Improving the Belle II Neural Track Trigger with Deep Neural Networks

Timo Forsthofer, Christian Kiesling, Kai Unger, Simon Hiesl

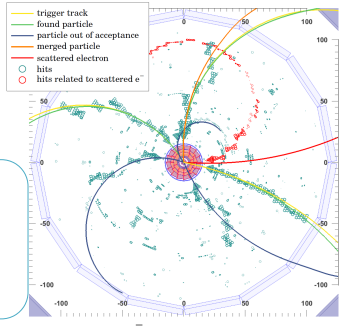
March 6, 2024

Members of the Belle II Trigger Group

KIT ITIV
- Marc Neu
- Kai Unger
- Jürgen Becker

KIT ETP
- Lea Reuter
- Greta Heine
- Slavomira Stefkova
- Torben Ferber

MPI & TUM & LMU
- Christian Kiesling
- Felix Meggendorfer
- Simon Hiesl
- Timo Forsthofer
- Alois Knoll

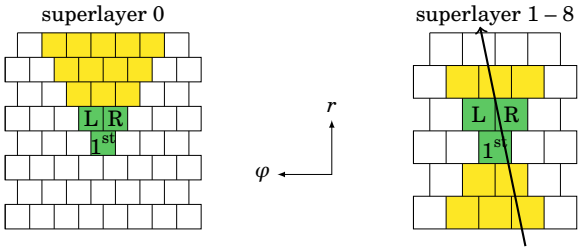


Focus of the MPI/LMU/TUM group:

Track Triggers

Preprocessing of Input Variables

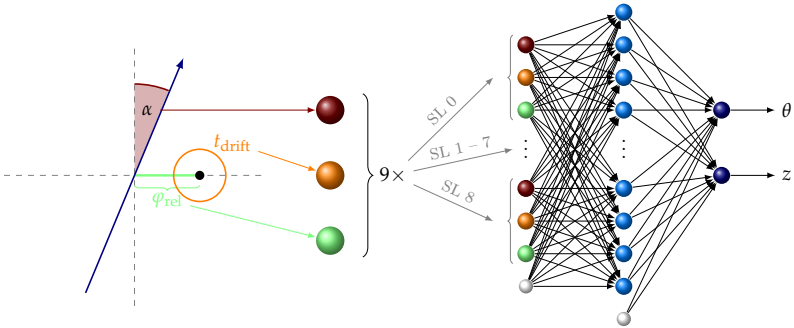
- Central Drift Chamber (CDC) at Belle II with 54 layers of wires organized in nine superlayers
- Hour-glass-shaped Track Segments passed from Track Segment Finder (TSF) over Hough-Finder to Neural Network

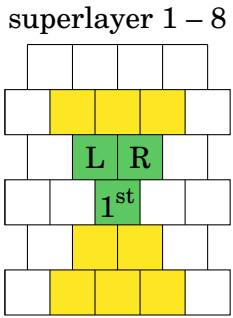


Structure of a Track Segment

Current Network

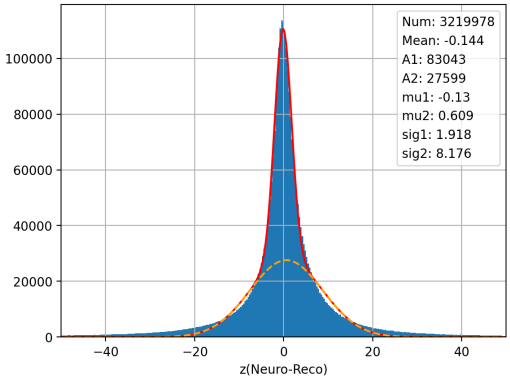
- Three input parameters for each superlayer
- 81 hidden nodes and two output nodes (z and theta)
- Five experts for different configurations of missing Stereo Layers

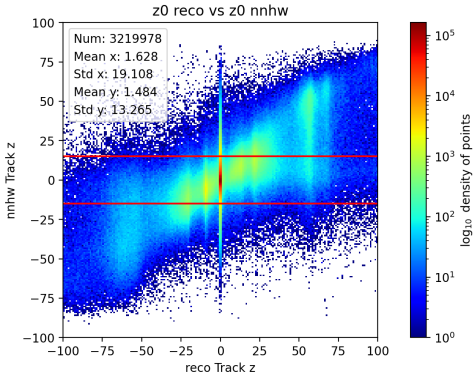




Current Network

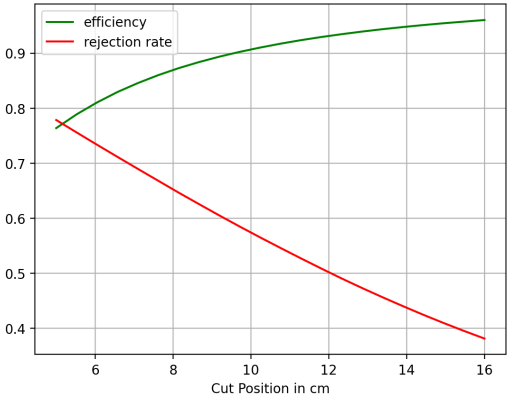
- Data in previous runs had less background
- Retraining on noisy data alone increases the resolution
- Strong bias in input data leads to Feed-Up and Feed-Down



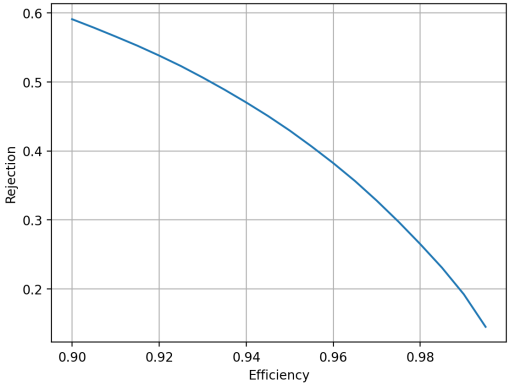


Current Network

- Present trigger condition: Accept all tracks with $|z_{neuro}| < 15\,\text{cm}$



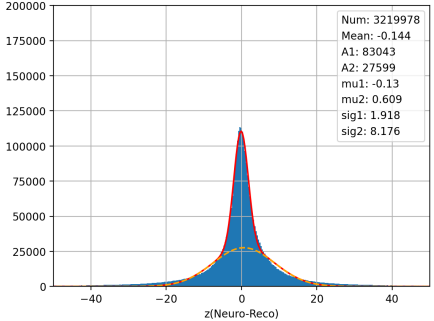
Efficiency and Rejection for Different Cuts



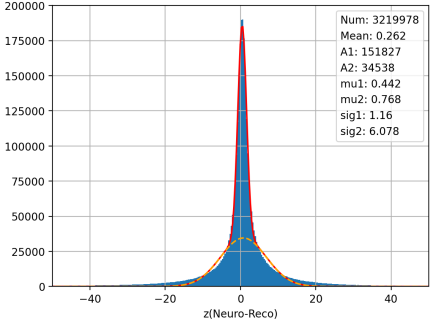
Rejection for Different Efficiencies

Deep Learning Architectures

- New, more powerful FPGAs allow for bigger networks
- Three or four hidden layers beneficial for resolution
- More hidden layers better than more nodes per layer



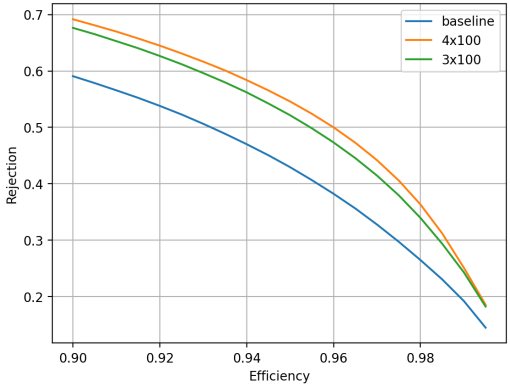
1HL with 81 Nodes



4HL with 100 Nodes per HL

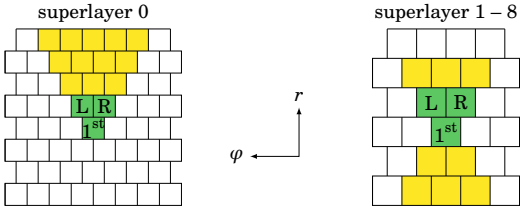
Deep Learning Architectures

- Rejection rate up to 10 percentage point higher compared to present architecture (baseline) for same efficiency
- More hidden layers better than more nodes per layer

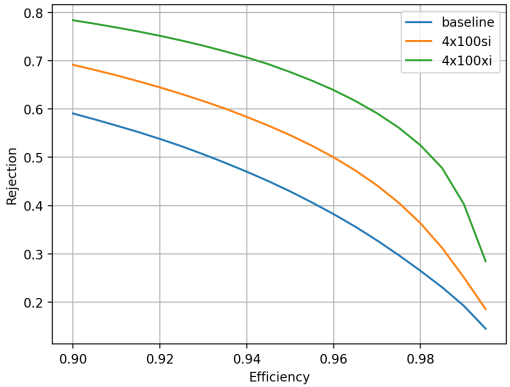


Extended Input

- New hardware can also pass more information to the network
- Drift times for all 11 wires in track segment passed on to network
- Particularly effective for improving the background rejection rate

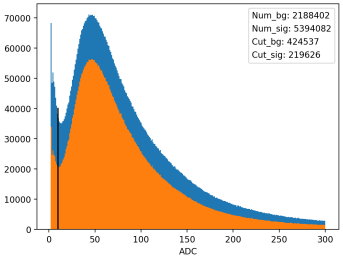


ADC-Distribution on Signal and BG

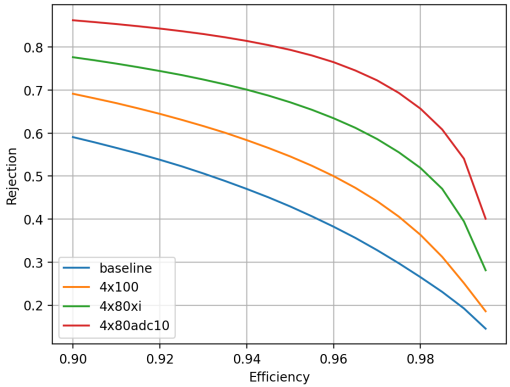


ADC-Cut

- ADC-count indicates strength of electric signal
- Potential for further decrease of noise in the CDC wires
- Low counts often not real hits, so it's better to ignore them

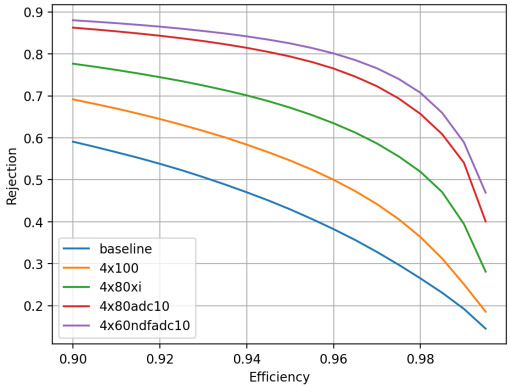


ADC-Distribution on Signal and BG



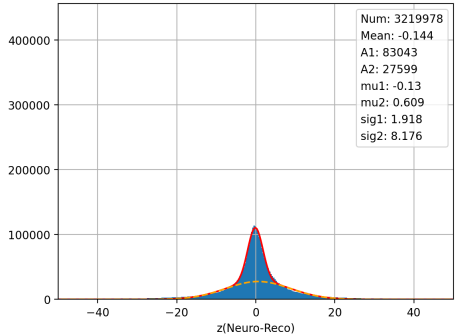
Combining ADC-Cut and 3D-Hough-Finder

- ADC-Cut works well with 3D-Hough Finder (see presentation by Simon Hiesl)
- 3D-Hough Finder already rejects a lot of background and fake tracks, so the performance is underrepresented here

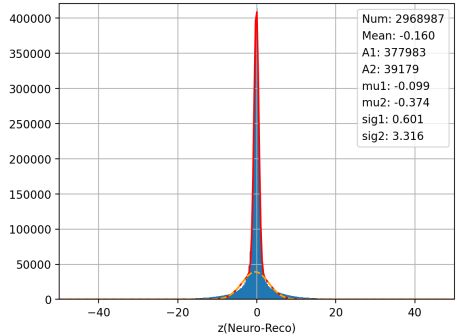


Final Performance Evaluation

- Combination of all advances leads to increase in accuracy by almost a factor of three
- z-Cut can be reduced from 15cm to under 10cm



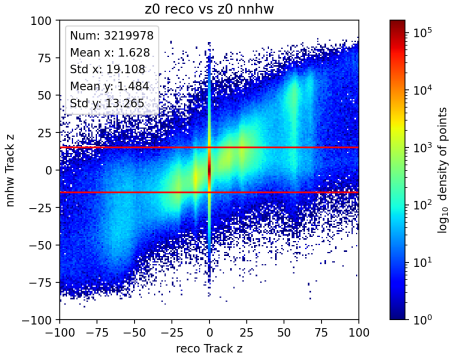
Present Network Architecture



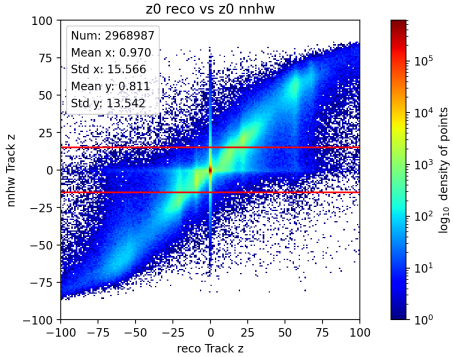
Deep Neural Network with Extended Input, ADC-cut and 3D-Input

Final Performance Evaluation

- Especially extended input helpful in reducing Feed-Up and Feed-Down



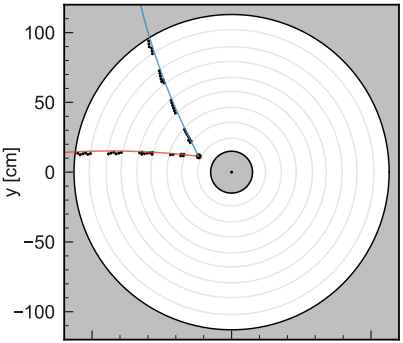
Present Network Architecture



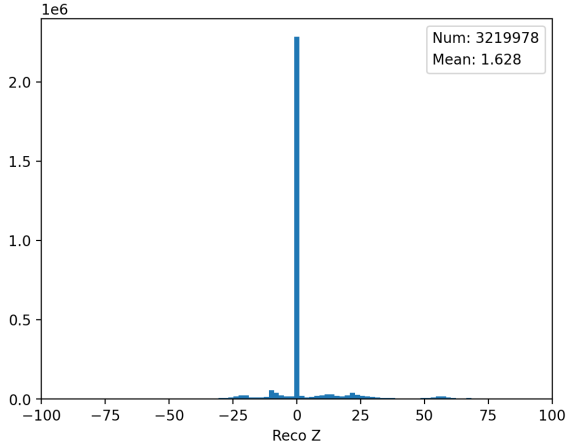
Deep Neural Network with Extended Input, ADC-cut and 3D-Input

Summary and Outlook

- Deep Neural Networks, Extended Input, ADC-Cut and the 3D-Hough Finder each bring significant improvements and work well together
- Implementation into hardware in cooperation with KIT ITIV
- Future training on unbiased input data
- Displaced Vertex Trigger



Backup


1e6
Num: 3219978
Mean: 1.628
2.0
1.5
1.0
0.5
0.0
−100
−75
−50
−25
0
25
50
75
100
Reco Z