UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Graph Networks for Track Reconstruction

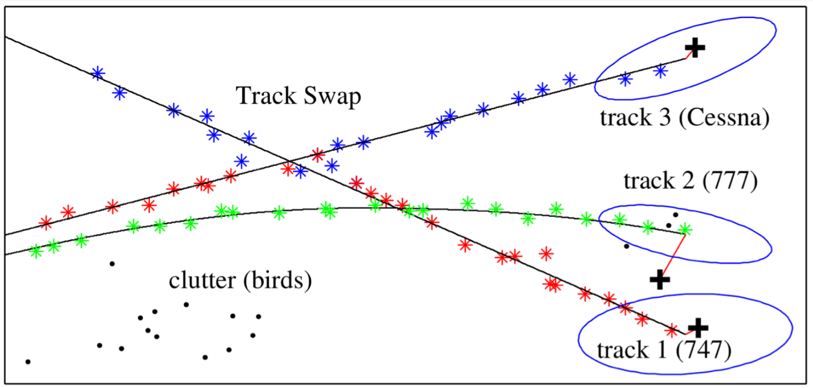
Mini Workshop on Machine Learning, ICNFP 2021

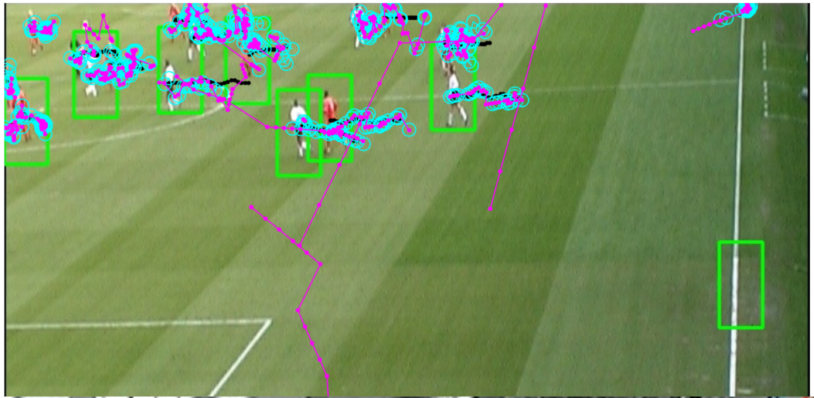
Marcel Kunze, Heidelberg University

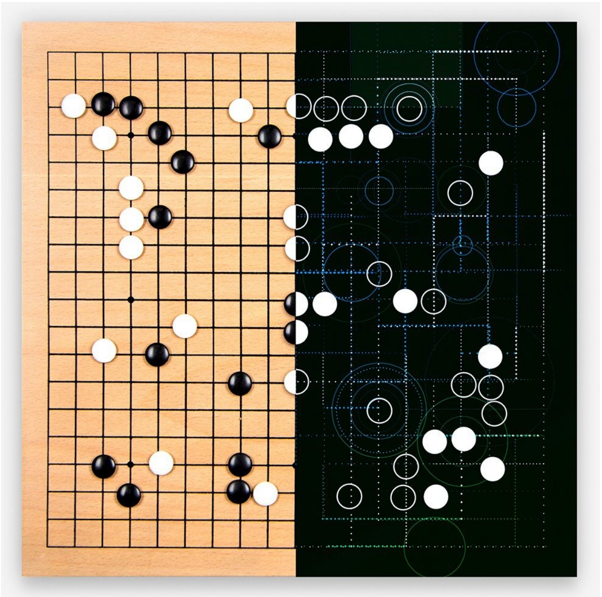
Pattern Recognition/Tracking

Pattern recognition/tracking is a very old, very hot topic in Artificial Intelligence, but very varied

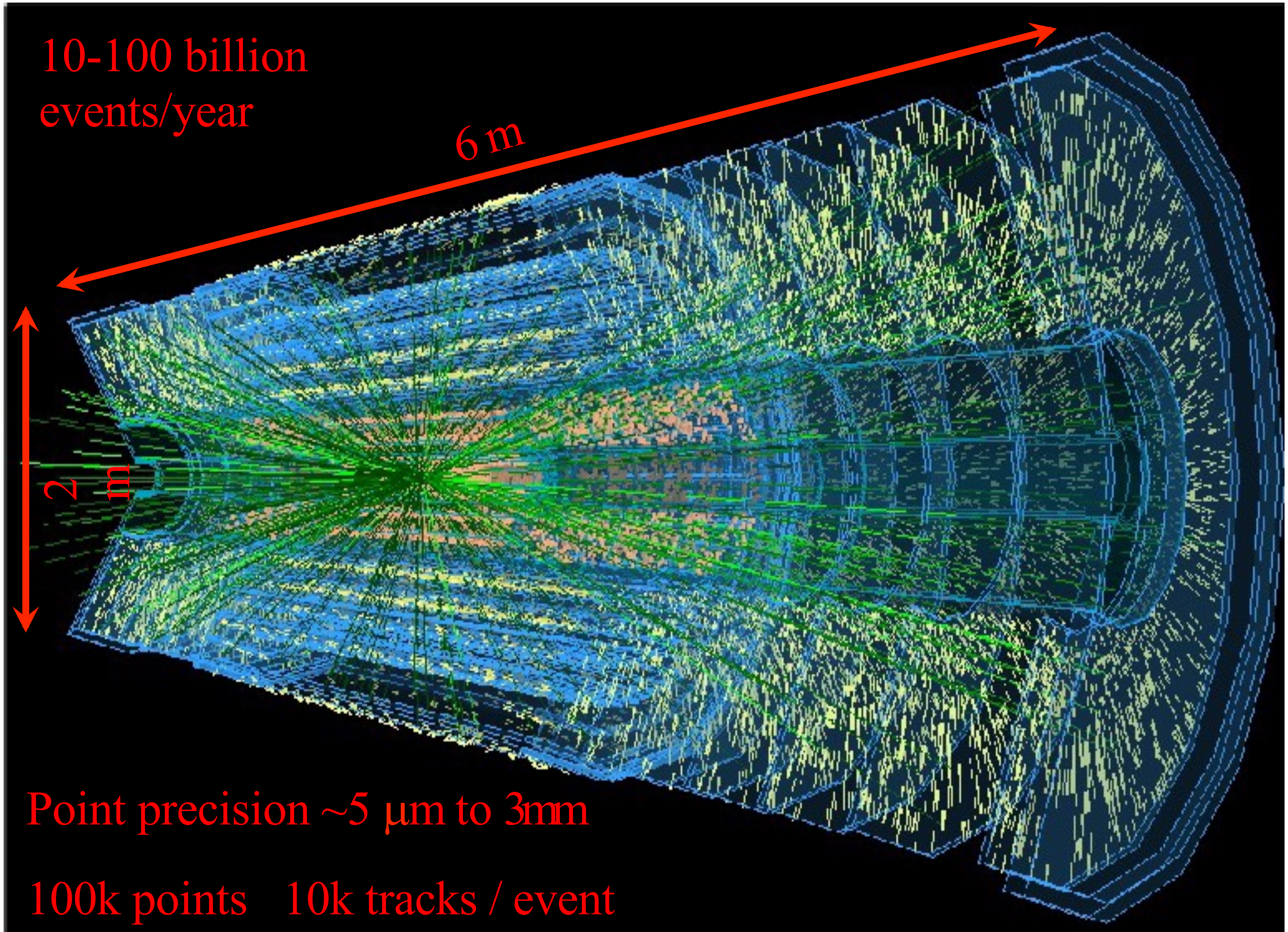
Note that these are real-time applications, with CPU constraints





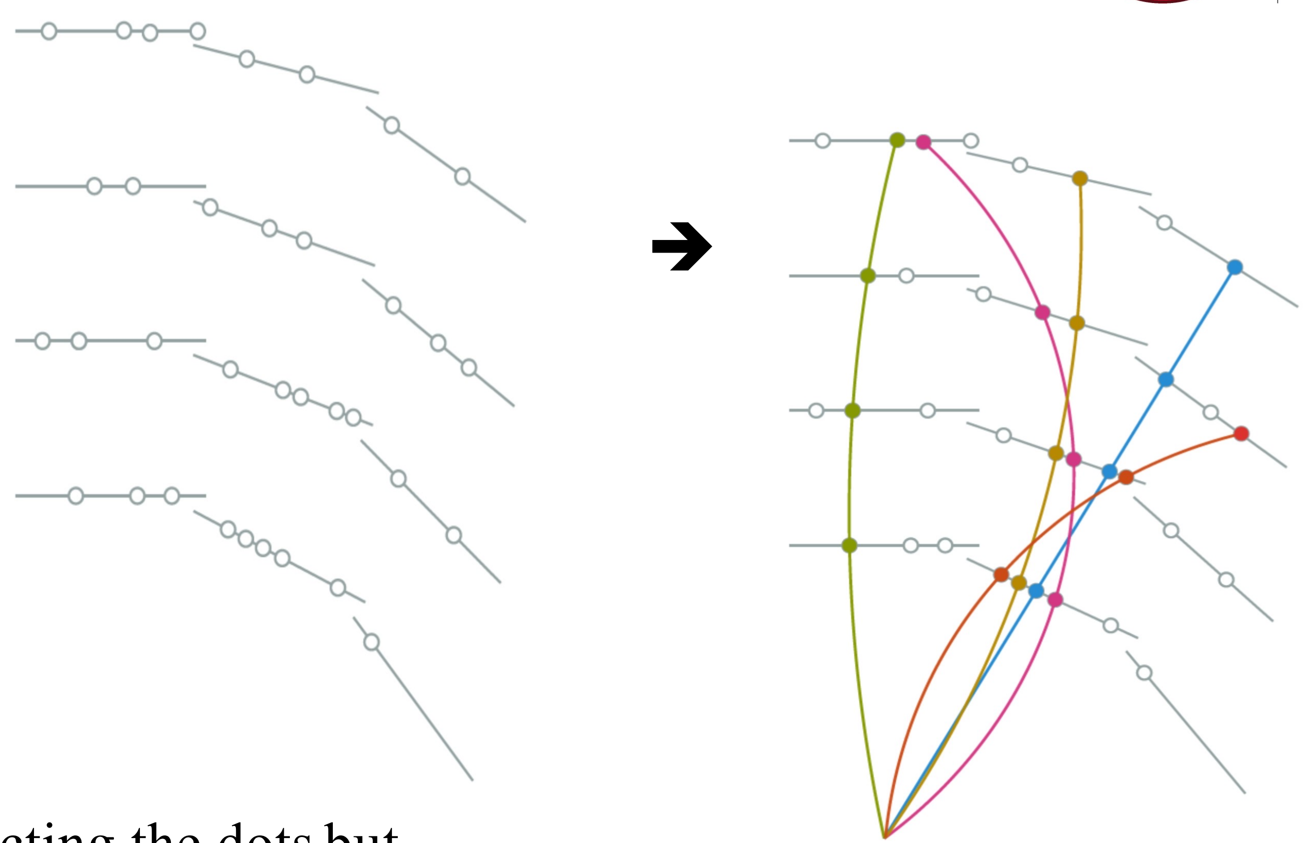


HL-LHC All Silicon Tracker (Simulation)



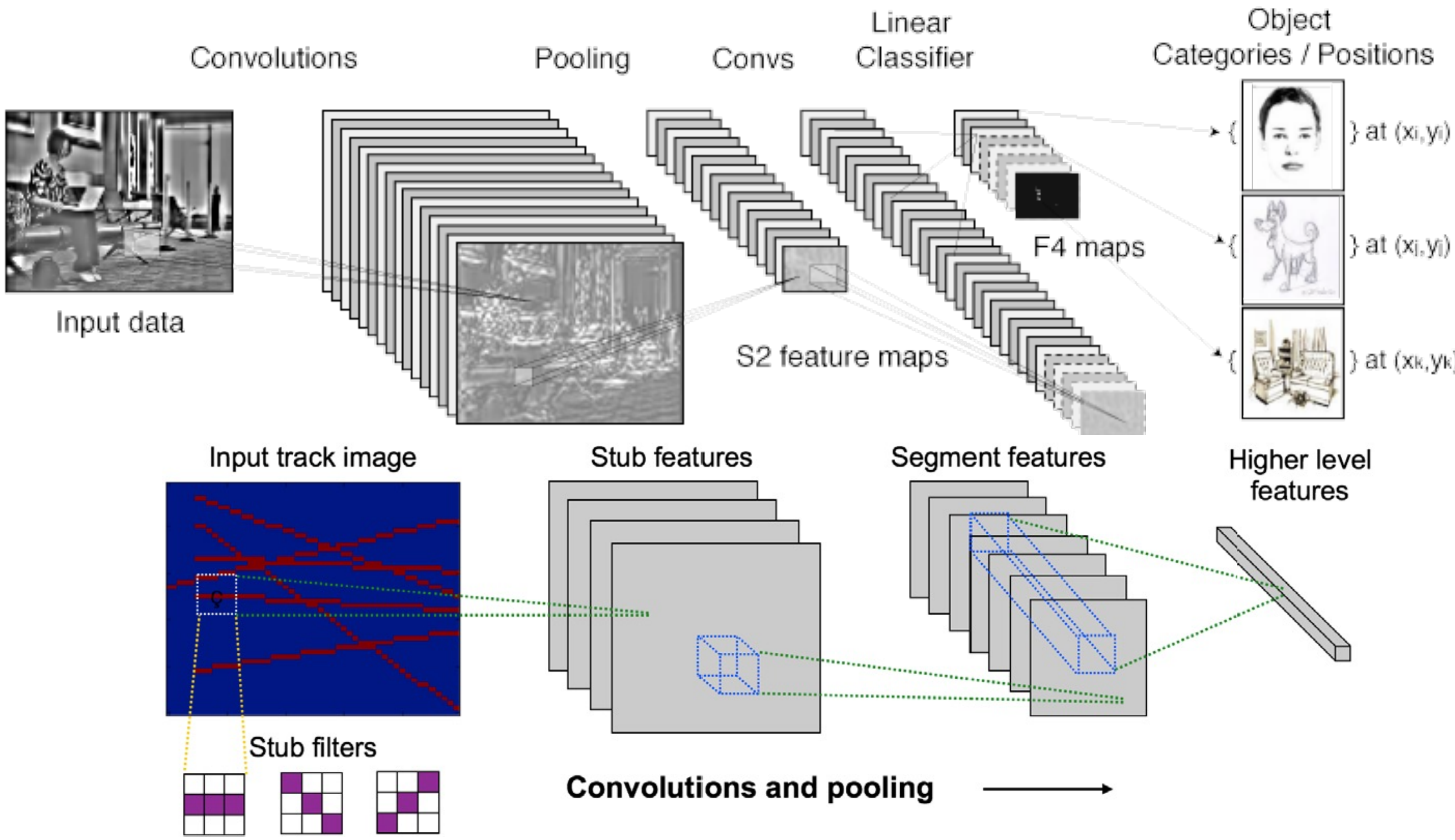
Problem Definition

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386



Connecting the dots but
- 3 dimensions
- 100’000 points into 10’000 tracks

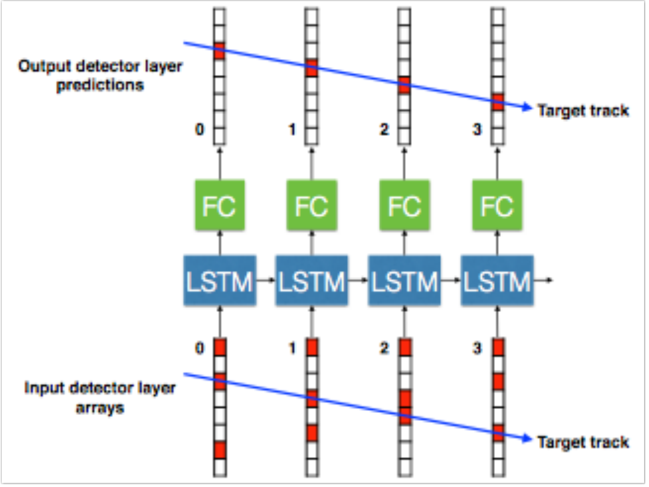
Convolution NN ?
Image based Approach

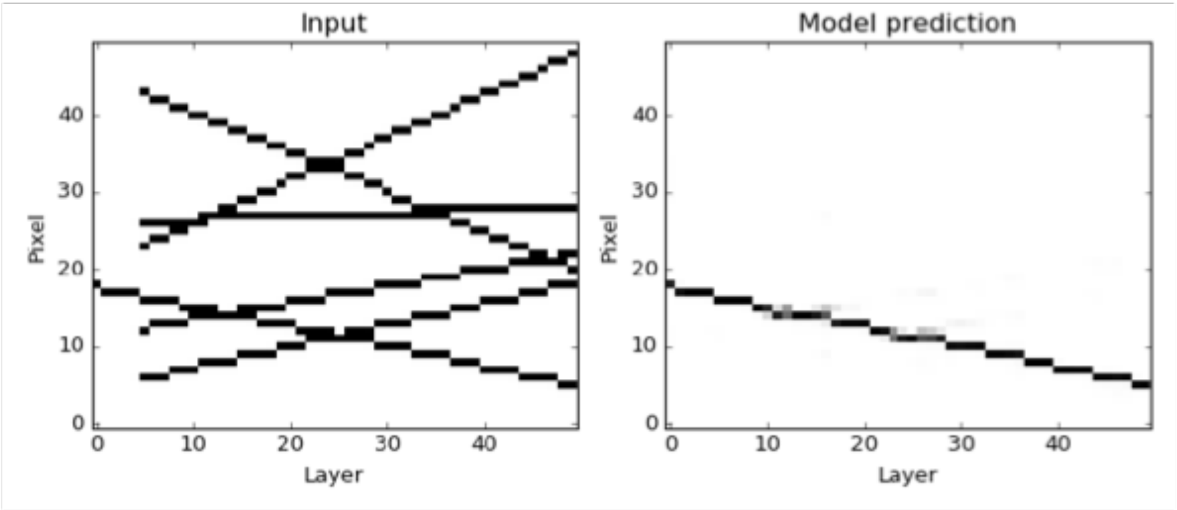


Recurrent Neural Networks (RNN) ? Point based Approach

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Long Short Term Memory (LSTM)

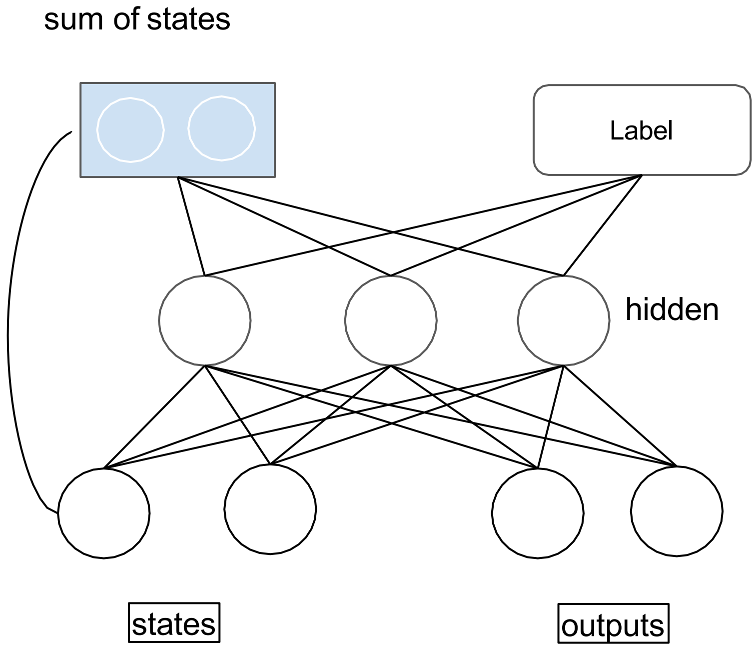




Graph Neural Network

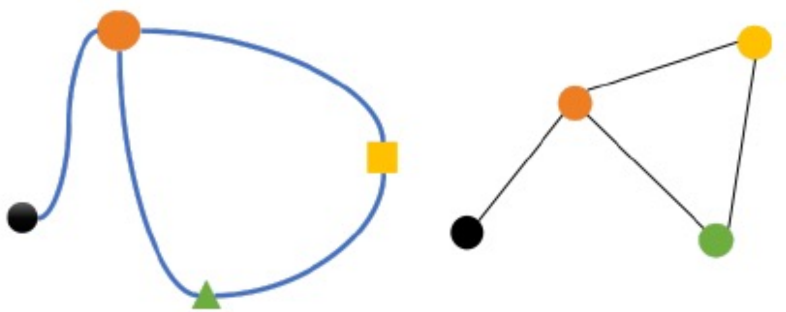
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

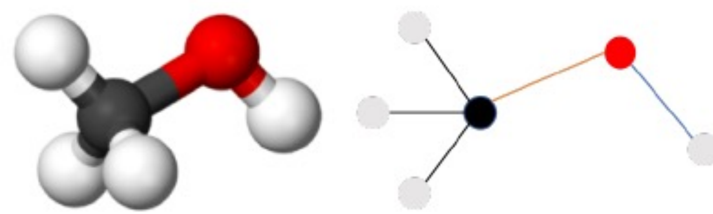
- A graph is processed node by node in a random order
- For a node in graph, the sum of the state vectors of neighboring nodes are computed and concatenated to its own label vector
- The algorithm guarantees a convergence of the state nodes to a stable and unique solution



Application Domains

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386





Computer Science:
Routing + tracking algorithms

Chemistry:
Molecular engineering

TrackML Throughput phase 3rd place

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Phase 2 cloudkitchen

Author: Marcel Kunze

ROOT · C++ CPP

Accuracy: 0.93
Time/event: ~7 sec
Memory: 0.7 Gb

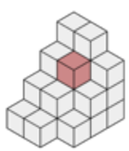
partly based on top quarks Phase 1 solution

1	—	**Top Quarks**	0.92182	10	2mo

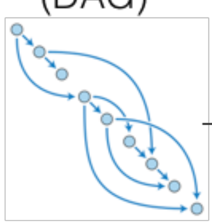
Algorithm outline

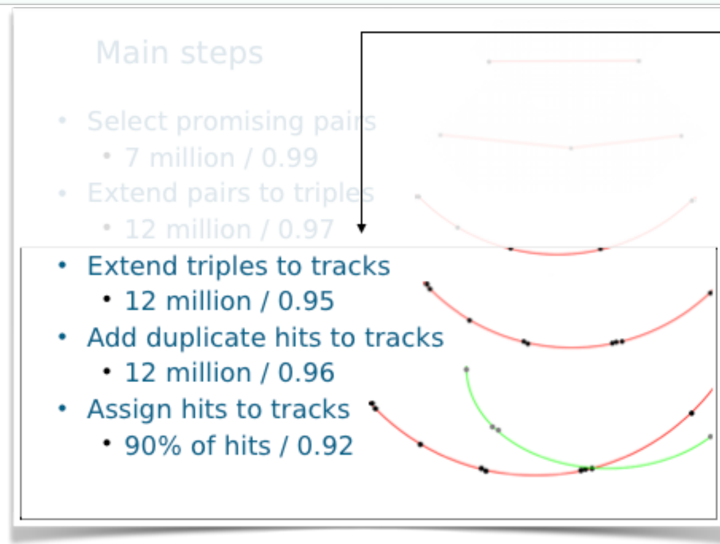
hits

↓ sorted in voxels



↓ organised in direct acyclic graphs (DAG)





DAGs are pre-trained on ~25 events ground truth

DAGs are used to fast navigate through voxel space → Disc section / Radial section

Triplet finder: NN3

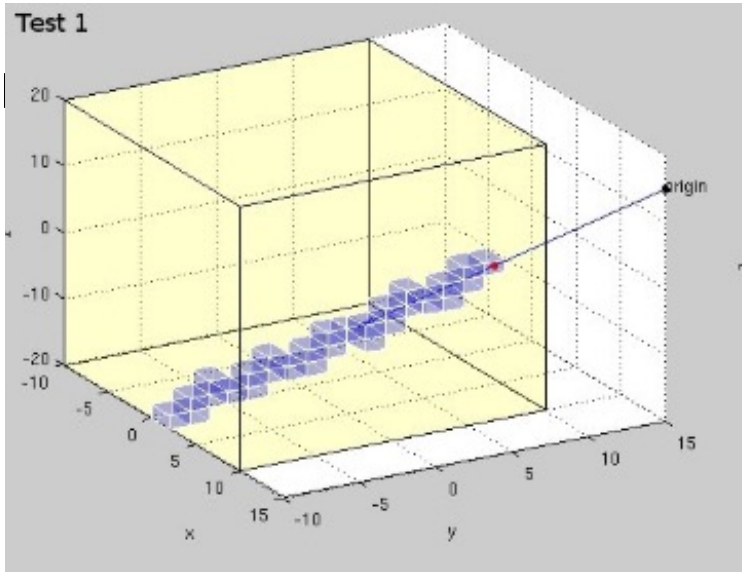
doublet finder: NN1, NN2

Threaded

Graph Network Training

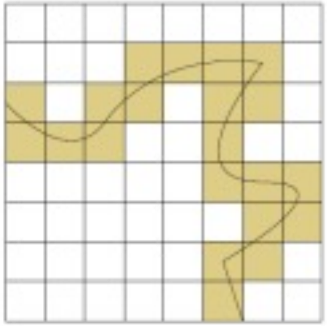
Define spatial elements in $\phi*\theta$ (voxel)

- Organize the voxels in DAGs according to track evolution in radial direction
- Flexible to model even arbitrary paths (kinks, missing hits, outliers, random walk, ..)
- Training is done with MC tracks of typically 15-25 events



Multiscale resolution

- Multiple networks for doublet finding (θ)
- Multiple networks for triplet finding (ϕ, θ)



Path finding

- Sort event hits into the trained networks
- Seed and follow the path strategy
- Which is the right direction to go?
- Classify doublet and triplet patterns of points

Pattern Recognition with Machine Learning

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Intuition

- Model free estimator
- Start with basic quantities
- Coordinates, simple derived values
- Only very basic detector specific information

Feature space

- Polar coordinates (R_t, ϕ, z)
- Directional cosines wrt. points
- Simple helix calculation (score)

} In principle not needed, but speeds the things up !

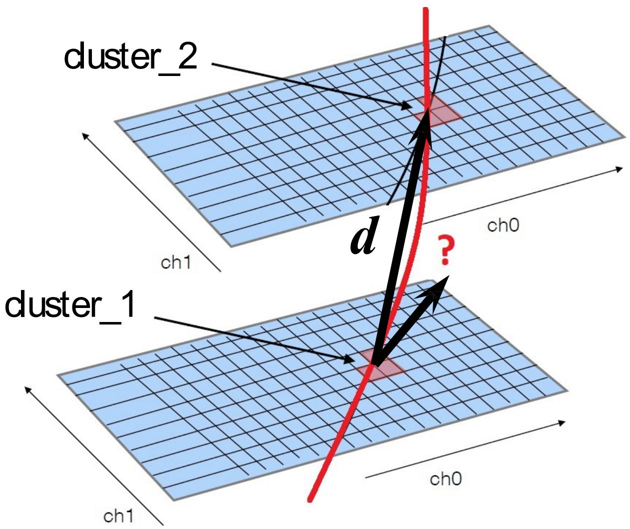
Training

- Supervised: presenting MC ground truth

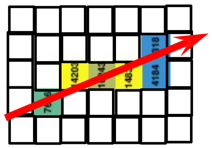
Feature Space

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

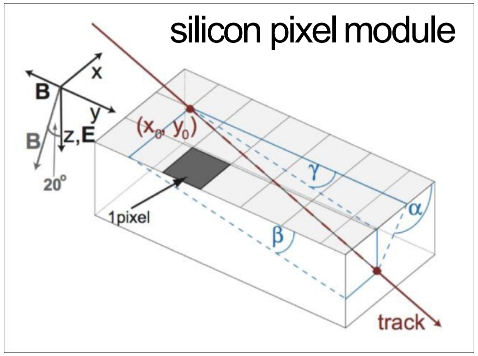
Given two hits (clusters of silicon cells): predict if they belong to the same track



- Estimate track direction from the cluster shape:



eigenvector of covariance matrix of the silicon cells



Features for the training

- Polar coordinates of the hit doublet: (r_1,ϕ_1,z_1), (r_2,ϕ_2,z_2)
- Triplet finder works the same with a hit triplet
- Simple helix score
- Angle/length deviations of the vector ***d*** projection from the values predicted by the shape of cluster 1
- Angle/length deviations of the vector ***d*** projection from the values predicted by the shape of cluster 2

Feature Folding

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

The tracking problem is symmetric wrt. polar coordinates

- Fold the input parameter space into an octagon slice using “abs” function
- Considerable improvement of the separation strength of the parameters
- Need less statistics / yield better results

Rank	Variable	Separation
1	log(score)	5.039e-01
2	rz3	5.491e-04
3	phi3	7.552e-05
4	z3	4.986e-05
5	rz2	1.519e-05
6	rz1	9.568e-06
7	phi2	4.101e-06
8	z1	1.967e-06
9	z2	1.965e-06
10	phi1	1.503e-06

→

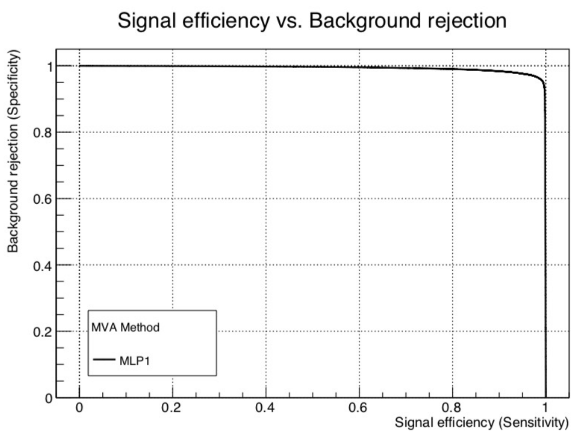
Rank	Variable	Separation
1	log(score)	5.978e-01
2	rz3	6.329e-04
3	abs(abs(phi3)-1.57079632679)	1.317e-04
4	abs(z3)	5.522e-05
5	rz2	2.067e-05
6	rz1	1.675e-05
7	abs(abs(phi2)-1.57079632679)	4.335e-06
8	abs(z1)	3.592e-06
9	abs(abs(phi1)-1.57079632679)	3.038e-06
10	abs(z2)	2.963e-06

Hit Doublet / Triplet Classification: MLP

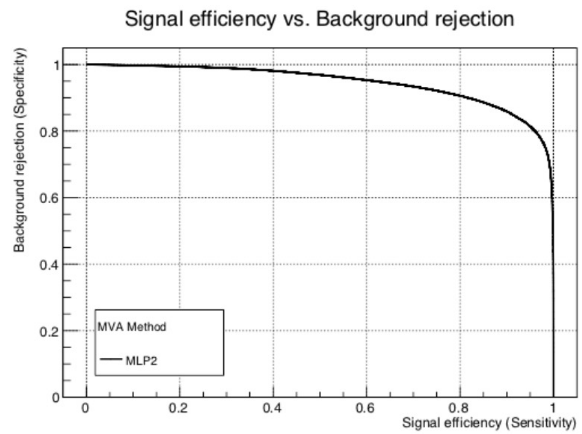
“Shallow learning” ;)

- **Classify the doublets and triplets with neural networks**
  - Multi Layer Perceptron: MLP1 8-15-5-1 / MLP2 9-15-5-1 / MLP3 10-15-5-1
  - Input: hit coordinates, directional cosines towards the clusters, helicity score wrt. origin
  - Output: doublet/triplet quality, supervised training with Monte-Carlo ground truth
  - Training: Typically 10 events, O(Mio) patterns, 500 epochs, one hour on standard PC
  - “*Receiver Operation Characteristics*” (ROC) curves indicate good quality

Doublet finder (disc)

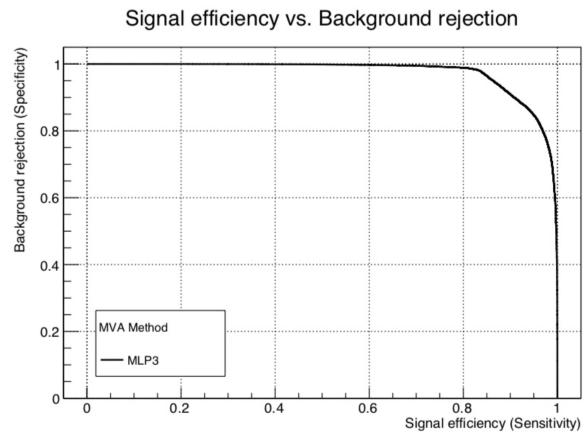


Doublet finder (tube)



Worse due to vertex shift !

Triplet finder

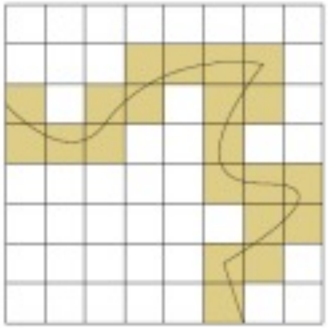


Machine Learning Advantage

Model free estimator (Basic features)
- Solution may be easily transferred to a different context

Graceful degradation in presence of changes
- Geometry
- Dead channels
- Calibration
- …

The graphs may represent arbitrary tracking paths
- Inhomogeneous magnetic field
- Kinks
- …



Takeaway Messages

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Graph Networks

- Very natural data representation for a lot of scientific problems.
- Promising performance in pattern recognition tasks.

Work has been published:
Amrouche, S. et al. *The Tracking Machine Learning challenge: Throughput phase* (2021). 2105.01160.

Dr. Marcel Kunze
marcel.kunze@uni-heidelberg.de
Im Neuenheimer Feld 293 / 106
D-69120 Heidelberg

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386