



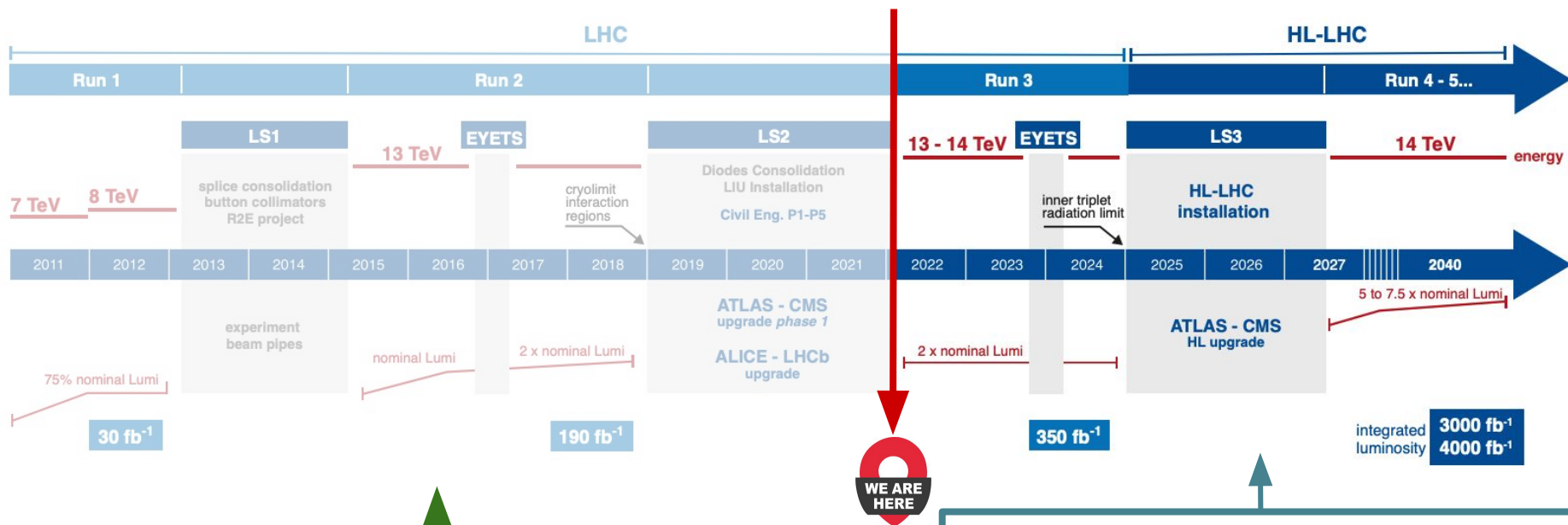
# Challenges and novel reconstruction techniques for the CMS High Granularity Calorimeter for HL-LHC



**Shamik Ghosh (LLR)**  
on behalf of the CMS Collaboration  
Lepton - Photon 2021 (Manchester)  
January 11, 2022



# Motivation: HL-LHC

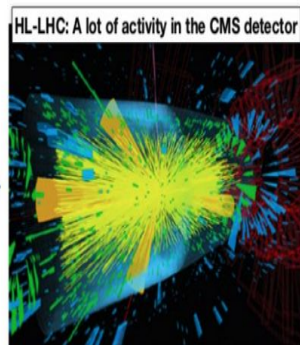
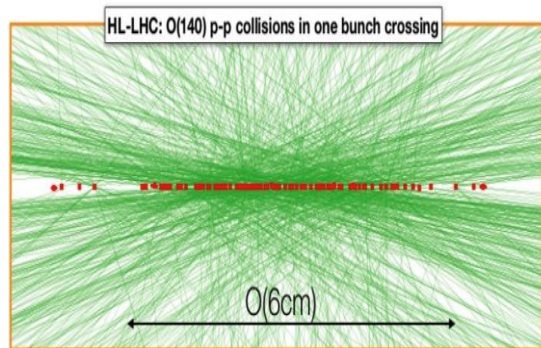
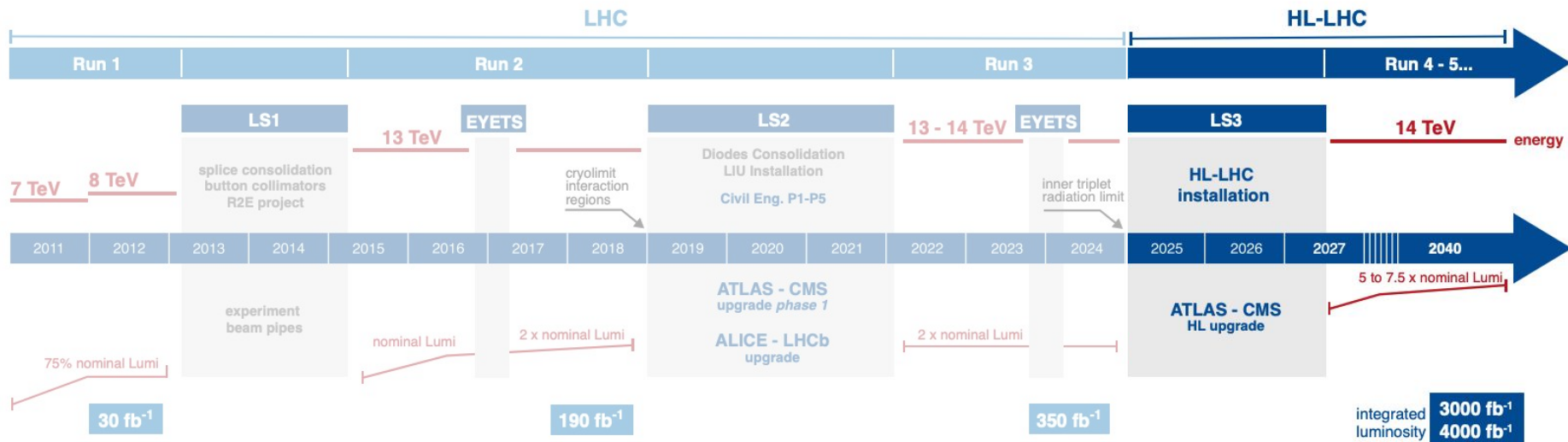


- Remarkable performance so far exceeding initial expectations
- But, things have just begun with ~5% of total expected data collected

## Significant increase in instantaneous luminosity

- $5 \times 10^{34}$  ( $7.5 \times 10^{34}$ ) cm<sup>-2</sup> s<sup>-1</sup> for 140 (200) PU in Run 4 (Run 5)
- Opportunity for Higgs boson precision studies, precision SM tests and BSM searches

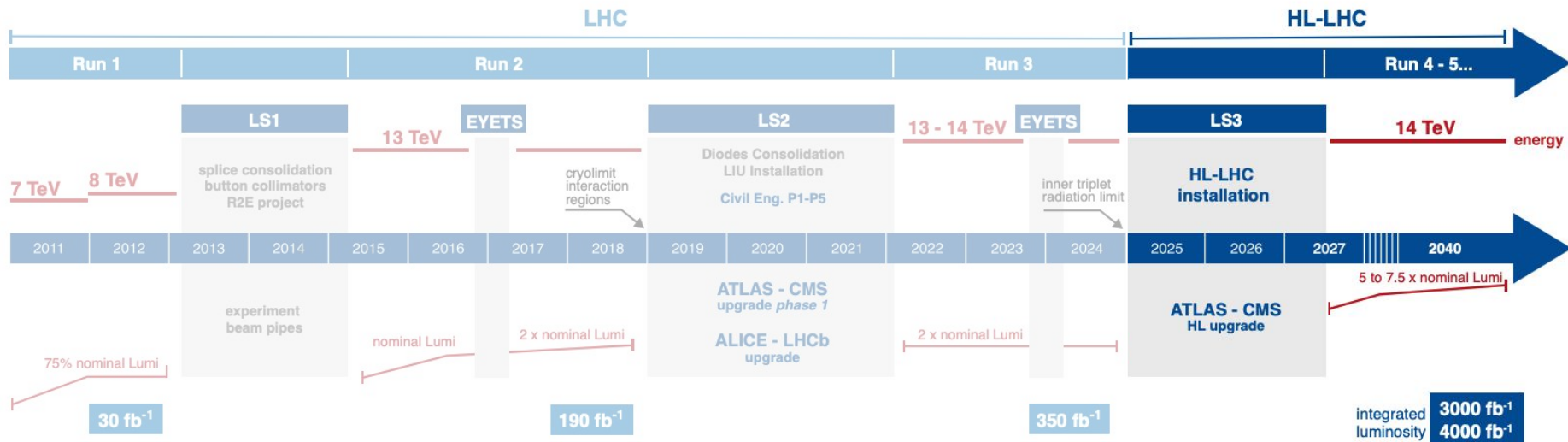
# Motivation: HL-LHC



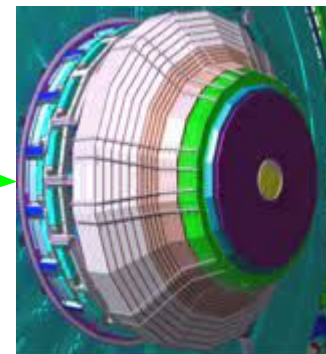
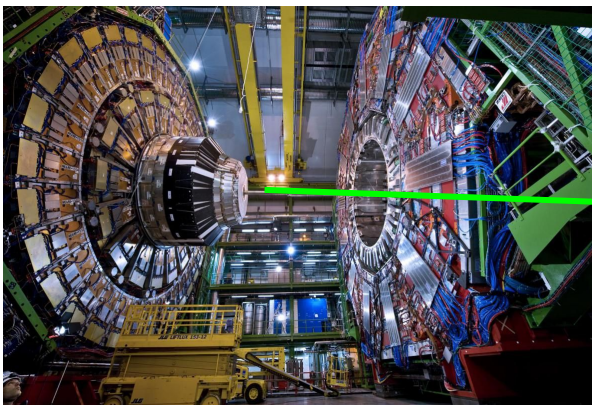
## But we have to pay to play!!

- High Pileup
  - ~200 collisions/BX (4-5x LHC)
- High Radiation Level
  - 1y @HL-LHC ~ 10 y @LHC

# Motivation: HL-LHC

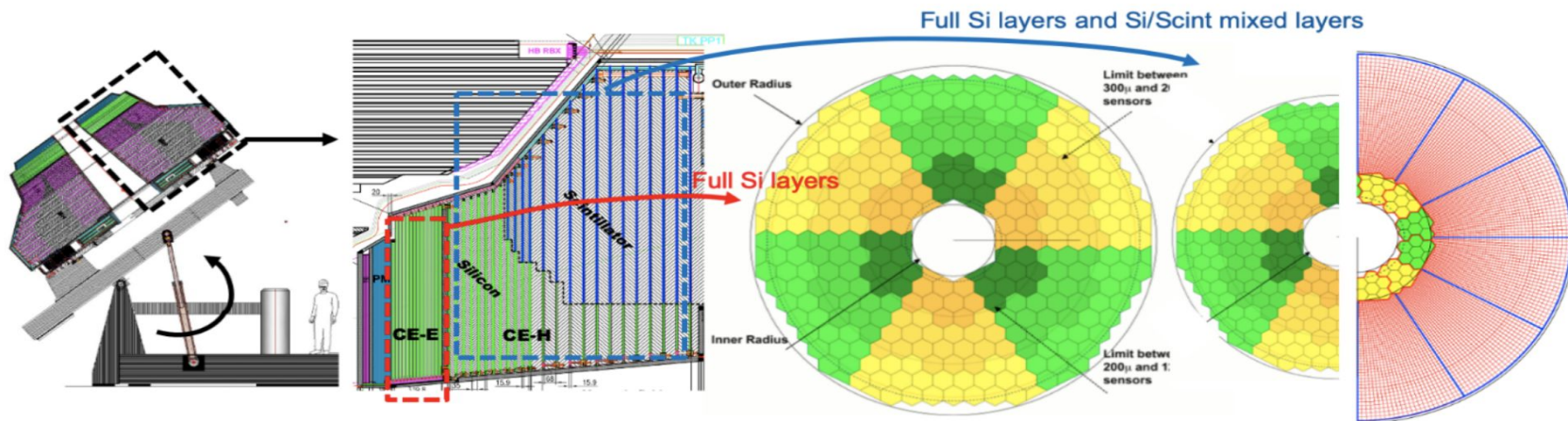


Existing endcap calorimeters will suffer the most → Replace with HGCal



# CMS Phase-II Upgrade Endcap Calorimeter

**High Granularity Calorimeter (HGCal):** granular and radiation hard endcap calorimeter replacement



## Calorimeter Endcap Electromagnetic (CE-E)

- EM focused part

### Active material

- 26 Layers of Si (cell size:  $0.5\text{-}1\text{ cm}^2$ )

### Passive material

- Pb, CuW, Cu
- $27.7\text{ }X_0$

## Calorimeter Endcap Hadronic (CE-H)

- HAD focused part (hybrid structure)

### Active material

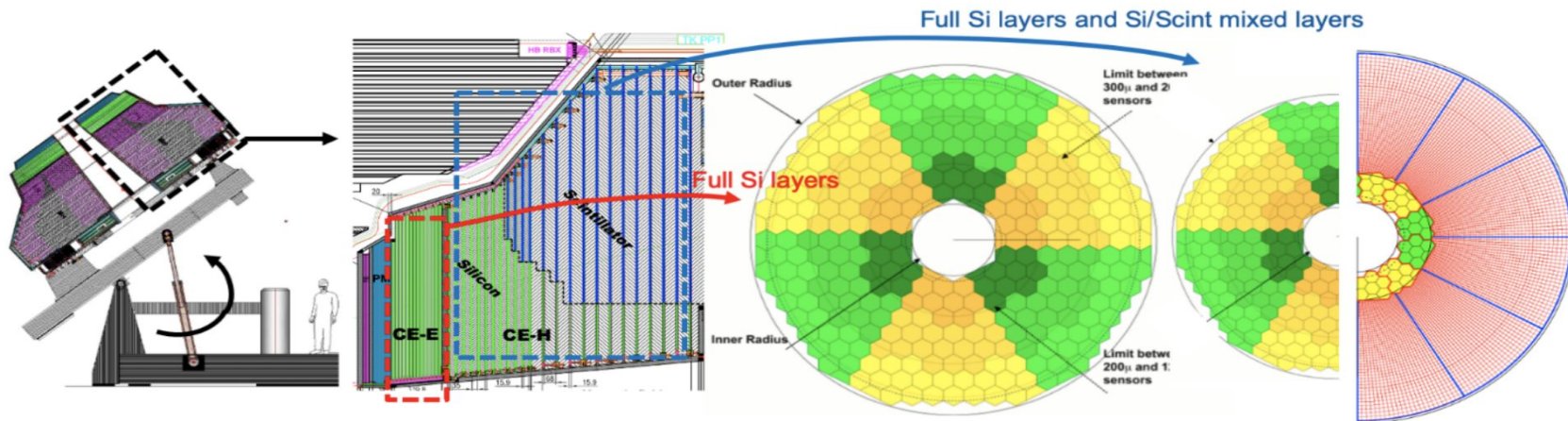
- 7 Layers of Si (cell size:  $0.5\text{-}1\text{ cm}^2$ )
- 14 Layers of Si and plastic scintillator

### Passive material

- Stainless Steel, Cu
- $10.0\text{ }\lambda$

# CMS Phase-II Upgrade Endcap Calorimeter

**High Granularity Calorimeter (HGCal):** granular and radiation hard endcap calorimeter replacement



## Calorimeter Endcap Electromagnetic (CE-E)

- EM focused part

### Active material

- 26 Layers of Si (cell size:  $0.5-1 \text{ cm}^2$ )

### Passive material

- Pb, CuW, Cu
- $27.7 X_0$

**~ 6M Si sensor channels**

## Calorimeter Endcap Hadronic (CE-H)

- HAD focused part (hybrid structure)

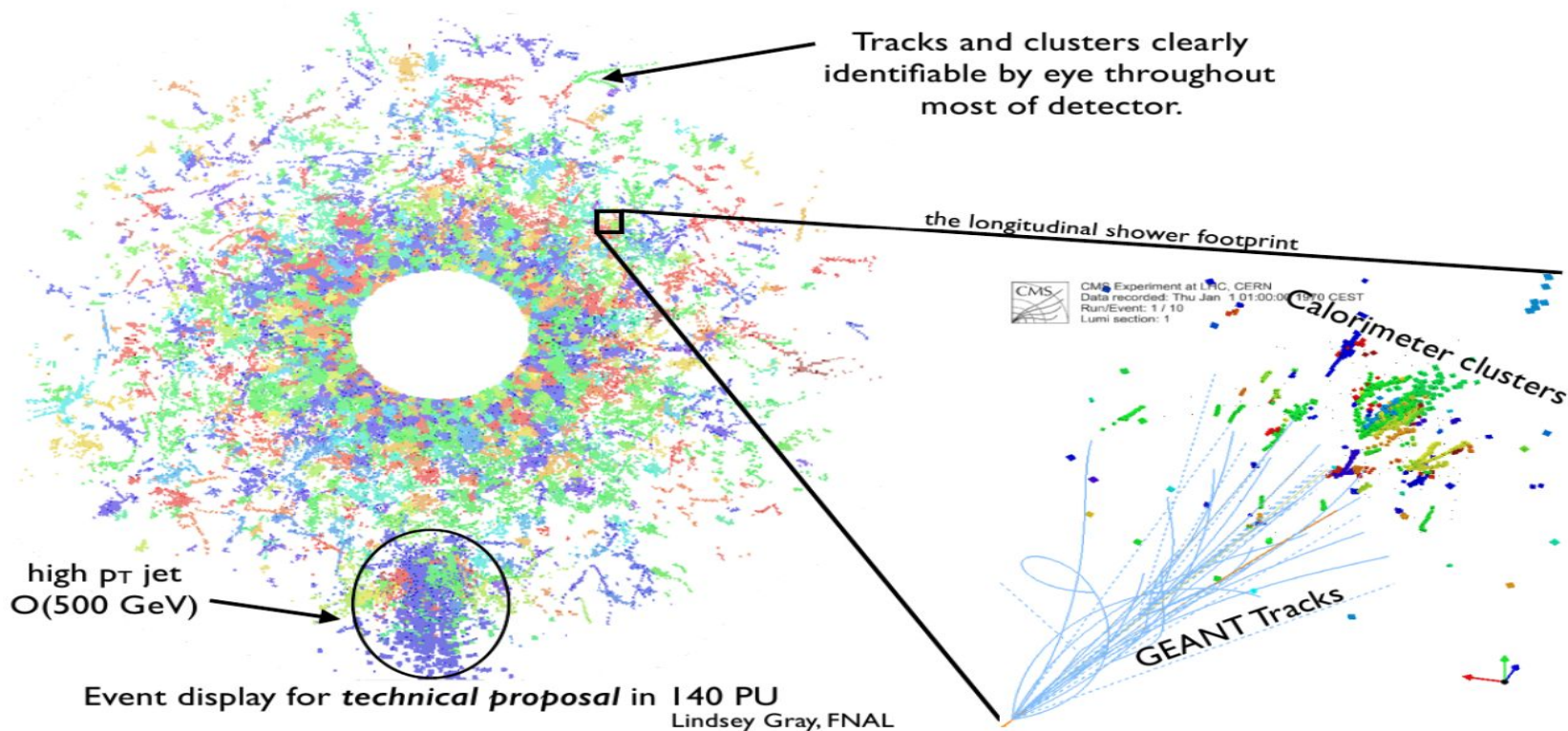
### Active material

- 7 Layers of Si (cell size:  $0.5-1 \text{ cm}^2$ )
- 14 Layers of Si and plastic scintillator

### Passive material

- Stainless Steel, Cu
- $10.0 \lambda$

# Reconstruction in HGCAL



# Reconstruction in HGCal

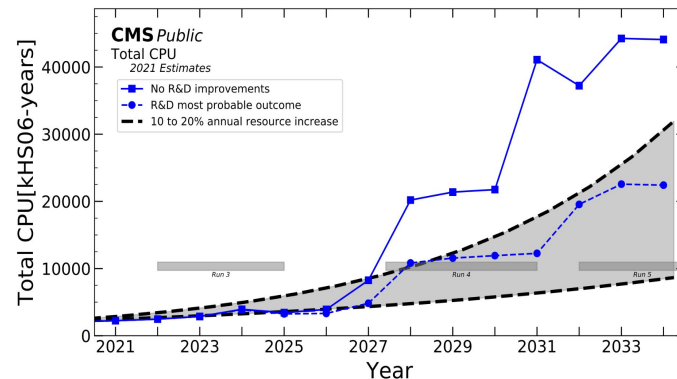
Reconstruction in HGCal is a real challenge due to the granularity and high PU environment

- **Imaging calorimeter** with very fine lateral and longitudinal segmentation, and precision timing capabilities
- Naive reconstruction algorithms based on considering all possible combinatorics lead to **memory/timing explosion**
- Overlapping showers are frequent in high PU and require efficient algorithms to **disentangle** them

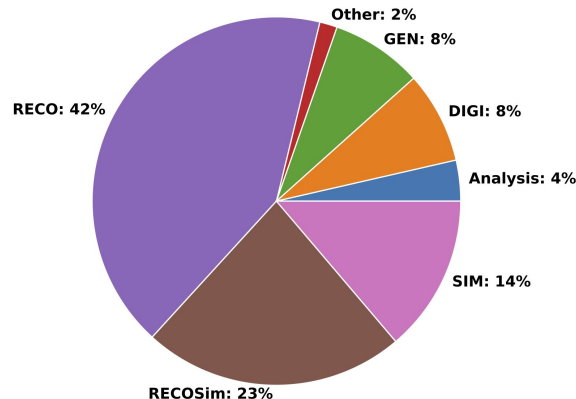
New techniques and algorithms to extract signals belonging to individual showers and properly identify them (clustering, linking, particle identification)

- Utilise modern computer architectures, graph theory, machine learning etc

Efficient workflows to utilise information from the tracking and timing detector

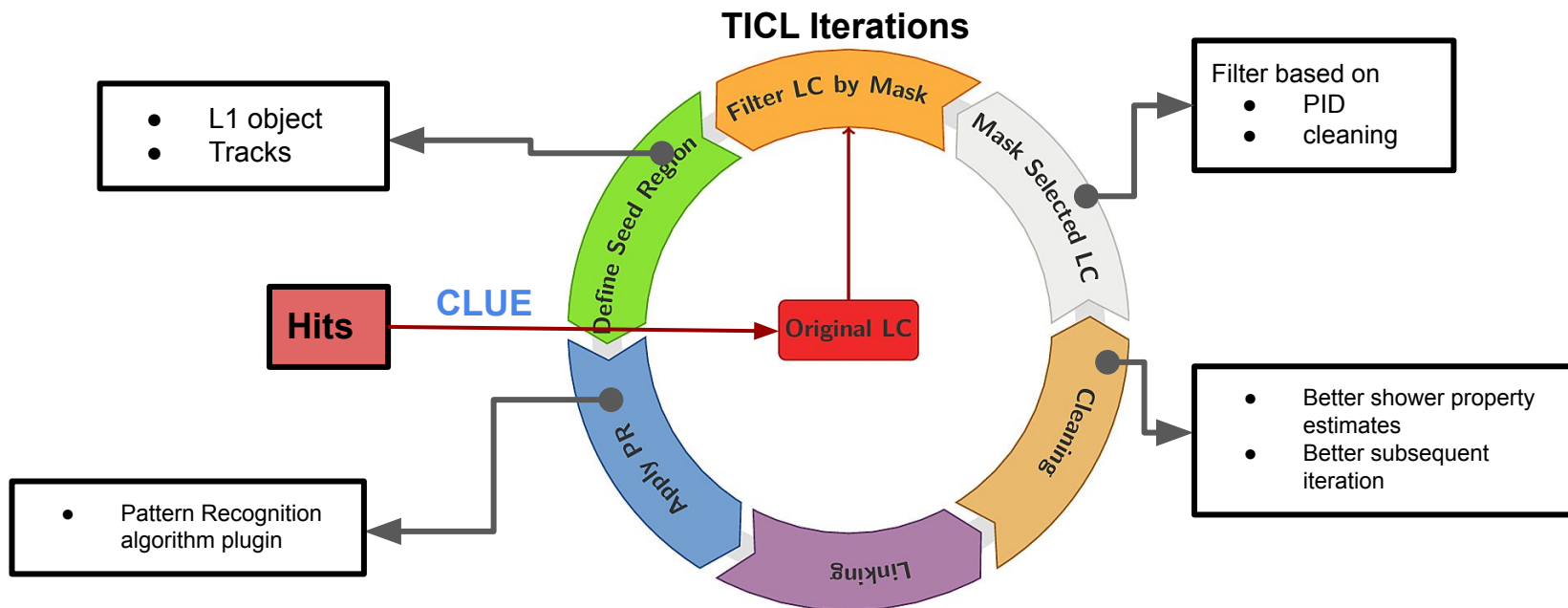


**CMS Public**  
Total CPU HL-LHC (2029/No R&D Improvements) fractions  
2021 Estimates



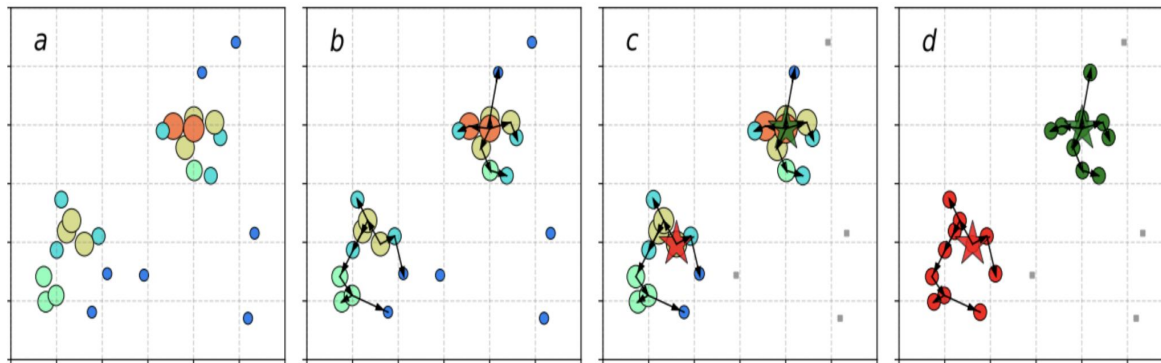
# TICL - The Iterative CLustering

- Framework to produce 3D clusters and particle properties starting from HGCAL hits (x, y, z, E, t)
  - Inspired from the successful CMS Run1 iterative tracking reconstruction strategy
- Framework is modular allowing for swapping of algorithms according to particle type
- Separate subdetector-based iterations and iterations using information from other subdetectors



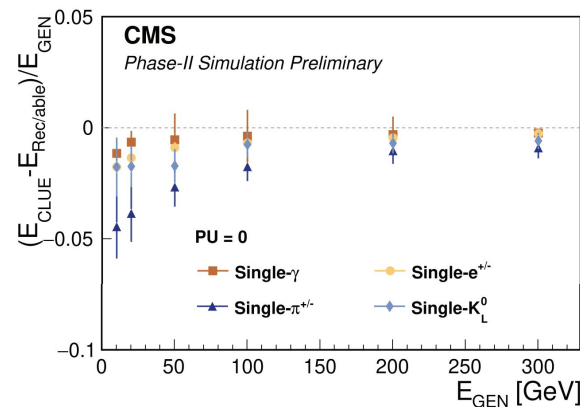
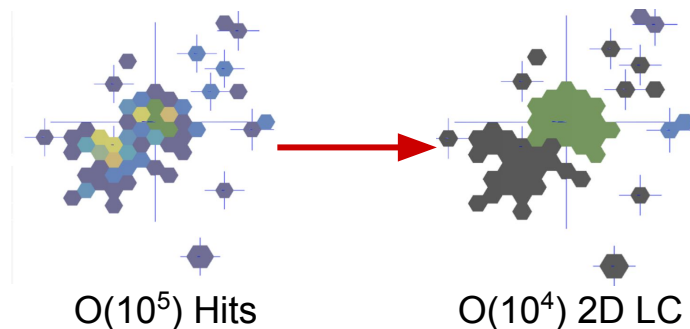
# CLUE [1] - CLUstering of Energy

- Energy density based 2D clustering algorithm
  - Removes noise
- Produces “Layer Clusters (LCs)” starting from hits
  - Dimensionality reduction by an order of magnitude



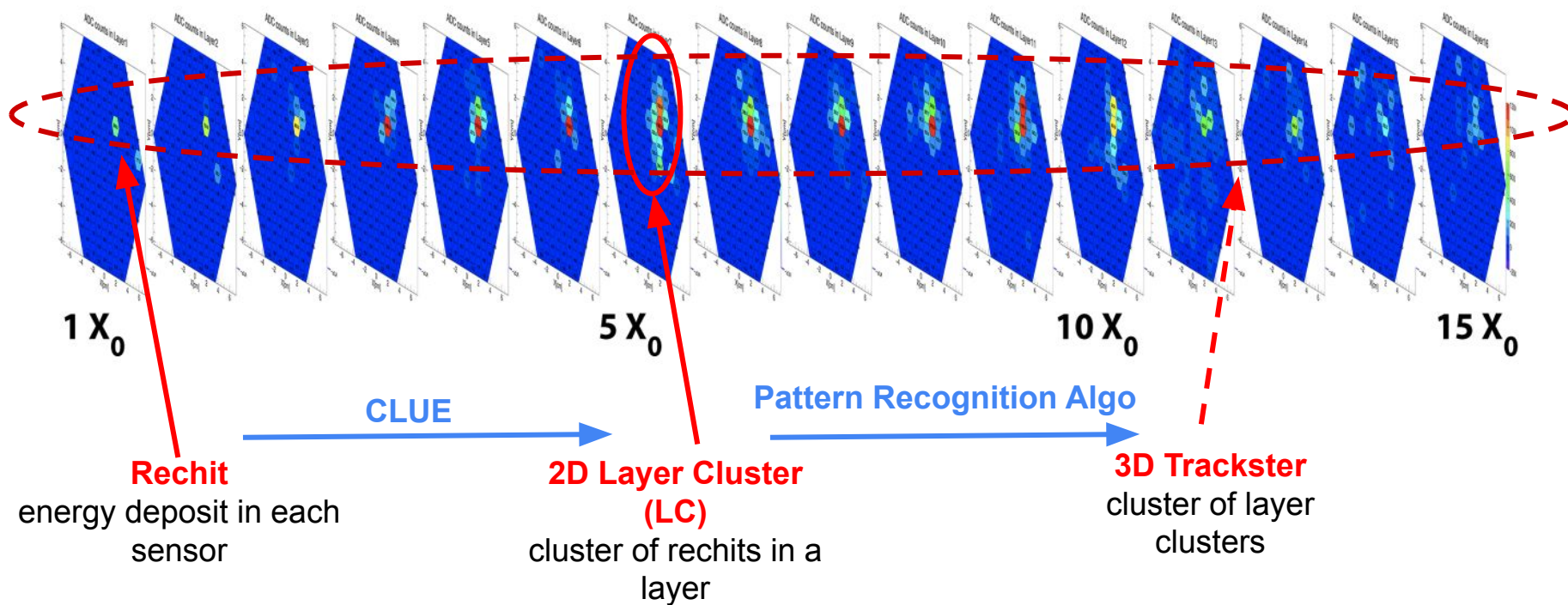
## Fully parallelizable GPU-ready algorithm

- 0.8% of total offline reconstruction on single CPU core @PU200
- 2 orders of magnitude faster on small GPU NVIDIA T4



# Pattern Recognition Algorithms

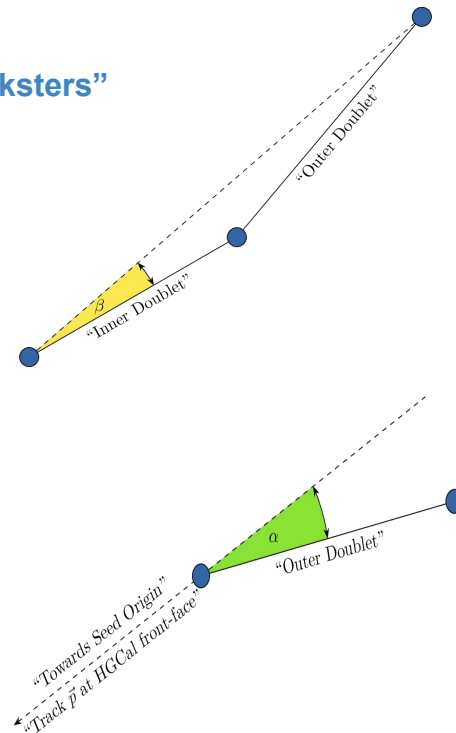
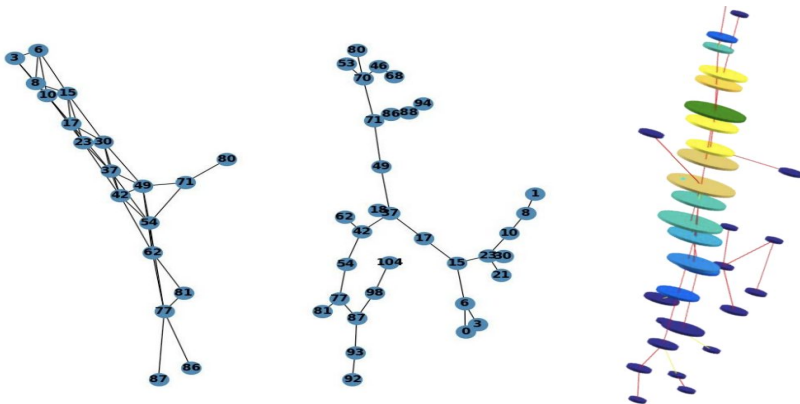
Pattern Recognition algorithms connect 2D LC's to form 3D clusters called “tracksters”



# Pattern Recognition Algorithms

Pattern Recognition algorithms connect 2D LC's to form 3D clusters called “tracksters”

- Tracksters are Direct Acyclic Graphs
  - Nodes are layer clusters
  - Edges are defined according to connecting algorithm
- Currently available connecting algorithms
  - Cellular Automaton
  - CLUE3D (energy density based clustering using 2D layer clusters)
  - FastJet
- Tracksters are linked to form high level particles



Pattern Recognition by CA

Shower reconstructed as graphs using CA(center) and CLUE3D (middle/right)

# Putting it all together : Object Reconstruction

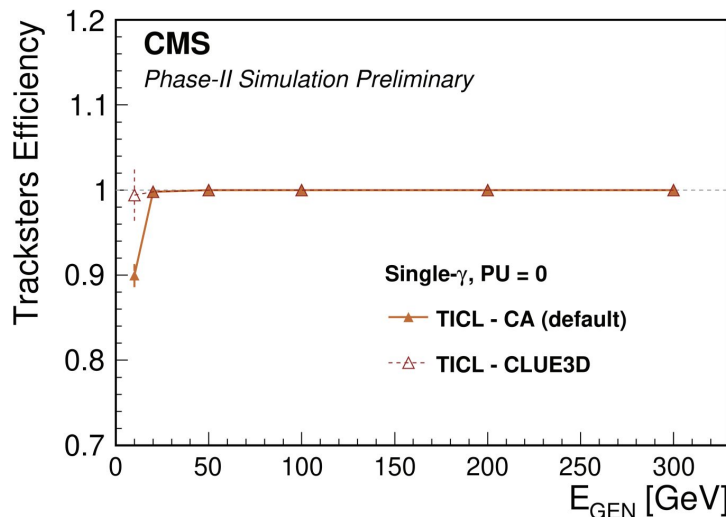
## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector

# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters

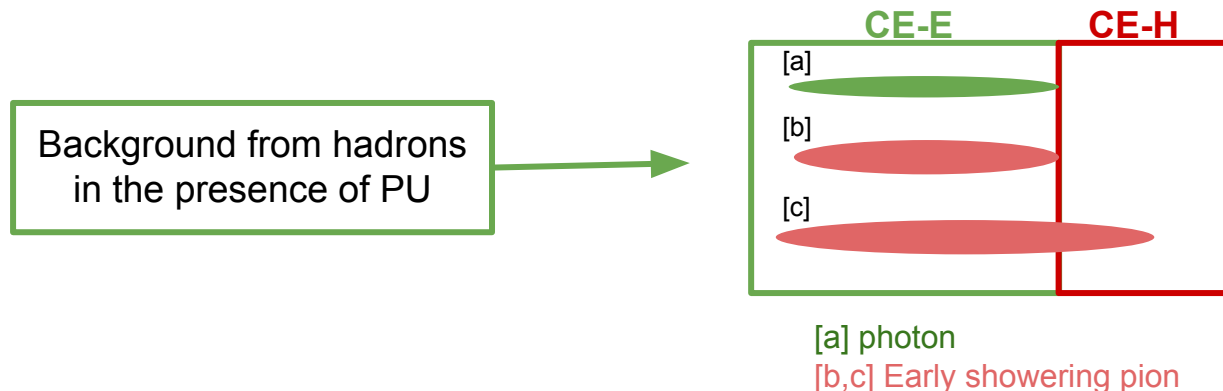


Tracksters efficiency for different pattern recognition algorithms

# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters

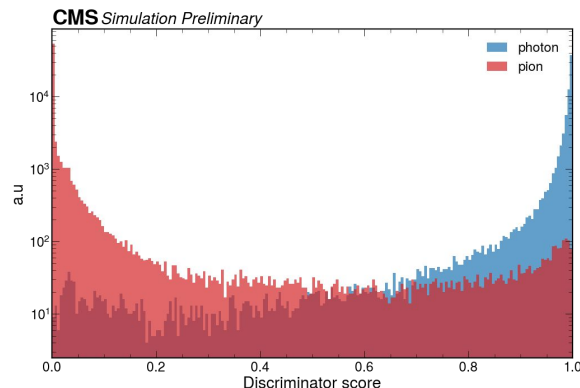
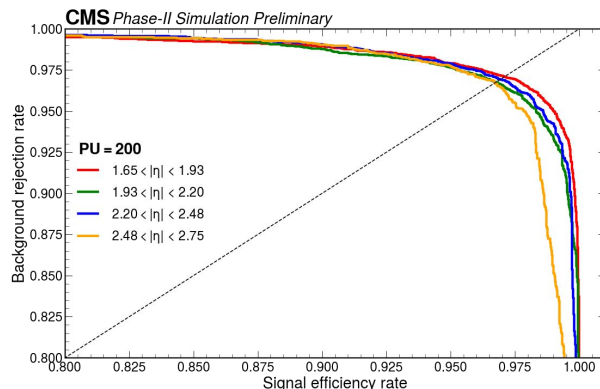


# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions

Unconverted Photons and early showering pions in **200 PU**



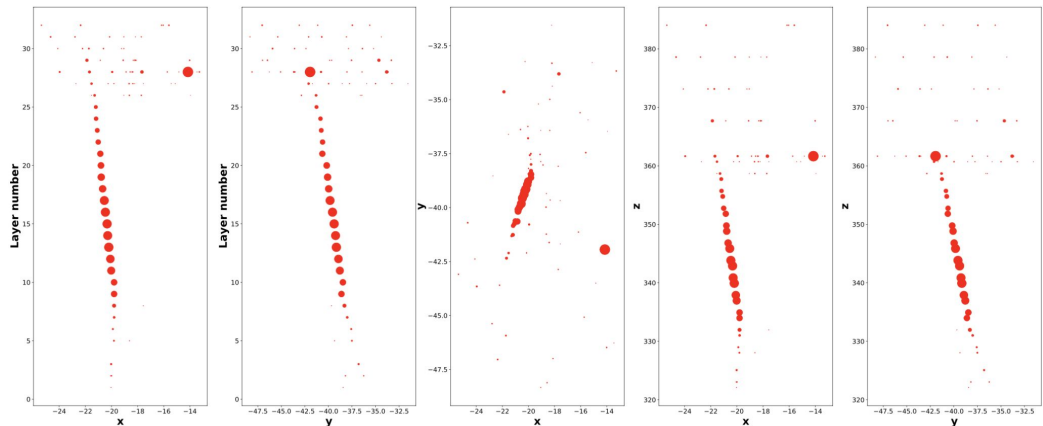
ROC curve (left) and discriminator scores (right) for PID based on Edge-convolution and greedy clustering based pooling [1]

# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions
- Cleaning tracksters to get better estimates of properties

Unconverted Photons and early showering pions in **200 PU**



Cleaning tracksters to get rid of PU/nearby particle contributions based on shower geometry

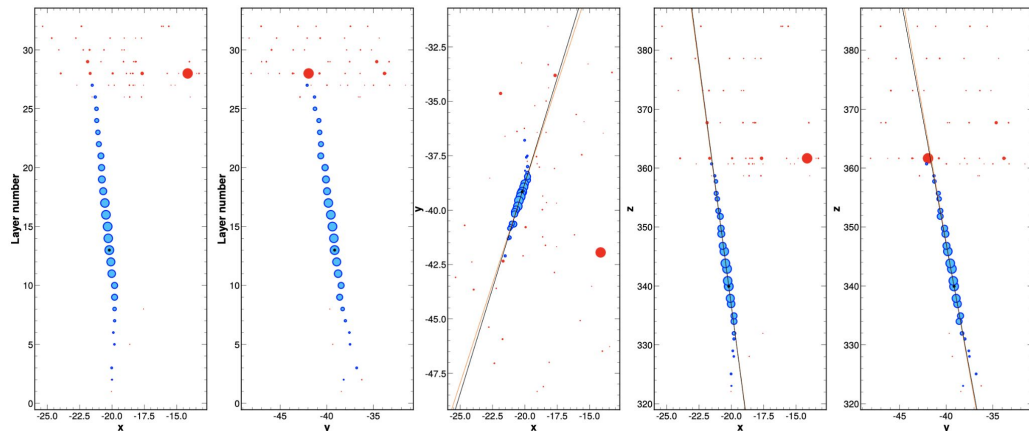
Tracksters before cleaning [all hits in red are one object]

# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions
- Cleaning tracksters to get better estimates of properties

Unconverted Photons and early showering pions in **200 PU**



Cleaning tracksters to get rid of PU/nearby particle contributions based on shower geometry

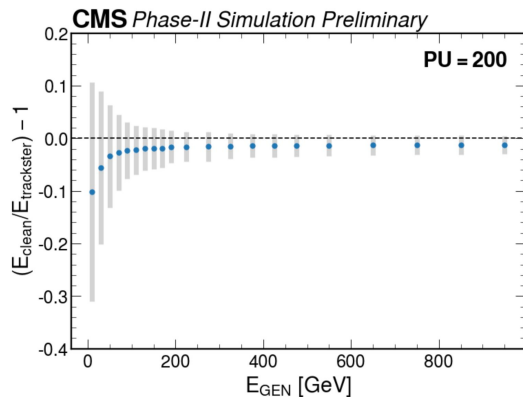
Tracksters after cleaning [only hits in blue remain after cleaning]

# Putting it all together : Object Reconstruction

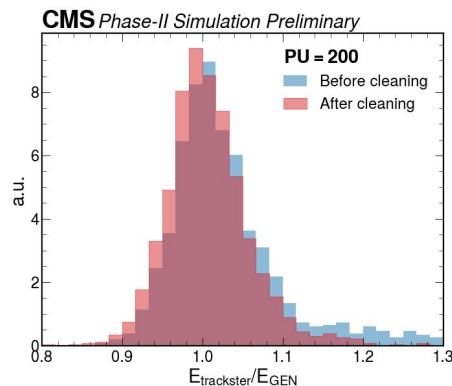
## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions
- Cleaning tracksters to get better estimates of properties

Unconverted Photons and early showering pions in **200 PU**



Ratio of trackster energy before and after cleaning as a function of generated particle energy shows very little difference



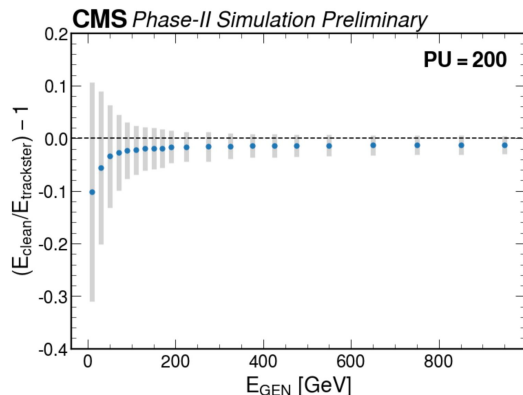
Ratio of trackster and generated particle energy for  $60 < E_{\text{GEN}} < 80 \text{ GeV}$ . Cleaning removes tails in resolution

# Putting it all together : Object Reconstruction

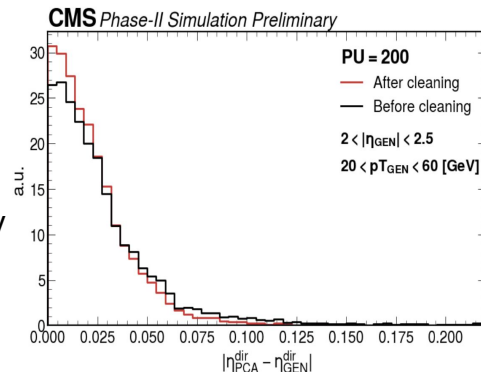
## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions
- Cleaning tracksters to get better estimates of properties

Unconverted Photons and early showering pions in **200 PU**



Ratio of trackster energy before and after cleaning as a function of generated particle energy shows very little difference



Trackster direction estimates before and after cleaning shows improvement after cleaning

# Putting it all together : Object Reconstruction

## Demonstrating by example : Electromagnetic iteration

- Iteration aimed at extracting EM objects first
  - Relatively less complex than HAD objects
  - Useful for electron/photon reconstruction
- Keep all LCs in the CE-E part of the detector + first few CE-H layers to capture leakage
  - Mask LCs deeper inside the detector
- Connect LCs using Cellular Automaton to form tracksters
- State of the art Graph Neural Networks based particle ID to distinguish from early showering pions
- Cleaning tracksters to get better estimates of properties

Tracksters and LCs filtered by the PID and/or cleaning fed back to LC collection for further HAD/MIP iterations

# Conclusions

- Reconstruction in CMS High Granularity Calorimeter poses **unprecedented challenges**
  - “**Tracking**” detector with high granularity
- **TICL** is a highly modular and flexible framework developed in CMS for HGCal reconstruction
  - Variety of pattern recognition algorithms can be **plugged in** and out
  - Different strategies for different particles
- **CLUE** is an “**imaging**” density based GPU friendly density-based algorithm
  - Provides building blocks for pattern recognition algorithms
  - Reduces hit multiplicity without sacrificing performance
  - Designed with **parallelism** in mind
- Variety of strategies being actively explored for best performance in **200 PU**
  - Optimal Particle Flow interpretation requires robust particle ID/ energy regression/ PCA
  - Utilise **novel machine learning** ideas like Graph Neural Networks
  - Strategy for **purifying** objects from PU contributions
- Next steps:
  - Improve strategies for hadron reconstruction and PF-objects interpretations

# References

- Trackster ID and cleaning for EM iterations, CMS DP -2022/002, [https://cds.cern.ch/record/2805638/files/DP2022\\_002.pdf](https://cds.cern.ch/record/2805638/files/DP2022_002.pdf)
- CLUE: A Fast Parallel Clustering Algorithm for High Granularity Calorimeters in High Energy Physics, M. Rovere, Z. Chen, A. Di Pilato, F. Pantaleo, C. Seez , <https://arxiv.org/abs/2001.09761>
- A Dynamic Reduction Network for Point Clouds, <https://arxiv.org/pdf/2003.08013.pdf>



# Thank You