technische universität
dortmund

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Track classification for high density beams in beam tests

Christopher Krause
christopher2.krause@tu-dortmund.de

Kevin Kröninger, Jens Weingarten

April 21. 2023

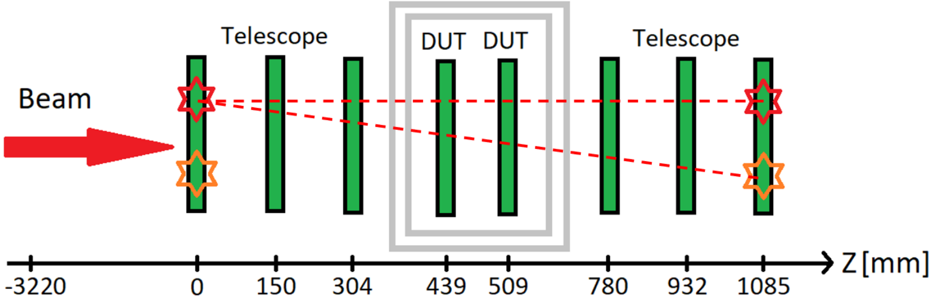
TU Dortmund
Department of Physics

11th BTTB Workshop

ap2
allpix squared

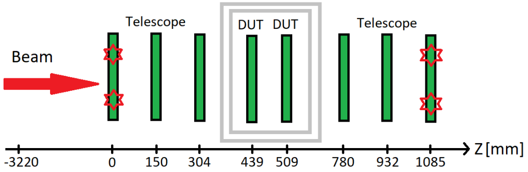
Track Reconstruction

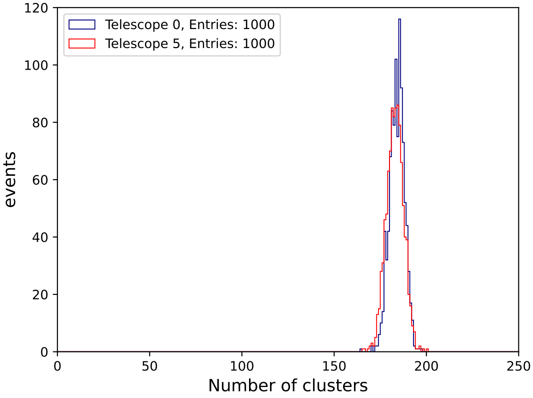
- Track reconstruction of high density beam causes combinatorial problems
- All cluster combinations on first and last plane are taken into account
- Wrong combinations possible $\longrightarrow$ False tracks



Cluster Multiplicity

- Allpix^2 simulation in air, 200 pions per event (1000 events)
- Telescope consists of 6 Mimosa26 sensors
- No sensor misalignment
- Beam opening angle: 0.1mrad
- Beam energy: (120.0 ± 0.5) GeV

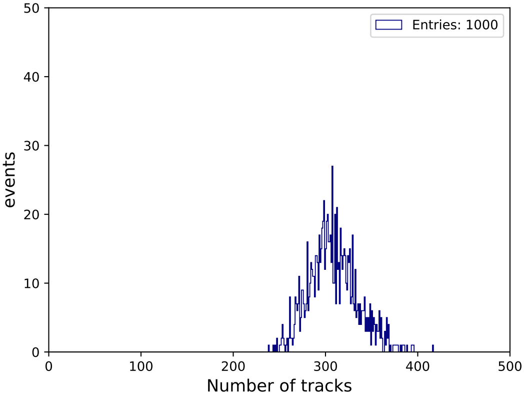




Track Multiplicity

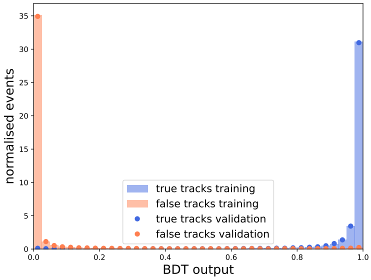
- Reconstruction configurations:
  - Demand hit on all planes for track
  - Matching Radius: 10 $\times$ spatial resolution
- **364k** tracks in total, less than **200k** clusters on each plane
- Number of tracks significantly larger than expected number of true tracks

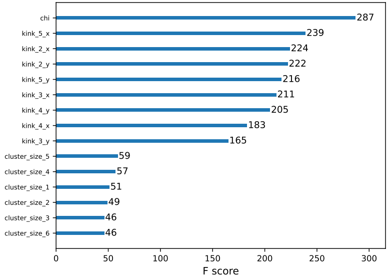
Can true and false tracks be distinguished?



XGBoost

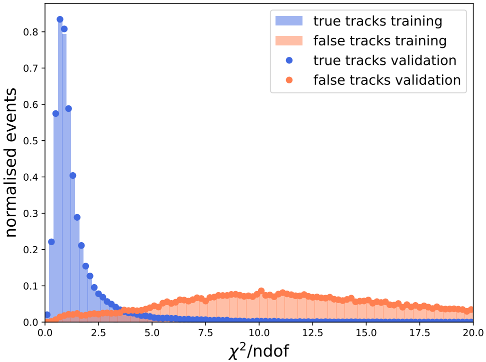
- Using machine learner (XGBoost) to determine true and false tracks
- Training data set: 200k π^+, Validation/Test data set: 50k π^+
- Input quantities: χ^2/ndof, kink angles, cluster sizes

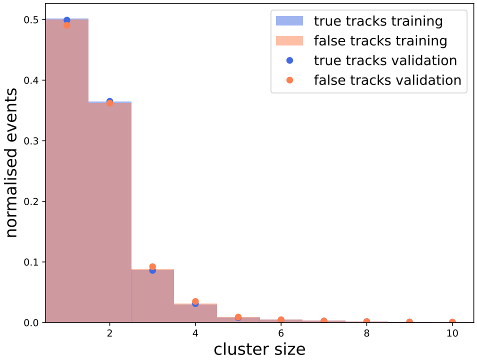




Separating Features

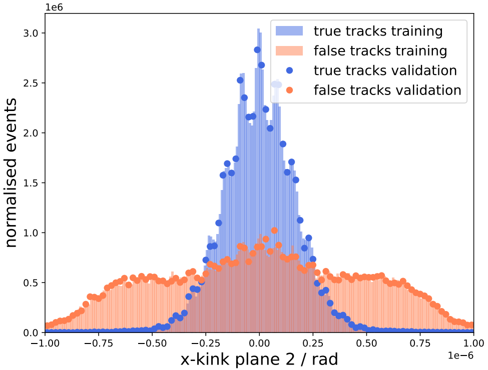
- True and false track distributions of the input features

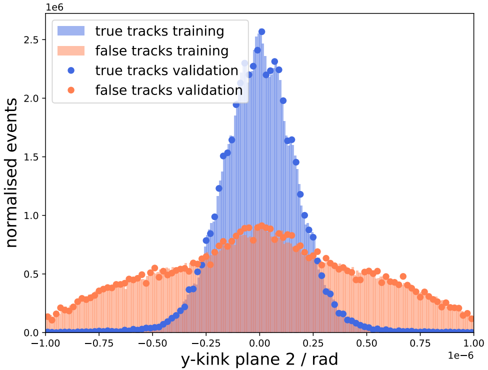




Separating Features

- True and false track distributions of the input features



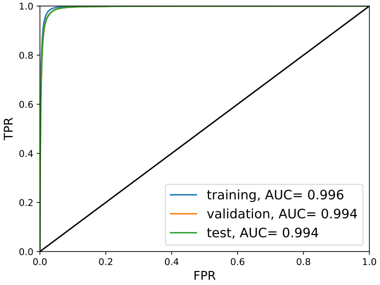


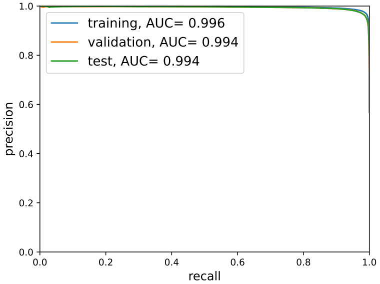
ROC and PR curve

- Using ROC curve and precision-recall curve for performance evaluation
- FPR = $f_p/(f_p + t_n)$
  Probability for false tracks to be labeled as true
- TPR = Recall = $t_p/(t_p + f_n)$
  Ability to find all true tracks
- Precision = $t_p/(t_p + f_p)$
  Ability to not label false tracks as true

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

[C. Dilmegani, Machine Learning Accuracy: True-False Positive/Negative, 2023]



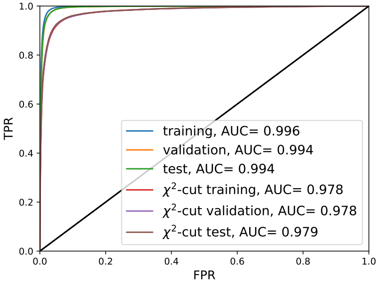


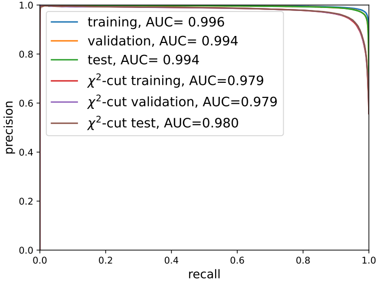
χ^2/ndof Cut

- FPR = $f_p/(f_p + t_n)$
  Probability for false tracks to be labeled as true
- TPR = Recall = $t_p/(t_p + f_n)$
  Ability to find all true tracks
- Precision = $t_p/(t_p + f_p)$
  Ability to not label false tracks as true

- Comparison to most important quantity:
  χ^2/ndof

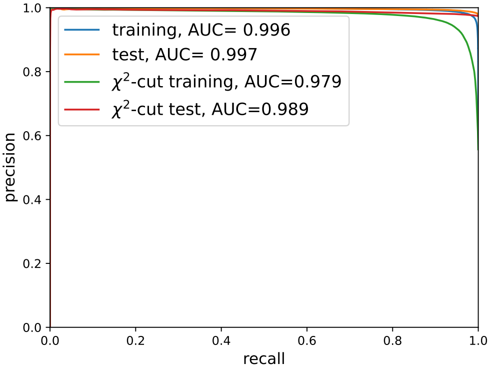
Cut on χ^2/ndof almost as good

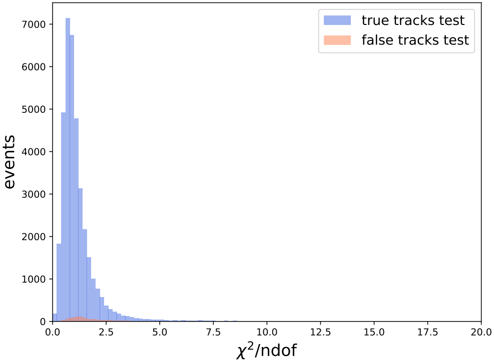




Unique Cluster Usage

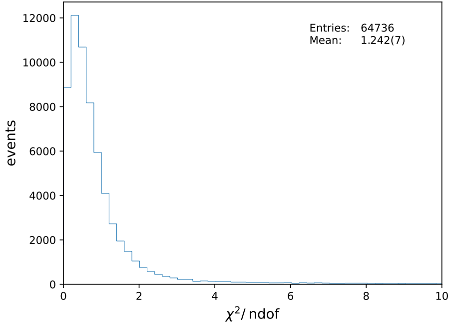
- Only using one track per cluster (keeping track with lowest χ^2/ndof)

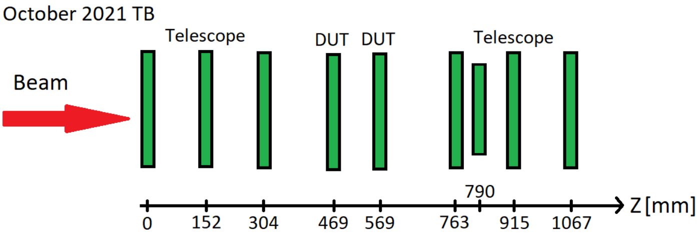




Real Testbeam

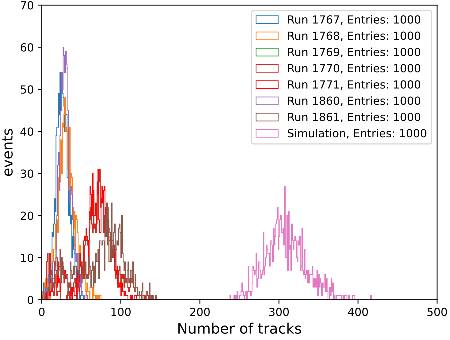
- Using Cern Testbeam data from October 2021 (High track densities)
- Telescope precisely aligned
  Masking: Frequency cut = 20
- Matching Radius: 10 × spatial resolution

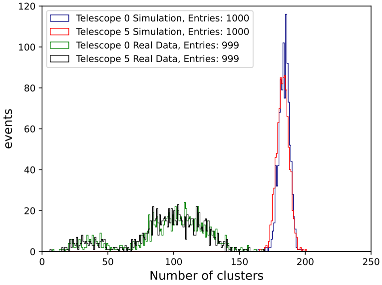


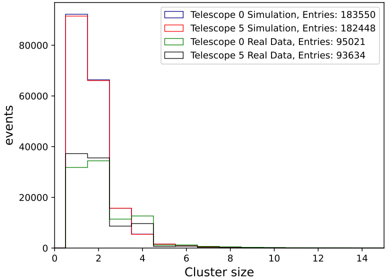


Simulation vs Reality

- Using only 1k events from real data (Run 1771)
- Simulated data has significantly higher track density
  $\longrightarrow$ Less false tracks in real data

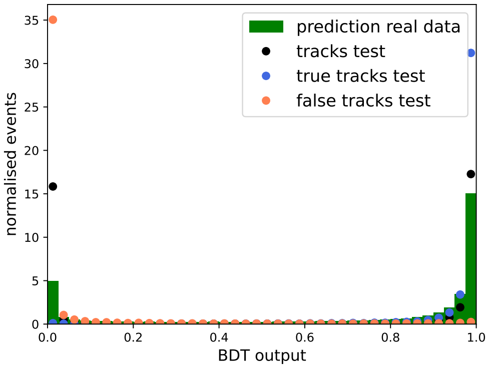


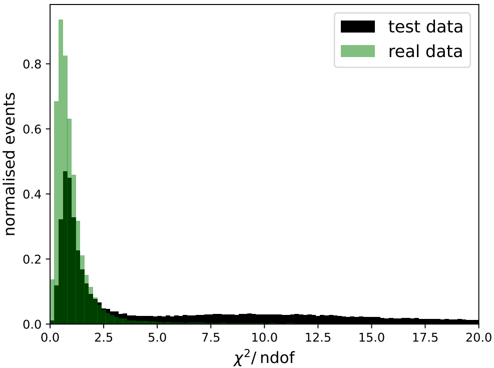




ML Results

- Clear distinction between tracks with low and high probabilities





Summary

- High track density beams in past campaigns had ≈ 2x larger track multiplicity
- Applying χ^2-cut yields high precision and recall values
- False tracks have significantly larger χ^2-values on average
- Using adequate χ^2-cut/unique clusters filters out most false tracks

Track reconstruction of high density beams in beam tests possible

Summary

- High track density beams in past campaigns had ≈ 2x larger track multiplicity
- Applying χ^2-cut yields high precision and recall values
- False tracks have significantly larger χ^2-values on average
- Using adequate χ^2-cut/unique clusters filters out most false tracks

Track reconstruction of high density beams in beam tests possible

Thank You!