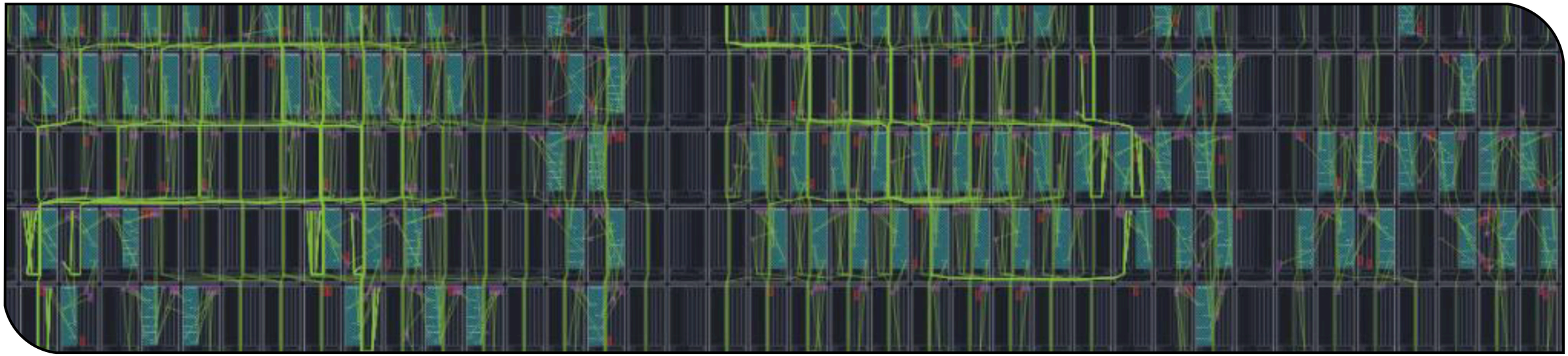
Karlsruhe Institute of Technology

Institut für Experimentelle Teilchenphysik

GNNETM $e^+e^- \to \mu^+\mu^-\gamma$ and $e^+e^- \to e^+e^-\gamma$ in data

Belle II TRG Meeting, B2GM, 24.02.2025

Torben FERBER (torben.ferber@kit.edu) (alphabetically: Jürgen Becker, Valdrin Dajaku, Isabel Haide, Timo Justinger, Marc Neu, Till Rädler)
Institute of Experimental Particle Physics (ETP) and Institut für Technik der Informationsverarbeitung (ITIV)



ECL GNN clustering on AMD Versal (by Till Rädler/KIT)

Updates since last presentation (https://indico.belle2.org/event/14065/)

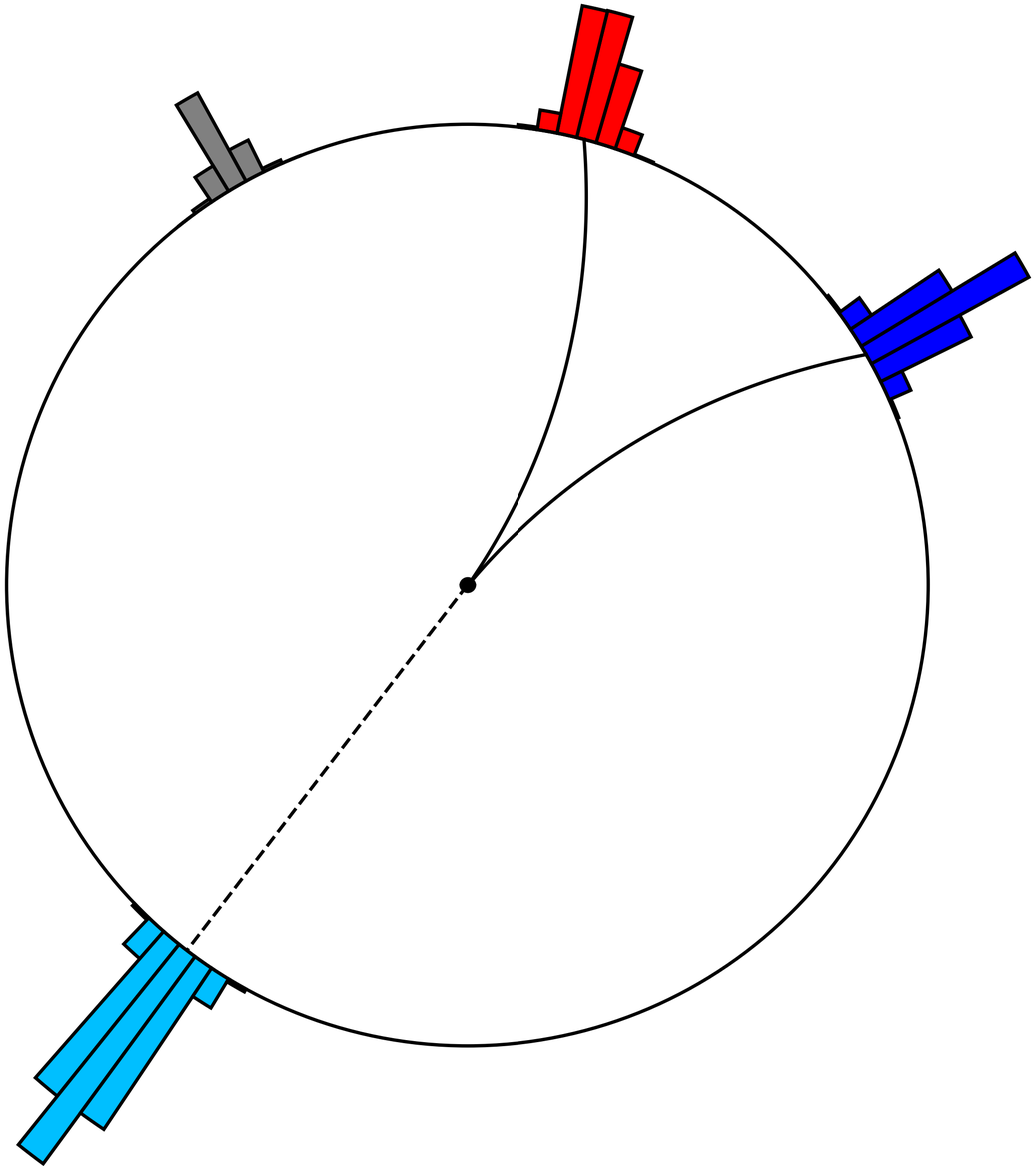
- Timing selection updated (see Isabel's presentation for all details)
- ICN position resolution after TC correction
- (In)Efficiency details for muons
- Energy resolution for low energy clusters data vs MC for muons Signal classifier performance

Technical details

- I am analysing exp35 run2882 mdst data (produced by Isabel) and KKMC ee→μμ(γ) events
  - 20 MeV seed and LM threshold in ECL offline clustering
  - No calibration or corrections available yet, tracking probably needs better global tags
  - Beam backgrounds in exp35 run2882 are very high…
- I am using offline GNN predictions from model **fine-gorge-1256** with GNN TC inputs from data (produced by Isabel)

Pure isolated cluster selection with radiative lepton events

- Goal: Validate GNN-identification of signal (blue, red, lightblue) and background clusters (gray)
- Select events with three isolated (at least 0.3rad separation) calorimeter clusters
  - **two of the clusters (0.1 < E < 0.4 GeV)** or **(E > 0.4 GeV)** must be matched to charged tracks (**red** and **blue**) with pt > 0.2 GeV
  - one cluster (E>0.2 GeV) must not be matched to a charged track (**blue**), if multiple clusters, choose the highest energy cluster
  - all three clusters must be isolated from any other cluster in the event with E > 0.08 GeV (charged or neutral) by at least 0.3 rad
  - event must be triggered by a track trigger (fyo, syo, ffy, or stt)
  - kinematic fit probability consistent (prob>0) with known initial four-momentum of colliding e^+e^- beams
  - best candidate selection: highest m(μμ) or m(ee)
  - all events with a photon $110 < \theta < 130°$ and $124 < \phi < 145°$ are removed (ECL crate hot)
  - next highest energy cluster that is not $110 < \theta < 130°$ and $124 < \phi < 145°$ is considered “background” (**gray**)
- Use only clusters that have an offline time in the trigger window
- Only use events that have GNN tc_timewindow==2 and event_window==3 (see Isabel’s talk)



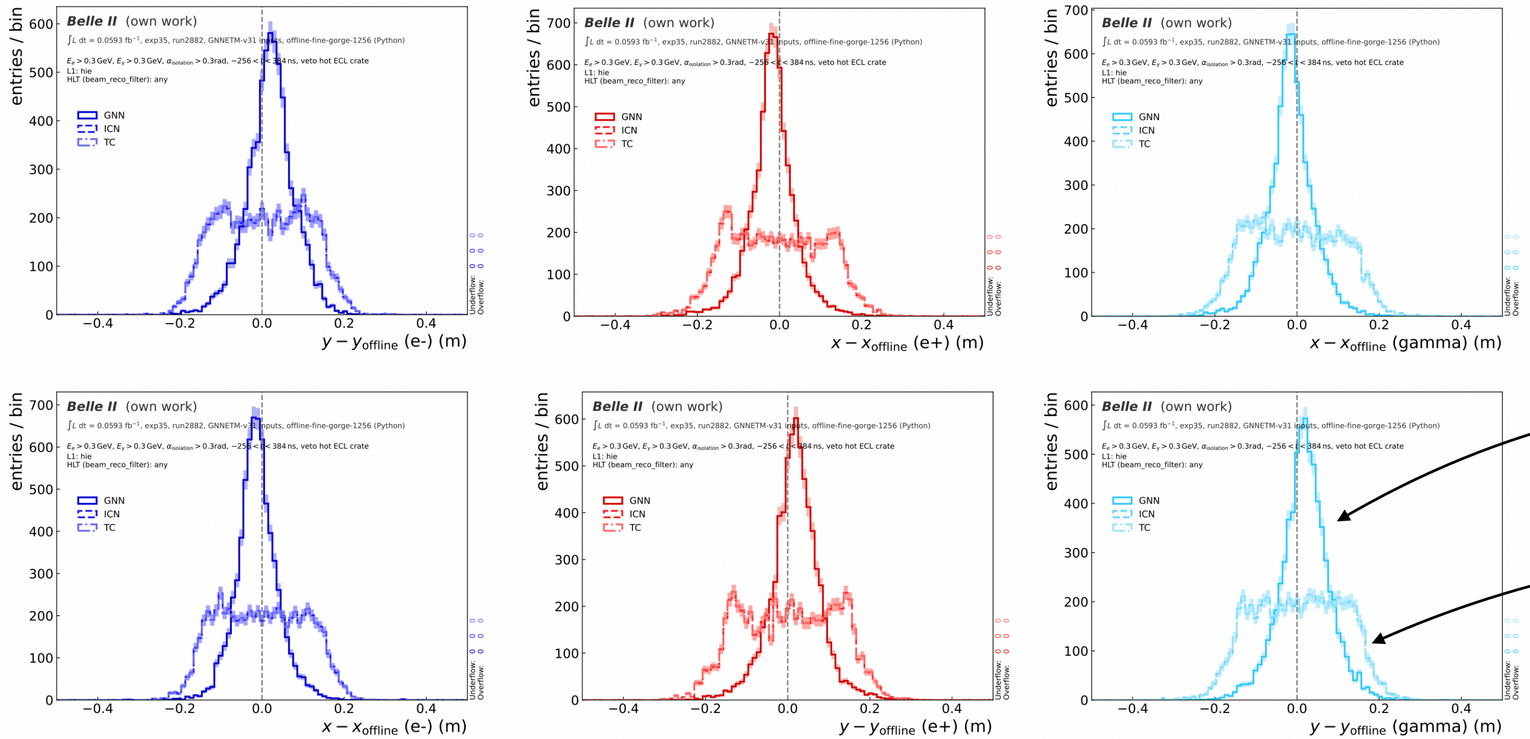
GNN-ETM, ICN-ETM, and TC matching

- For each offline cluster (at least three per event), I check every masked GNN cluster (or ICN cluster, or input TC):
  - spatial distance $\sqrt{dx^2 + dy^2 + dz^2} < 0.4\,\text{m}$ (with $dx = x_{\text{trigger}} - x_{\text{offline}}$
  - energy ratio $0.01 < E_{\text{trigger}}/E_{\text{offline}} < 2$
- If both conditions are fulfilled, the GNN cluster (or ICN cluster, or input TC) is **matched** to the offline cluster
- If multiple GNN clusters (or ICN clusters, or input TCs) are matched, I keep the one that has $E_{\text{trigger}}/E_{\text{offline}}$ closest to 1
- In principle multiple offline clusters can still be matched to the same GNN cluster (or ICN cluster, or input TC). Due to the isolation selection, this is very rare in my sample.

ICN position resolution as shown at a previous meeting

GNN-ETM, ICN-ETM, and TC resolution: x and y

ICN position is the highest energy TC in the cluster, and usually the highest energy TC is the one that is matched to the offline cluster.



GNN

ICN

10 Torben Ferber Institute of Experimental Particle Physics (ETP)

TRG weekly meeting
23.01.2025
https://indico.belle2.org/event/14065/

$ee \to ee\gamma$

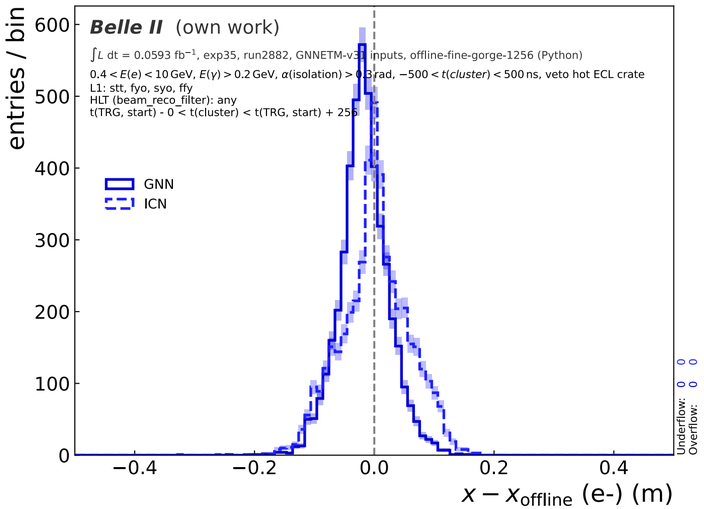
Position resolution

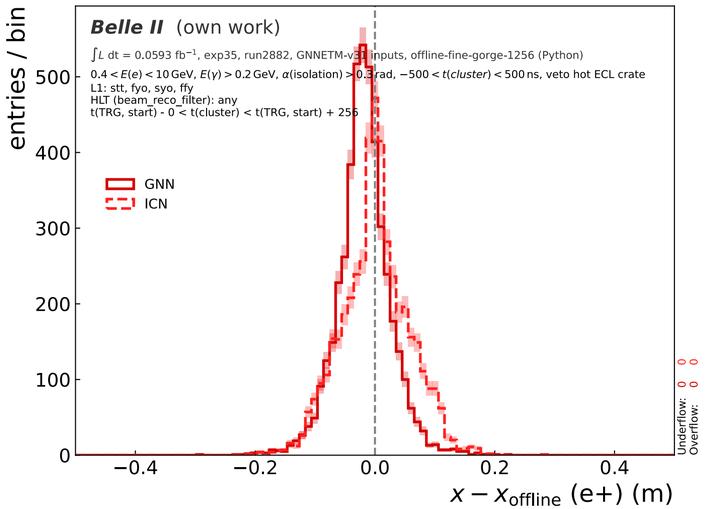
- For ICN clusters, the position of the ICN is calculated using the highest energy TC in the cluster
- It uses a mapping of TC to position that is not fully consistent with the offline shower positions, mostly a difference between crystal front face and crystal center
- After correcting this (offline), the ICN resolution is much better, but the GNN still outperforms the ICN
- In addition, we are using now track triggers for selection, and select the correct timing window (does not change the qualitative discussion, but the events are not identical)

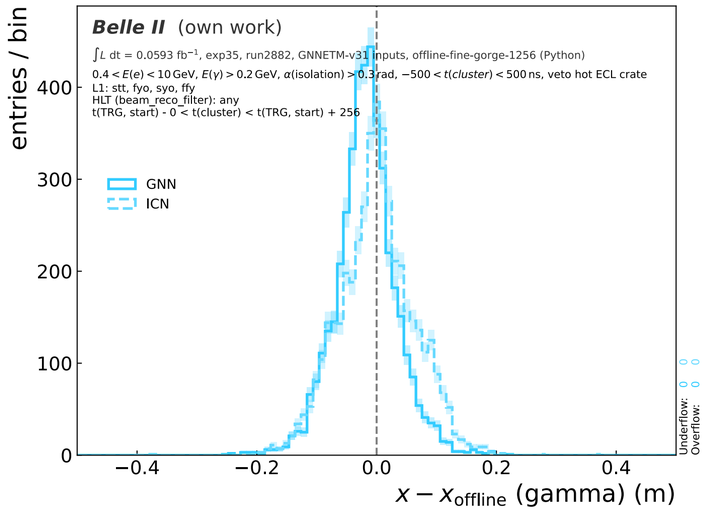
ICN position resolution after correction

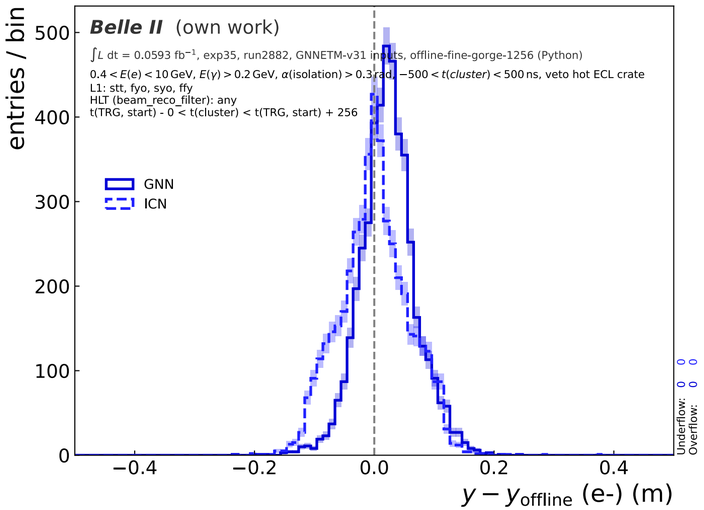
$e^+e^- \rightarrow e^+e^-\gamma$

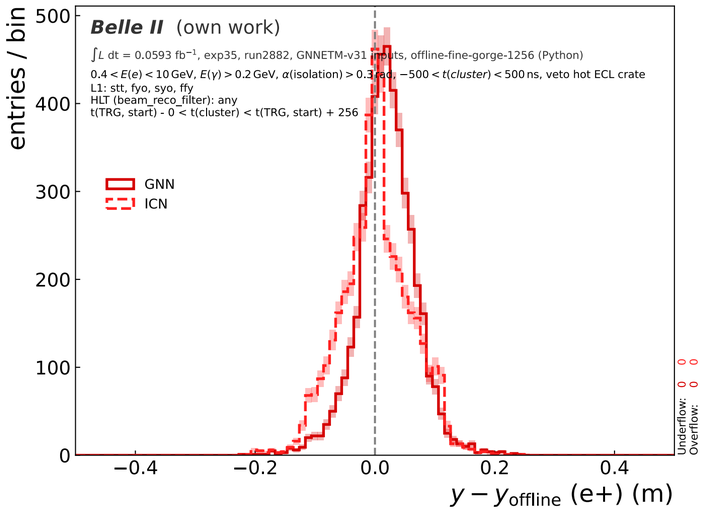
data

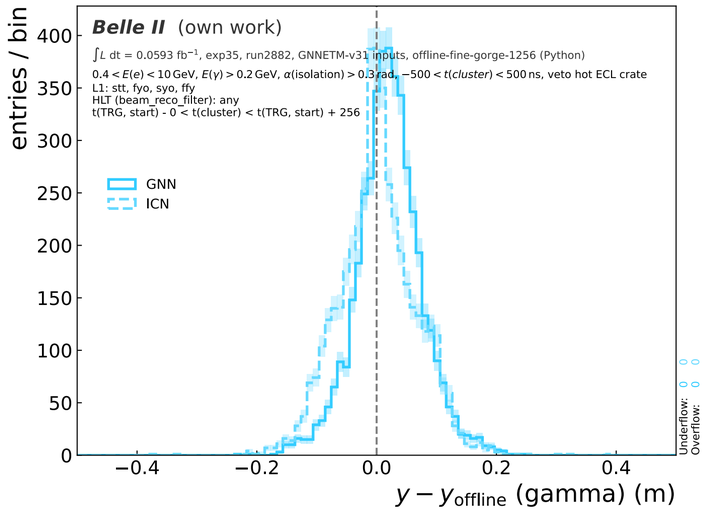






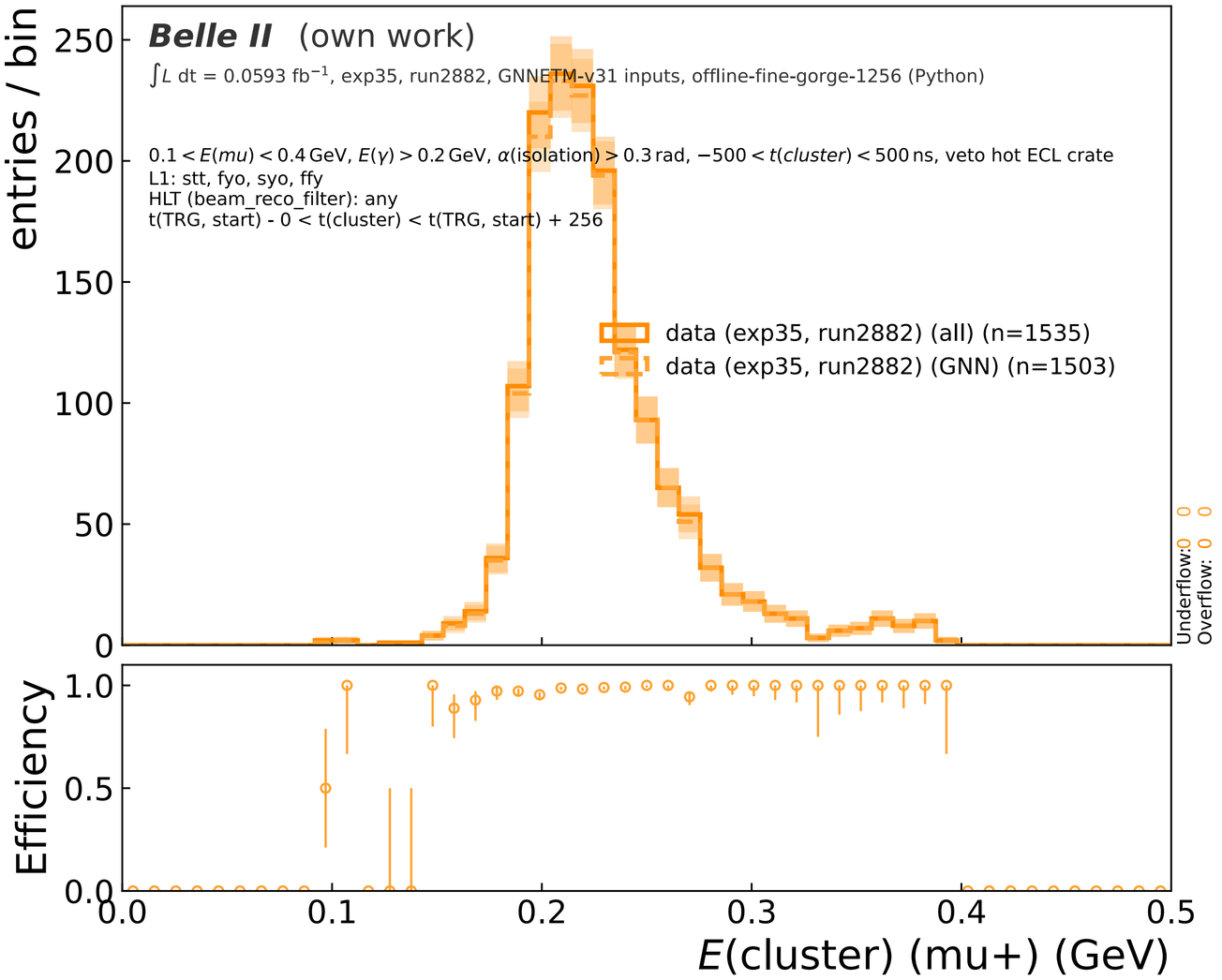






Efficiency

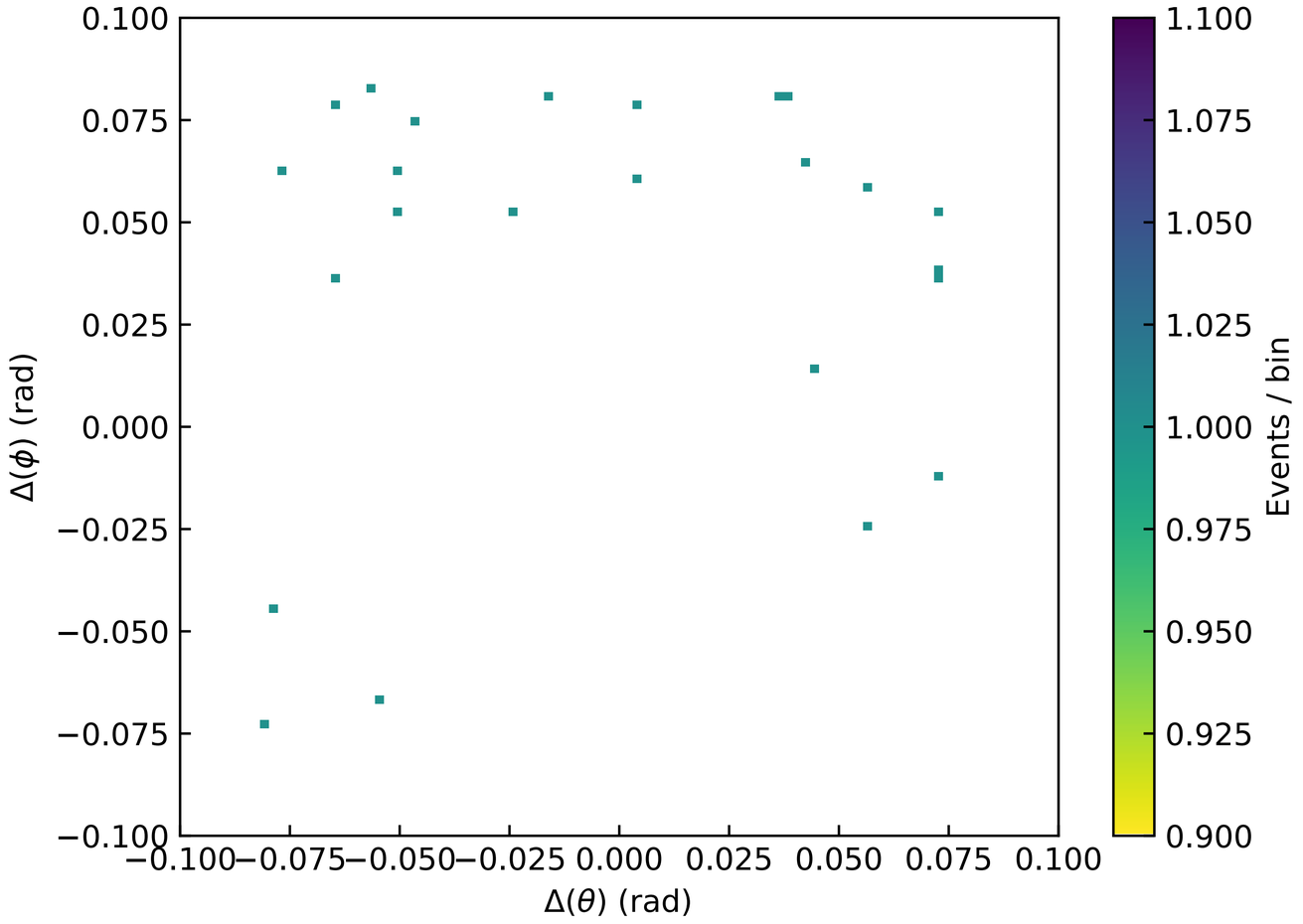
$$e^+e^- \to \mu^+\mu^-\gamma$$
$$data$$



- Of all 1535 events after selections, 32 are missing an matched GNN cluster (30 for ICN)
- Extrapolating the offline reconstructed track (!) into the ECL, and restricting the polar angle of the track to $0.6 < \theta < 2.0$ ($34 < \theta < 115°$), we are missing 24 GNN clusters (22 for ICN)

Efficiency

$e^+e^- \to \mu^+\mu^-\gamma$
data



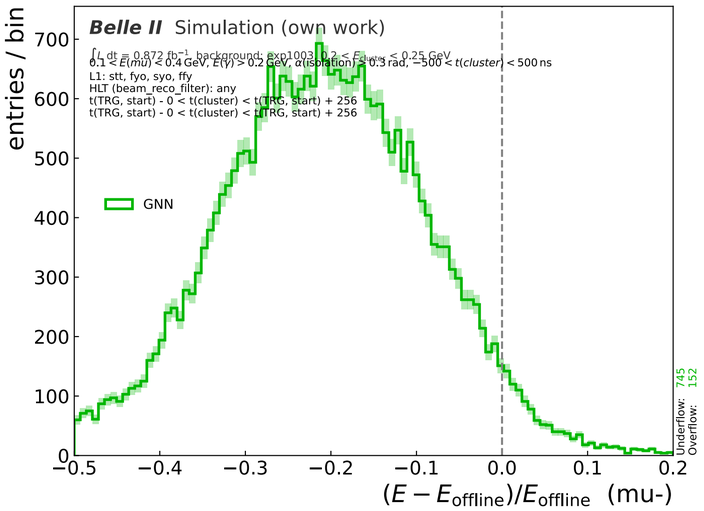
0.2 rad are about 4 crystals (==1 TC width) in the barrel

- The missing 24 GNN clusters are all located close to the edge of the trigger cell and the TC energy is below threshold
- This inefficiency is about 1-2% (for both GNN and ICN) and appears to be an intrinsic limitation for muons because of our TC energy thresholds

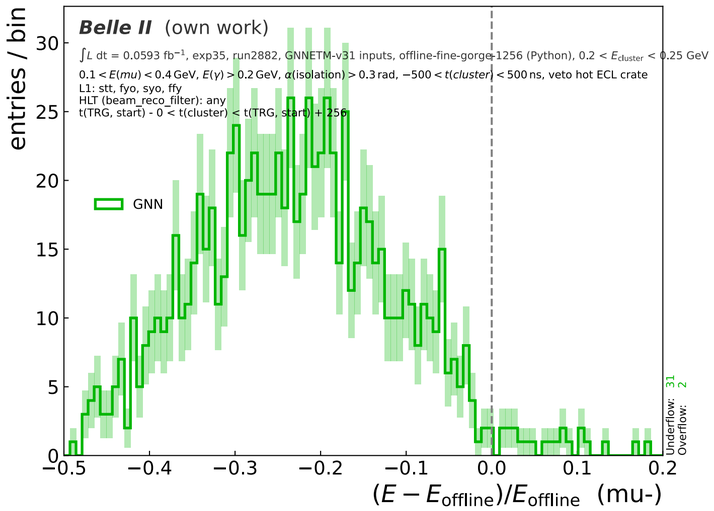
Energy resolution

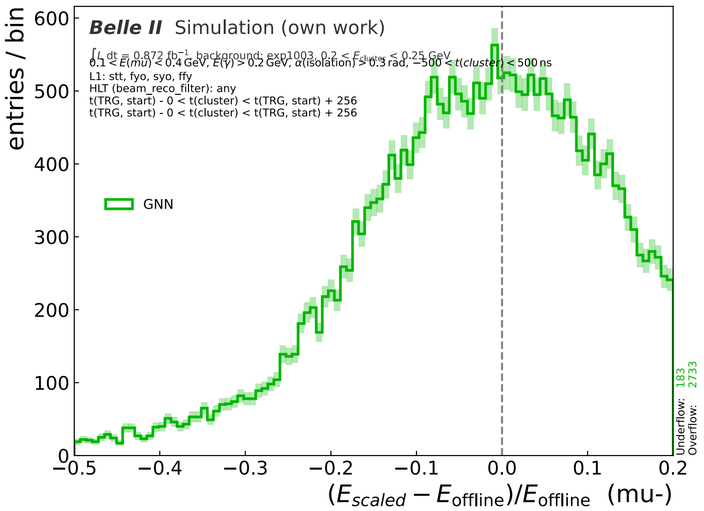
$e^+e^- \to \mu^+\mu^-\gamma$

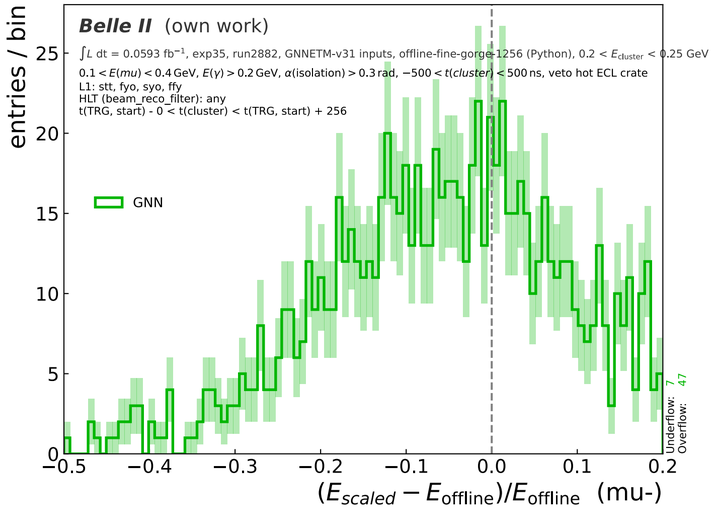
simulation (*exp*1003)



data



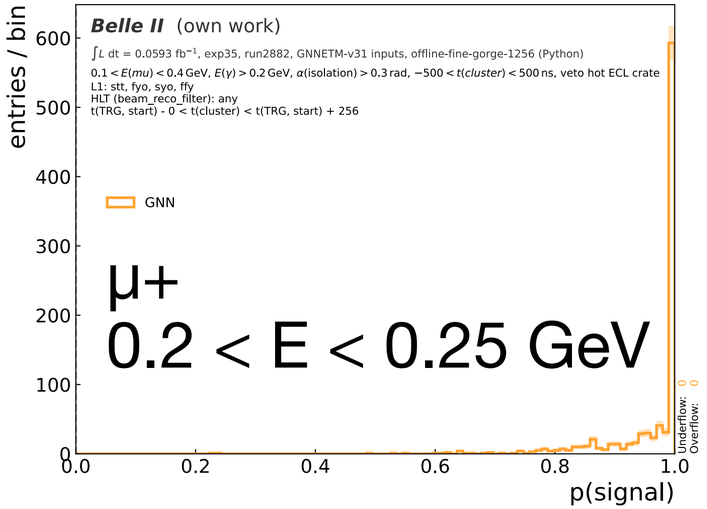


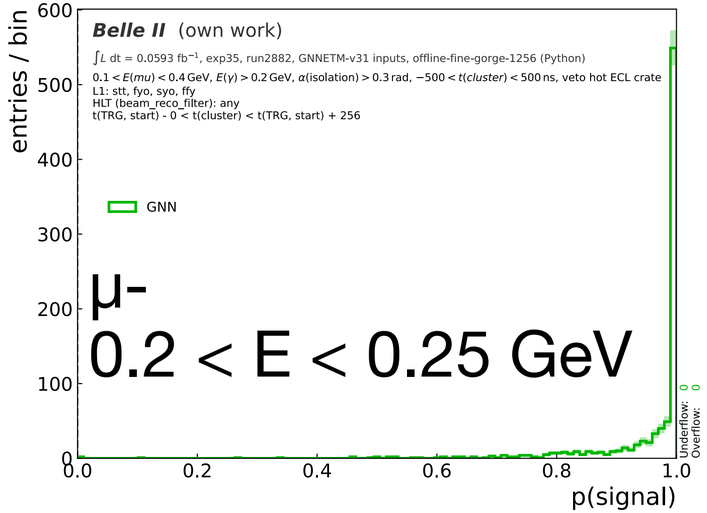


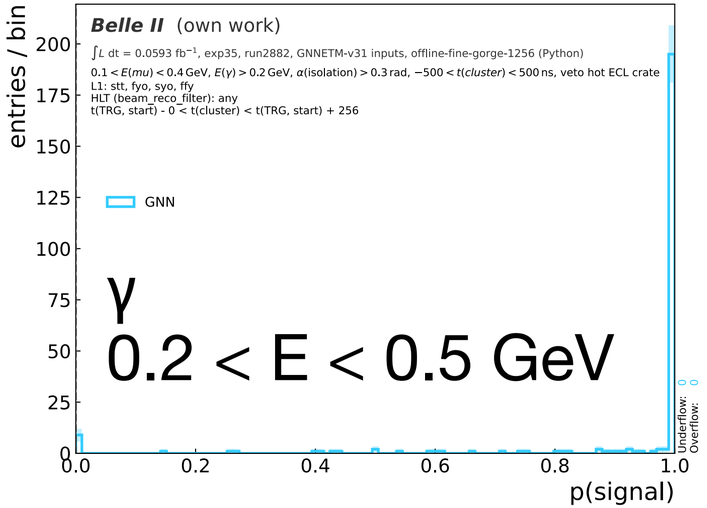
- Both data and MC show a comparable energy bias and energy resolution
- The energy resolution is poor compared to the ICN
- Highest priority on the algorithm side to get this under control until 2025a

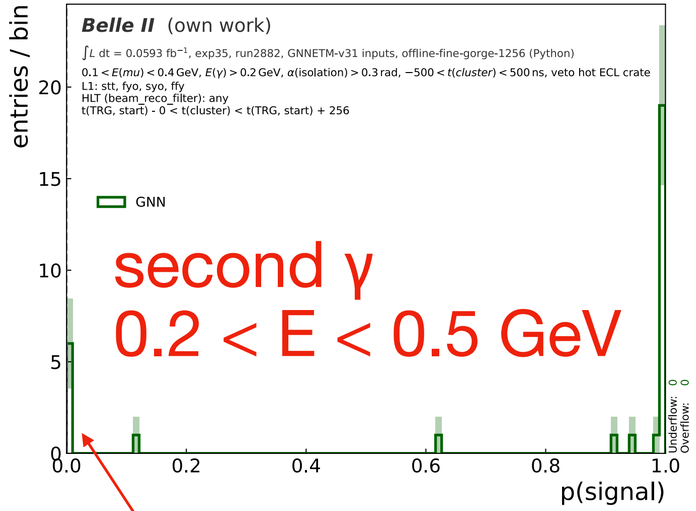
Signal classifier

$e^+e^- \to \mu^+\mu^-\gamma$









- The GNN has a signal classifier (→0 for background, → 1 for signal)
- Given the large 4×4 TCs, the algorithm probably mostly learns timing and position to distinguish between background and signal
- The “second gamma” does show indications of background 🤩

Conclusion

- The ICN position encoding needs a fix which improves offline-TRG agreement and hence our resolution metric. GNN clearly outperforms ICN for high energy position resolution even with the fix.
- The current TC threshold introduce a 1-2% inefficiency for muons entering close to TC edges.
- GNN energy resolution for low energy clusters is poor, but comparable in data and MC.
- Signal classifier validation is difficult, but very first studies do not look not promising (I don’t want to say they look promising just yet).
- Frank and Thomas, two new MSc-students at KIT are going to look deeper into this.