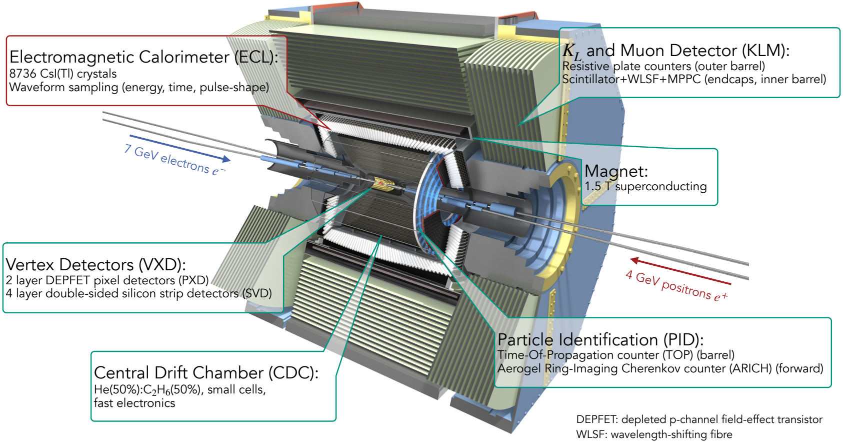
Real-time Clustering with GNNs for the Electromagnetic Calorimeter Trigger at Belle II

Workshop on Fast Realtime Systems and Realtime Machine Learning 2024

Isabel Haide, Marc Neu, Torben Ferber | Wednesday 10th April, 2024

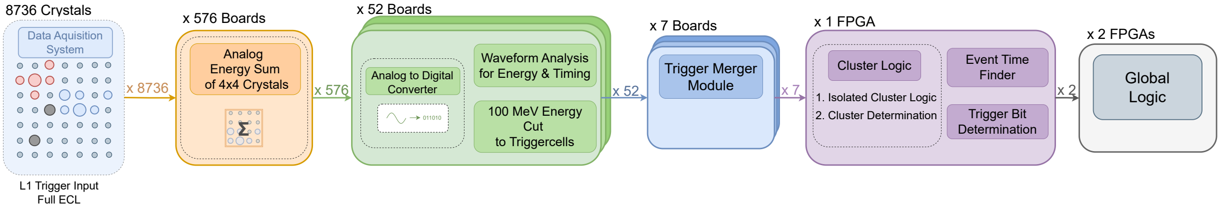
The Belle II Detector

Karlsruhe Institute of Technology



DEPFET: depleted p-channel field-effect transistor
WLSF: wavelength-shifting fibre
MPPC: multi-pixel photon counter

Current Electromagnetic Calorimeter (ECL) Trigger



1.) All 8736 crystals are read out

2.) The energy of 16 crystals is analogously summed into one trigger cell (TC)

3.) A 100 MeV energy cut is applied to each TC to reduce input size

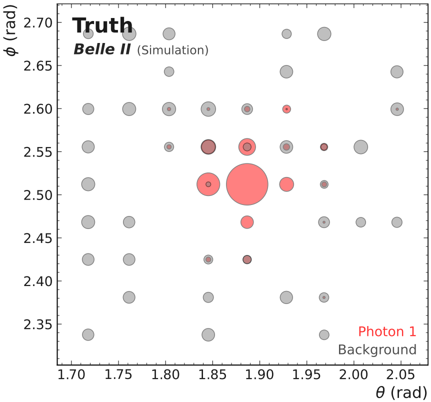
4.) Clusters are found through isolating cluster logic

5.) Event Timing for all triggers is defined through highest energetic cluster

6.) Trigger lines for Bhabha rejection and low multiplicity events are calculated

Current Clustering Logic

- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals

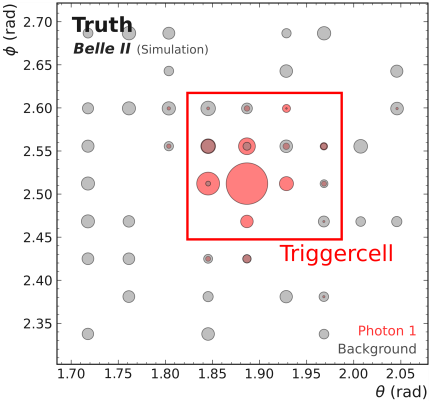


arXiv:2306.04179

Current Clustering Logic

Karlsruhe Institute of Technology

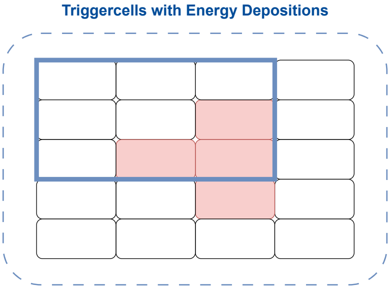
- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals



arXiv:2306.04179

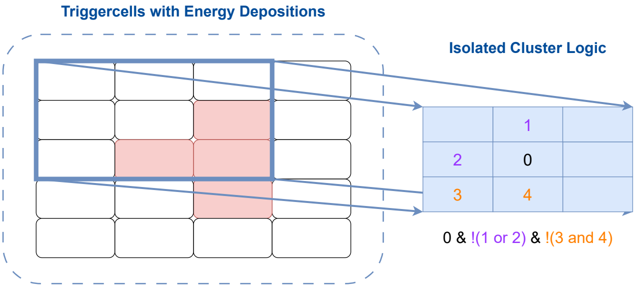
Current Clustering Logic

- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals
- FPGA logic selects TCs with an energy deposit over 100 MeV for clustering to decrease noise



Current Clustering Logic

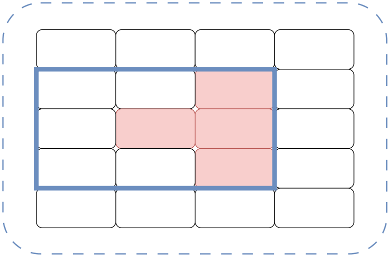
- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals
- FPGA logic selects TCs with an energy deposit over 100 MeV for clustering to decrease noise
- Current clustering logic finds clusters through isolated cluster logic



Current Clustering Logic

- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals
- FPGA logic selects TCs with an energy deposit over 100 MeV for clustering to decrease noise
- Current clustering logic finds clusters through isolated cluster logic

Triggercells with Energy Depositions



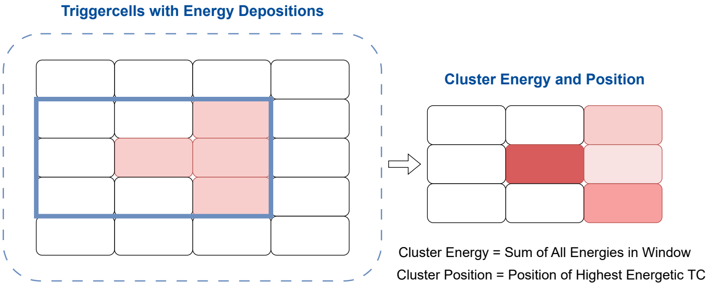
Isolated Cluster Logic

	1	
2	0	
3	4	

0 & !(1 or 2) & !(3 and 4)

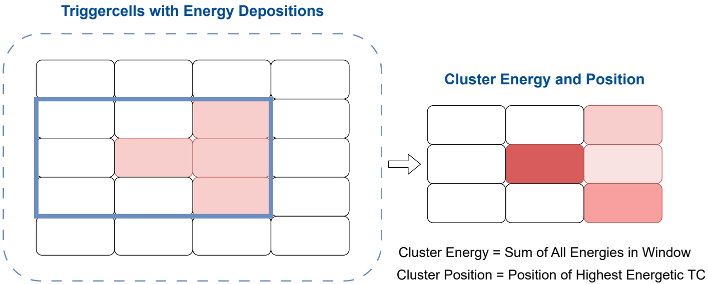
Current Clustering Logic

- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals
- FPGA logic selects TCs with an energy deposit over 100 MeV for clustering to decrease noise
- Current clustering logic finds clusters through isolated cluster logic

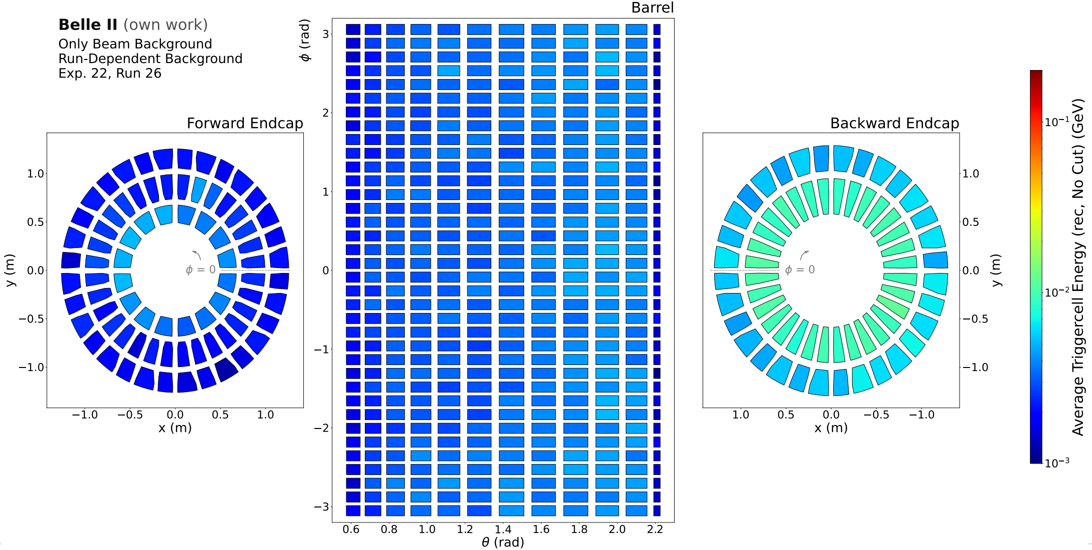


Current Clustering Logic

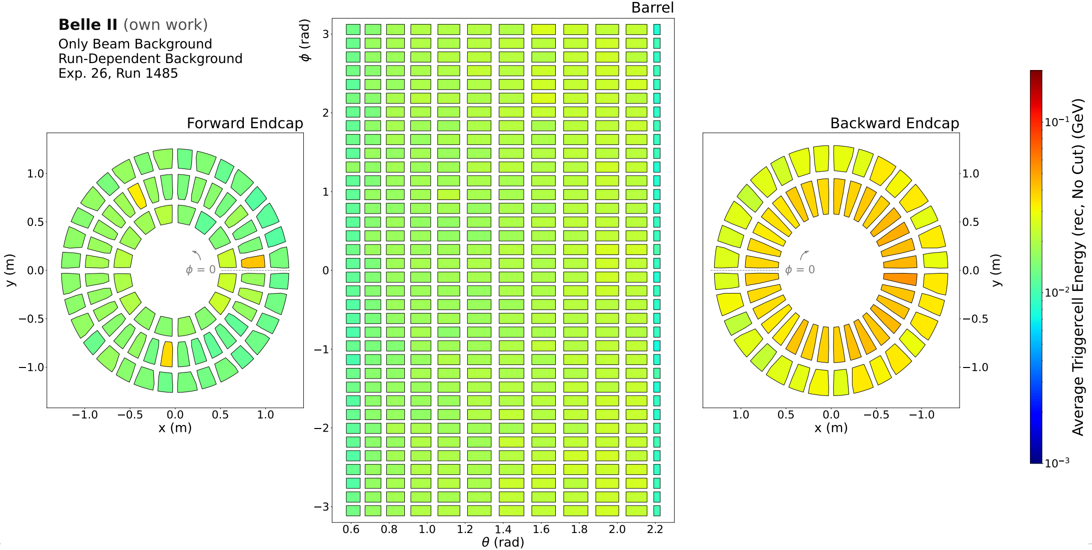
- Typical photon cluster is contained within a 5x5 grid of crystals
- Current ECL Trigger reads out analog energy sum of triggercells (TCs) consisting of up to 16 crystals
- FPGA logic selects TCs with an energy deposit over 100 MeV for clustering to decrease noise
- Current clustering logic finds clusters through isolated cluster logic
- Current algorithm returns up to 6 isolated clusters in the order Barrel -> Forward Endcap -> Backward Endcap
- Overlapping or adjacent clusters can by design not be distinguished
- High beam background energy depositions look like clusters
  $\rightarrow$ high trigger rate



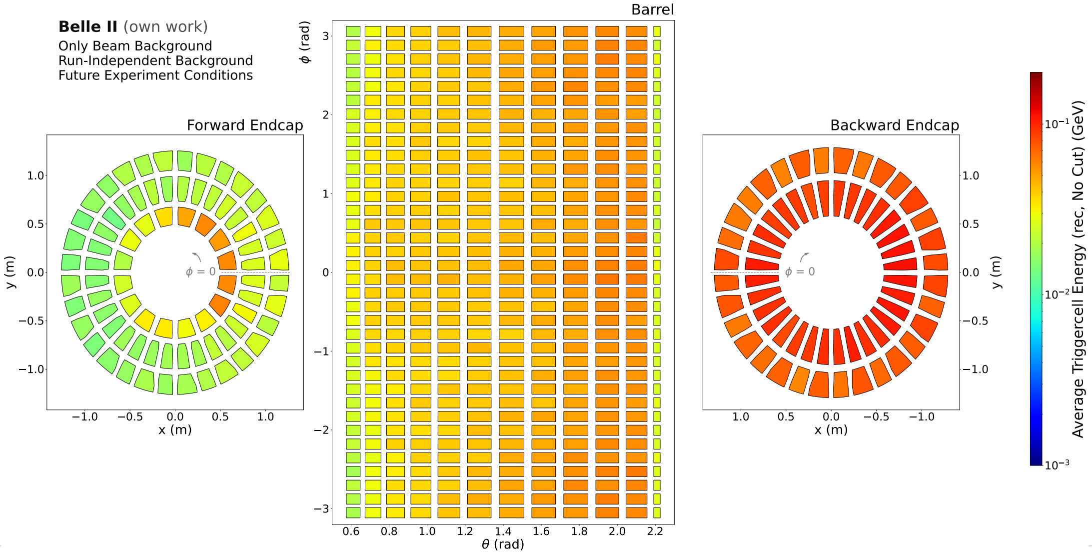
Beam Background Level in 2021



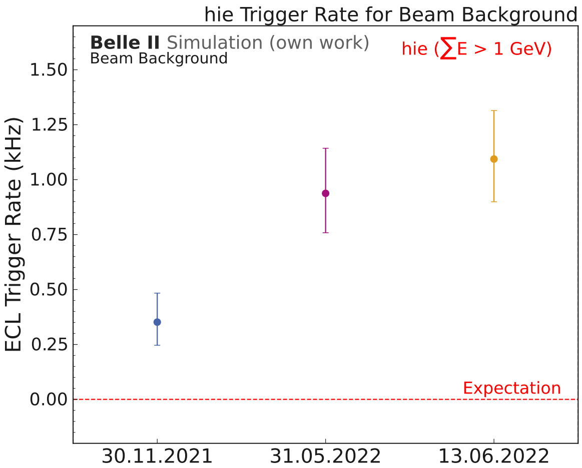
Beam Background Level in 2022



Estimated Future Beam Background Level



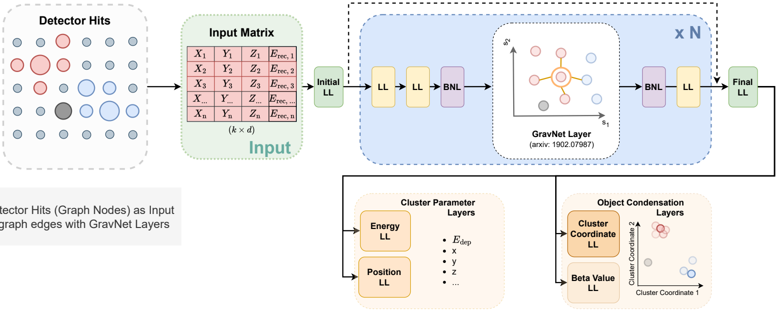
Beam Background in the ECL



- hie: Total energy sum in inner ECL > 1 GeV
- Simulation containing only beam background taken from data already increases trigger rate of single line to > 1 kHz
- Maximum possible total trigger rate = 30 kHz

⇒ **Trigger algorithm has to be adapted to rising background conditions**

Object Condensation for the ECL Trigger



Use Detector Hits (Graph Nodes) as Input
Learn graph edges with GravNet Layers

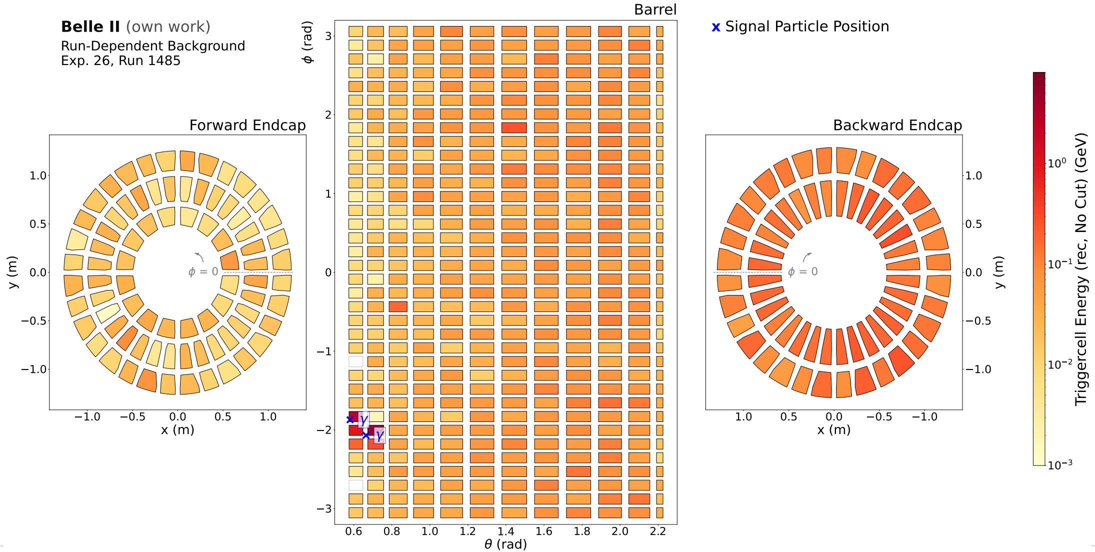
Hyperparameter	Value
Nr. of Neighbors	9
Nr. of GravNet Blocks	2-4
Nr. of Nodes for LL 1	16/32
Nr. of Nodes for LL 2	16/32
Momentum	0.6
Cluster Coordinates	3

$\Rightarrow$ Between 12 000 and 60 000 parameters

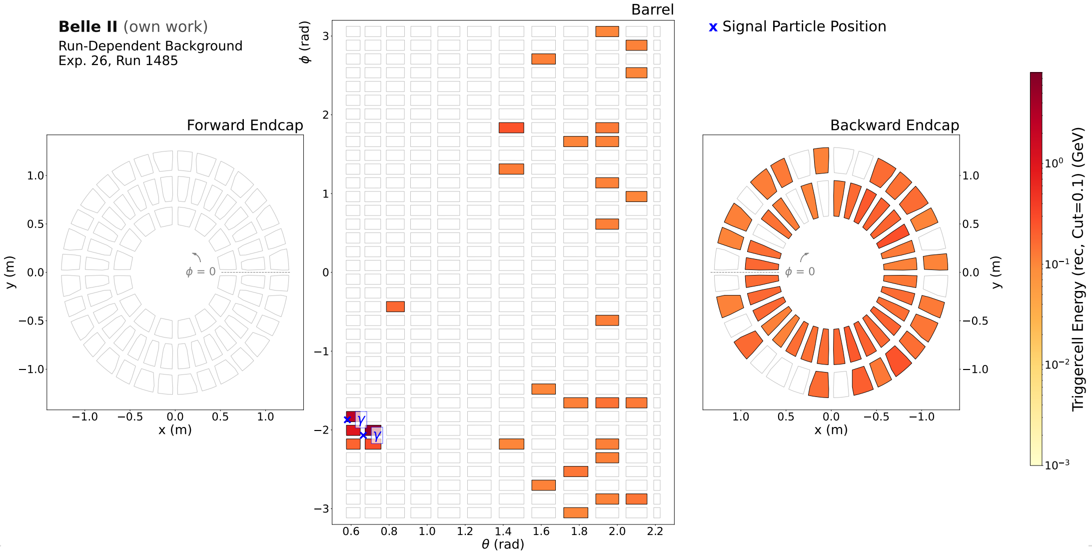
- Object Condensation (OC): One-shot algorithm for both detection and reconstruction of clusters (arXiv:2002.03605)
- Irregular geometry and varying input sizes in the ECL $\rightarrow$ **Graph Neural Networks (GNN)**
- OC algorithm is adaptable to **different beam backgrounds** and can be (for future upgrades) adapted to different inputs

- **Inputs**: All TCs with energy deposition > threshold with position of TC, reconstructed energy, reconstructed timing
- **Outputs**: Existence of cluster, cluster position, cluster energy
- **All training and evaluation is done on CPU at the moment**

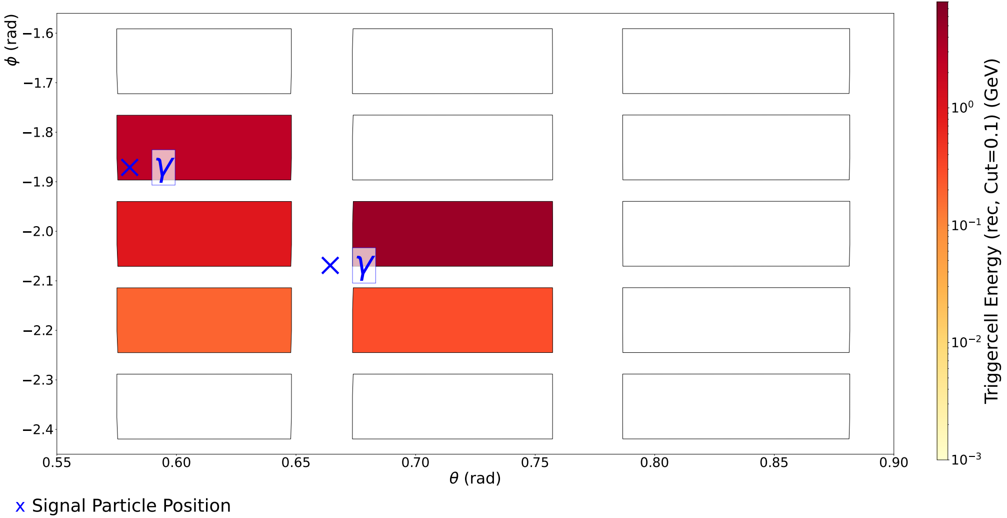
Example Event - ECL Triggercells



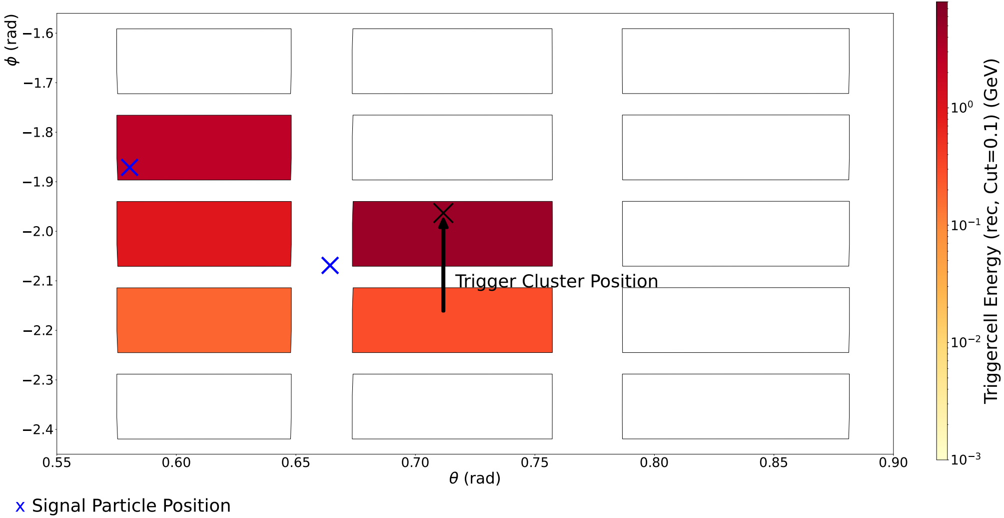
Example Event - Triggercells with 0.1 GeV Cut



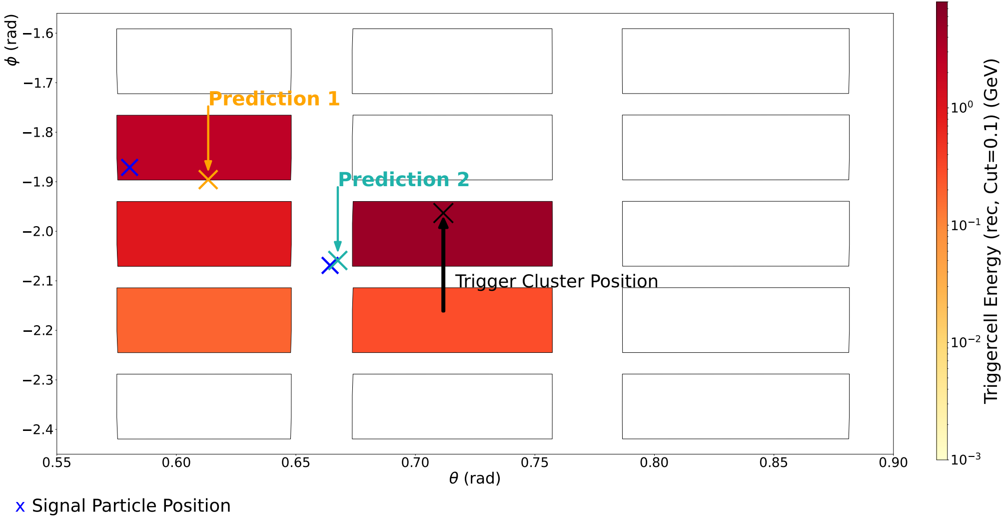
Example Event - Triggercells, Zoomed



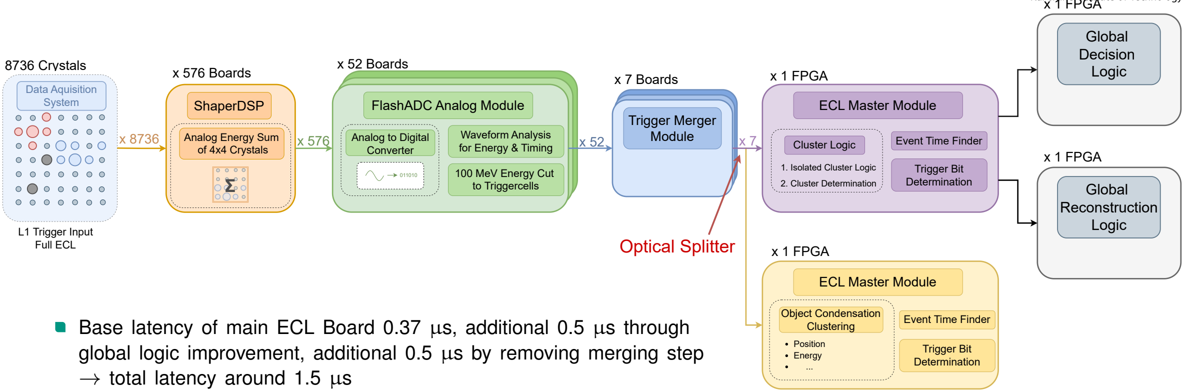
Example Event - Triggercells, Trigger Clusters



Example Event - Triggercells, Predictions



OC for the ECL Trigger - Implementation

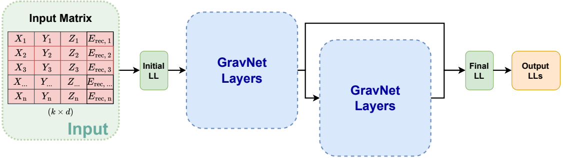


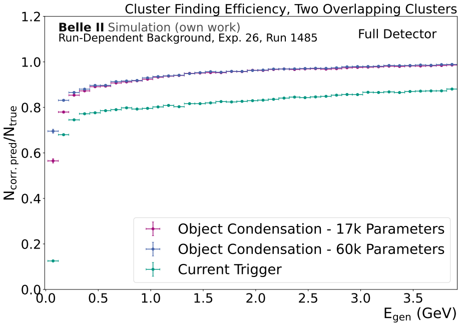
- Base latency of main ECL Board 0.37 µs, additional 0.5 µs through global logic improvement, additional 0.5 µs by removing merging step → total latency around 1.5 µs
- Testing Hardware: XCUV160
- GravNet module employs KNN algorithm → implemented on FPGA
- We plan to implement this in parallel to current trigger algorithm for testing in deployment

OC for the ECL Trigger - Network Design

Karlsruhe Institute of Technology

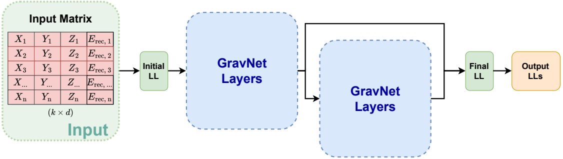
- Trigger Input Signal: 1 bit hit + 7 bit timing + 12 bit energy
- For kNN algorithm, limit of 64 input TCs and 16 nearest neighbors
  - For $B\bar{B}$ events, 99% of events have less than 28 TCs above 100 MeV for current background levels (< 103 TCs for future background levels)
- Testing hardware limits network size to 2 GravNet blocks
  - Offline testing with small networks sizes still outperforms trigger
- First quantized implementation of GravNet layer on FPGA successfully done





OC for the ECL Trigger - Network Design

- Trigger Input Signal: 1 bit hit + 7 bit timing + 12 bit energy
- For kNN algorithm, limit of 64 input TCs and 16 nearest neighbors
  - For $B\bar{B}$ events, 99% of events have less than 28 TCs above 100 MeV for current background levels (< 103 TCs for future background levels)
- Testing hardware limits network size to 2 GravNet blocks
  - Offline testing with small networks sizes still outperforms trigger
- First quantized implementation of GravNet layer on FPGA successfully done



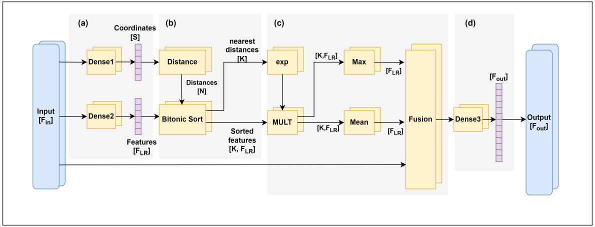
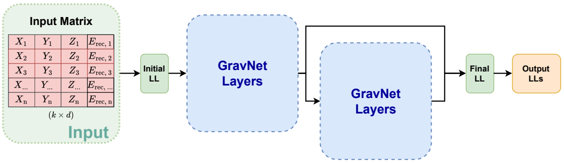


Figure: Shen Shao

OC for the ECL Trigger - Network Design

- Trigger Input Signal: 1 bit hit + 7 bit timing + 12 bit energy
- For kNN algorithm, limit of 64 input TCs and 16 nearest neighbors
  - For $B\bar{B}$ events, 99% of events have less than 28 TCs above 100 MeV for current background levels (< 103 TCs for future background levels)
- Testing hardware limits network size to 2 GravNet blocks
  - Offline testing with small networks sizes still outperforms trigger
- First quantized implementation of GravNet layer on FPGA successfully done



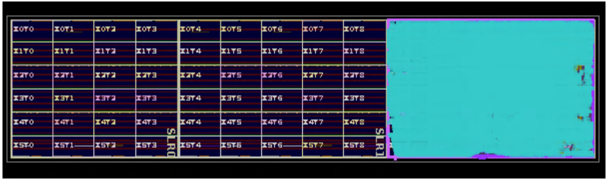
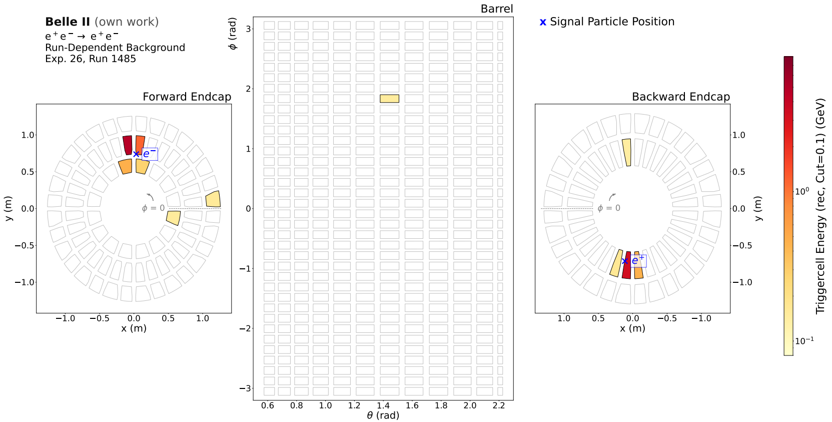


Figure: Marc Neu

ECL Bhabha Veto

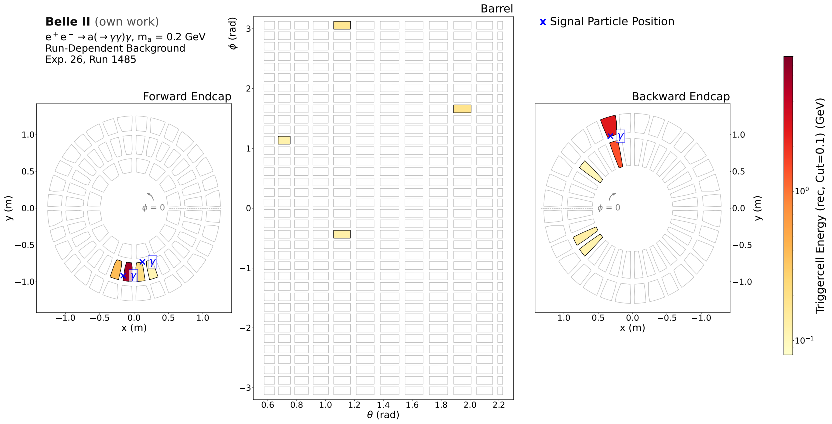
$e^+e^- \rightarrow e^+e^-$ Event:



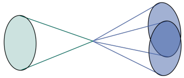
- One main task of the ECL Trigger is to veto Bhabha-like events
- Bhabha vetoes on trigger level do not use track information but only clusters
  - Two clusters above 3 GeV and one cluster above 4.5 GeV (in center-of-mass system)
  - Angle requirement to have back-to-back clusters in center-of-mass system
- Many physics trigger lines depend on no Bhabha veto

ECL Bhabha Veto - ALP Decays

Low Mass ALP Event (0.2 GeV):

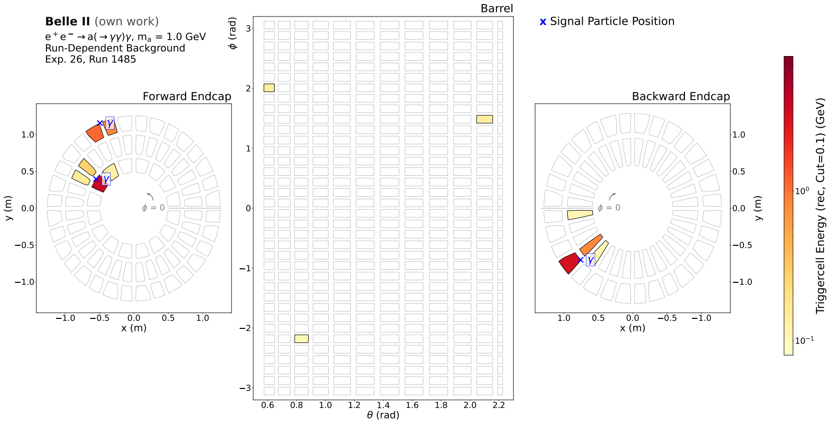


- Decays with signatures mimicking Bhabha events in the calorimeter have very low efficiency
- Example: ALP-Strahlung $e^+e^- \to a\,(\to \gamma\gamma)\,\gamma$, highly overlapping clusters for light ALPs (arXiv:1709.00009)



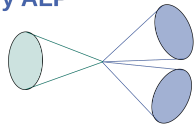
ECL Bhabha Veto - ALP Decays

Higher Mass ALP Event (1.0 GeV):



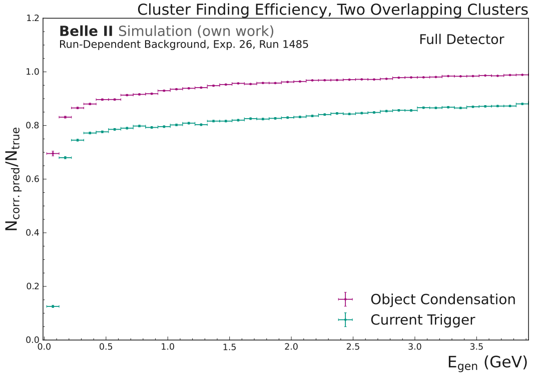
- Decays with signatures mimicking Bhabha events in the calorimeter have very low efficiency
- Example: ALP-Strahlung $e^+e^- \to a\,(\to \gamma\gamma)\,\gamma$, highly overlapping clusters for light ALPs (arXiv:1709.00009)

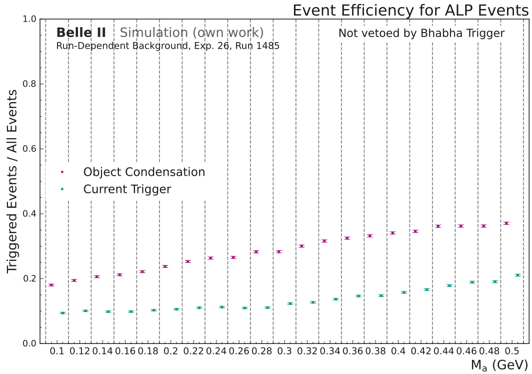
Heavy ALP



ECL Bhabha Veto - Improvement

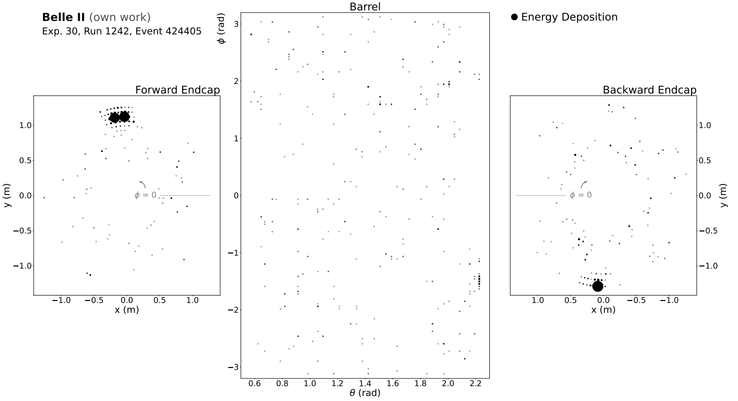
- OC algorithm can separate overlapping clusters with higher efficiency than current trigger algorithm
- Visible improvement in estimated trigger efficiency for e.g. ALP analyses due to improved separation
- Currently fully run on CPU





Summary

- High beam background challenge can be tackled with GNN trigger algorithm
- Small GNN network (16k parameters) and input sizes (< 30 TCs) for application on FPGAs
  - Implementation of kNN and GravNet algorithm on FPGAs already successful
- Improvement on low multiplicity efficiency and Bhabha rejection (on CPU)
- End of year goal: Full implementation on FPGA and parasitic running at Belle II

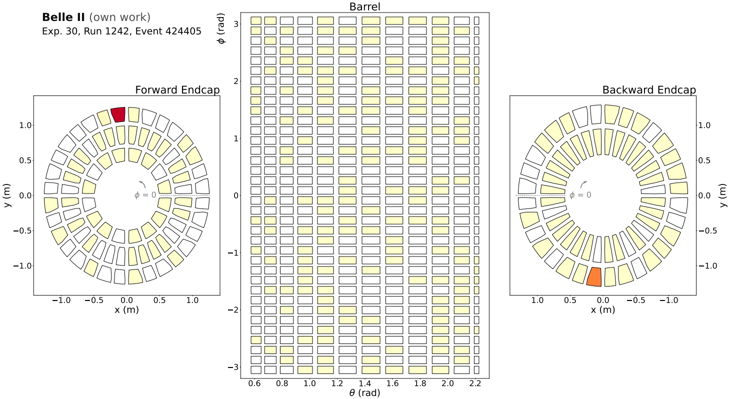


Event recorded in Feb. 24 - Crystals

Summary

Karlsruhe Institute of Technology

- High beam background challenge can be tackled with GNN trigger algorithm
- Small GNN network (16k parameters) and input sizes (< 30 TCs) for application on FPGAs
  - Implementation of kNN and GravNet algorithm on FPGAs already successful
- Improvement on low multiplicity efficiency and Bhabha rejection (on CPU)
- End of year goal: Full implementation on FPGA and parasitic running at Belle II

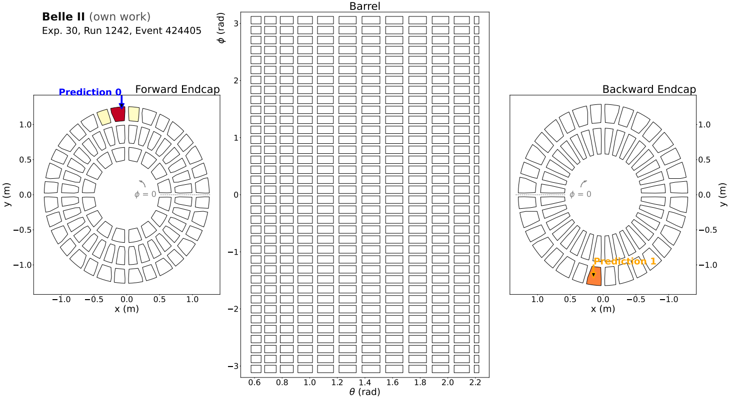


Event recorded in Feb. 24 - TCs

Summary

Karlsruhe Institute of Technology

- High beam background challenge can be tackled with GNN trigger algorithm
- Small GNN network (16k parameters) and input sizes (< 30 TCs) for application on FPGAs
  - Implementation of kNN and GravNet algorithm on FPGAs already successful
- Improvement on low multiplicity efficiency and Bhabha rejection (on CPU)
- End of year goal: Full implementation on FPGA and parasitic running at Belle II



Event recorded in Feb. 24 - Predicted Clusters (evaluated on CPU)