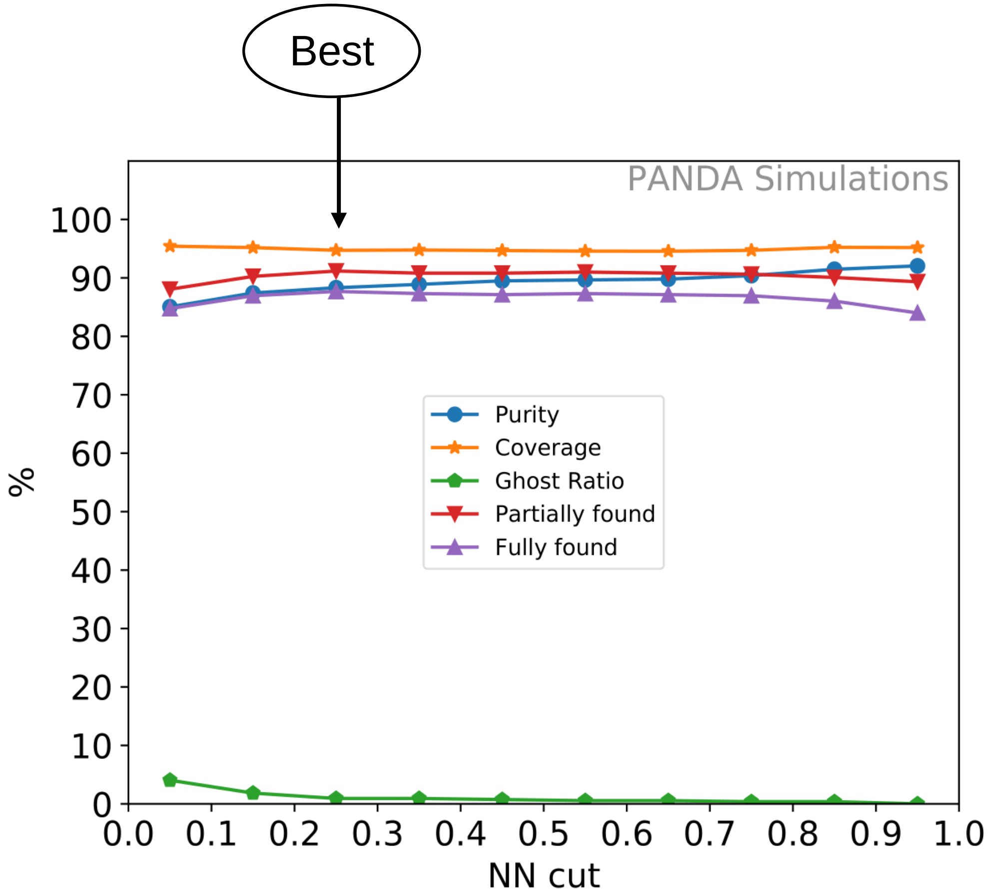
Quality Criteria

- Finding tracks is finding graph connected components (subgraphs)
- A traversal algorithm, starting at vertex v_i then visit all vertices.
- **Depth First Search DFS**

- Evaluation on DPM background generator

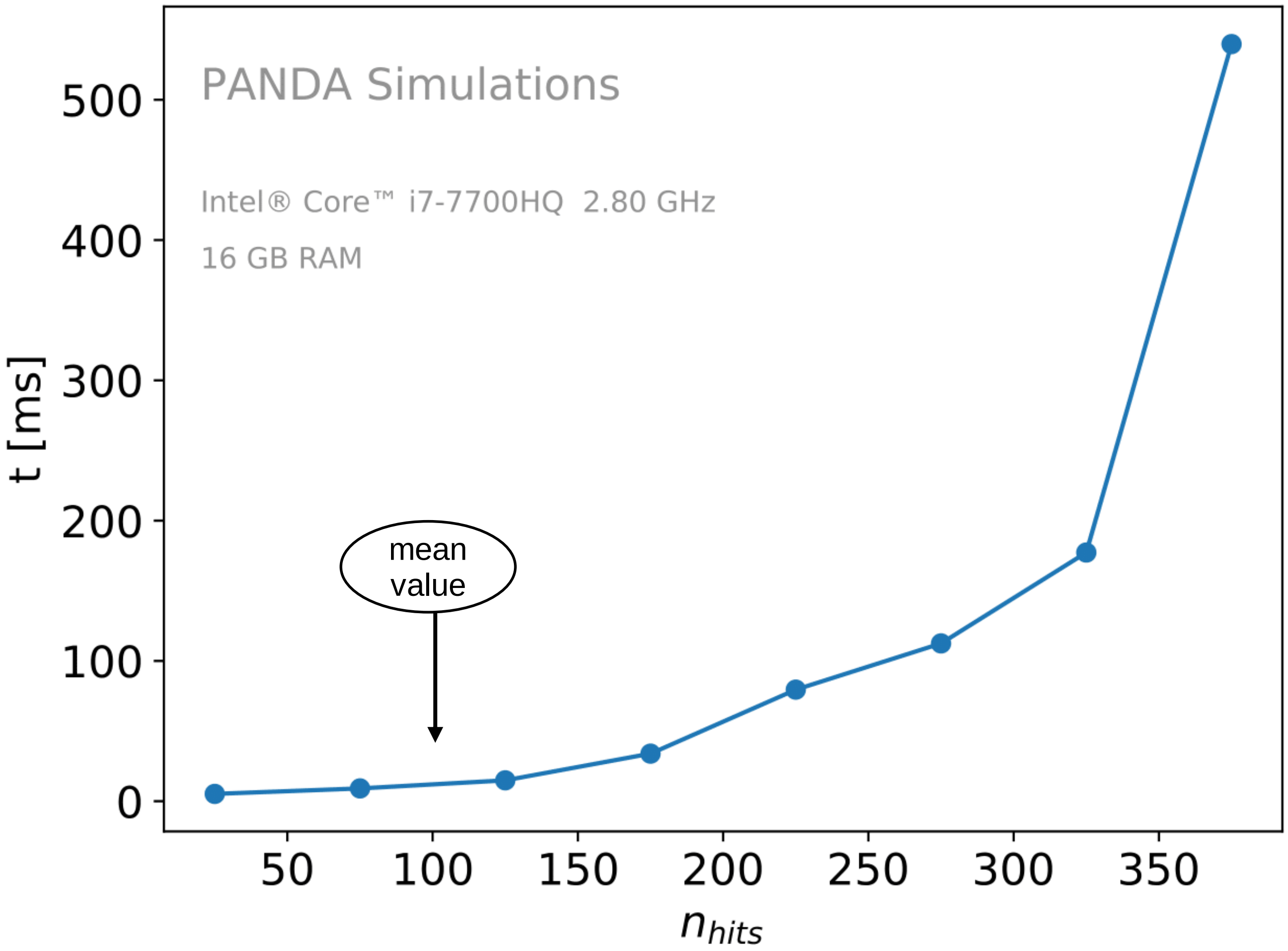
- At O(NN) > 0.25

  - Purity = 87%
  - MC Coverage = 97%
  - Efficiency = 86 - 91 %
  - Ghost ratio = 0.5%



DL Global Approach 9

How much time needed to create a graph



Conclusion:

- A variety of ML and DL algorithms can be applied to the track reconstruction problem
- The simplest approach a clustering algorithms k-means, DBScan, …
- NN can be used to build seeds for conventional algorithms
- RNN can be used to connect tracks from different sub-detectors
- RNN can also be used as track follower
- GNN shows the best results, however, it needs a large sample for training and powerful resources

- Training is done in Python using PyTorch
- libTorch (PyTorch C++ front-end) is used to interface PANDA framework with PyTorch

Thank You

Back Up

Performance

Quality Criteria

- To quantify the quality of the tracking algorithm:

1) Find the matching particle
2) Calculate the purity and MC coverage

$$\text{purity} = \frac{\mathrm{N}_{\text{hits-reco}} \cap \mathrm{N}_{\text{hits-MC}}}{\mathrm{N}_{\text{hits-reco}}} \qquad \text{coverage} = \frac{\mathrm{N}_{\text{hits-reco}} \cap \mathrm{N}_{\text{hits-MC}}}{\mathrm{N}_{\text{hits-MC}}}$$

3) Calculate the Efficiency and Ghost ratio

$$\text{partially found} = \frac{\mathrm{N}_{\text{reconstrcuted}}\ (\text{purity} > 80\ \%)}{\mathrm{N}_{\text{reconstructable}}}$$

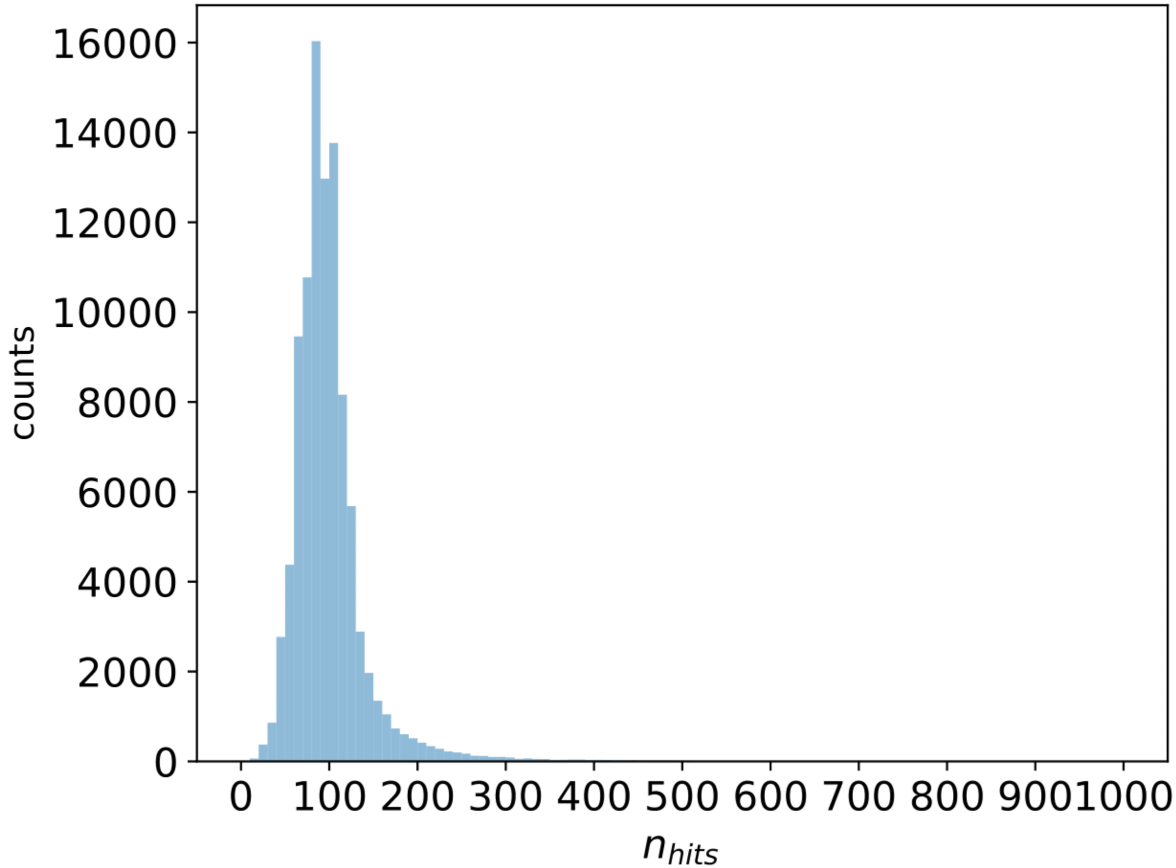
Fully found is partially found with coverage 100%

4) Ghost is fraction of reconstructed tracks with a purity < 80 %

- **Track is reconstructable if $\mathrm{N}_{\text{hits-MC}}$ > 4 before/after and inside the field**

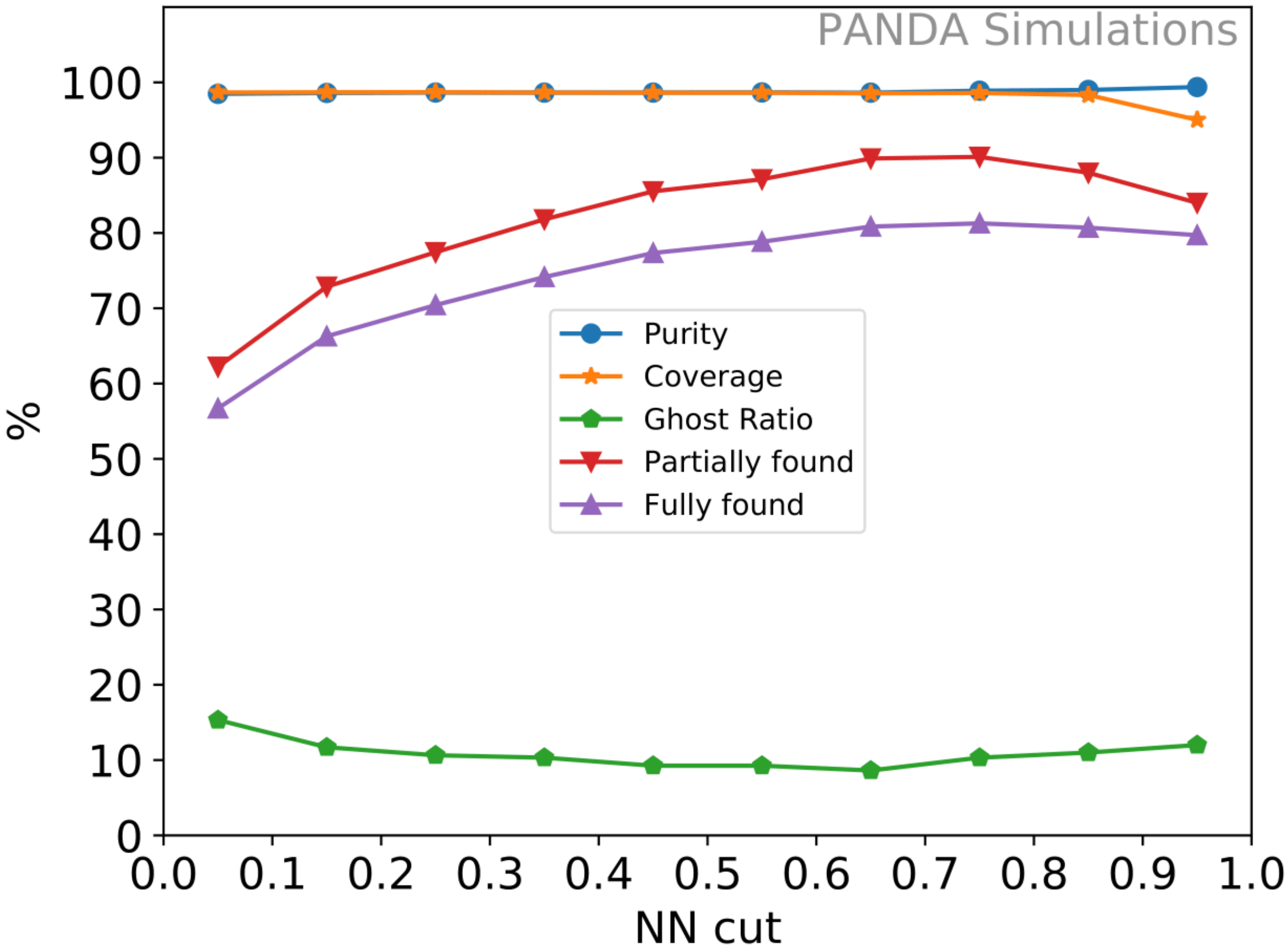
Training Set

- Particle gun with momentum range 0.1 – 10 GeV/c
- Multiplicity ≈ 5 tracks/event



DL Local Approach

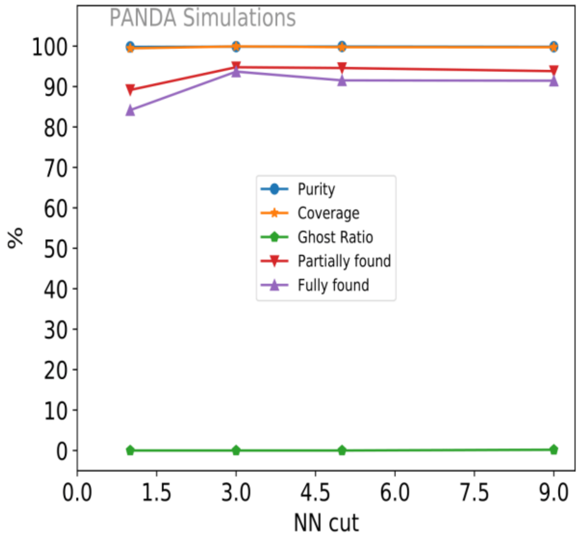
Quality Criteria NN+RNN



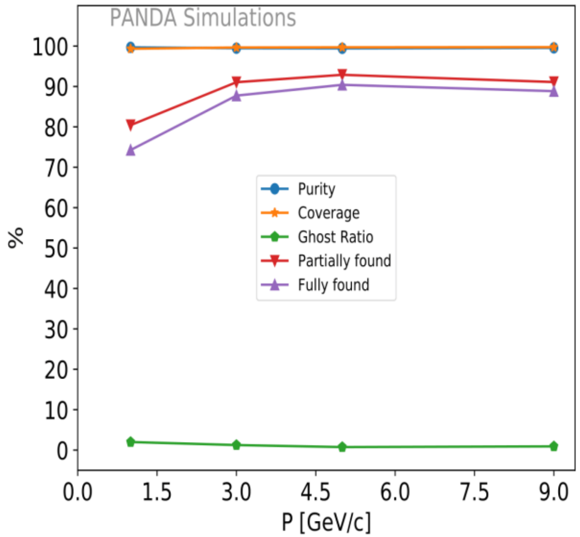
DL Global Approach

Momentum Dependence

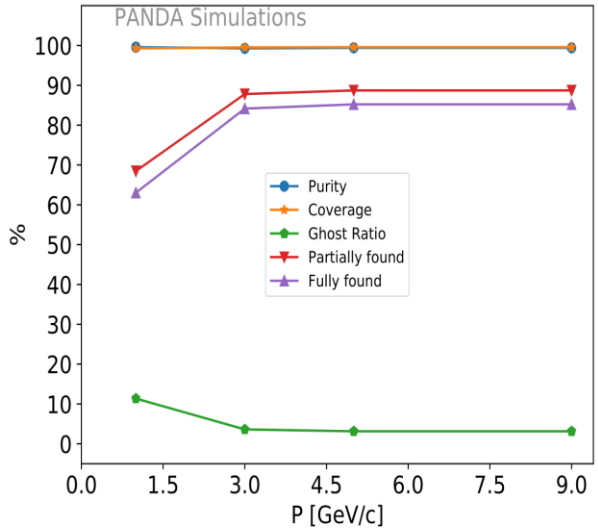
1 track/event



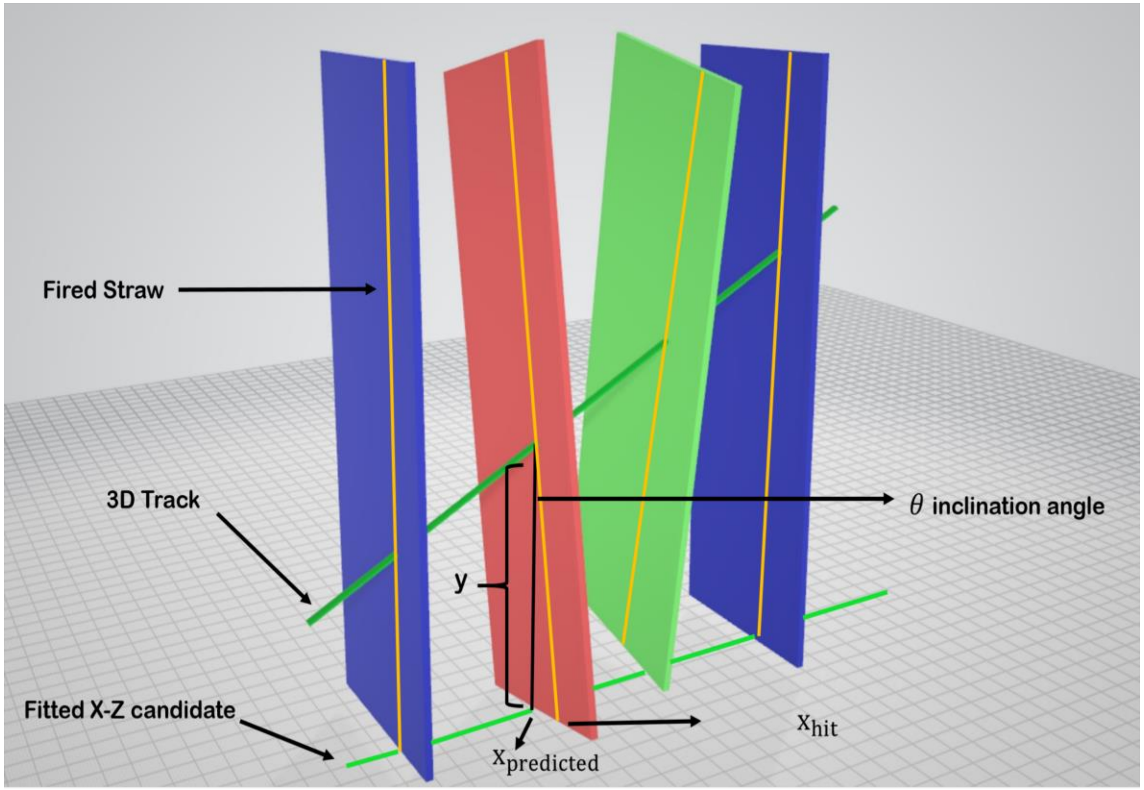
3 track/event



5 track/event



Build 3D track



1. Fit with line or circle

2. $\Delta x_{\pm} = x_{predicted} - (x_{hit} \pm r)$

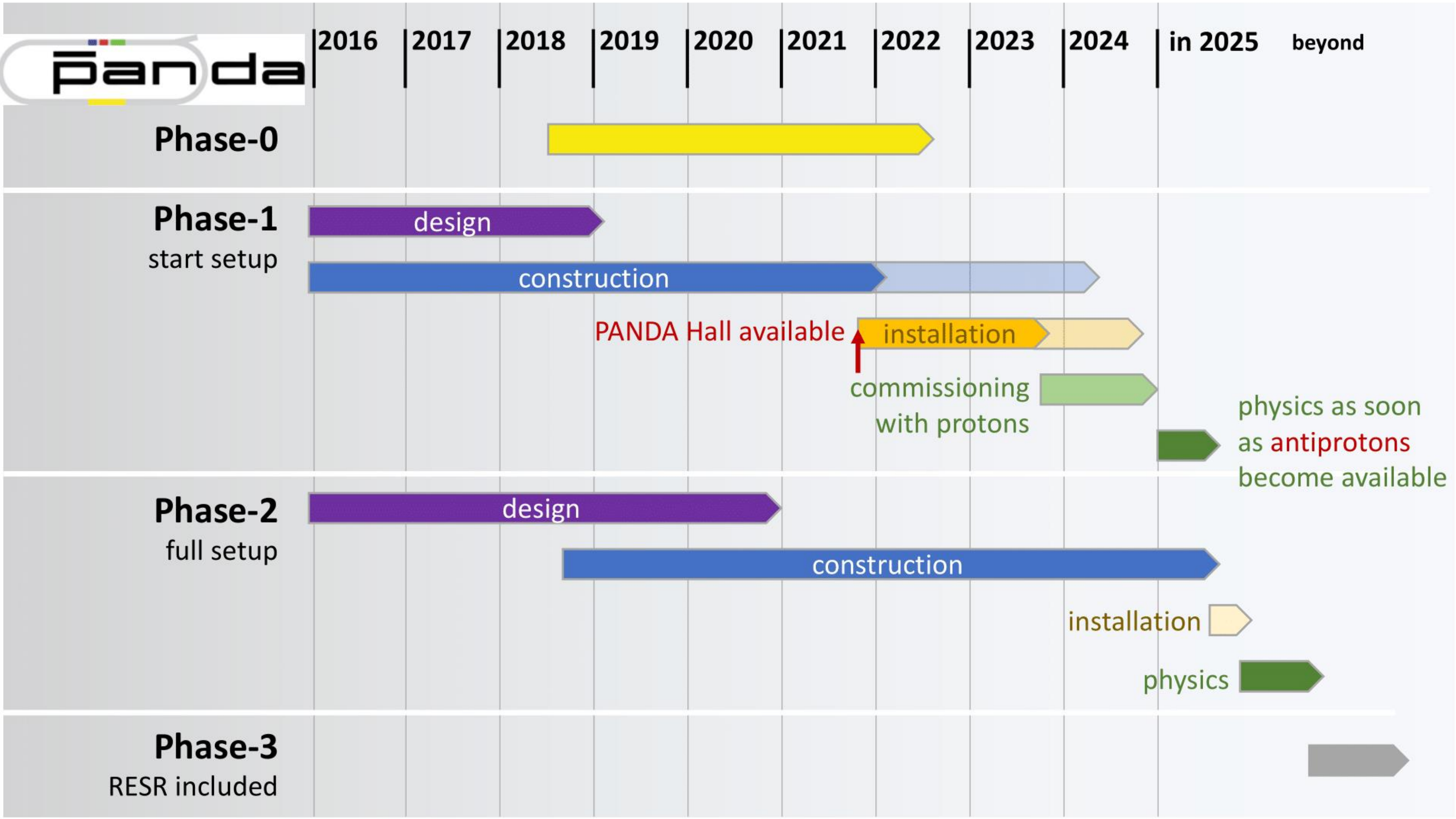
3. $y_{\pm} = \frac{\Delta x_{\pm}}{\tan\theta}$

4. $s_{\pm} = \frac{y_{\pm}}{z}$

5. Histogramming slopes

➢ Momentum is estimated from the ***kick angle***

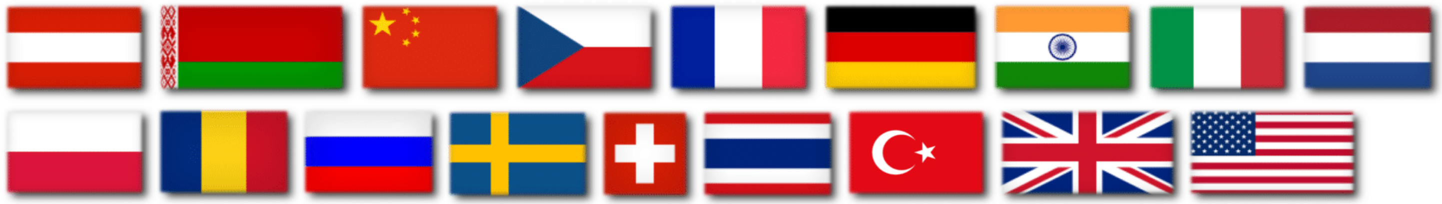
PANDA map



2016
2017
2018
2019
2020
2021
2022
2023
2024
in 2025
beyond
Phase-0
Phase-1
start setup
design
construction
PANDA Hall available
installation
commissioning with protons
physics as soon as antiprotons become available
Phase-2
full setup
design
construction
installation
physics
Phase-3
RESR included

PANDA Collaboration

Collaboration



UP Marche Ancona
U Basel
IHEP Beijing
U Bochum
Abant Izzet Baysal
U Golkoy, Bolu
U Bonn
U Brescia
IFIN-HH Bucharest
AGH UST Cracow
IFJ PAN Cracow
JU Cracow
Cracow UT
FAIR Darmstadt
GSI Darmstadt
JINR Dubna
U Erlangen

NWU Evanston
U Frankfurt
LNF-INFN Frascati
U & INFN Genova
U Gießen
Giresun U
U Glasgow
KVI-CART Groningen
Gauhati U, Guwahati
USTC Hefei
URZ Heidelberg
Doğuş U, Istanbul
Okan U, Istanbul
FZ Jülich
IMP Lanzhou
INFN Legnaro

Lund U
HI Mainz
U Mainz
RINP Minsk
ITEP Moscow
MPEI Moscow
U Münster
BINP Novosibirsk
Novosibirsk State U
U Wisconsin, Oshkosh
U & INFN Pavia
PNPI St. Petersburg
West Boh. U, Pilzen
Charles U, Prague
Czech TU, Prague

IHEP Protvino
Irfu Saclay
KTH Stockholm
Stockholm U
SUT, Nakhon Ratchasima
SVNIT Surat-Gujarat
S Gujarat U, Surat-Gujarat
FSU Tallahassee
Nankai U, Tianjin
U & INFN Torino
Politecnico di Torino
U Uppsala
SMI Vienna
NCBJ Warsaw
U York

more than 420 physicists from
from more than 65 institutions in 18 countries