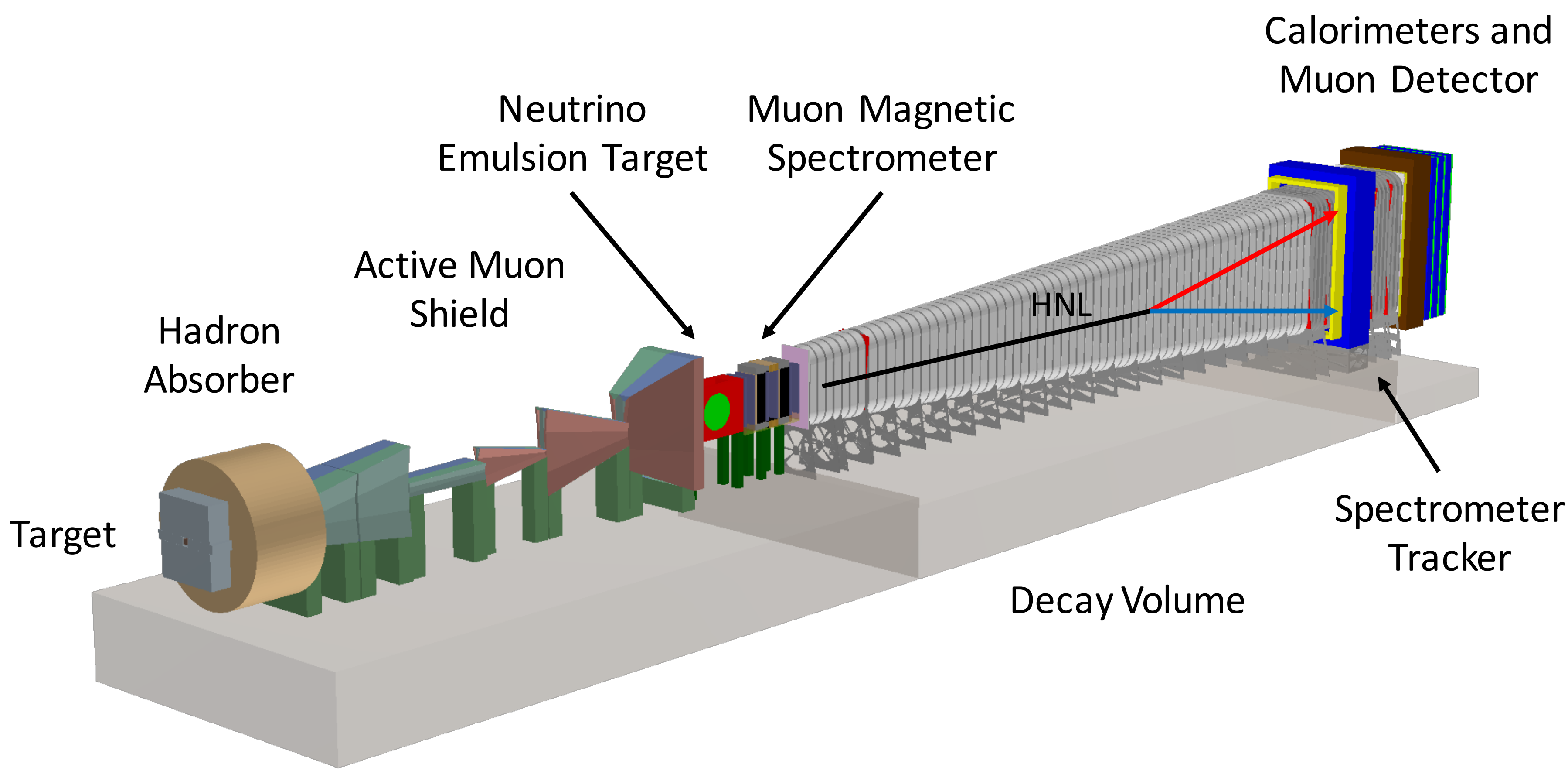
ML Approaches in Track Pattern Recognition for the SHiP Spectrometer Tracker

Mikhail Hushchyn, Andrey Ustyuzhanin, Eric Van Herwijnen

National Research University Higher School of Economics
Moscow Institute of Physics and Technology
Yandex School of Data Analysis
SHiP Collaboration

CHEP2018, Sofia, Bulgaria
9-14 July, 2018

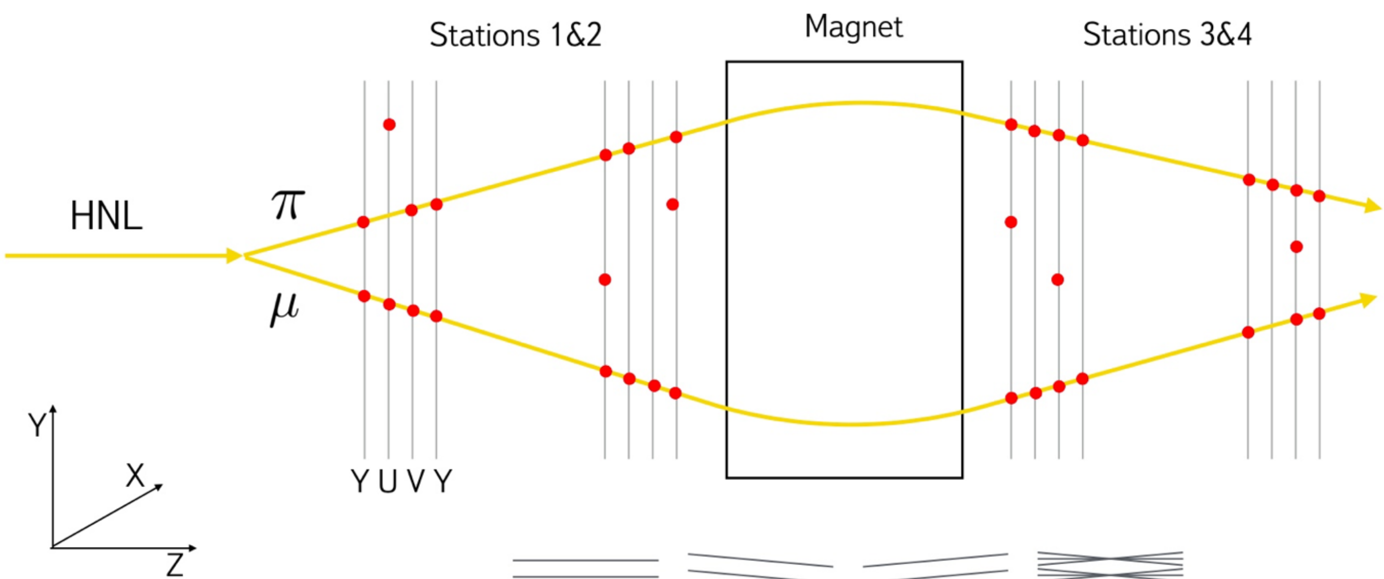
SHiP Experiment

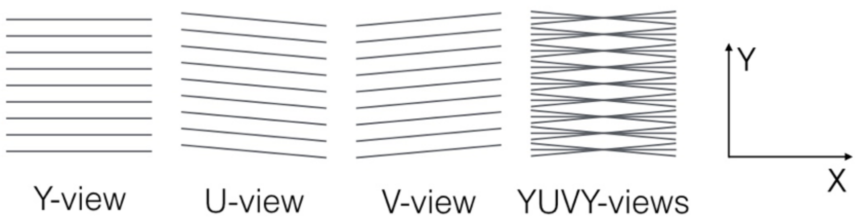


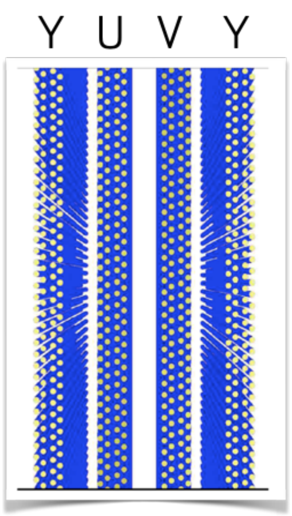
The SHiP Experiment is a new general purpose fixed target facility proposed at the CERN SPS accelerator which is aimed at exploring the domain of hidden particles (Heavy Neutral Leptons (HNL)) and make measurements with tau neutrinos.

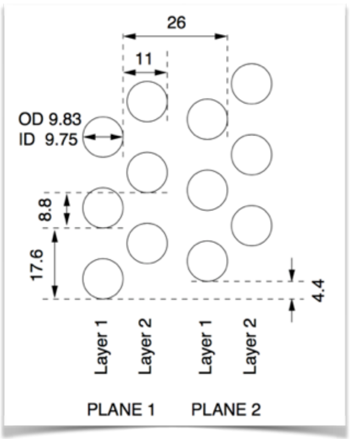
SHiP Spectrometer Tracker

SHiP Spectrometer Tracker recognizes tracks of a Heavy Neutral Lepton (HNL) decay products.









- 2 stations before and after the magnet
- 4 views in each station: 2 Y-views and 2 Stereo-views
- 4 straw tube layers in each view
- $\pm 5^o$ between Y and Stereo views

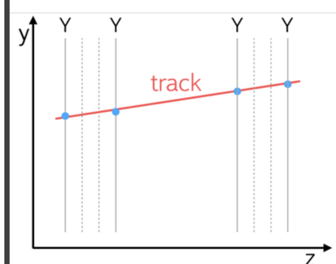
Pattern Recognition Overview

Several different algorithms were applied for SHiP Spectrometer Tracker:

- Hough Transform,
- Artificial Retina,
- Template Matching.

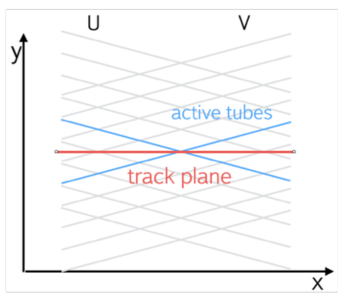
See backup slides for details

1



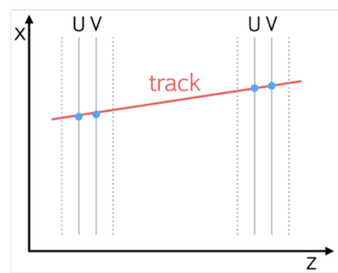
Look for tracks in the 16 layers with horizontal straw tubes (Y-views, y-z plane).

2



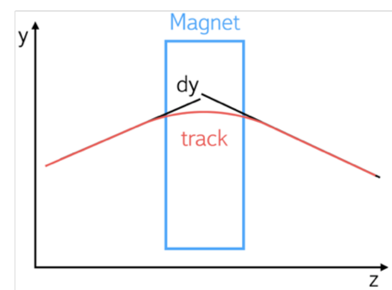
For each track found in the Y-view, intersect the plane defined by this track and the x-axis with the stereo (U,V) hits. This gives (z,x) coordinates.

3



Look for tracks in 16 layers of stereo-views in x-z plane using (z, x) coordinates of the intersections.

4



Combine tracks before and after the magnet. Reconstruct a particle momentum and charge.

Machine Learning for Tracking

ML application in tracking:
- Hits classification to reduce fake/noise hits
- Track seeds classification to reduce bad seeds
- Tracks classification to reduce ghosts
- Replacing Kalman filter with RNNs and CNNs
- Hits and links classification in graph based methods

There are two large research collaborations:
- TrackML
- HEP.TrkX

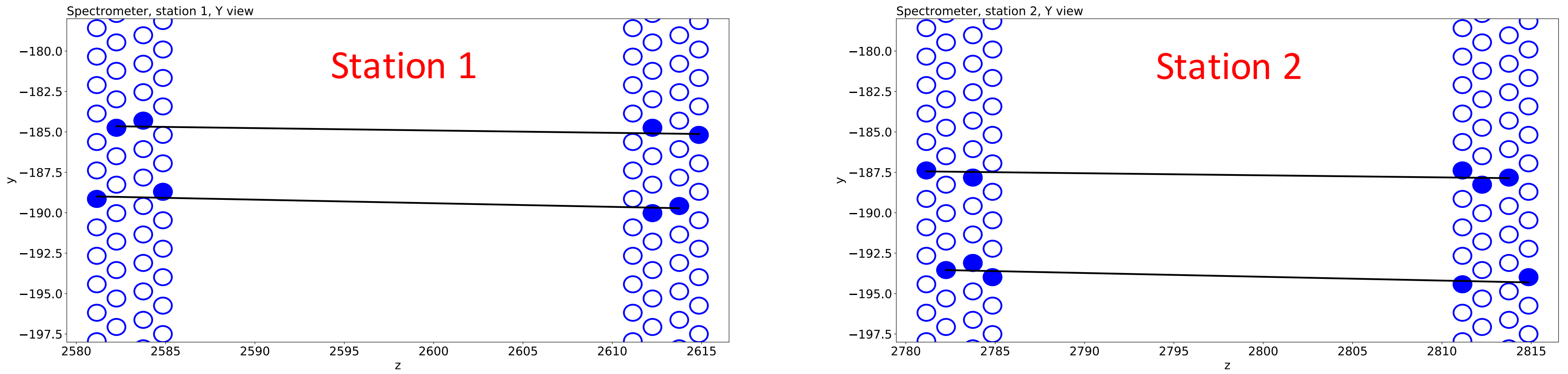
These collaborations develop new ML approaches for track pattern recognition. See their talks on CHEP, ACAT, CTD and IML conferences. They are very cool ☺

In SHiP we also implement ML approaches for tracking.

Problem Statement

The goal of this talk is to recognize 2D tracks in Y-views of stations 1&2 of the SHiP Spectrometer.

An event:

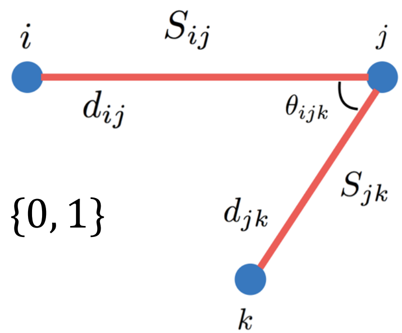


- Station width is 40 cm
- Distance between the stations is 200 cm
- 16 straw tube layers in total
- Straw tube diameter is 1 cm

Inspiration

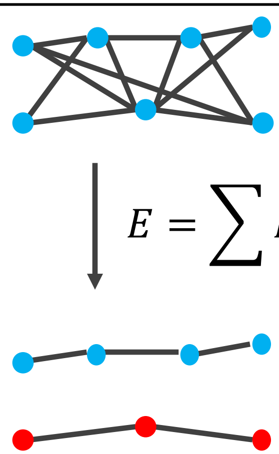
The presented solution was inspired by Denby-Peterson and Cellular Automaton methods of track recognition. They are based on hits relations in triplets.

Denby - Peterson



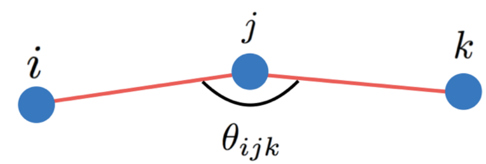
$$S_{ij} = \{0, 1\}$$

$$E_{ijk} = -\frac{-\cos^m(\theta_{ijk})}{d_{ij}+d_{jk}} S_{ij} S_{jk}$$

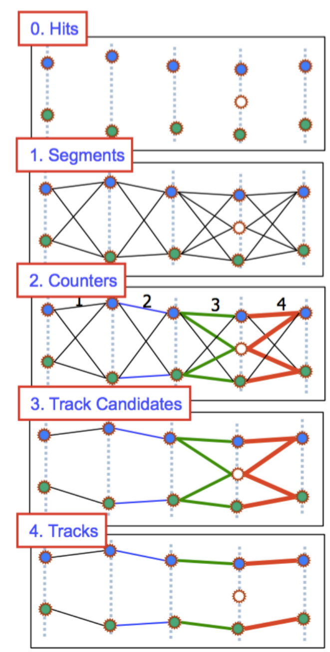


$$E = \sum E_{ijk} \rightarrow min$$

Cellular Automaton



$$S_{jk} = max\{S_{ij} | \theta_{ijk} > \theta^{min}\} + 1$$



Idea

Idea of the presented track pattern recognition approach:

- Take all hits of an event
- Generate hit triplets
- Classify the triplets on good and bad using ML
- Use the classifier to select good triplets
- Combine good triplets into tracks

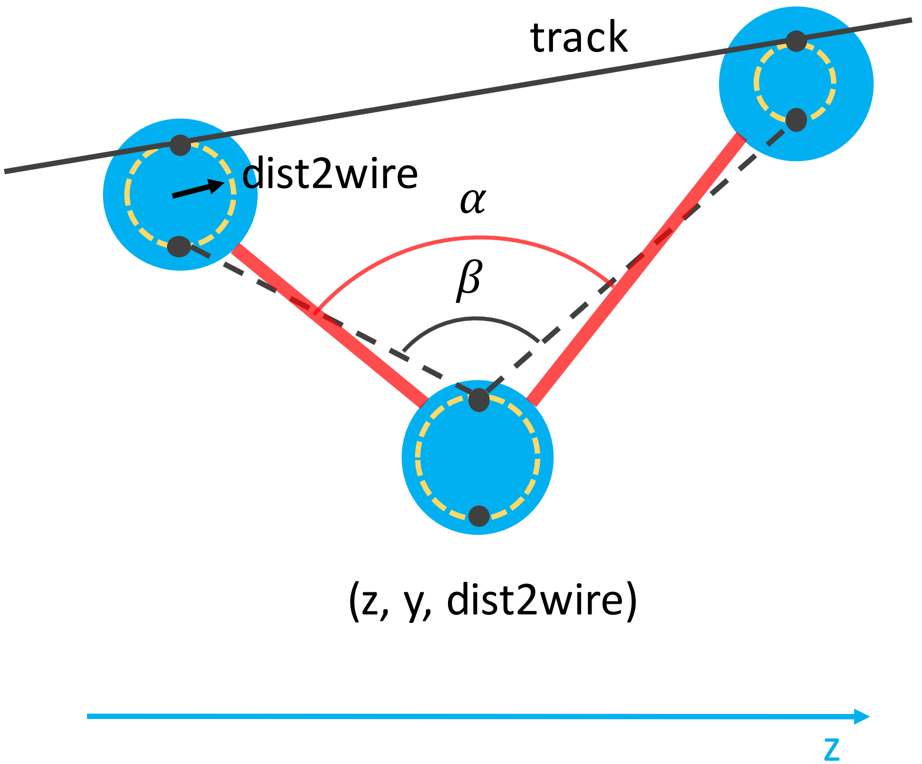
Good triplets: triplets of hits of the same track
Bad triplets: triplets of hits from different tracks

Triplets Generation & Description

We have the small number of tracks, so we can generate all possible triplets.

A triplet is described by the set of features:

- (z, y) coordinates of straw tube centers
- Dist2wire – distance from a track to a straw tube center
- Cos α (baseline)
- Cos β
- Slopes of the segments

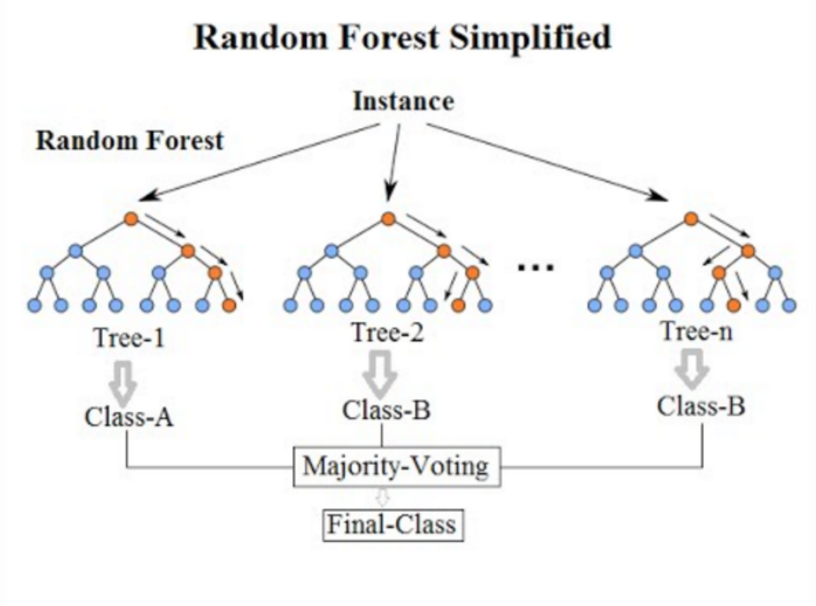


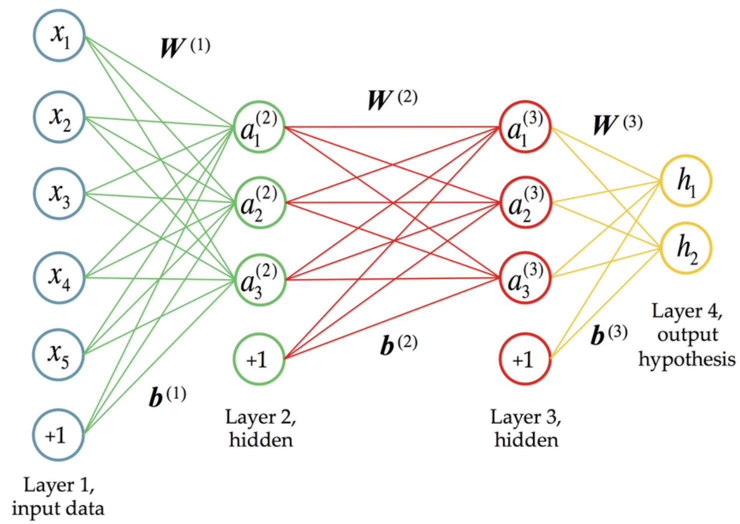
Additional conditions:

- Hits in a triplet are ordered by z coordinate of straw tube centers.
- Only one hit per straw tube layer.

Triplets Classification #1

Signal: triplets of hits of the same track
Background: triplets of hits from different tracks
Methods: Random Forest (RF), One-layer NN, Deep NN



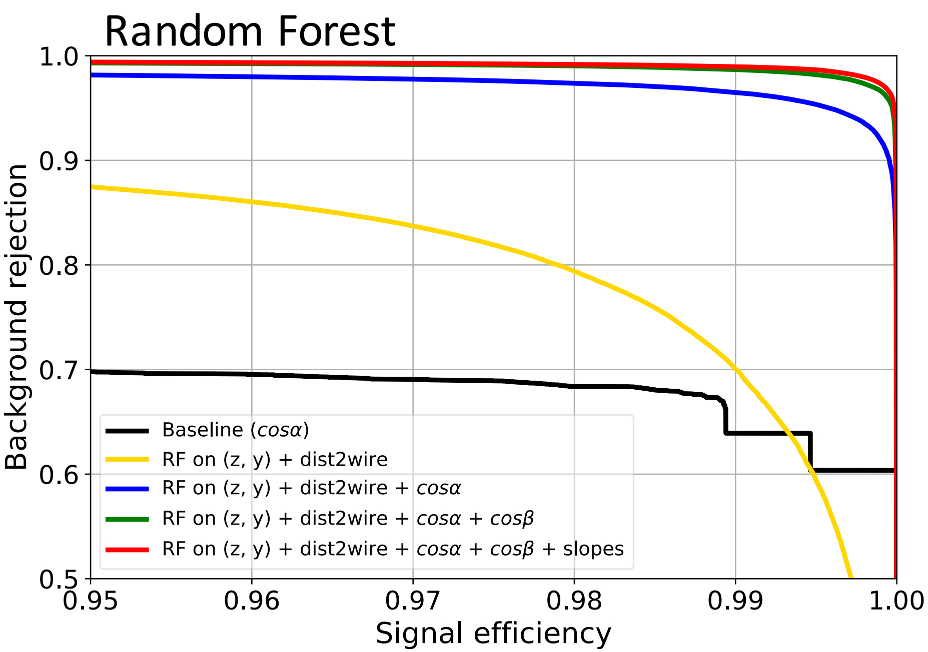


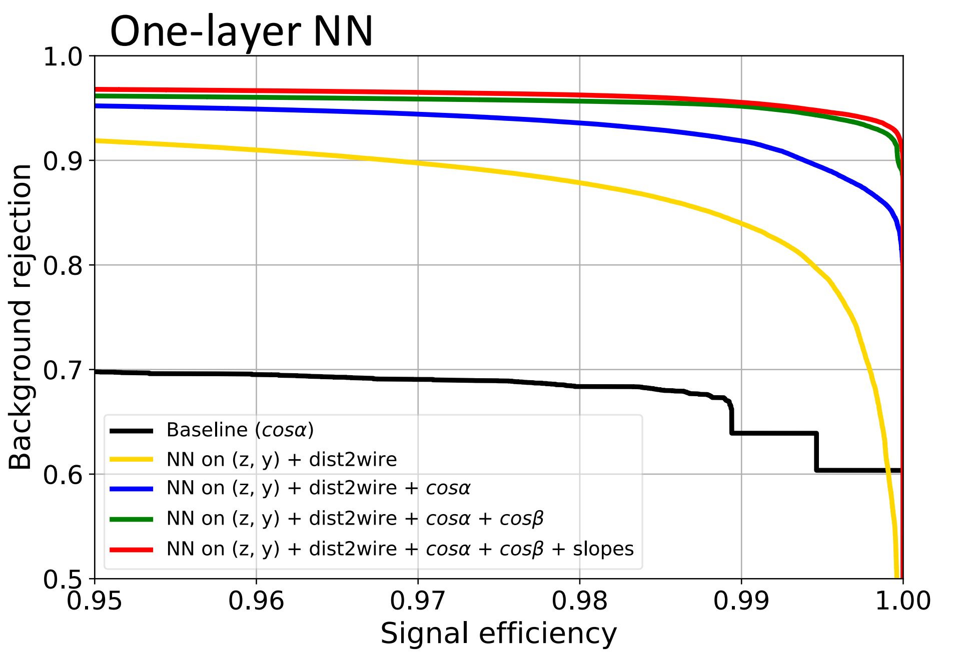
Motivation:

- RF is more robust to outliers produced by particle scattering
- NN is able to find strong features => better classification quality

Triplets Classification #2

Signal: triplets of hits of the same track
Background: triplets of hits from different tracks
Methods: Random Forest (RF), One-layer NN, Deep NN

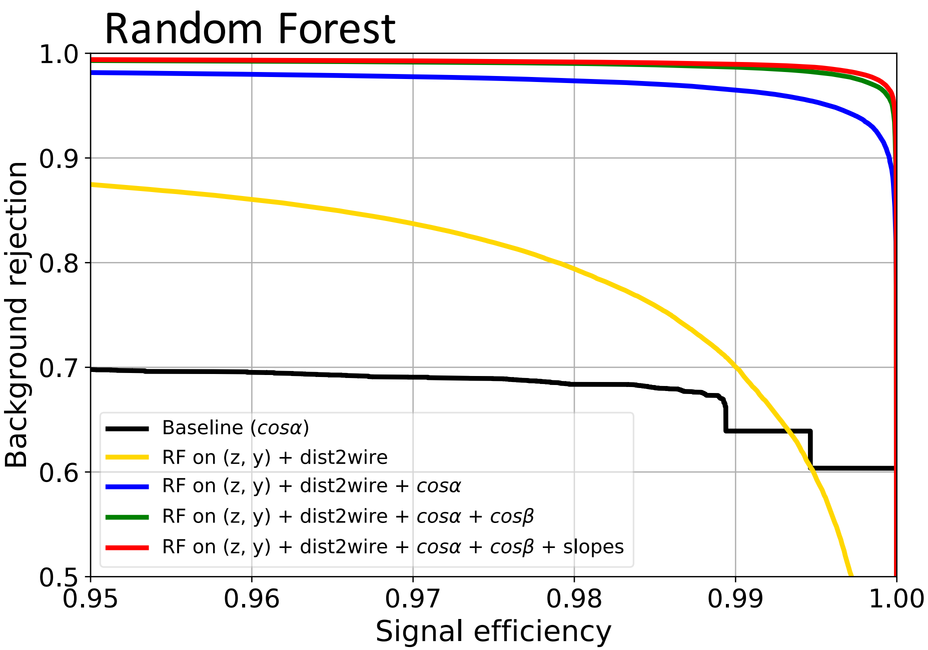


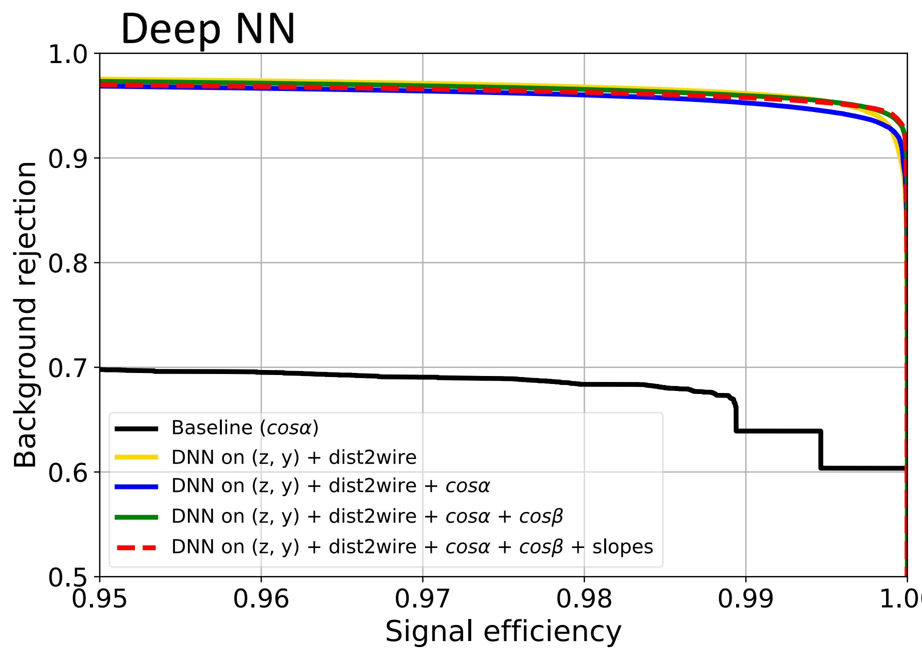


- Baseline (cos α) is used in Cellular Automaton
- NN is better on (z, y) + dist2wire
- RF is better with additional strong features

Triplets Classification #3

Signal: triplets of hits of the same track
Background: triplets of hits from different tracks
Methods: Random Forest (RF), One-layer NN, Deep NN

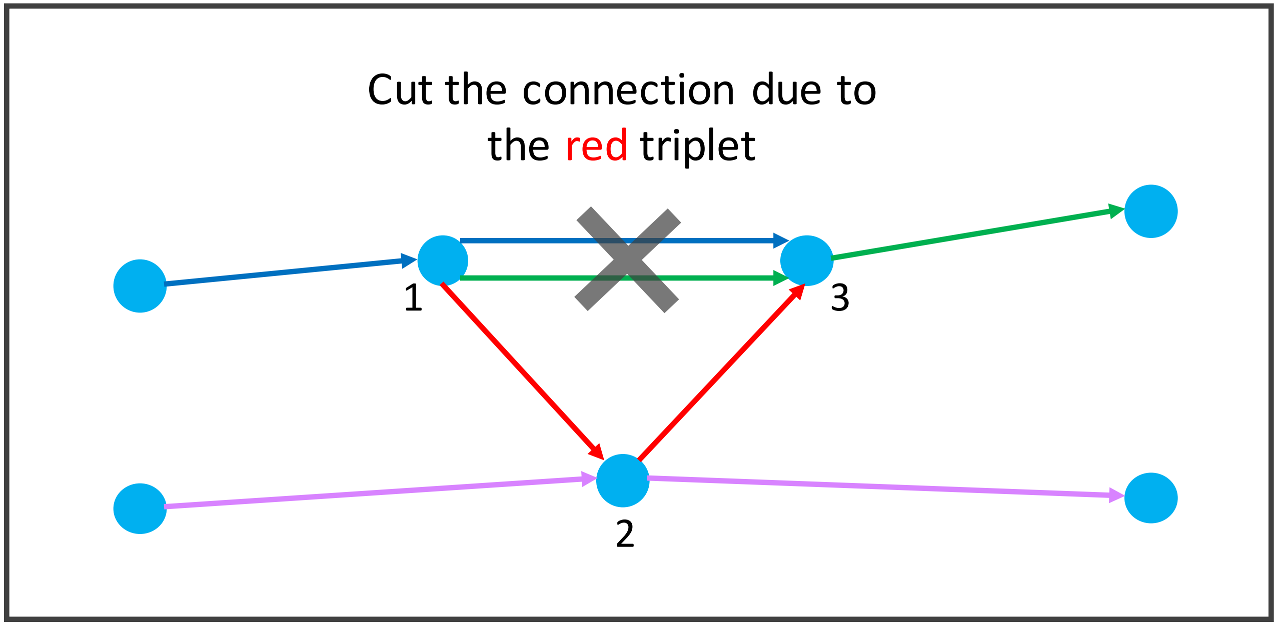




- Deep NN reconstructs strong features, so they do not help much
- RF is still better with additional strong features

Graph Creation

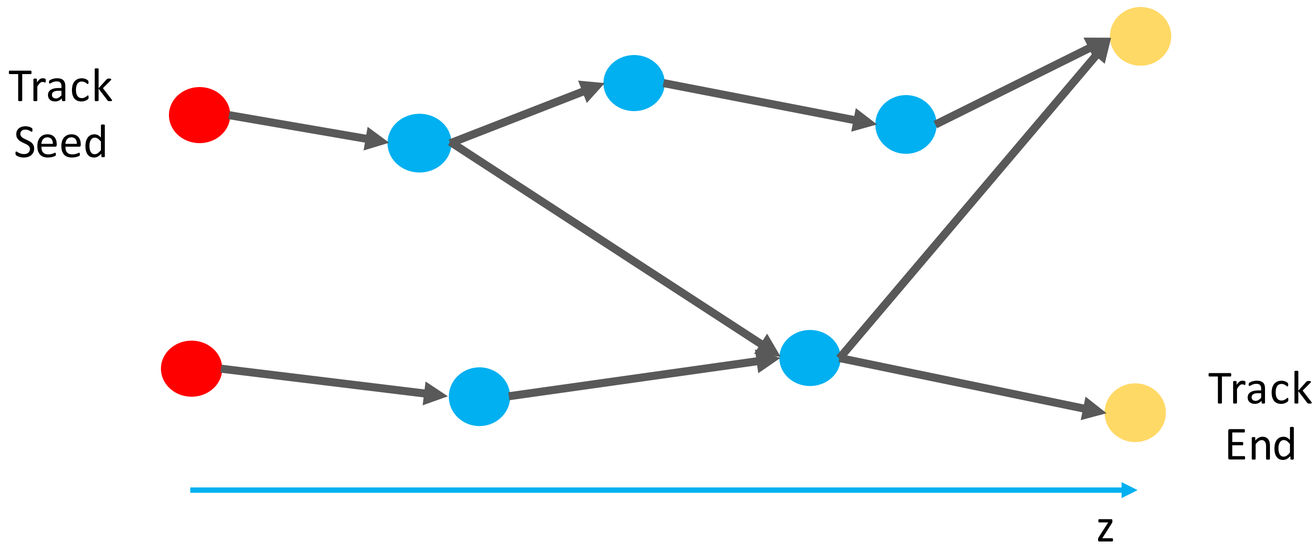
- Select good triplets based on the classifier output
- Create directional graph using the selected triplets:
    - Connect hits in the triplets
    - If there is 1 -> 2 -> 3 triplet, remove 1 -> 3 connection



Motivation:

- Reduces the number of possible paths => less combinatorics, clones and ghosts
- Creates long paths => track with the large number of hits

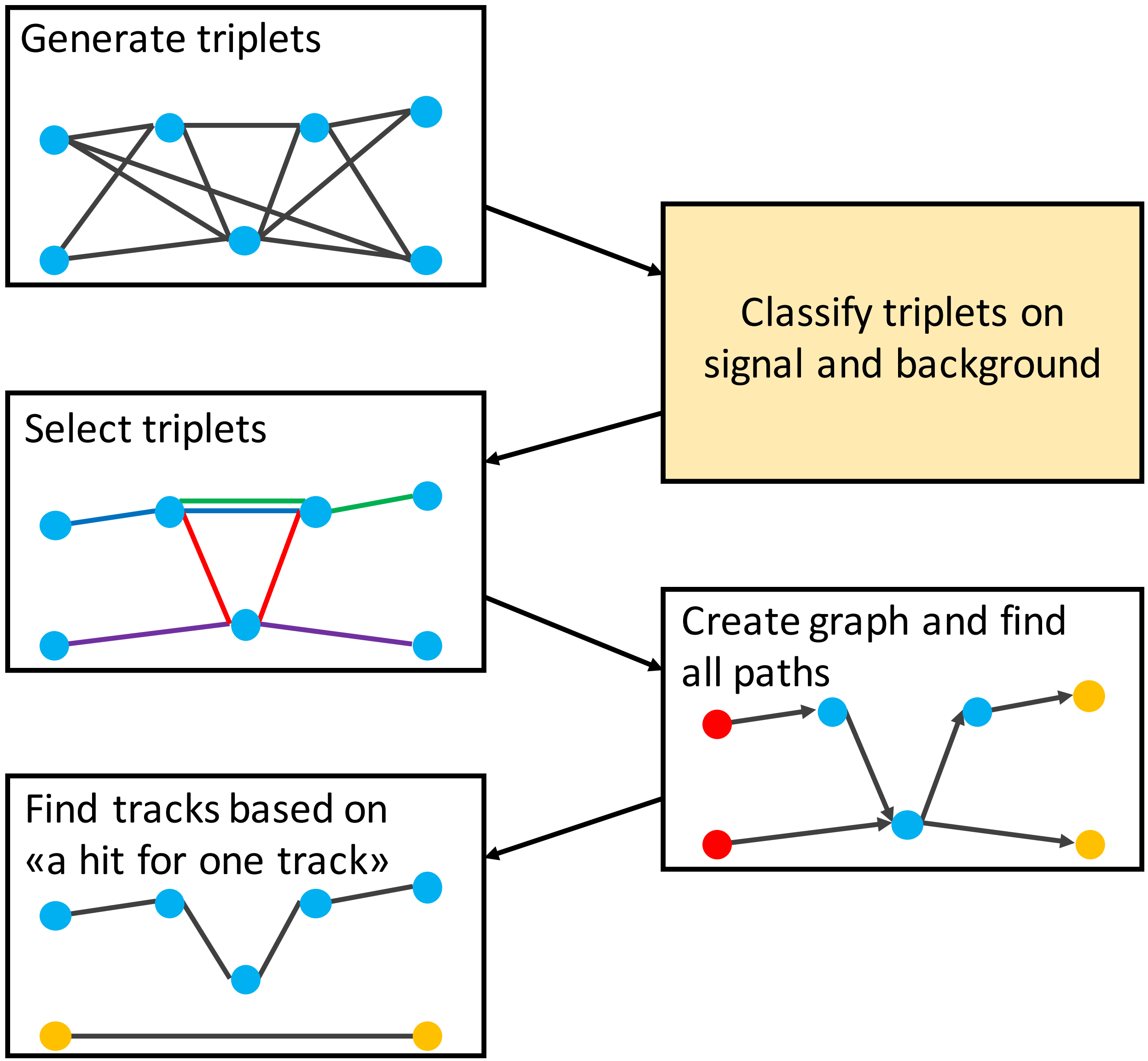
Graph Properties



The graph structure represents the main physical properties of tracks:

- Directional graph: a particle flies in z direction
- One path – one track candidate
- One hit per straw tube layer in a path
- The small number of track candidates
- Long paths: the large number of hits in track candidates

Track Pattern Recognition Overview



Results

With clones and ghosts reduction (one hit is just for one track)

Method	ROC AUC	Recognition Efficiency	Ghost Rate	Clone Rate
RF on (z, y) + dist2wire	0.9580	0. 946	0.040	0.017
RF on (z, y) + dist2wire + $cos\alpha$	0.9937	0. 990	0.026	0.015
RF on (z, y) + dist2wire + $cos\alpha + cos\beta$	0.9977	0.995	0.020	0.016
RF on (z, y) + dist2wire + $cos\alpha + cos\beta$+ slopes	0.9981	0.995	0.018	0.014

The better classifier:

- the higher recognition efficiency
- the lower ghost rate
- the lower clone rate

Conclusion

- New track pattern recognition method for SHiP
- It is triplets-based method
- Tested on SHiP MC

- The best triplets classifier has ROC AUC = 0.9981
- Selected triplets are combined using a graph
- The graph represents main physical properties of tracks

- 99.5% of tracks are recognized
- Ghost and clone rates < 2%
- Ideal triplets classifier => ideal tracks recognition

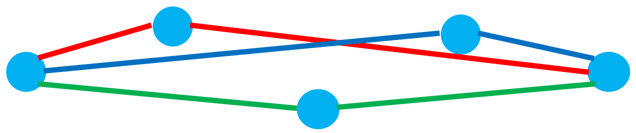
Backup slides #1

Triplets Combining

Good triplets are combined into tracks. There are several possible ways to do this:

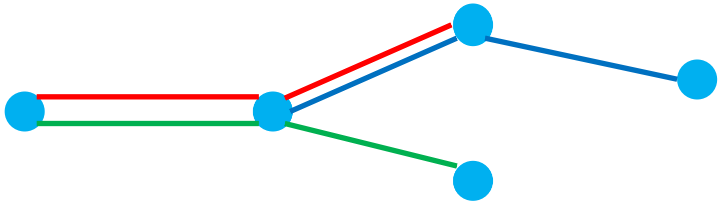
Combine all triplets with the same two hits:

- Other triplets are not taken into account
- A lot of track candidates => a lot of clones



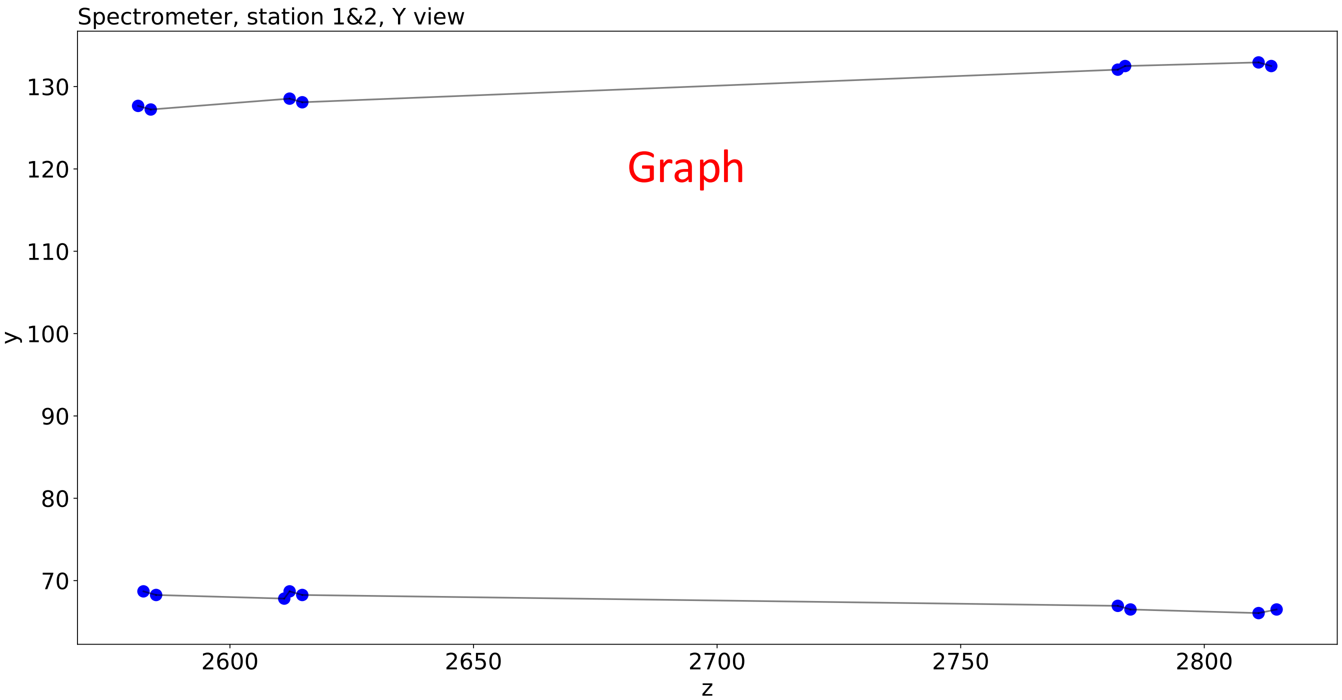
Merge triplets with shared any two hits:

- All triplets can be merged into one cluster
- Sensitive to bad triplets
- A lot of track candidates

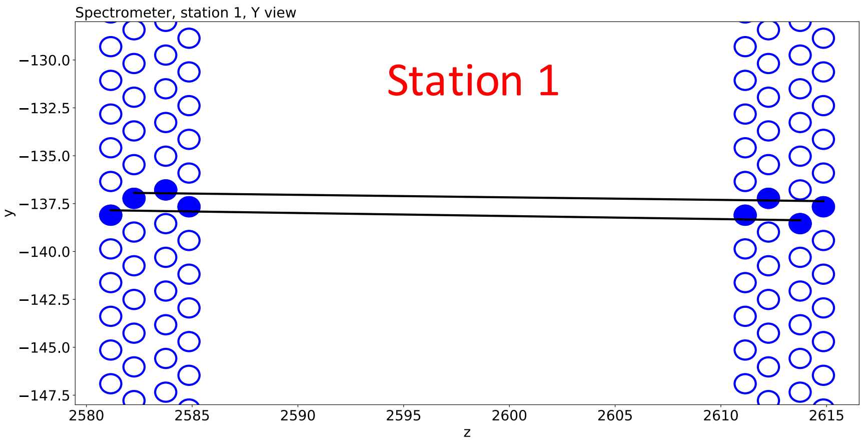


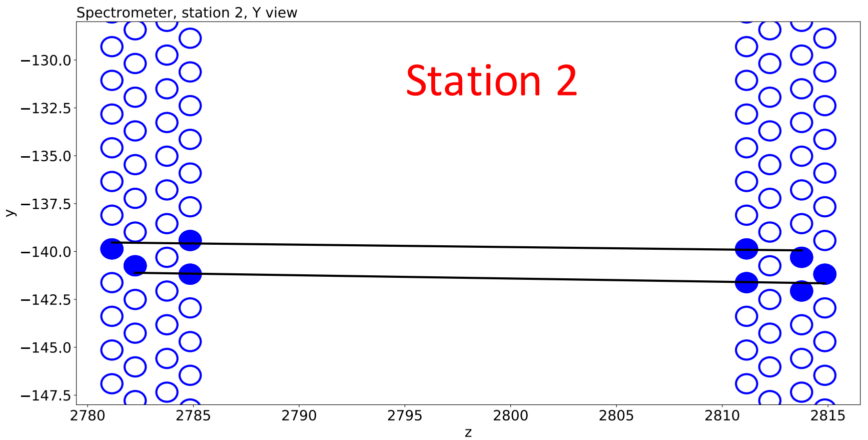
We need to combine triplets in a way that takes into account all triplets and which is robust to the background triplets.

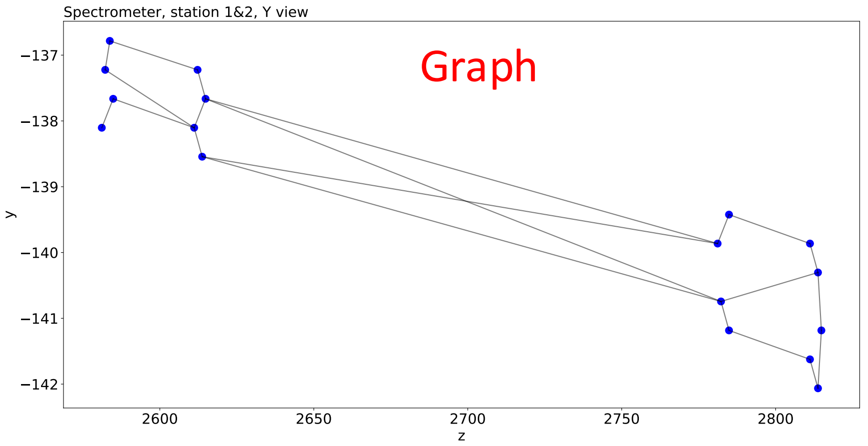
Simple Event



Difficult Event







Results

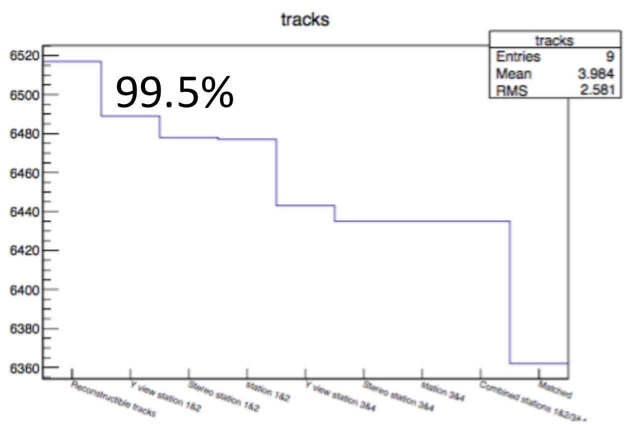
Without clones and ghosts reduction (one hit is for several tracks)

Method	ROC AUC	Recognition Efficiency	Ghost Rate	Clone Rate
RF on (z, y) + dist2wire	0.9580	0. 973	5.771	4.766
RF on (z, y) + dist2wire + $cos\alpha$	0.9937	0. 997	0.497	1.125
RF on (z, y) + dist2wire + $cos\alpha + cos\beta$	0.9977	0.998	0.467	0.541
RF on (z, y) + dist2wire + $cos\alpha + cos\beta$+ slopes	0.9981	0.998	0.404	0.398

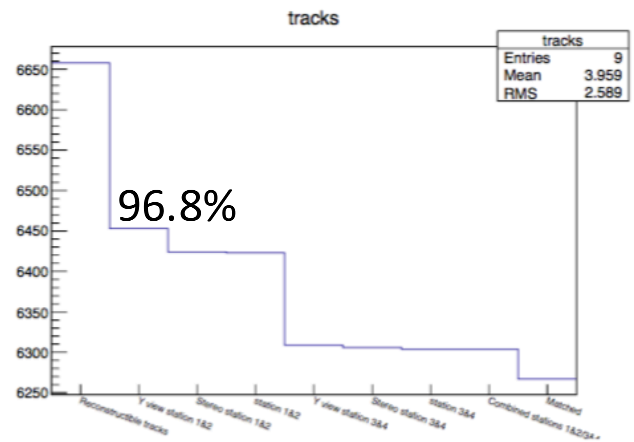
The better classifier:

- the higher recognition efficiency
- the lower ghost rate
- the lower clone rate

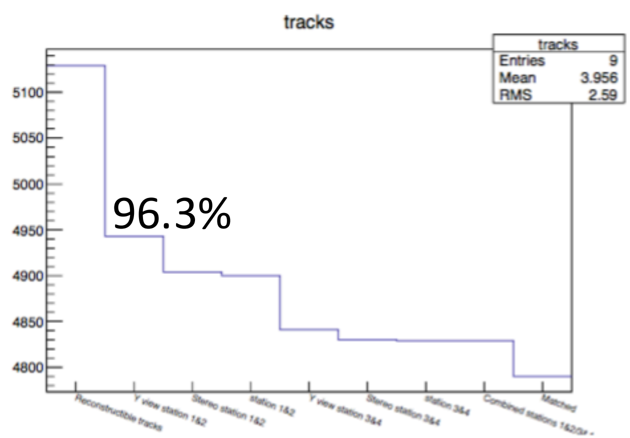
Baseline in SHiP



(a) Events with 1 track



(b) Events with 2 tracks



(c) Events with 3 prong

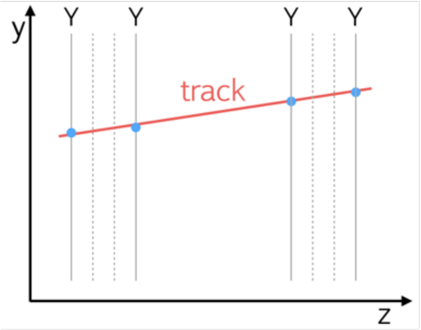
Figure 19: Pattern recognition efficiency per event

[1] H. Dijkstra, M. Ferro-Luzzi, E. van Herwijnen and T. Ruf Simulation and pattern recognition for the SHiP Spectrometer Tracker, SHiP-PUB-2015-002, March 31, 2015, CERN-SHiP-NOTE-2015-002

Backup slides #2

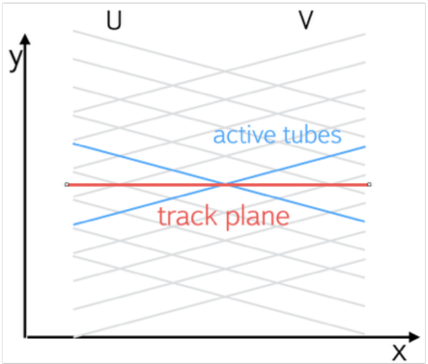
Pattern Recognition Overview

1



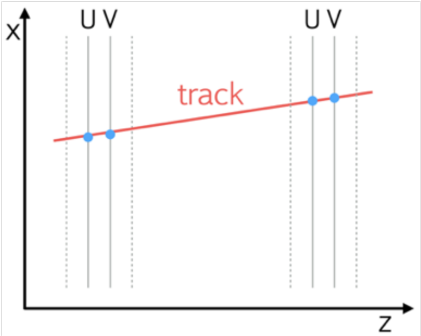
Look for tracks in the 16 layers with horizontal straw tubes (Y-views, y-z plane).

2



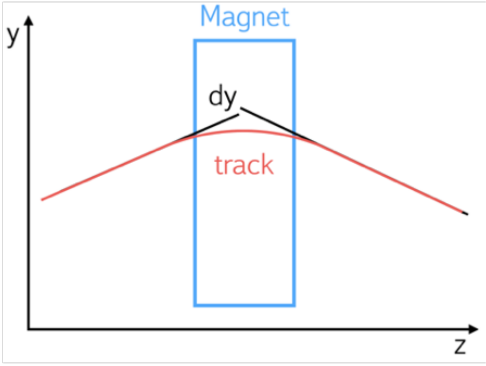
For each track found in the Y-view, intersect the plane defined by this track and the x-axis with the stereo (U,V) hits. This gives (z,x) coordinates.

3



Look for tracks in 16 layers of stereo-views in x-z plane using (z, x) coordinates of the intersections.

4

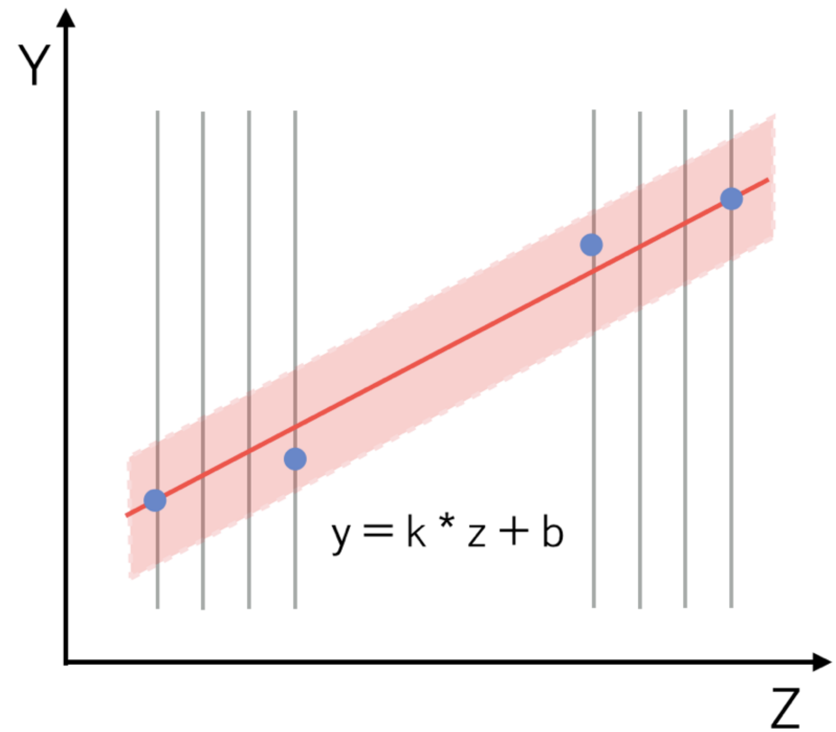


Combine tracks before and after the magnet. Reconstruct a particle momentum and charge.

Baseline [1]

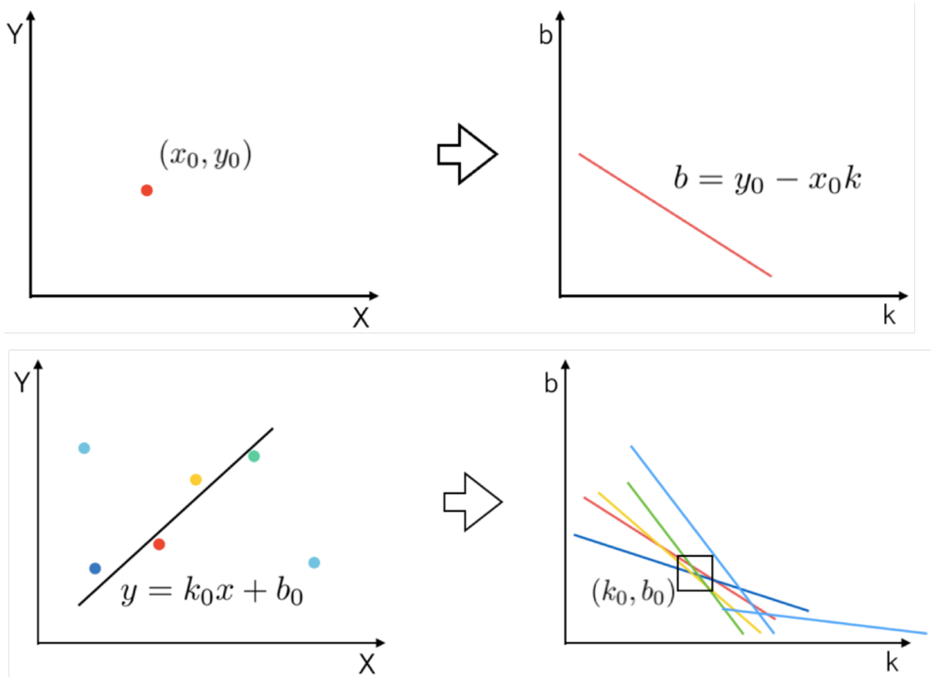
Track PR procedure in a plane:

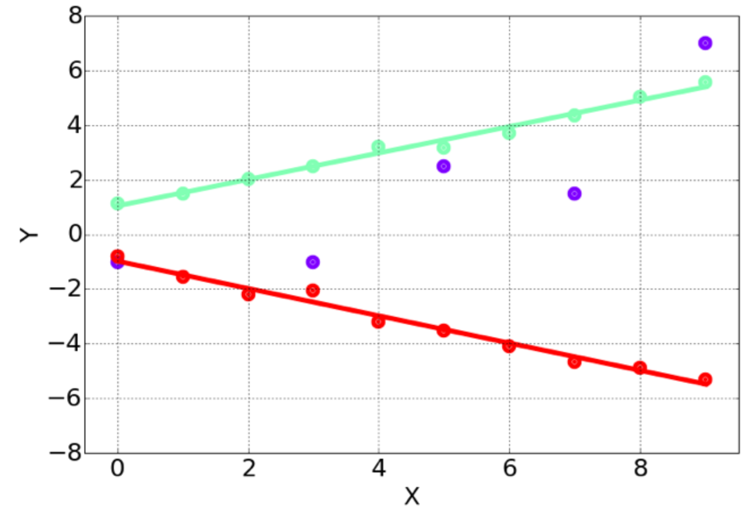
1. Estimate line parameters (k, b) for two hits.
2. Find track hits within a window width from the line.
3. Repeat steps 1 and 2 for each pair of hits.
4. Select two tracks with the largest number of unique hits.

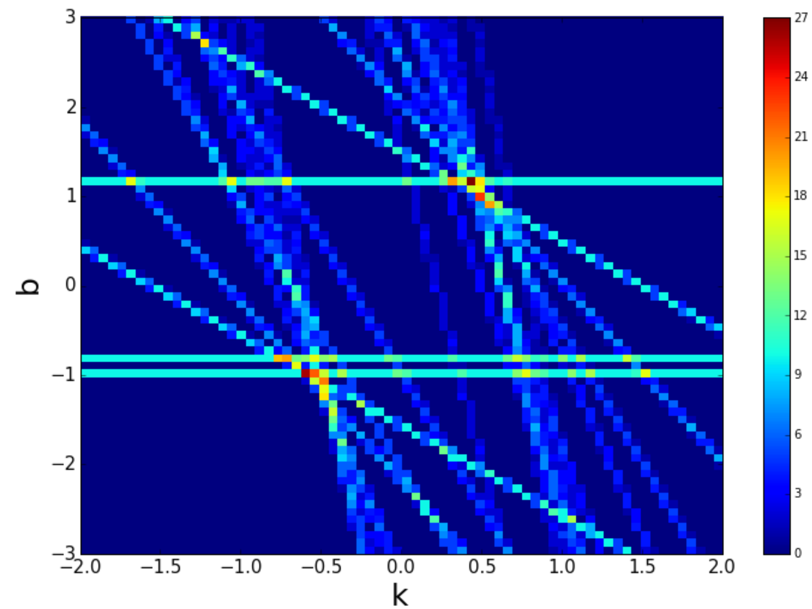


[1] H. Dijkstra, M. Ferro-Luzzi, E. van Herwijnen and T. Ruf Simulation and pattern recognition for the SHiP Spectrometer Tracker, SHiP-PUB-2015-002, March 31, 2015, CERN-SHiP-NOTE-2015-002

Hough Transform







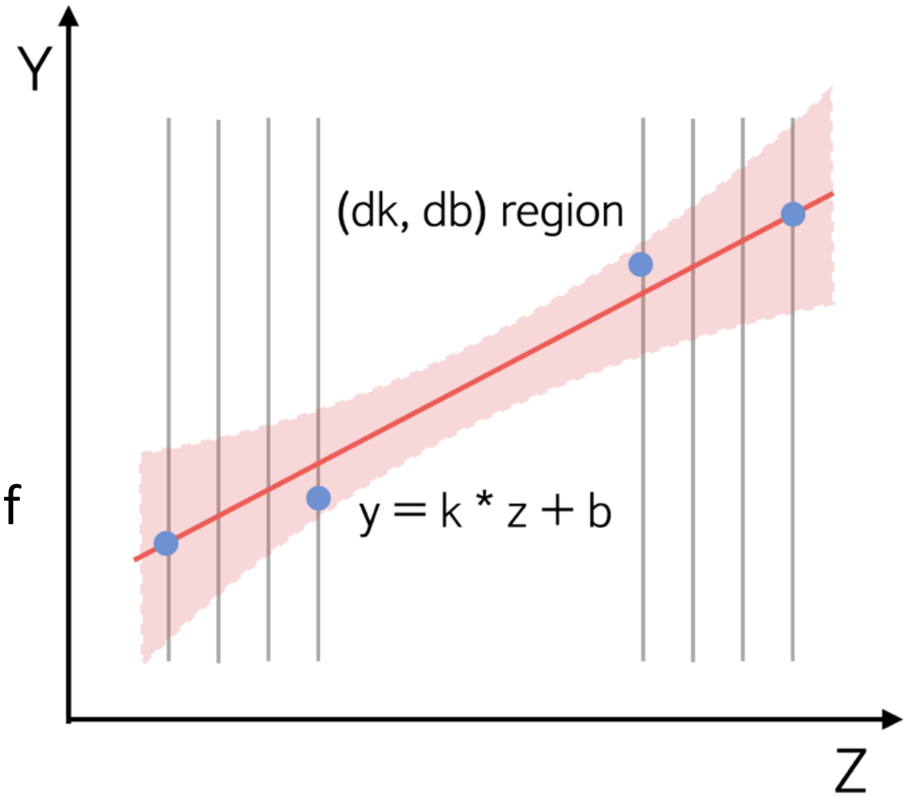
Histogramming technique:

- A lot of bins to process
- Number of bins depends on the number of tracks per event
- Not reasonable for the small number of tracks

Hough Transform

Track PR procedure in a plane:

1. Estimate line parameters (k, b) for two hits (a seed).
2. Use Hough Transform to find track hits within (dk, db) parameters region (one bin).
3. Repeat steps 1 and 2 for each pair of hits.
4. Select two tracks with the largest number of unique hits.



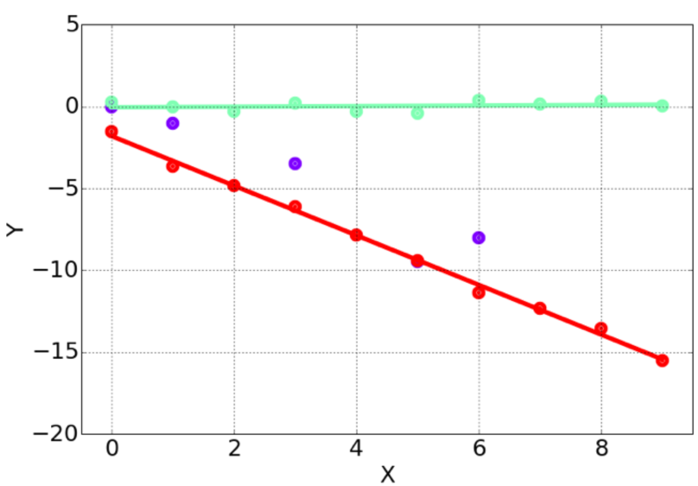
Artificial Retina

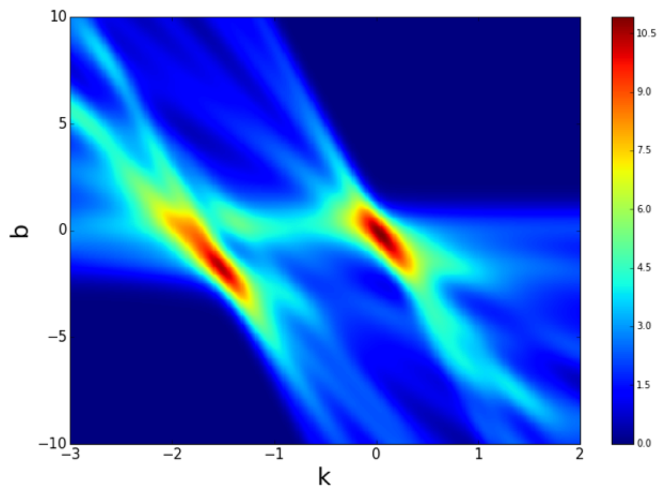
The artificial retina function is defined as:

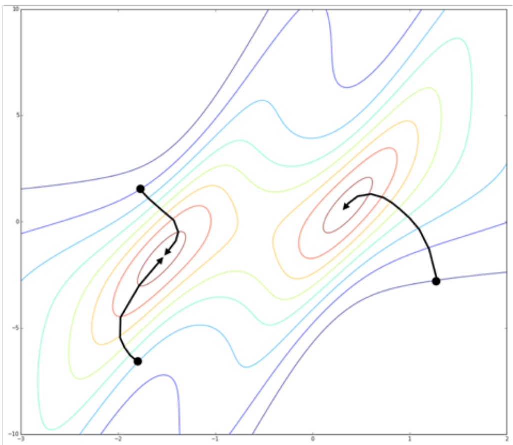
$$R(\theta) = \sum_i e^{-\frac{\rho^2(\theta, x_i)}{\sigma^2}}$$

where $\rho(\theta, x_i)$ is distance between the i-th hit and a track with parameters θ.

For 2D tracks: $\rho(\theta, x_i) = y_i - (kx_i + b)$ $\theta = [k, b]$



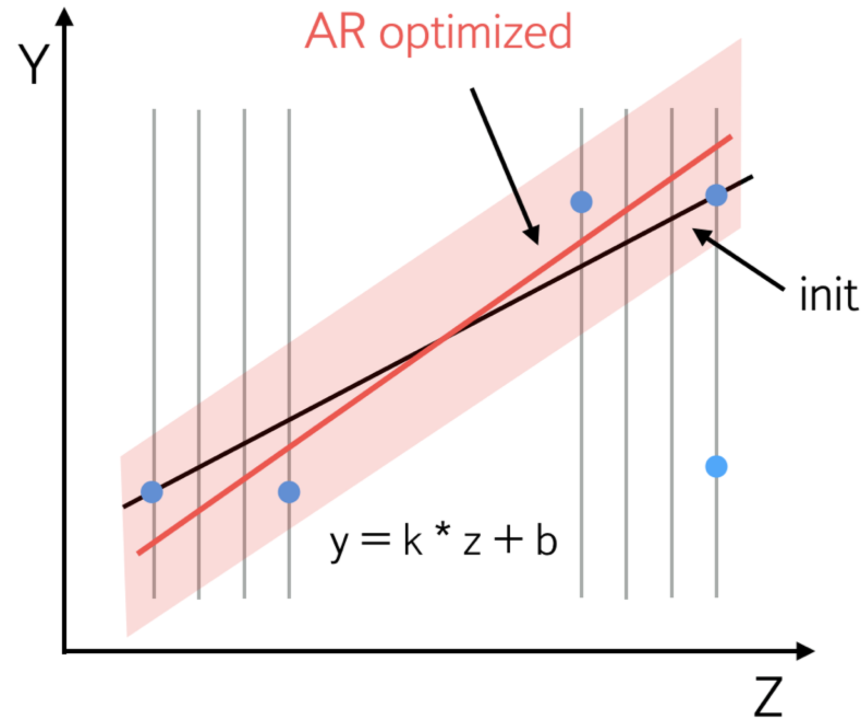




Artificial Retina

Track PR procedure in a plane:

1. For each pair of hits estimate line parameters (k, b) and calculate AR function value.
2. Use pair with the highest value as init point of AR function maximization procedure.
3. Fit track parameters by maximizing AR function value.
4. Select hits within a window width from the fitted track. Exclude the hits from the sample.
5. Repeat steps 1-4.



Results

Test sample: 30k events

Method	Events Passed PR, %	Track Efficiency*, %	Left-Right Amb. Res., %
Baseline	94.1	99.2	94.5
Hough Transform	98.7	99.3	94.9
Artificial Retina	99.3	99.8	98.3

* Track Efficiency is N_correct_reco_hits / N_reco_hits

- Tests only for PatRec scripts without any analysis steps. So, the results are biased.
- Artificial Retina is the best, but it worse scaled with number of tracks.