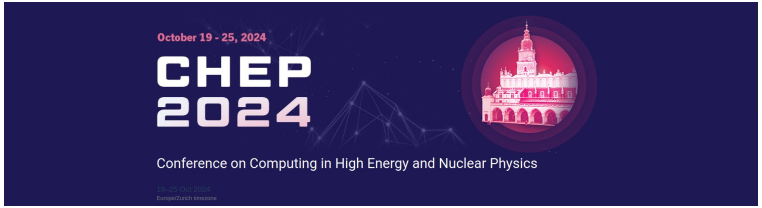
The CMS Phase-2 High Granularity Calorimeter Endcap Event Reconstruction with the TICL Framework

Wahid Redjeb (Rheinisch Westfaelische Tech. Hoch. (DE))
on behalf of the CMS collaboration

October, 22, 2024

CHEP 2024, Kraków, Poland



Bundesministerium
für Bildung
und Forschung

CMS

CERN

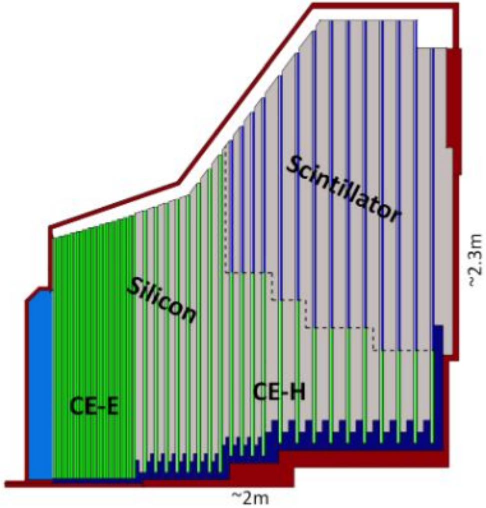
Physics
Institute III A

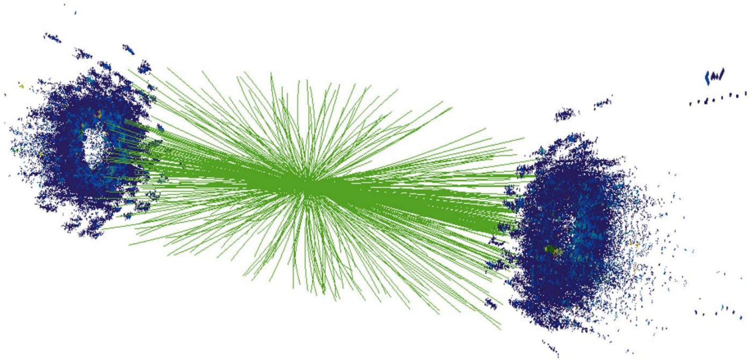
RWTH AACHEN UNIVERSITY

HGCAL

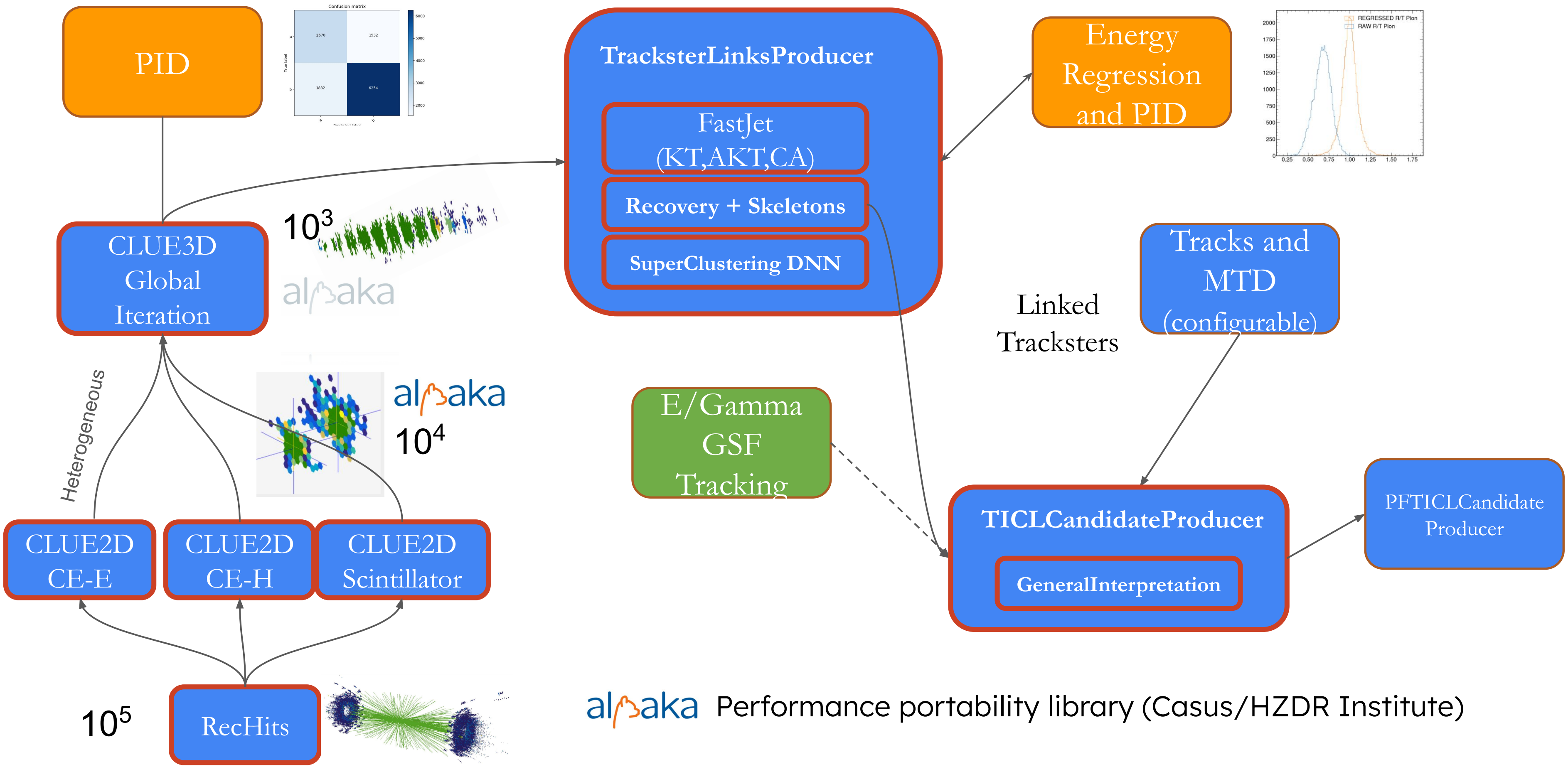
- **High Granularity Calorimeter** for 5D (t,E,x,y,z) particle shower reconstruction
  - For CMS Phase-2 Upgrade
- Electromagnetic Section (CE-E)
  - Si, Cu & CuW & Pb absorbers, 26 layers
- Hadronic Section (CE-H)
  - Si & Scintillator, stainless steel & Cu absorbers, 21 layers
- Cover 1.5 < eta < 3.0
  - Silicon Sensors (120/200/300 μm).
  - Plastic Scintillators with Silicon Photomultiplier (SiPM)

> 500k rechits per event!

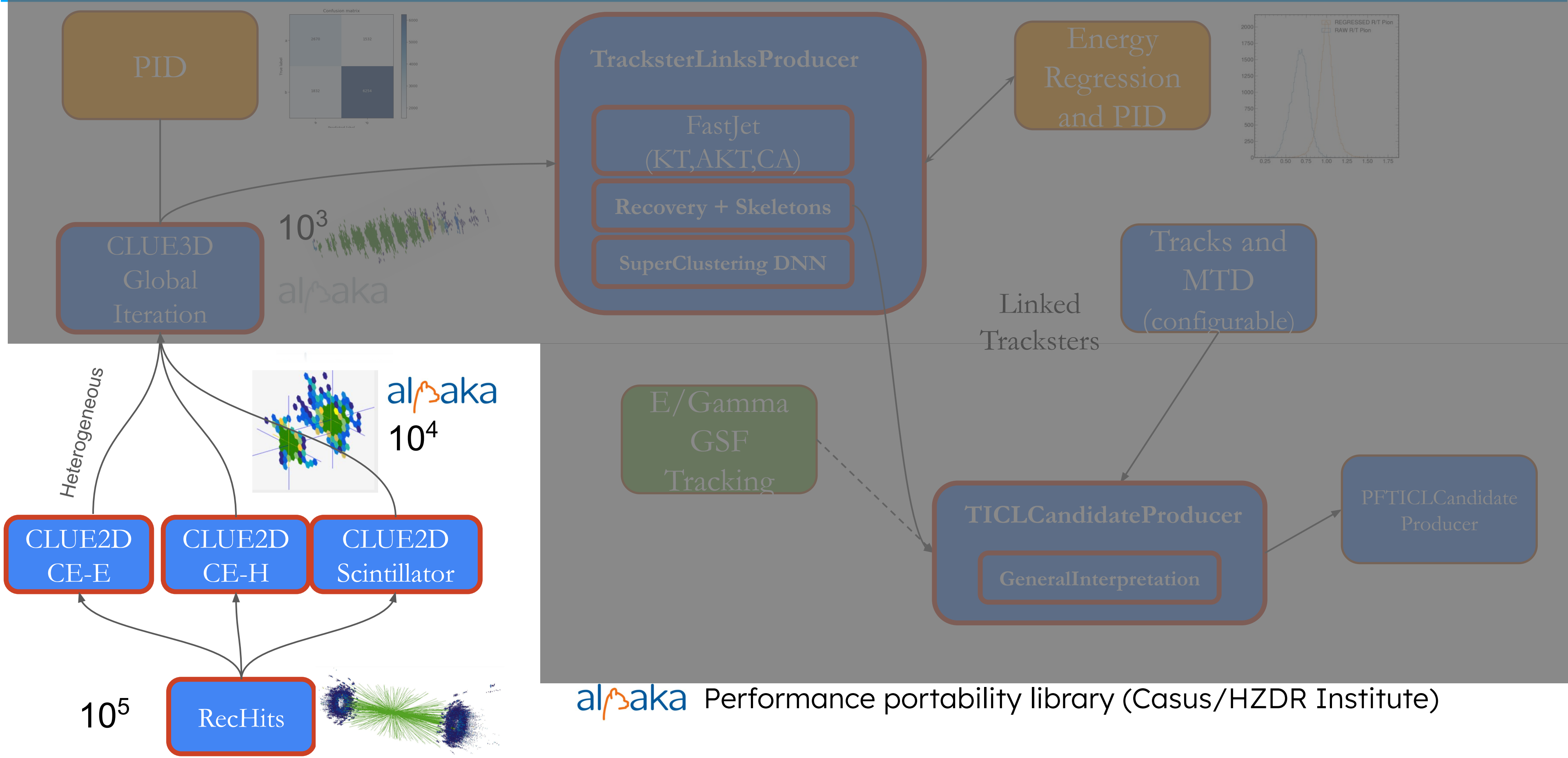




The Iterative CLustering (TICL) v5



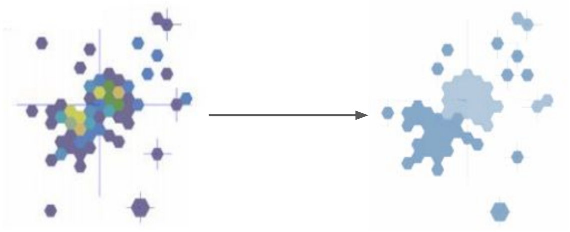
Layer Clustering by CLUE



alpaka Performance portability library (Casus/HZDR Institute)

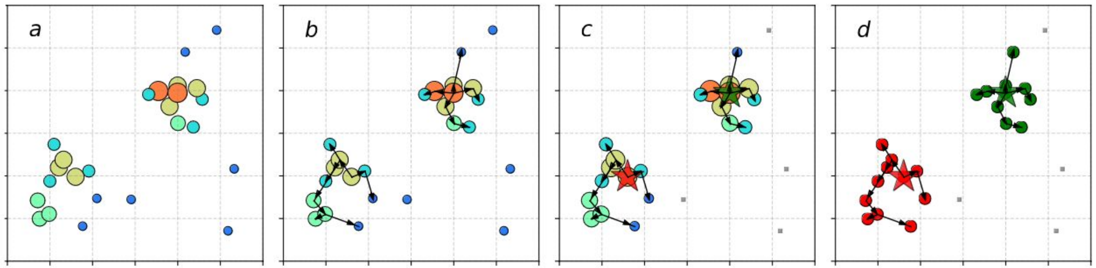
Layer Clustering by CLUE

- Removes Noise
- Reduces dimensionality of the problem
  - (10^5 hits to 10^4 layer clusters)
- Minimal loss of information
- GPU Friendly
  - Alpaka implementation
  - Build a chain of followers (b) and identify seeds and outliers based on density and distances values (c)
  - Cluster indices passed through this list (d)

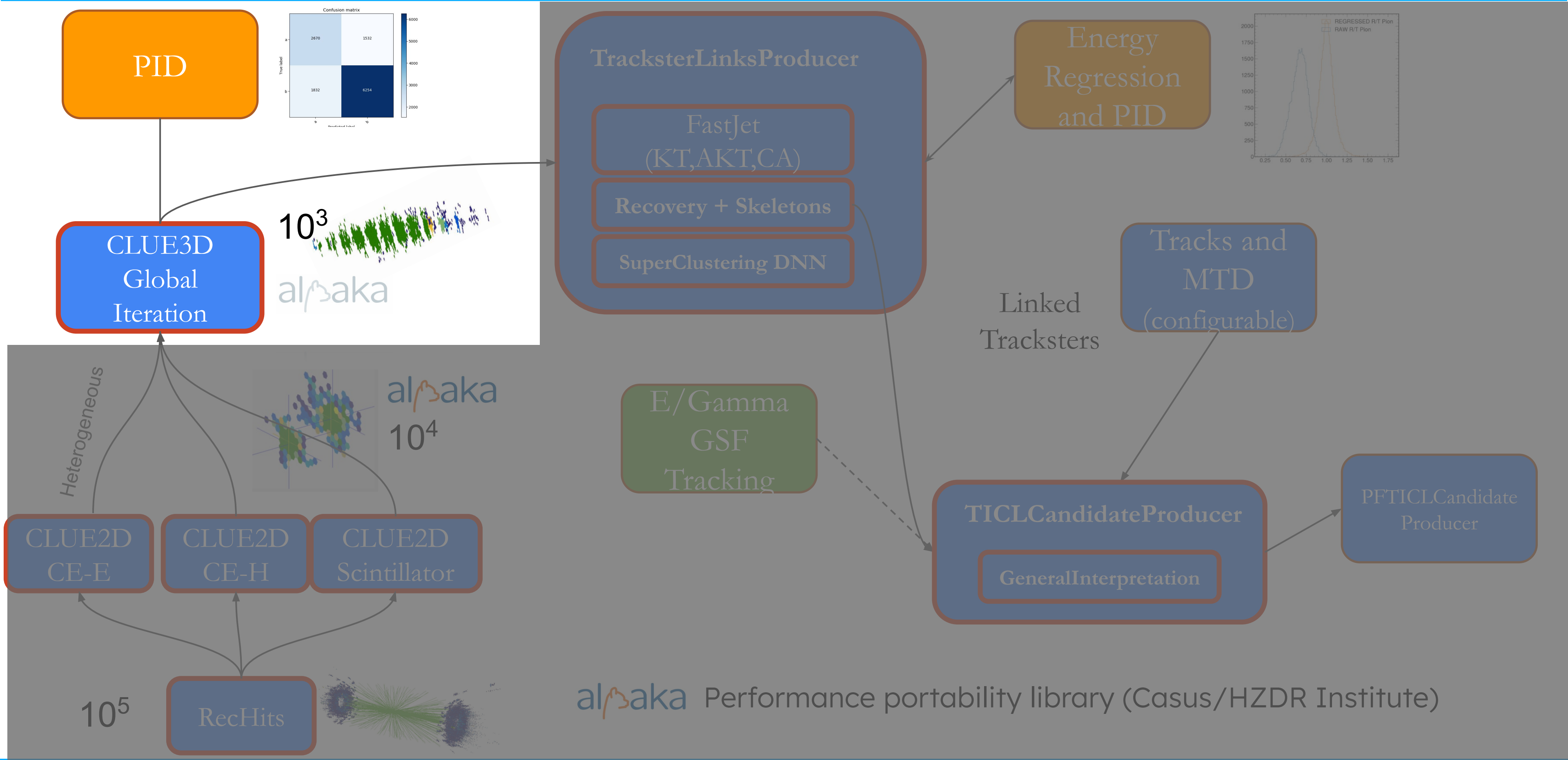


Rechits **Layer Clusters**

alpaka

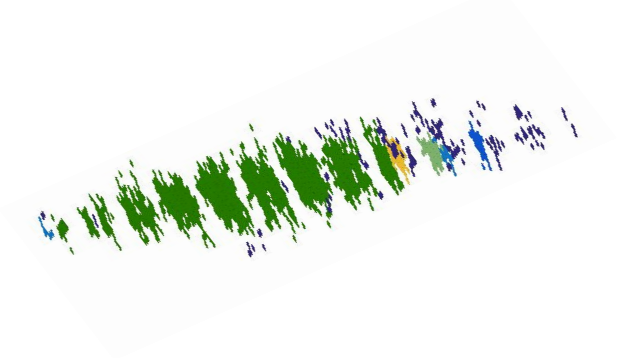


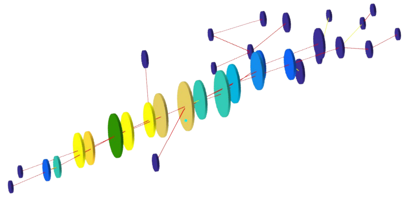
Pattern Recognition by CLUE3D



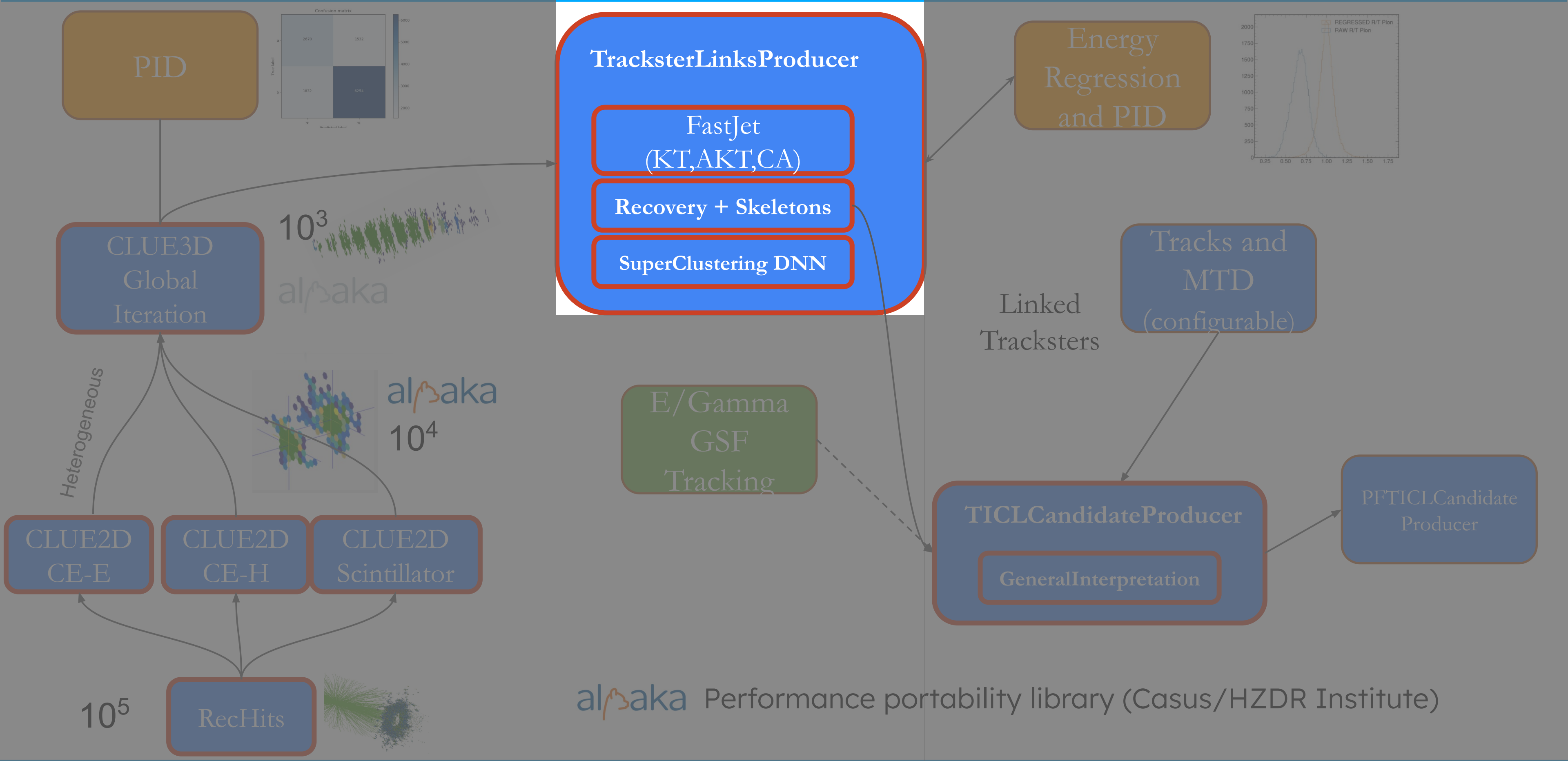
Pattern Recognition by CLUE3D

- **Pattern recognition** is another core part of the TICL Framework
  - Clusters Layer Clusters in 3D objects → Particles
- **Tracksters**: Direct Acyclic Graphs of Layer Clusters
- Several algorithms available:
  - Cellular Automaton
  - **CLUE3D: default one**
  - FastJet
- **CLUE3D**:, same logic as CLUE, but:
  - Input: Layer clusters instead of RecHits
  - Output: 3D Clusters instead of 2D
    - Considering longitudinal dimension
- Pattern Recognition can consume GPU-LCs
  - Reduce impact of legacy conversion (copy to CPU)





Linking Step



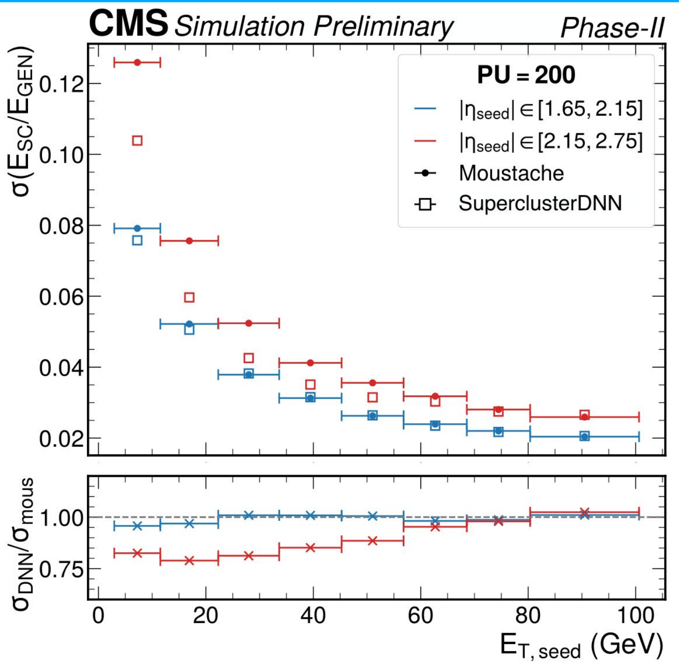
Linking Step - E/Gamma SuperClustering

Bundesministerium für Bildung und Forschung

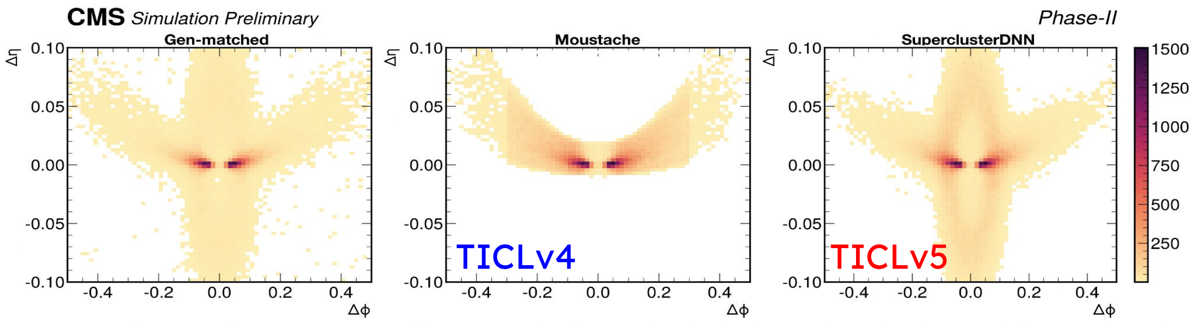
- A PID model is applied on Tracksters from CLUE3D
- Tracksters identified as Electromagnetic object are fed to the SuperClustering
- **TICLv4**: SuperClustering algorithm used in ECAL: **Mustache SuperClustering**
  - Prone to collect pileup (PU) if it falls in the seed $\Delta\eta$-$\Delta\Phi$ window
  - Designed for homogeneous calorimeter
- **TICLv5**: **DNN** to keep or reject Tracksters within the seed window
  - Seed chosen as Mustache algorithm → highest p_T trackster
    - Min seed p_T = 4GeV
    - Min trackster energy = 2GeV
    - Cut on explained variance on input trackster
  - Exploit **HGCAL shower variables**, mostly Trackster direction given by Principal Component Analysis (PCA)

Linking Step - E/Gamma SuperClustering

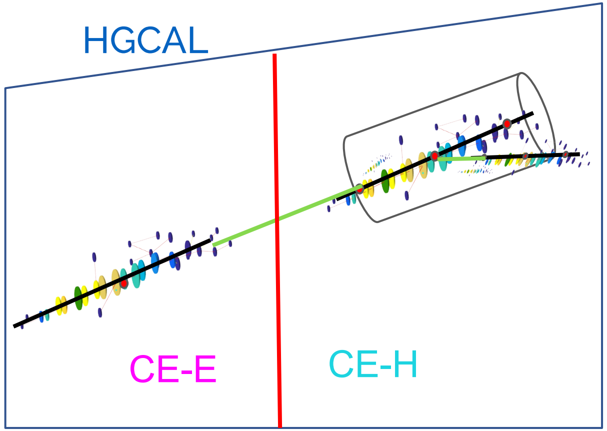
- **TICLv4**: SuperClustering algorithm used in ECAL: Mustache SuperClustering [6]
  - Classical Moustache Shape
- **TICLv5**: DNN to keep or reject Tracksters within the seed window
  - Clusters more objects around the seed
  - **More resilient against PU**
  - **More energy collection**
  - **Better energy resolution**
    - Specially in high η region where the PU occupancy is higher



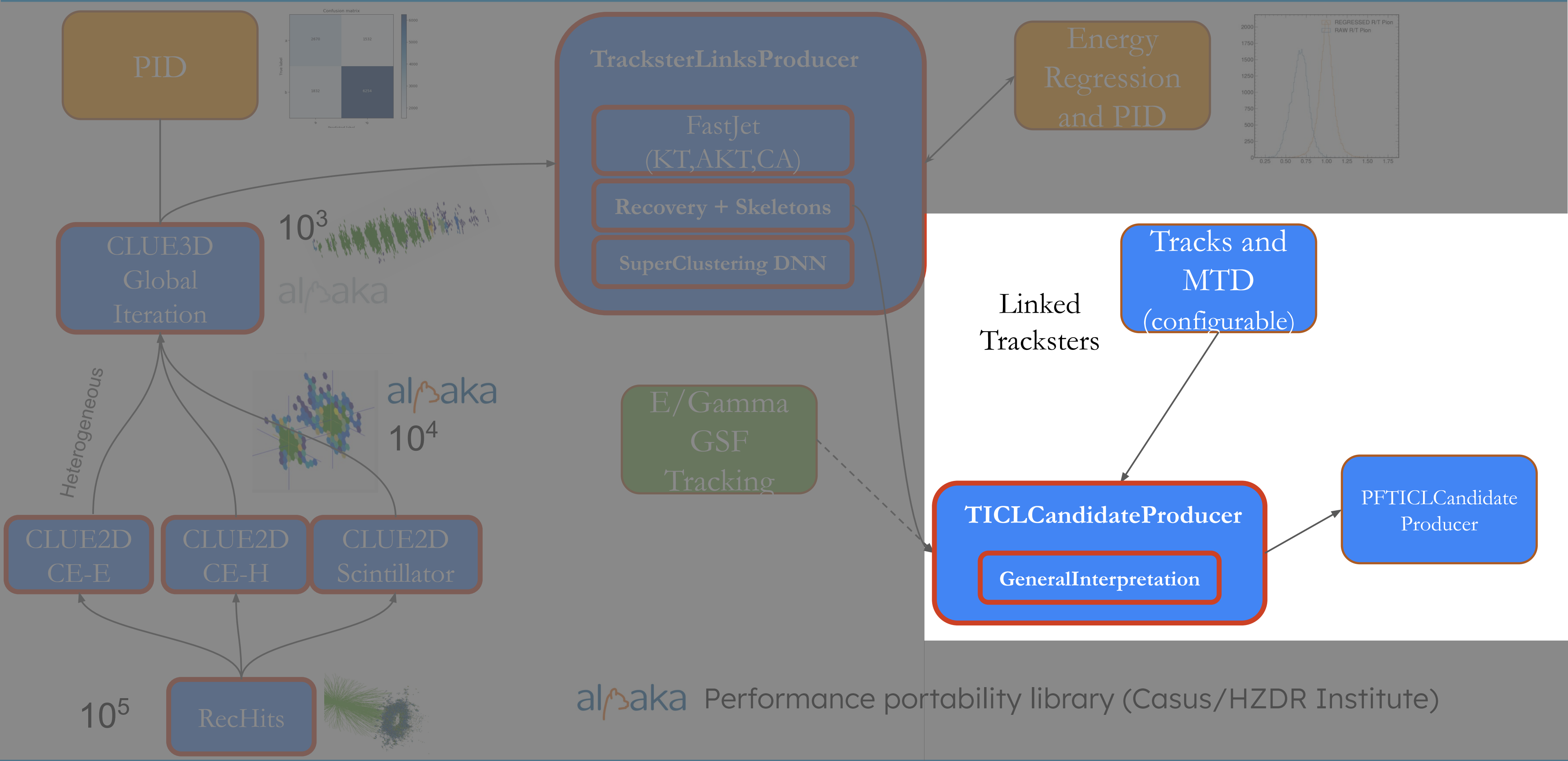
PU200



- CLUE3D tuned for High Purity Clusters, minimizing contamination of PU
  - Linking: Additional step needed to recover splitting of CLUE3D
- **TICLv4**: Introduction of Geometrical Linking
  - Works well in Linking big aligned clusters, but struggles in linking misaligned objects
- **TICLv5**: New Linking plugin system, with the following algorithms available
  - **Linking by FastJet**
  - **Linking by Skeletons** (default): Improved Geometrical Linking
    - Exploit topology of CLUE-3D Tracksters and PCA axes
      - Allow links between tracksters
      - Allow recovery of unclustered energy deposits
      - Looks around shower to cluster small fragments
- Energy Regression and PID on final object

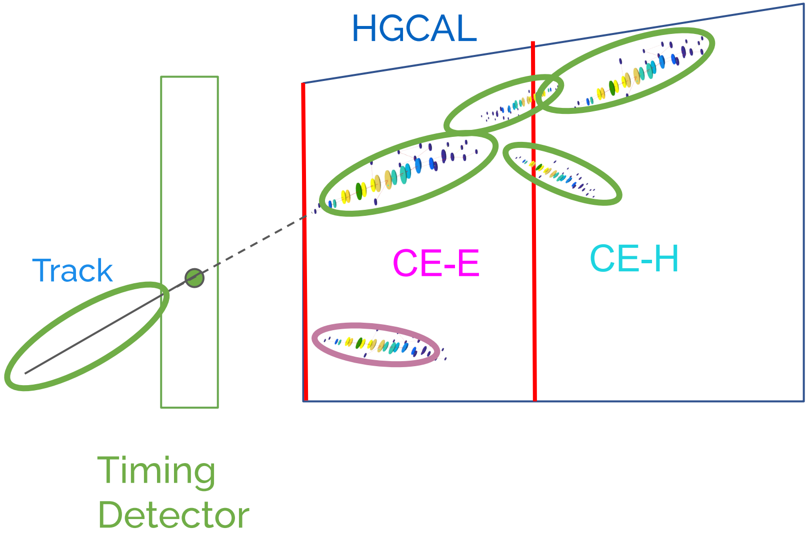


Particle Flow Interpretation

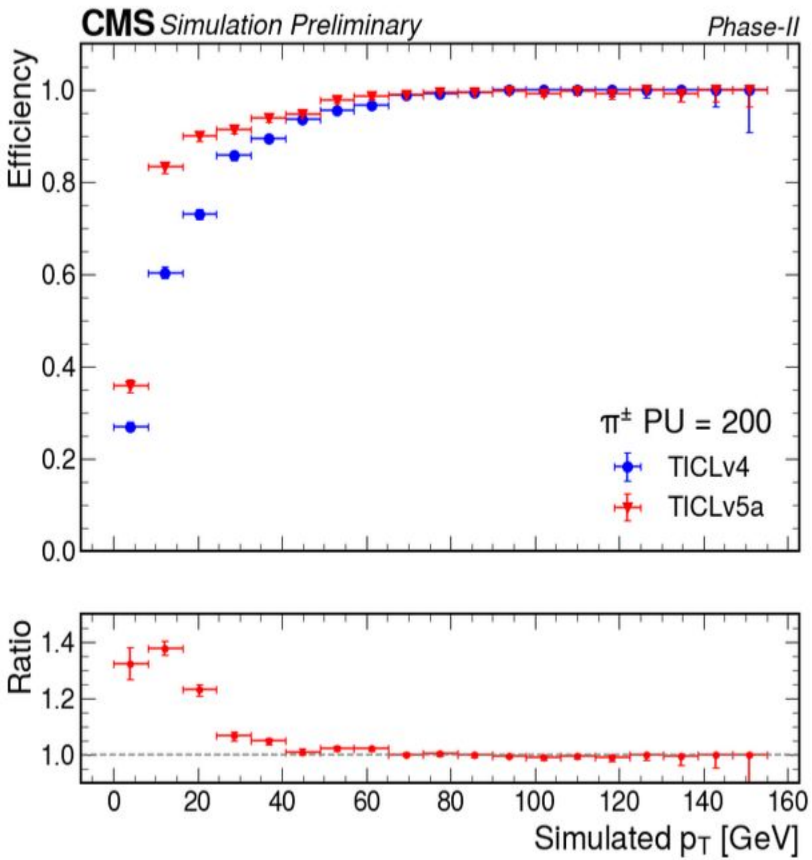


Particle Flow Interpretation

- Given the linked Tracksters we can also exploit the information from the Tracks
  - **Combine calorimeter and track energy**
  - **Exploit time information associated with Track**
- Tracks entering in the calorimeter are checked for position/energy and time compatibility with a calorimeter object
  - If multiple Tracksters are linked to the same Track
    - Link them together
- **Create Final Particle Flow Candidates**
  - All the kinematic information
  - Particle probabilities
  - Timing



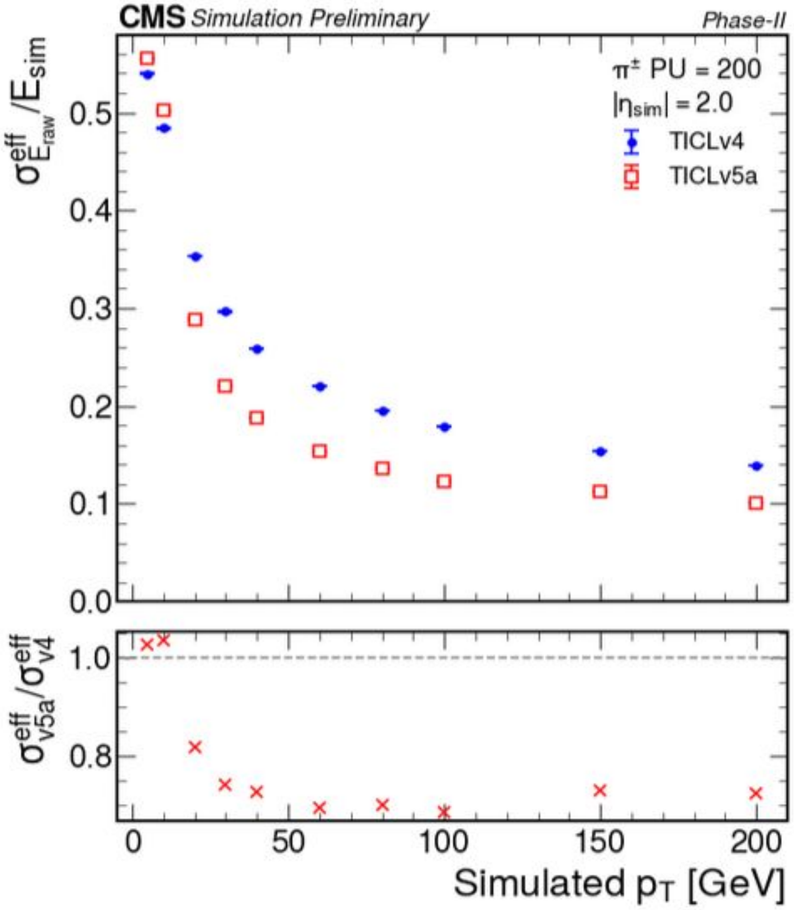
Hadronic Linking - Physic Performance



New Linking algorithm brings a better energy collection for hadronic showers

- quite big recovery in terms of efficiency

Hadronic Linking - Physic Performance

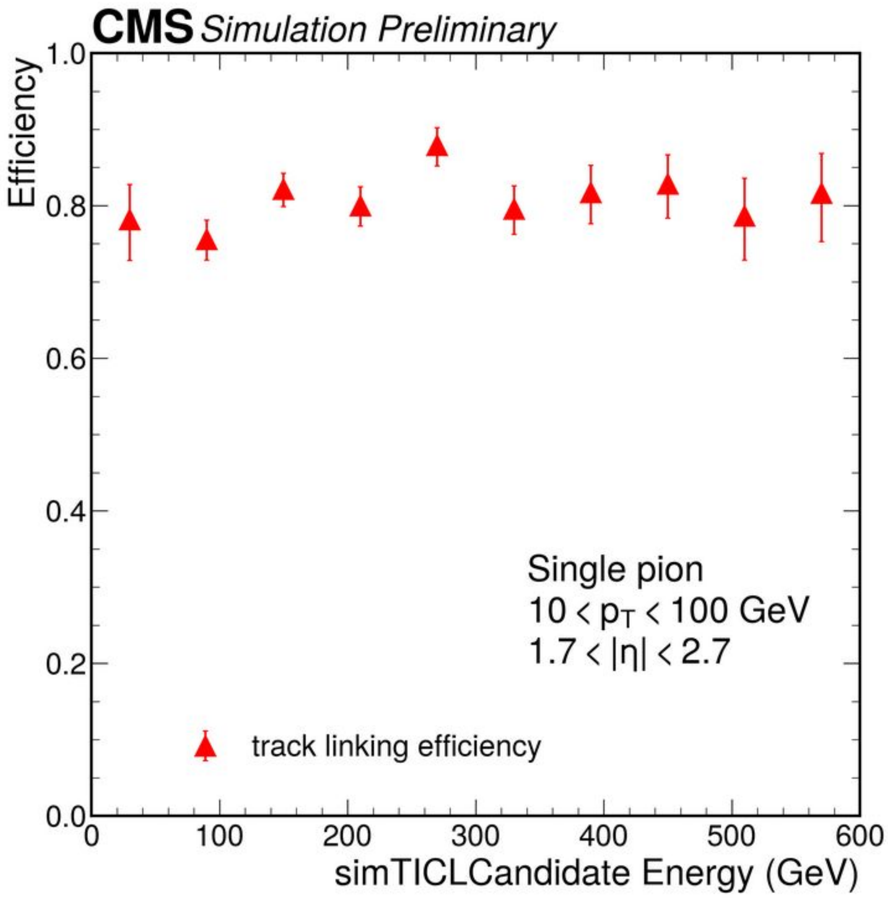


New Linking algorithm brings a better energy collection for hadronic showers

- quite big recovery in terms of efficiency
- better raw energy resolution

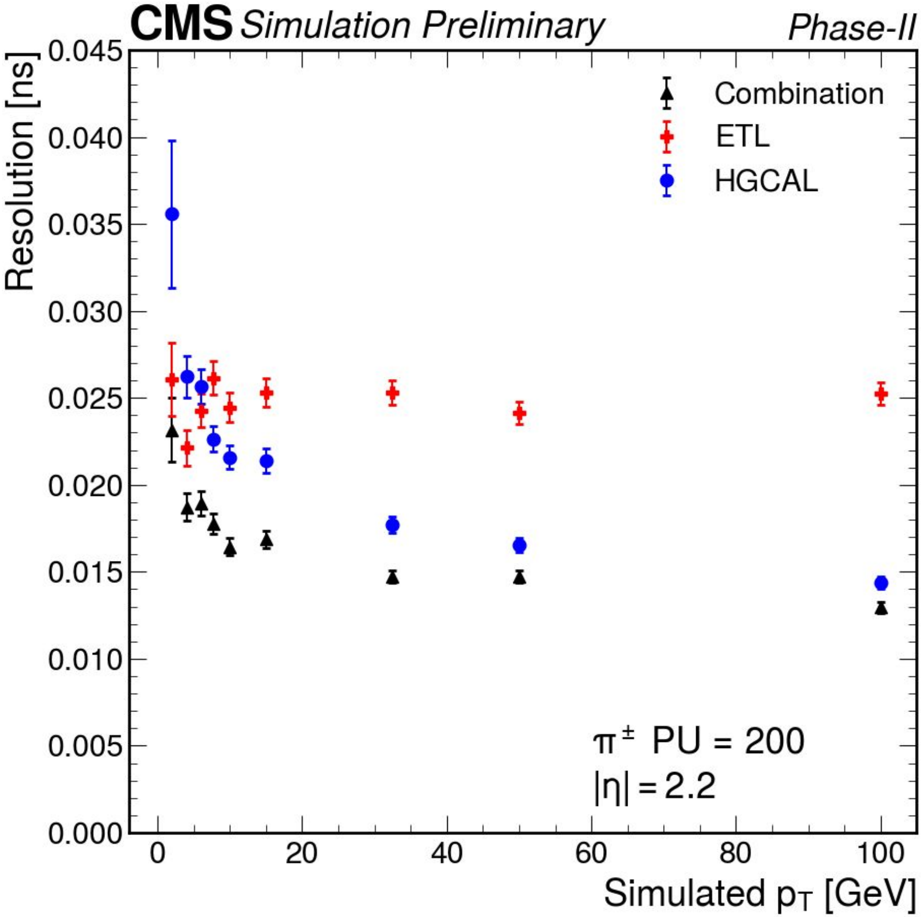
Time information

- With the Track-Trackster linking step we have
  - **HGCAL time available**
  - **Endcap Timing Layer Time available**
- Combination of HGCAL Time and ETL Time
  - Time is then **propagated back to the primary vertex**
    - For Charged: Following Track trajectory and using speed hypothesis given by ETL
    - For Neutral: Straight line at speed of light
- **~80% Track-Trackster linking efficiency**

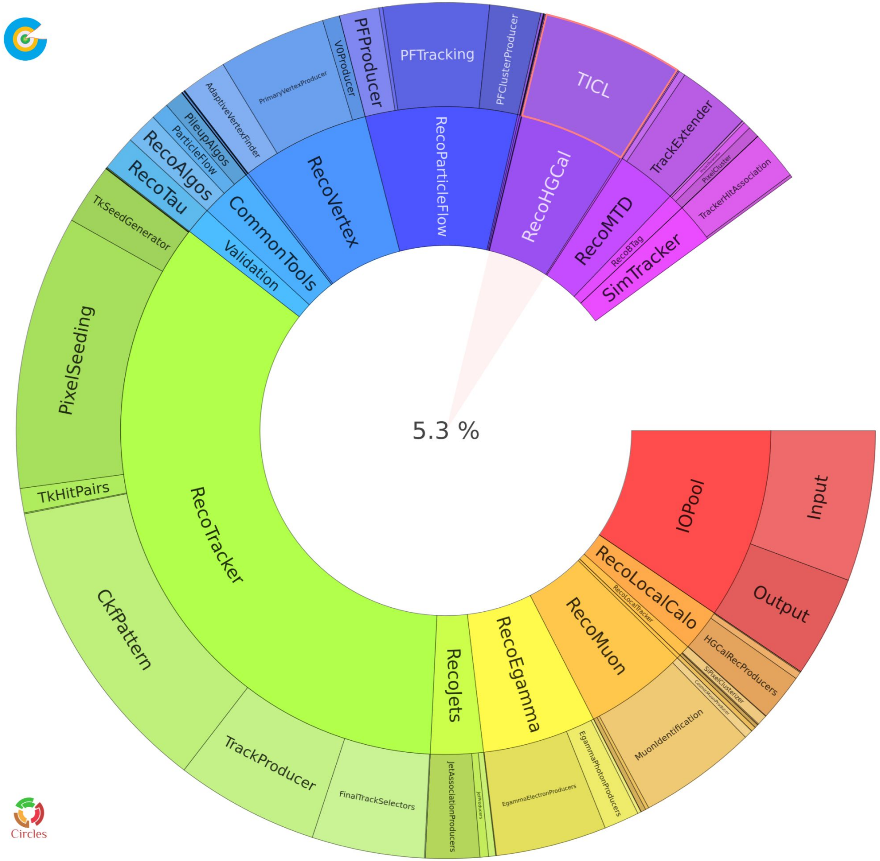


Time information

- With the Track-Trackster linking step we have
  - **HGCAL time available**
  - **Endcap Timing Layer Time available**
- Combination of HGCAL Time and ETL Time
  - Time is then **propagated back to the primary vertex**
    - For Charged: Following Track trajectory and using speed hypothesis given by ETL
    - For Neutral: Straight line at speed of light
- **~15ps time resolution in 200PU**
  - Can be exploited for PU mitigation and isolation



Computing Performance



Real time per event measured with CMS software running on sample of ttbar events with PU 200 on AMD EPYC 9454.

→ Currently taking only 5% of the entire Phase-2 CMS Reconstruction

- Core algorithms have been already ported to GPU and will be offloaded on GPUs
  - Expected additional reconstruction time reduction

Summary and Conclusions

- TICL is a framework developed for HGCAL event reconstruction and currently being extended to the barrel as new Particle Flow Framework for Phase-2
- Performs reconstruction from RecHits up to the Particle Flow Reconstruction and Interpretation
    - Different treatment of EM and Hadronic objects
- Designed with Heterogeneous Computing in mind
    - Alpaka version of the algorithms is being integrate in the CMS Software (CMSSW)
- Fast reconstruction!
    - And expected to be faster with more GPU offload

This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research (grant no. 13E18CHA)

Bundesministerium
für Bildung
und Forschung