GNN-based Track and Vertex Finding

B2GM TRG

Lea Reuter, Philipp Dorwarth, Slavomira Stefkova, Torben Ferber | Thursday 3rd November, 2022

Project members: Machine Learning for Trigger

KIT (ETP)

L. Reuter, I. Haide, P. Dorwarth,
P. Ecker, S. Stefkova, T. Ferber

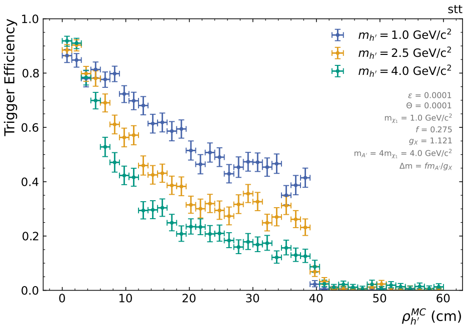
KIT (ITIV)

M. Neu, K. Unger, J. Becker

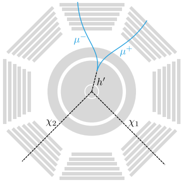
MPI (MPP)

E. Schmidt, F. Meggendorfer, C. Kiesling

Motivation



Credit: Patrick Ecker



Searches for displaced vertices

- Displaced vertices important signature in searches for new physics
- Example signal decay with dark photon A' and dark higgs h'

$$e^+e^- \to A'h',$$
$$h' \to \mu^+\mu^-,$$
$$A' \to \chi_1\chi_2,$$
$$\chi_2 \to \chi_1 e^+e^-$$

Problem:

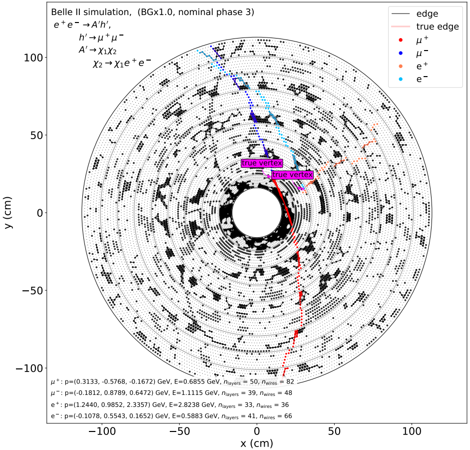
- Tracks with displacement larger than 40 cm are currently not triggered by **S**ingle **T**rack **T**rigger (stt)
- stt reconstruction efficiency decreases depending on displacement

Project Goal

- **Improve Track and Develop Vertex Finding using Graph Neural Networks (GNNs):**
  - Find events with displaced vertices (develop vertex finding)
  - Need to improve online L1 Trigger reconstruction

Project Goal

- **Improve Track and Develop Vertex Finding using Graph Neural Networks (GNNs):**
  - Find events with displaced vertices (develop vertex finding)
  - Need to improve online L1 Trigger reconstruction
- **Challenge:**
  - Tracks with low p_t (tracks curve)
  - Large occupancy due to beam-background hits (nominal phase 3)
  - Beam-background tracks (look like signal tracks)
  - Displaced vertices that are not pointing back to the interaction point



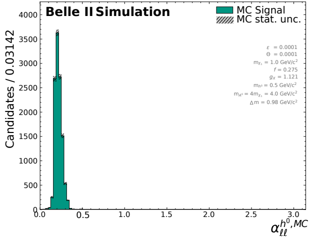
/group/belle2/dataprod/BGOverlay/nominal_phase3/prerelease-05-00-00a/overlay/phase3/BGx1/

MC Simulated Samples

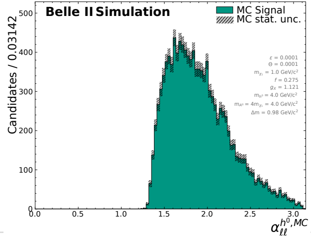
- Release:
  `feature/BII-9379-store-cdchit-relations-to-all-particle`
- Globaltags: `main_2022-01-27` and `patch_main_release-07`
- Starting with BGx0 and early-phase 3 BGx1
  `/group/belle2/dataprod/BGOverlay/early_phase3/`
  `release-05-01-15/overlay/phase31/BGx1/set0/`
- Signal samples:
  - Single displaced vertex samples
  - $e^+e^- \rightarrow A'h'$,
    $h' \rightarrow \mu^+\mu^-$,
    $A' \rightarrow \chi_1\chi_2$,
    $\chi_2 \rightarrow \chi_1 e^+e^-$ (outside of CDC)
  - on-shell (two-body)
  - $m(h)$ (0.5-4.0 GeV) in 0.1 GeV steps
- Background samples:
  - $e^+e^- \rightarrow e^+e^-$
  - $e^+e^- \rightarrow \mu^+\mu^-$

Distribution of opening angle $\alpha_{ll}^{h'}$

$m(h) = 0.5$ GeV

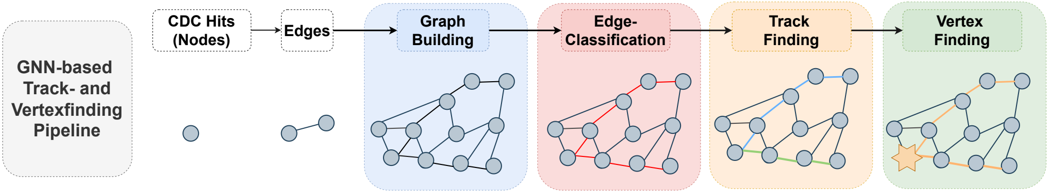


$m(h') = 4.0$ GeV



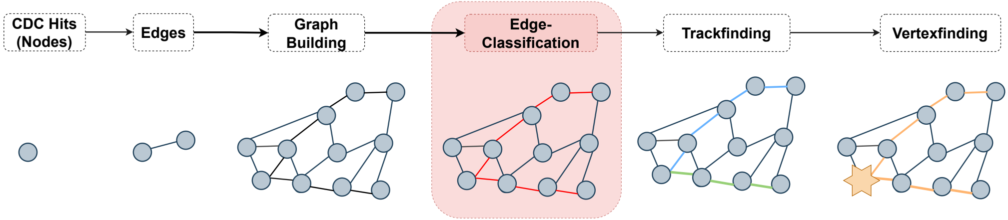
Approach with Graph Neural Networks

Variable number of CDC hits $\rightarrow$ utilize Graphs and Graph Neural Networks



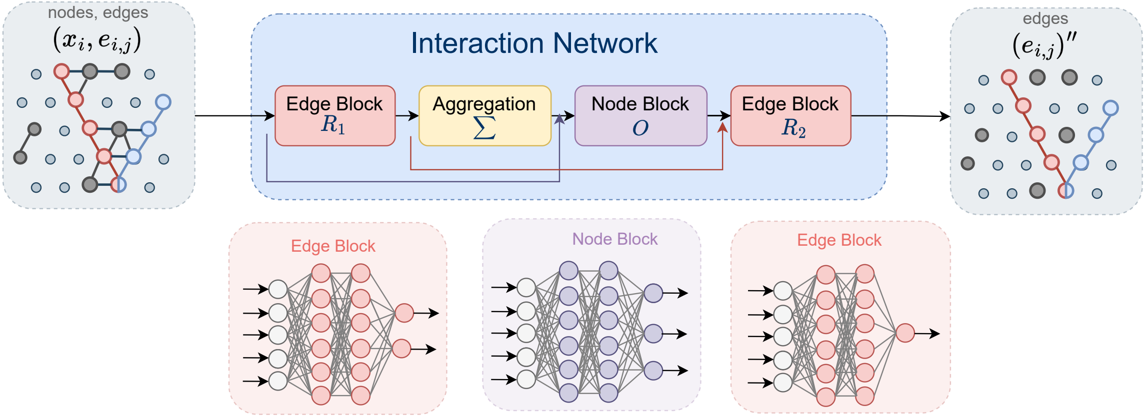
Approach with Graph Neural Networks: Edge Classification

Classify True and False edges of the graph



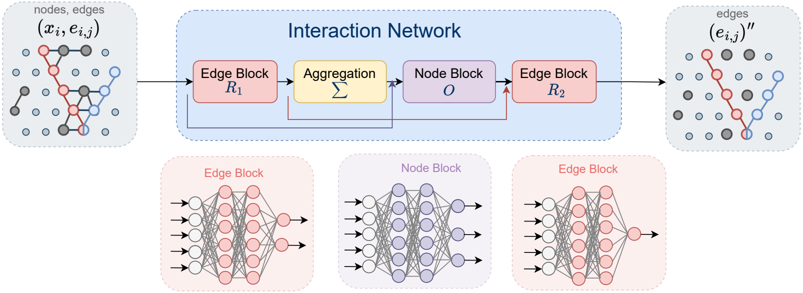
GNN: Interaction Network

Interaction Network: Graph Neural Networks for Charged Particle Tracking on FPGAs (arxiv:2112.02048)



GNN: Interaction Network

Graph Neural Networks for Charged Particle Tracking on FPGAs (arxiv:2112.02048)



Layer.Parameter	Param.Shape	Param #
R1.layers.0.weight	[30, 8]	240
R1.layers.0.bias	[30]	30
R1.layers.2.weight	[30, 30]	900
R1.layers.2.bias	[30]	30
R1.layers.4.weight	[2, 30]	60
R1.layers.4.bias	[2]	2
O.layers.0.weight	[30, 5]	150
O.layers.0.bias	[30]	30
O.layers.2.weight	[30, 30]	900
O.layers.2.bias	[30]	30
O.layers.4.weight	[3, 30]	90
O.layers.4.bias	[3]	3
R2.layers.0.weight	[30, 8]	240
R2.layers.0.bias	[30]	30
R2.layers.2.weight	[30, 30]	900
R2.layers.2.bias	[30]	30
R2.layers.4.weight	[1, 30]	30
R2.layers.4.bias	[1]	1

Total params: 3696

current input features:

- nodes: ρ, ϕ, TDC
- edges: $\Delta\rho$, $\Delta\phi$

Edge Classification GNN Evaluation

Determine binary threshold using maximal F_1 score:

- $purity = \frac{TP}{TP+FP}$
- $efficiency = \frac{TP}{TP+FN}$

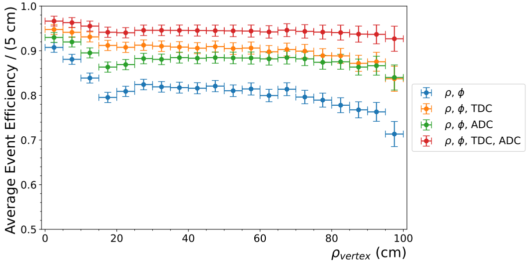
- $F_1 = 2 \cdot \frac{purity \cdot efficiency}{purity + efficiency} = \frac{TP}{TP+(FP+FN)/2}$

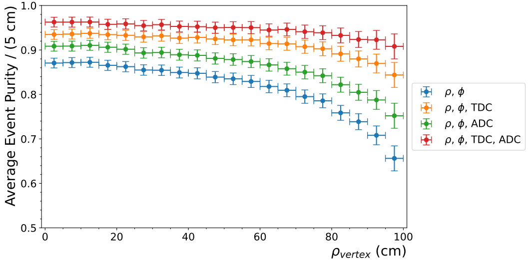
Using both TDC and ADC results in a

- Event classification efficiency of 94% and
- Event classification purity of 93%

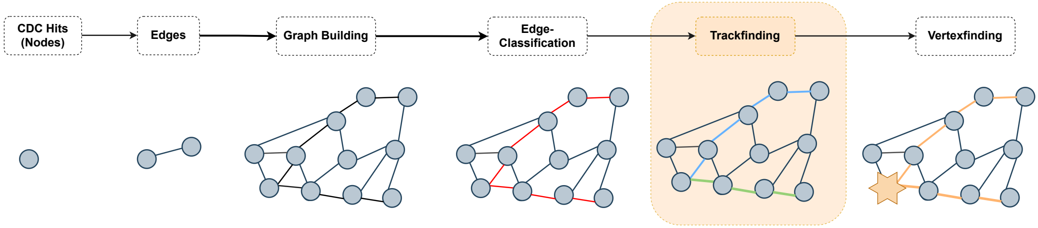
on signal with early-phase 3 beam-background

evaluated on 1000 samples per mass



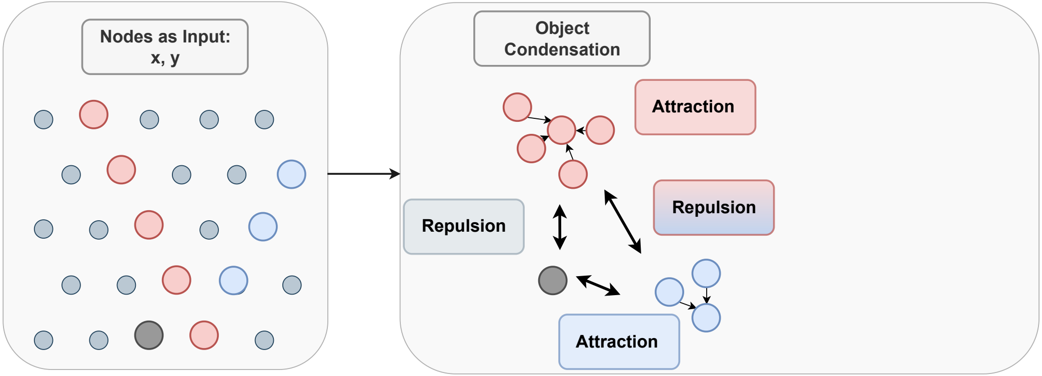


Approach with Graph Neural Networks: Trackfinding



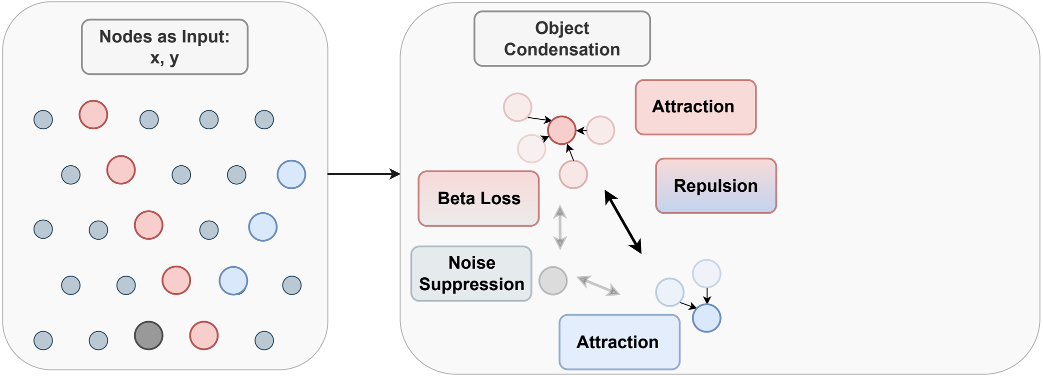
Use **Object Condensation** (arXiv:2002.03605)
→ Based on Isabel Haide: Improving ECL Clustering with Object Condensation

Object Condensation Trackfinding Approach



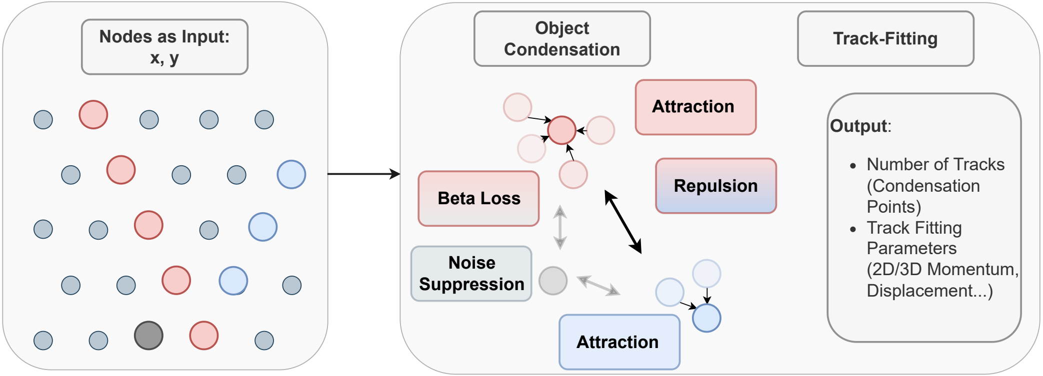
→ Use nodes as input to Object Condensation (arXiv:2002.03605)
→ **Goal:** predict track fitting parameters and find condensation points

Object Condensation Trackfinding Approach



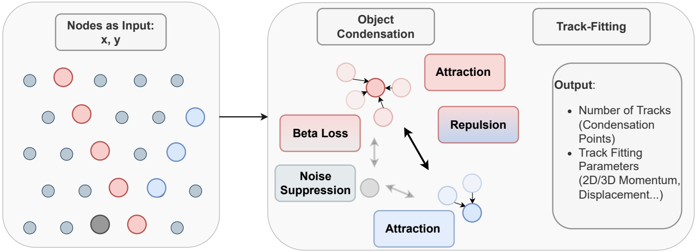
→ Use nodes as input to Object Condensation (arXiv:2002.03605)
→ **Goal:** predict track fitting parameters and find condensation points

Object Condensation Trackfinding Approach

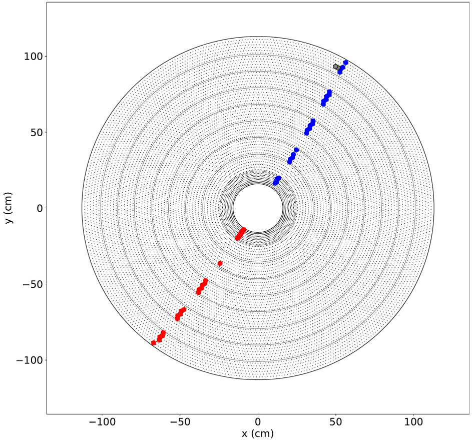


$\rightarrow$ Use nodes as input to Object Condensation (arXiv:2002.03605)
$\rightarrow$ **Goal:** predict track fitting parameters and find condensation points

Object Condensation Trackfinding Approach

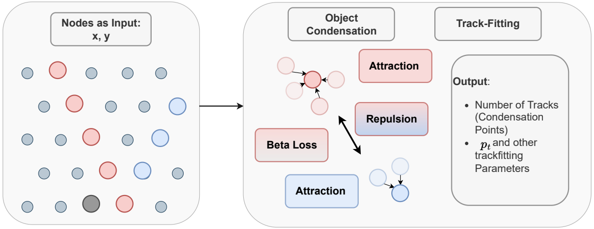


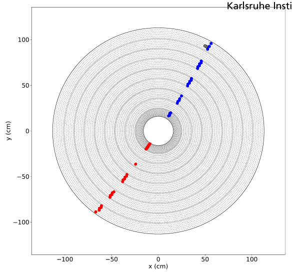
Start with very simple case:
$e^+e^- \to \mu^+\mu^-$
no beam-background

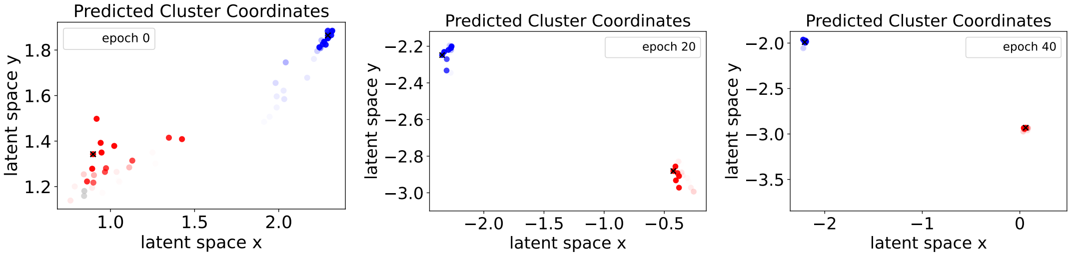


→ Use nodes as input to Object Condensation (arXiv:2002.03605)
→ **Goal:** predict track fitting parameters and find condensation points

Object Condensation Trackfinding Approach





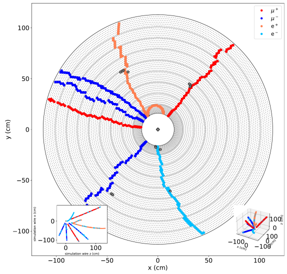


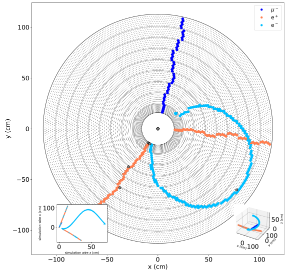
Object Condensation Trackfitting Approach:

Karlsruhe Institute of Technology

Create Samples with Particle Gun:

- number of particles [1,9]
- PDG code [-11,11,-13,13]
- BGx0 and BGx1 early-phase 3
- **not** displaced
- $\theta = [30, 120]$
- $p = [0.5, 5]$ GeV
- cleanup: remove events where primary particles have less than 5 CDC hits

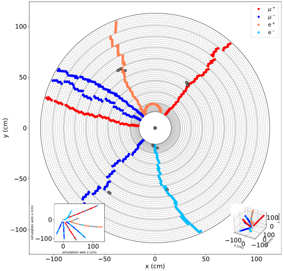


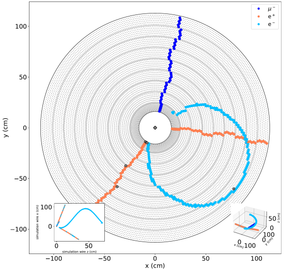


Object Condensation Trackfitting Approach:

Create Samples with Particle Gun:

- number of particles [1,9]
- PDG code [-11,11,-13,13]
- BGx0 and BGx1 early-phase 3
- **not** displaced
- $\theta = [30, 120]$
- $p = [0.5, 5]$ GeV
- cleanup: remove events where primary particles have less than 5 CDC hits





Training including stereo layers:

- Trained on 12 000 samples
- Number of particles [1,9]
- Input features per node: x, y, TDC, ADC
- Predicting: n_{tracks} p_x, p_y , p_z
  This can be changed (so for example more focus on track finding efficiency or predicting only pt)

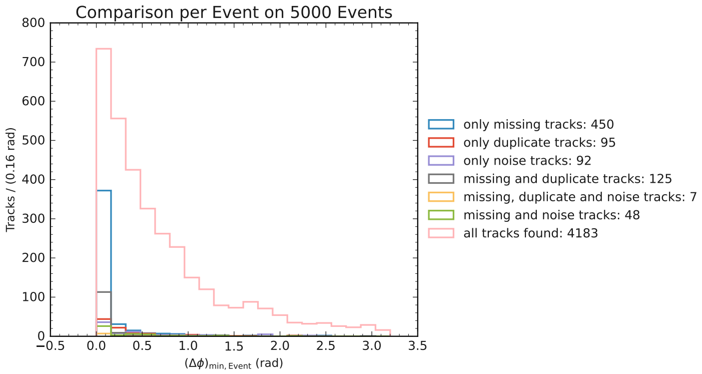
Object Condensation: BGx0 all Layers First Evaluation

Found Tracks	
All Tracks found (no duplicate/unmatched)	83.7%
Missing Tracks (no duplicate/unmatched)	9.0%
All Tracks and duplicate/unmatched	3.7%
Missing Tracks and duplicate/unmatched	3.6%

Object Condensation: BGx0 all Layers First Evaluation

Found Tracks	
All Tracks found (no duplicate/unmatched)	83.7%
Missing Tracks (no duplicate/unmatched)	9.0%
All Tracks and duplicate/unmatched	3.7%
Missing Tracks and duplicate/unmatched	3.6%

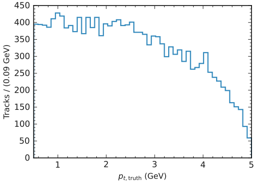
Calculate ϕ and p_t from p_x and p_y
$(\Delta\phi)_{min,Event}$ is the ϕ difference between the two nearest tracks per event:
$(\Delta\phi)_{min,Event} = min$(for all i,j: $\phi_i - \phi_{j\neq i}$)



Object Condensation: BGx0 all Layers First Evaluation

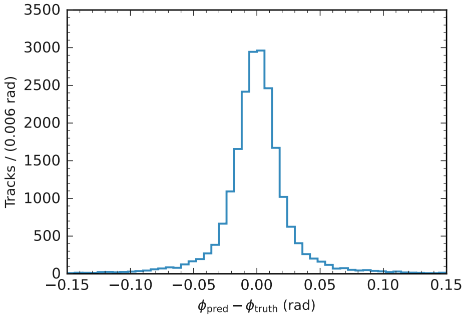
Karlsruhe Institute of Technology

truth p_t distribution:



Prediction:

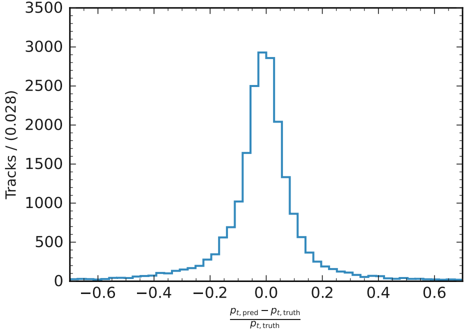
- number of particles
- used p_x, p_y to calculate p_t and ϕ



For CDC Track Finding
(and not only Displaced Vertex Finding)

- What prediction would be needed on trigger level?

Resolution of the transverse Momentum $_t$



$\rightarrow$ For the current triggers:

- is momentum calculated, whats the resolution?
- is there a cut on p_t?

Object Condensation: First Evaluation on early-phase3

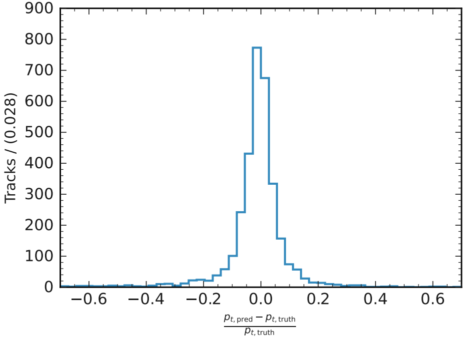
BGx0 all Layers Training

Found Tracks	
All Tracks found (no duplicate/unmatched)	83.7%
Missing Tracks (no duplicate/unmatched)	9.0%
All Tracks and duplicate/unmatched Tracks	3.7%
Missing Tracks and duplicate/unmatched	3.6%

BGx1 early-phase3 only axial layer Training

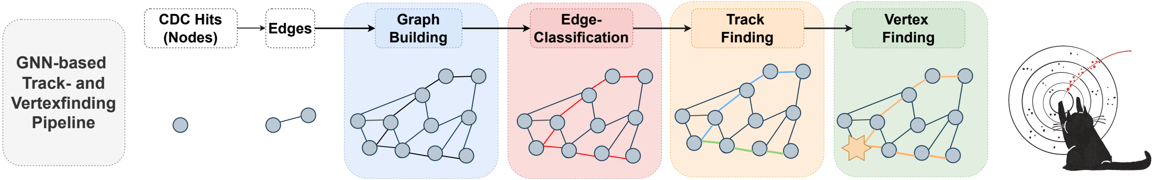
Found Tracks	
All Tracks found (no duplicate/unmatched)	73.8%
Missing Tracks (no duplicate/unmatched)	16.0%
All Tracks and duplicate/unmatched Tracks	7.4%
Missing Tracks and duplicate/unmatched	2.8%

BGx1 early-phase 3 relative p_t resolution



Similar to resolution with BGx0

Summary and Outlook



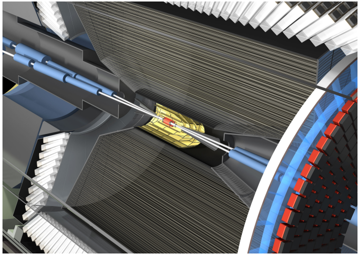
Current Status

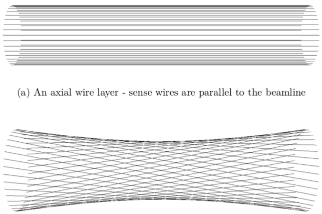
- Implemented Object Condensation for CDC Track Finding and testing BGx0 samples and BGx1 early-phase3 samples
- Evaluation Object Condensation Prediction
- Working on Object Condensation Truth Matching for Tracks

Outlook

- Implement Graph Building and Edge Classification Network on FPGA
- Data/MC comparison for Edge Classification
- Extend Object Condensation to predict Displaced Tracks
- Evaluate Object Condensation on Data/MC for Displaced Vertex Example
  $e^+e^- \to \Phi\gamma,$
  $\Phi \to K_{S^0} K_{L^0}$

Central Drift Chamber (CDC)





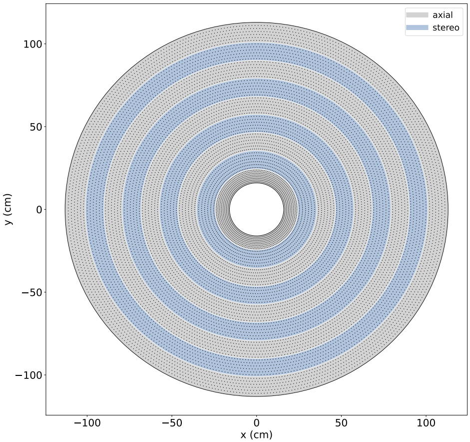
(a) An axial wire layer - sense wires are parallel to the beamline

(b) A stereo wire layer - sense wires are skewed to the beamline (exaggerated)

- Sense wires are arranged around the beamline (z-axis) to measure charged particles
- z Information gathered from stereo layers
- Events with displacement $\rho > 16.0\,\text{cm}$ start within the CDC

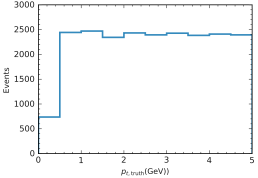
$\rightarrow$ Focus on track reconstruction using the CDC information

CDC x-y view



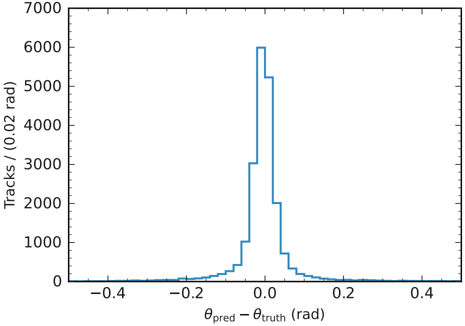
Object Condensation Trackfitting Approach

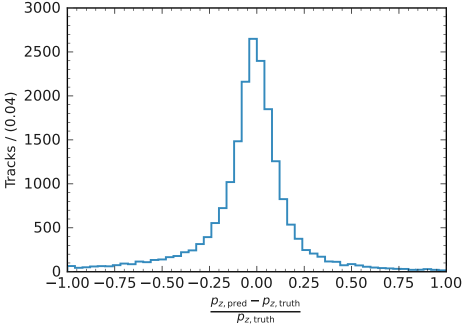
truth p distribution:



Prediction:

- number of particles
- using only axial hits
- input features
  x, y, TDC, ADC
- used p_x, p_y, p_z to
  calculate *theta*





GNN Evaluation: F_1: Input Feature Studies

Karlsruhe Institute of Technology

Determine binary threshold using maximal F_1 score: $F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} = \frac{TP}{TP + (FP+FN)/2}$

input features: $[\rho, \phi]$

Confusion matrix

Actual \ Predicted	True	False	Row Sum
True	36784 TP: 14.68%	7465 FN: 2.98%	44249 TPR: 83.13% FNR: 16.87%
False	9826 FP: 3.92%	196459 TN: 78.42%	206285 TNR: 95.24% FPR: 4.76%
Column Sum	46610 PPV: 78.92% FDR: 21.08%	203924 NPV: 96.34% FOR: 3.66%	250534 ACC: 93.10% MISS: 6.90%

$F_1 = 0.81$

input features: $[\rho, \phi, TDC]$

Confusion matrix

Actual \ Predicted	True	False	Row Sum
True	41499 TP: 16.56%	3414 FN: 1.36%	44913 TPR: 92.40% FNR: 7.60%
False	5111 FP: 2.04%	200510 TN: 80.03%	205621 TNR: 97.51% FPR: 2.49%
Column Sum	46610 PPV: 89.03% FDR: 10.97%	203924 NPV: 98.33% FOR: 1.67%	250534 ACC: 96.60% MISS: 3.40%

$F_1 = 0.91$

input features: $[\rho, \phi, ADC]$

Confusion matrix

Actual \ Predicted	True	False	Row Sum
True	40465 TP: 16.15%	5993 FN: 2.39%	46458 TPR: 87.10% FNR: 12.90%
False	6145 FP: 2.45%	197931 TN: 79.00%	204076 TNR: 96.99% FPR: 3.01%
Column Sum	46610 PPV: 86.82% FDR: 13.18%	203924 NPV: 97.06% FOR: 2.94%	250534 ACC: 95.16% MISS: 4.84%

$F_1 = 0.87$

input features: $[\rho, \phi, TDC, ADC]$

Confusion matrix

Actual \ Predicted	True	False	Row Sum
True	43741 TP: 17.46%	2760 FN: 1.10%	46501 TPR: 94.06% FNR: 5.94%
False	2869 FP: 1.15%	201164 TN: 80.29%	204033 TNR: 98.59% FPR: 1.41%
Column Sum	46610 PPV: 93.84% FDR: 6.16%	203924 NPV: 98.65% FOR: 1.35%	250534 ACC: 97.75% MISS: 2.25%

$F_1 = 0.94$

GNN Evaluation per Mass