Near-Time Upgrade of the Neural Network Track Trigger for Belle II

Simon Hiesl

- 3DHough Finder (3DF) + Deep Neural Network Trigger (DNN)
- Displaced Vertex Trigger (DVT)

Members of the Belle II Trigger Group

KIT ITIV

- Marc Neu
- **Kai Unger**
- Jürgen Becker
- Timo Justinger
- Valdrin Dajaku
- Felix Mältzer
- Till Rädler

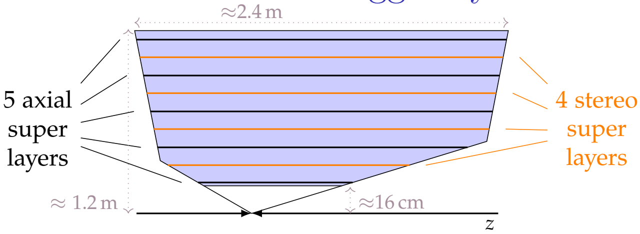
KIT ETP

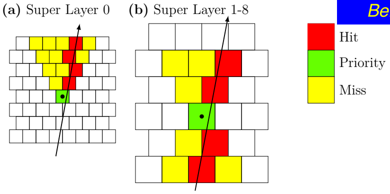
- **Lea Reuter**
- **Greta Heine**
- **Isabel Haide**
- **Torben Ferber**

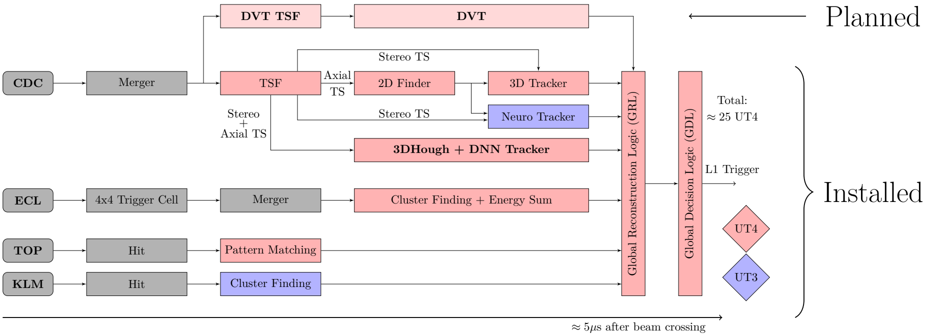
MPI

- **Timo Forsthofer**
- **Simon Hiesl**
- **Christian Kiesling**

The L1 CDC Track Trigger System

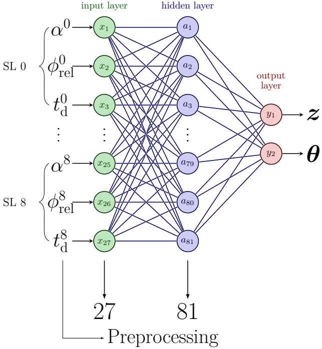


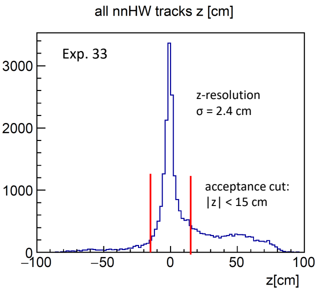


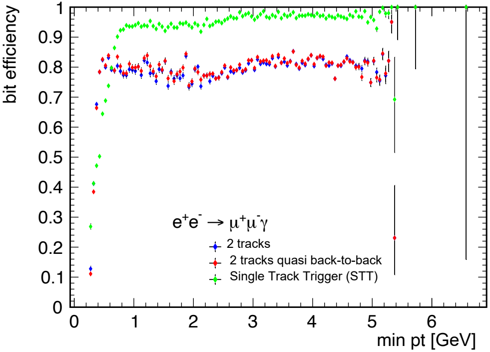


The Present Neural Network Track Trigger

2D track + Stereo TS $\Longrightarrow$ $z + \theta$ prediction







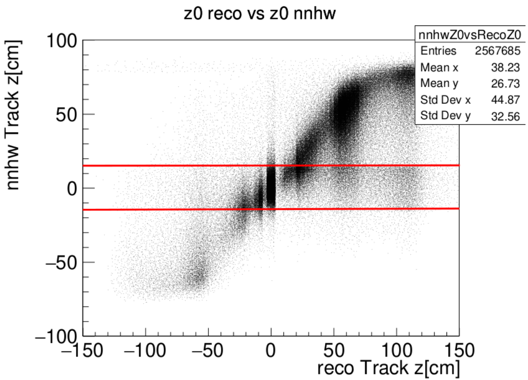
Network reoptimized with Exp. 26 data (NN24) running since 2024b

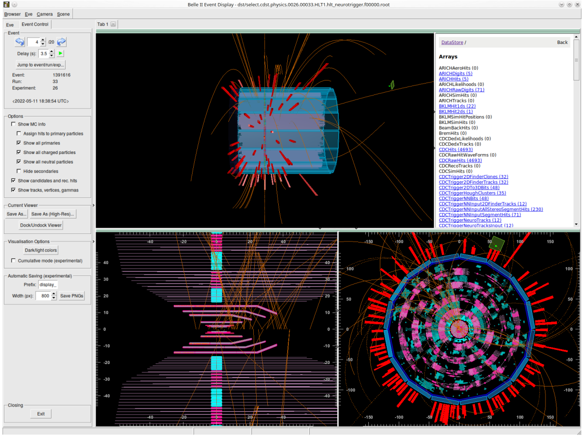
$\Longrightarrow$ improved z-resolution

Single-Track-Trigger:
$|z| < 15\,\text{cm}$ and $p > 0.7\,\text{GeV}$ (via θ)
(S. Bähr et al., arXiv:2402.14962, submitted to NIMA)

Problems with the L1 Neural Network Trigger

- High Background: Large number of fake 2D track candidates
  $\Longrightarrow$ Fake neuro tracks from combination with background stereo TS
- "Feed-Down" effect: Background tracks $\Longrightarrow$ Mapped into acceptance region



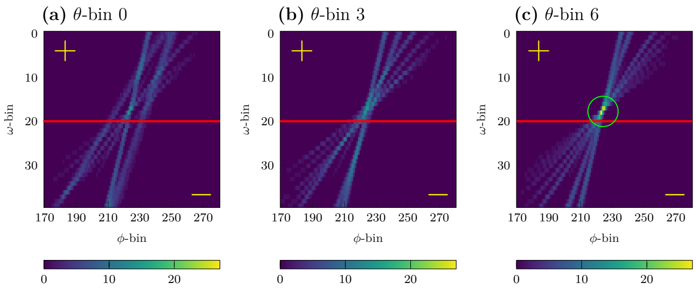


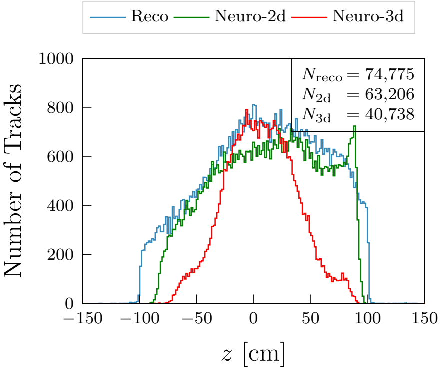
The 3DHough Finder

Third track parameter in addition to ω and ϕ: Polar scattering angle θ

Track vertex constraint: $(x, y, z) = (0, 0, 0)$

Particle Gun Single Tracks:





Positive (negative) tracks in upper (lower) ω half plane

$\Longrightarrow$ Intersection point yields ω, ϕ and θ (see green circle)

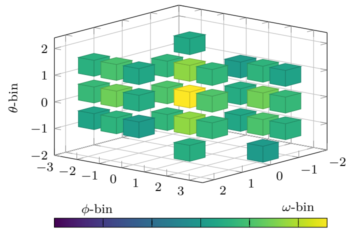
Natural suppression of tracks outside of IP region

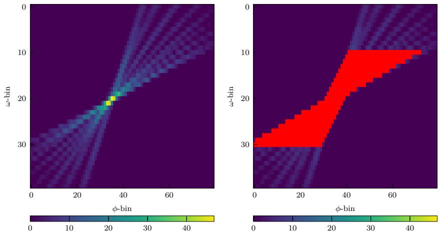
Clustering Algorithm in 3 Dimensions

$\Longrightarrow$ **Fixed Cluster Shape**

(motivated by independence of tracks' p_{T} and θ)

Average Cluster Shape





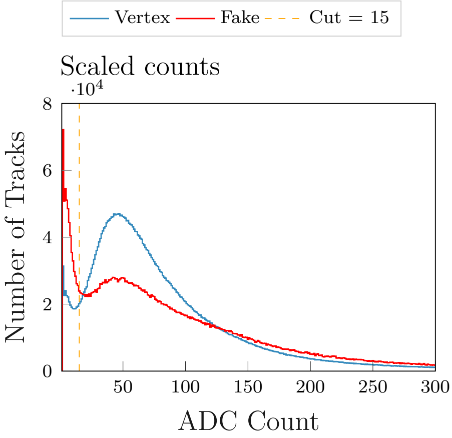
(a) Complete Cluster **(c)** Cutout

1. Search global maximum in 3D Hough space
2. Fixed cluster shape around maximum
3. Precise determination of ω, ϕ and θ by weighted average of cluster cells
4. Clear cells around global maximum ("Butterfly-Shape" cutout)
5. Search global maximum in remaining 3D Hough space $\rightarrow$ **2.** (≤ 2 track candidates per quadrant)

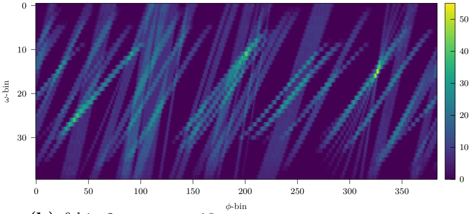
Implemented in basf2 (release 9)

Further Noise Reduction in TS

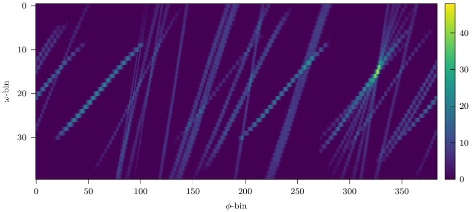
TSF on UT4: Suppress wires with low ADC count



(a) θ-bin 2: No `adccut`



(b) θ-bin 2: `adccut`=10

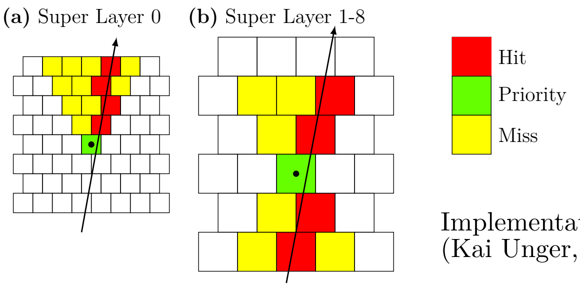


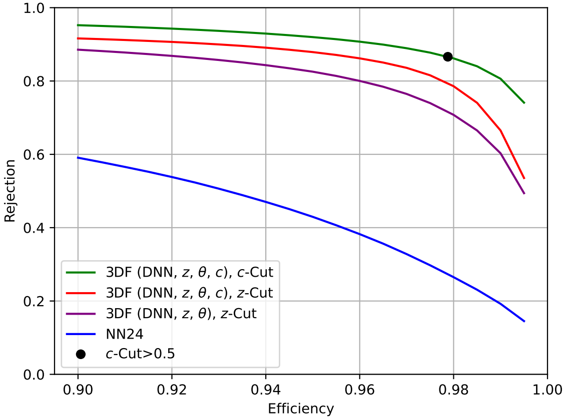
$\Longrightarrow$ Reduction of fake 3D track candidates

Extended Network Architecture

1. 3DFinder track candidates
2. DNN: 4 hidden layers, 60 nodes each
3. Extended input (27 $\Longrightarrow$ 126)
4. Additional classification output node (z, θ, c) (Yuxin Liu)

Extended Input:
Every wire passing ADC cut in TS

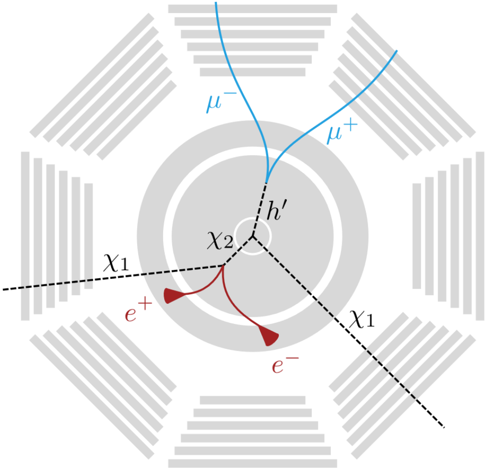




Implementation ongoing
(Kai Unger, KIT)

Timo Forsthofer

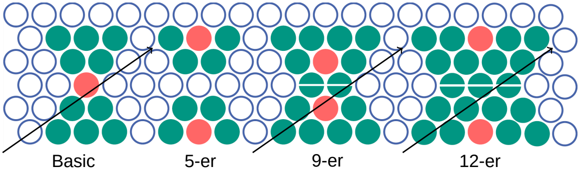
The Displaced Vertex Trigger



(M. Duerr et al, https://arxiv.org/abs/2012.08595)

Problem with the Neural Track Trigger: Tracks with displaced vertices only very inefficiently triggered

Non-pointing tracks: New track segments required (9-er or 12-er preferred)

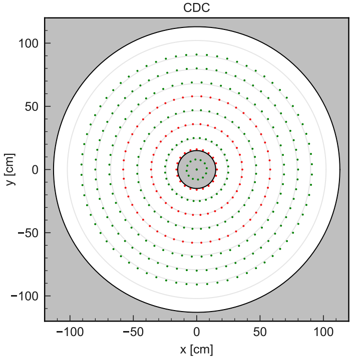


Track finding via Hough transform

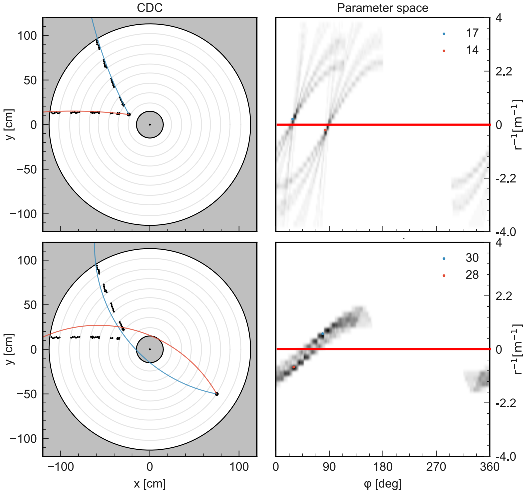
DVT Algorithm: Require two oppositely charged particles from same vertex within CDC volume

Hough Transform: Vertex Hypotheses

CDC partitioned in $\mathcal{O}(100)$ "Macro Cells"



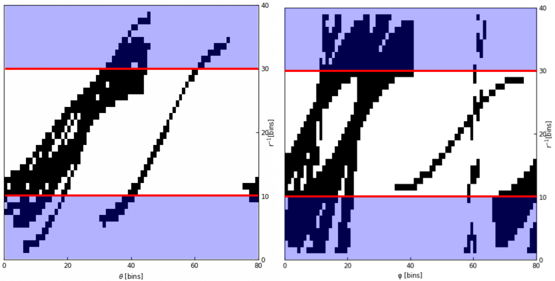
FPGA: All Hough transforms executed in parallel
$\implies$ $\mathcal{O}(100)$ track pair candidates!



Simple maximum search not possible

Find the Correct Track Pair

Select small set of track pair candidates ($\mathcal{O}(5)$)
Method: Reduce Hough matrix to bitmap

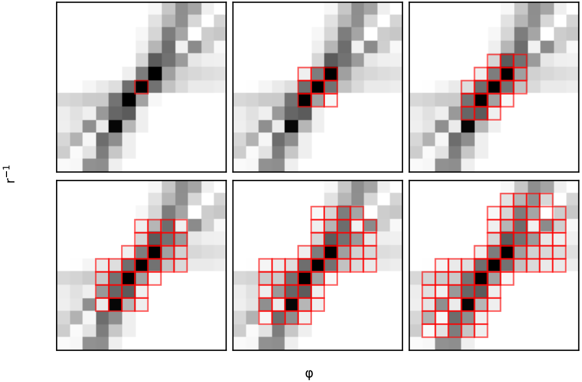


Wrong vertex assumption

Correct vertex assumption

For each Macro Cell i: $R_i = n_{\text{out}}/n_{\text{in}}$
$\Longrightarrow$ sort by R_i, take 5 largest

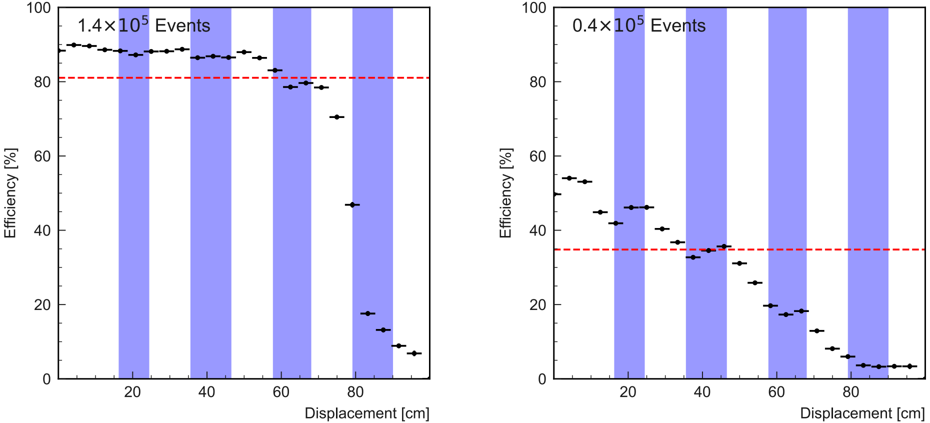
"Colored" Hough map for the 5 candidates



Iterative cluster shape algorithm
Calculate 10 cluster parameters, e.g.:
COG, size, orientation ...
$\Longrightarrow$ Neural network decision ($10 \times 40 \times 1$)

Expected Performance of DVT

Allowed trigger rate of < 1 kHz (Very preliminary optimization)



Early Phase-3 Nominal Phase-3 Elia Schmidt

Summary

3D Hough Finder + DNN:

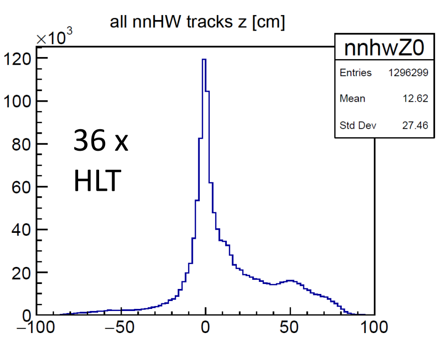
- Automatic suppression of tracks outside the interaction region (candidates implicitly originate from the IP)
- Better immunity against higher backgrounds
- Better track segment selection, smaller fake rate
- New network architecture (extended input + deep networks + classification)
  $\Longrightarrow$ better resolution in z and θ
- Higher efficiency for low charged multiplicity events

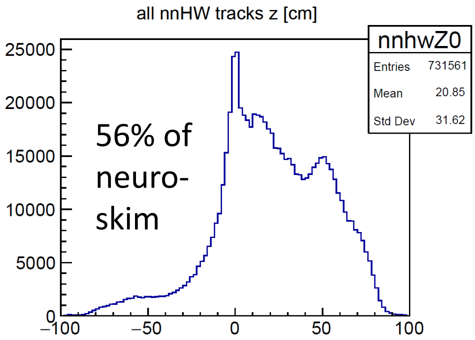
DVT:

- "Standard" neural trigger inefficient for displaced vertex events
- $\mathcal{O}(100)$ Hough transforms executed in Macro Cells in entire CDC volume
- Require 2 Hough maxima for tracks with opposite charge
- Correct Macro Cell for 2-track-vertex selected by neural network

Details on Implementation on UT4 see Kai Unger's presentation

Outlook: Training with unbiased data from “fstream”





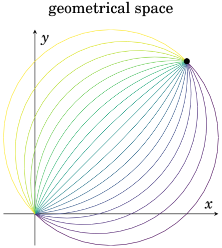
Backup

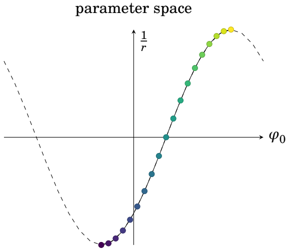
Preprocessing of the Network Input: Track Finding

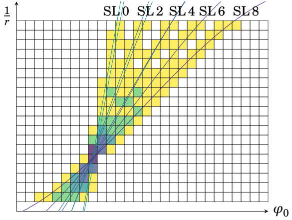
Which TS belong to a real track?

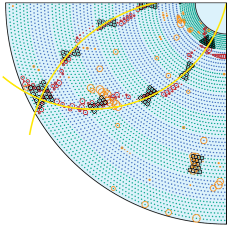
TS selection using a two-dimenisonal Hough transformation:

- Axial hit in CDC (TS) gets transformed to a curve in parameter (Hough) space
- Intersection point yields the track parameters ϕ and $r_{2d} \propto p_T$





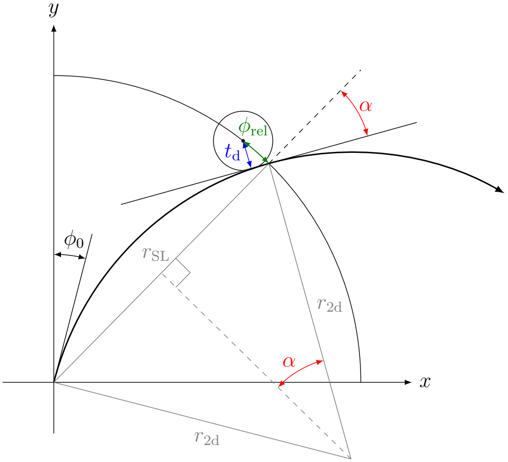




The Neuro Trigger has been running since January 2021 years with remarkable success.

Preprocessing for the L1 Neural Network Trigger

From the track finding (Hough transformation) we get: $\omega = \pm 1/r_{2d}$ and ϕ_0



With the TS information

$$\phi_{wire},\ n_{wire},\ r_{SL},\ \sigma_{LR},\ t_{d,wire}$$

we can calculate:

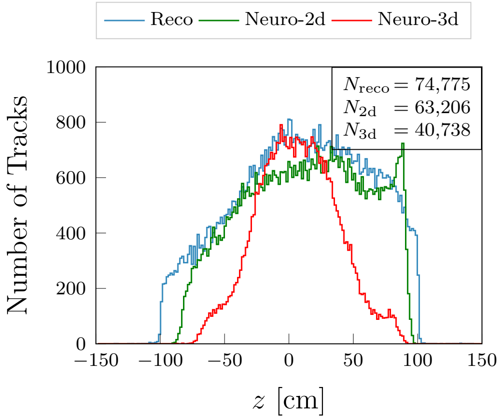
$$\alpha = \arcsin\left(\frac{1}{2}\frac{r_{SL}}{r_{2d}}\right)$$

$$\phi_{rel} = \phi_{wire} - n_{wire} \cdot \left(\frac{\phi_0 - \alpha}{2\pi}\right)$$

$$t_d = \sigma_{LR} \cdot (t_{d,wire} - t_{d,min})$$

3DHough: Natural Suppression of Displaced Tracks

- Promising results from **Particle Gun** single tracks ($z_{\text{reco}} \in [-100, 100]$ cm)
- The number of found tracks falls quickly with large $|z|$



Parameter optimization did not sufficiently solve the following problems:

- Resource heavy on the hardware, non-deterministic length, difficult to implement
- Clusters can get very large $\implies$ Very bad resolutions for some tracks
- Very high fake rates when using nominal phase-3 background (3 fakes for 1 signal)

(a) Cut at 10%   **(b)** Cut at 20%   **(c)** Cut at 30%

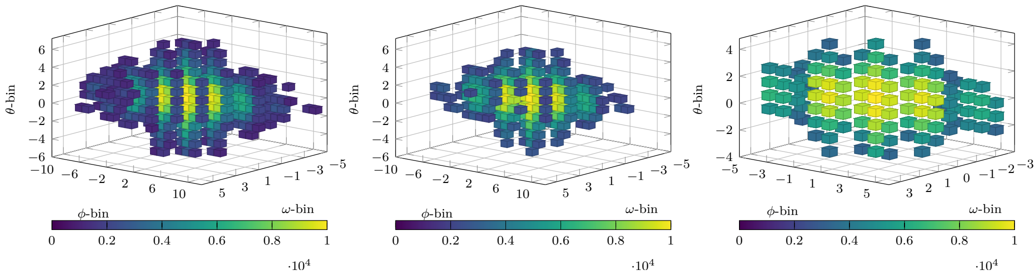


Figure: The cut percentage defines how often (on average) a cell was present in the clusters.

Cluster Statistics: Average cell weights

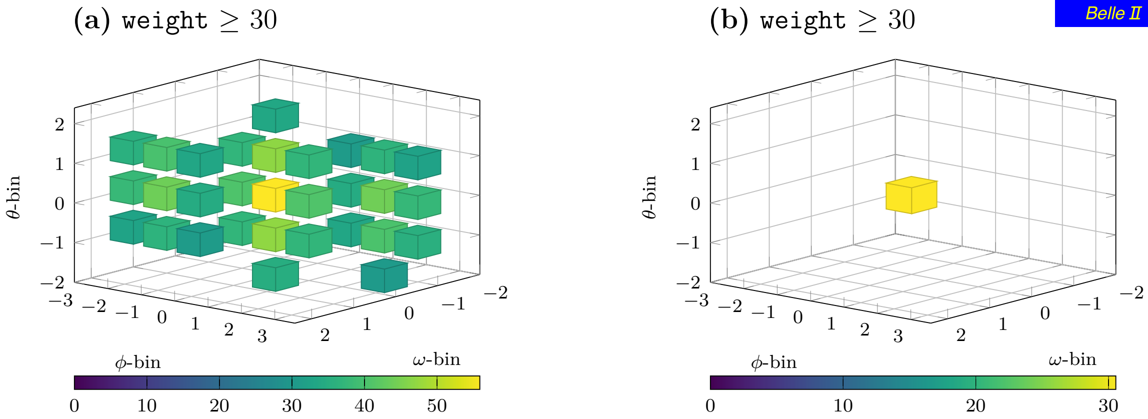
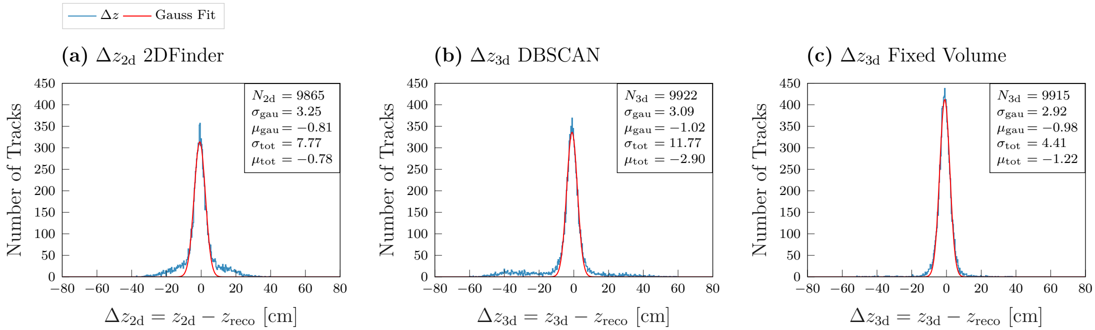


Figure: The average cluster weights (10000 clusters) above a weight of 30: **(a)** Real tracks with nominal phase-3 background, **(b)** Fake tracks from only nominal phase-3 background.

New idea: Fixed cluster shape clustering

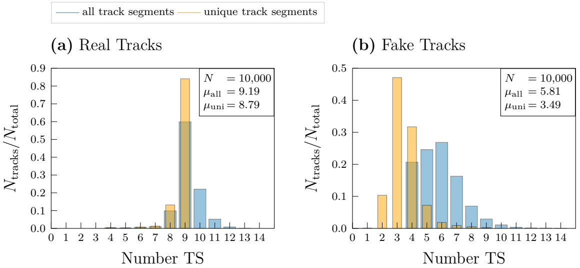
Monte Carlo Single-IP Tracks

- 10 000 single IP tracks
- The efficiency and resolution are high with the fixed volume clustering (plot **(c)**):



Nominal Phase-3 Background Studies

- Very high fake rates are observed $\Longrightarrow$ Solution: Cut on the number of hit super layers



Background rate per event (10 000 neutrinos):

Clustering	`minhits`	`minsuper`	N_{3d}^{all}
DBSCAN	4	0	29 424
DBSCAN	6	0	11 350
Fixed Volume	5	5	783

- Default DBSCAN: **290%**
- Fixed Volume with `minsuper` = 5 cut: **6%**

Classifier for Tracks from IP