ML for adaptive reconstruction and detector geometry optimization

Dolores Garcia, Michele Selvaggi, Brieuc Francois,
Gregor Krzmanc, Jan Kieseler, Philipp Zehetner

ML for very adaptive reconstruction

- Most detectors geometries are implemented → focus now on reconstruction

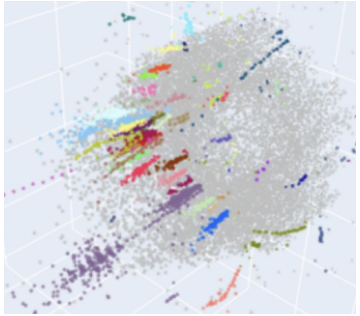
Adaptable reconstruction is crucial for a systematic design optimization:

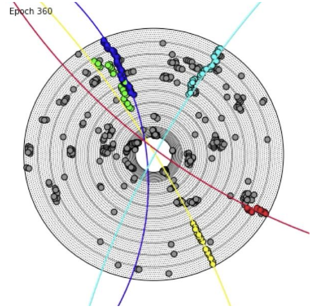
- No need for hand picked parameter tuning, the detector will still change many times
- Costly conventional implementation
- Reduction of person power, increased performance?

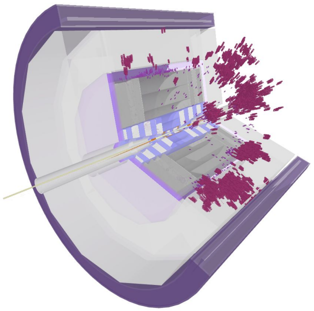
Problems that are similar:

- Calorimeters
- Tracking
- Combining information from multiple subdetectors
- Images of sea lions

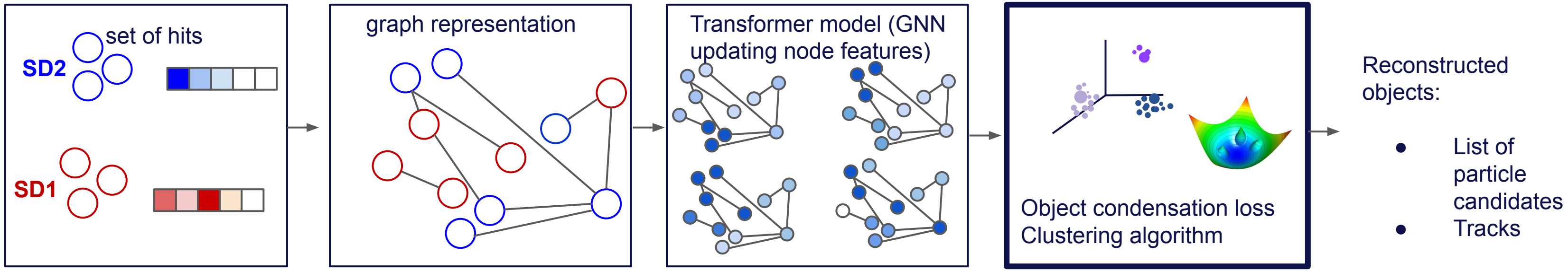
Adaptive reconstruction algorithms are working: Belle II, HGCAL, CMS







End-to-end reconstruction approach



Input:

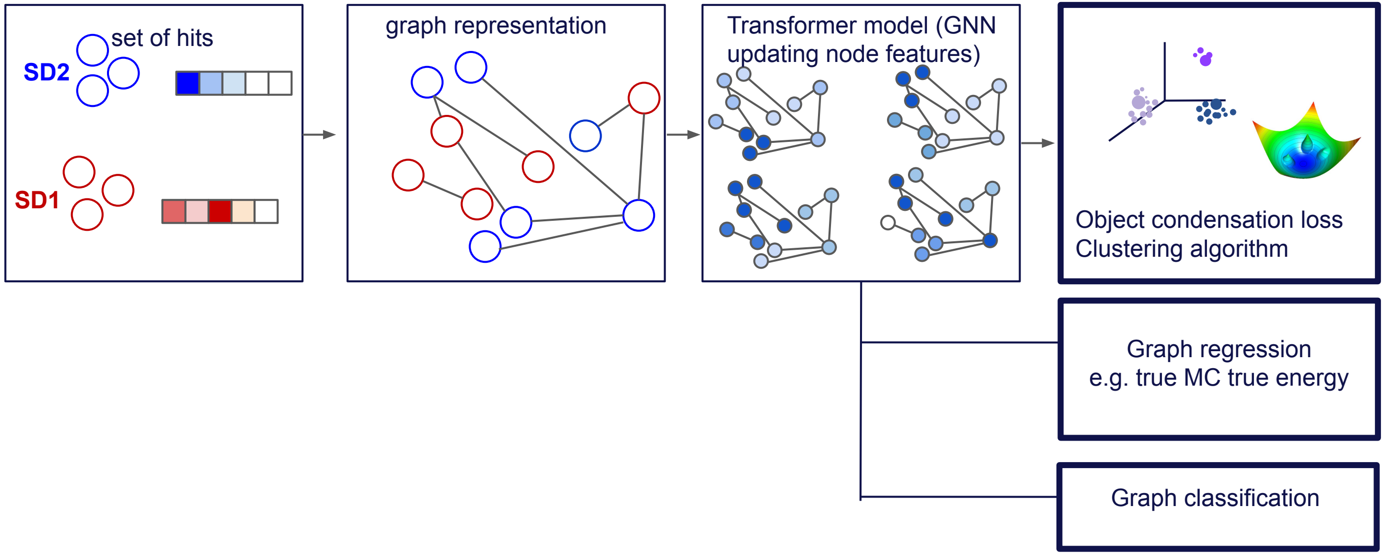
- A set of hits from different sensors (coordinates, type of hit, energy)
- Each hit is one node in the graph O(600) per particle

Output:

- Coordinate in embedding space (3D>)
- Beta (q)
- Use clustering space to build particle candidates

[1] Kieseler, J. (2020). Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *The European Physical Journal C*, *80*, 1-12.
[2] Qasim, Shah Rukh, et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks." *The European Physical Journal C* 79.7 (2019): 1-11.

End-to-end reconstruction approach



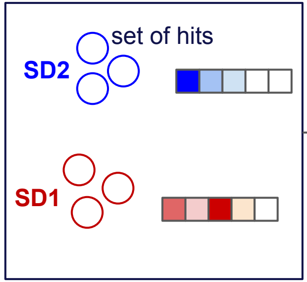
Pattern recognition

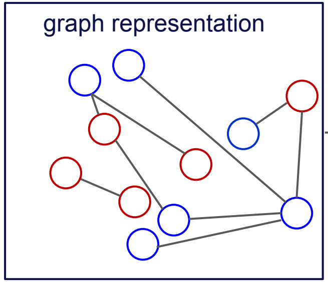
Energy calibration

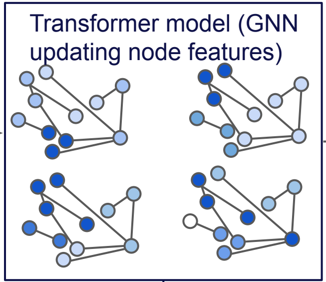
PID

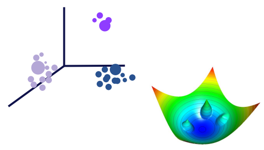
[1] Kieseler, J. (2020). Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *The European Physical Journal C*, *80*, 1-12.
[2] Qasim, Shah Rukh, et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks." *The European Physical Journal C* 79.7 (2019): 1-11.

End-to-end reconstruction approach









Object condensation loss
Clustering algorithm

Pattern recognition

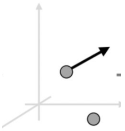
Graph regression
e.g. true MC true energy

Energy calibration

Graph classification

PID

- Vector space + geometric product
- Encode geometric objects and operations (points, planes,lines, rotations, translations, reflections)
- Inner product, wedge product, dual (empty to full dimensions)
- Equivariance to E(3), but symmetry breaking is part of the network (positional encodings)

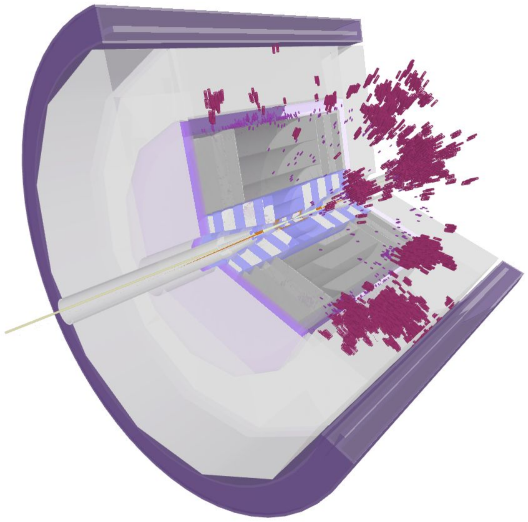


[1] Kieseler, J. (2020). Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *The European Physical Journal C*, *80*, 1-12.
[2] Qasim, Shah Rukh, et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks." *The European Physical Journal C* 79.7 (2019): 1-11.

ML for Particle Flow, a very adaptive reconstruction

The particle flow algorithm aims to **identify the produced particles in a collision through the combination of the information from the entire detector** and **provide best combined energy/momentum resolution**

- Main drivers of performance are:
  - Tracking and calo cluster efficiency
  - Track cluster matching
  - Resolution
- Asymptotic PF jet resolution ~ 1-2%, best achieved so far 3-4%
  - There is room for improvement
- Hoping to achieve higher reconstruction performance: cluster merging, arbitration of track vs cluster energy

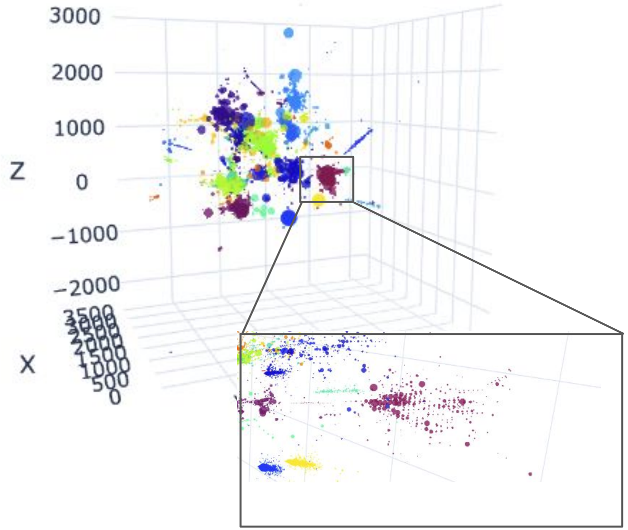


A Example of input data in the CLD detector

[1] Pata, J. Machine learning for particle flow reconstruction at CMS, presentation at CDS.
[2] Qasim, S. R., Chernyavskaya, N., Kieseler, J., Long, K., Viazlo, O., Pierini, M., & Nawaz, R. (2022). End-to-end multi-particle reconstruction in high occupancy imaging calorimeters with graph neural networks. *The European Physical Journal C*, *82*(8), 753.

Dataset

- Event generation:
  - Particle gun (10-15/40-50 particles) → ~7/24 k hits
  - E ∈ [0.5, 50] GeV
  - p, n, K_L, π, e+-, γ
  - FullSim CLD
  - Truth from gen
  - Training on 400k events
- Target of training is:
  - clustering
  - and the energy of the MC particle

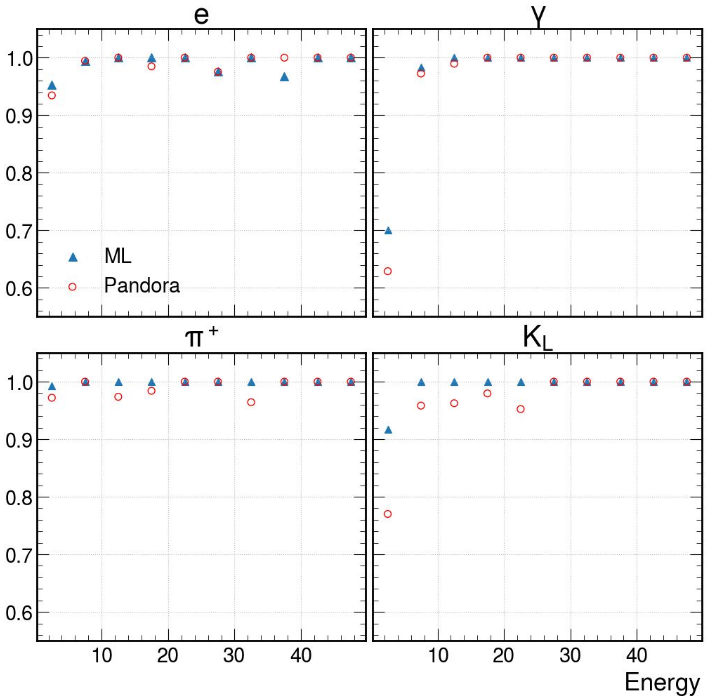


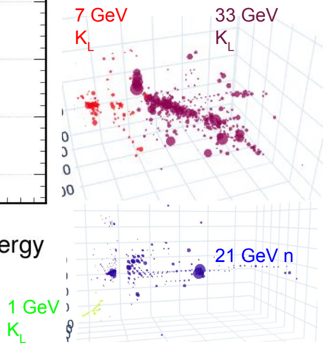
A Example of event from training set

[1] Bacchetta, Nicola, et al. "CLD--A Detector Concept for the FCC-ee." *arXiv preprint arXiv:1911.12230* (2019).

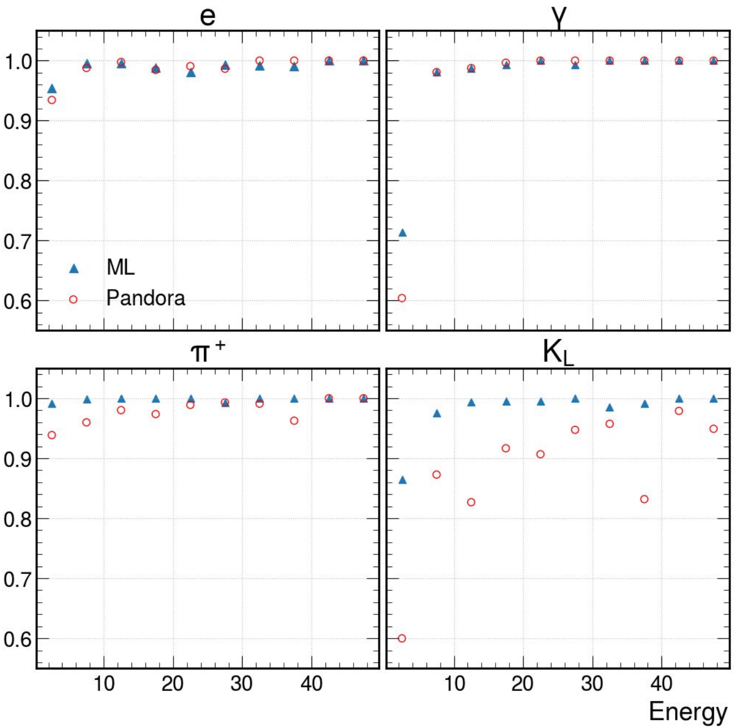
Efficiency (CLD Fullsim):

Dataset with jet-like 10-15 particles

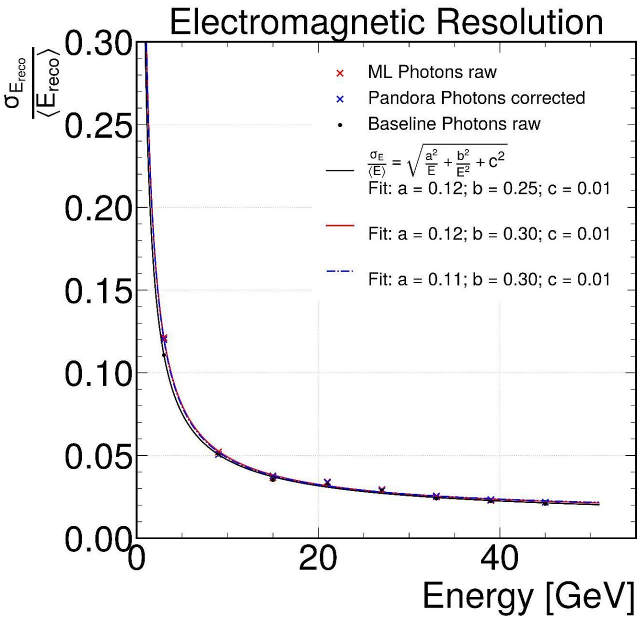


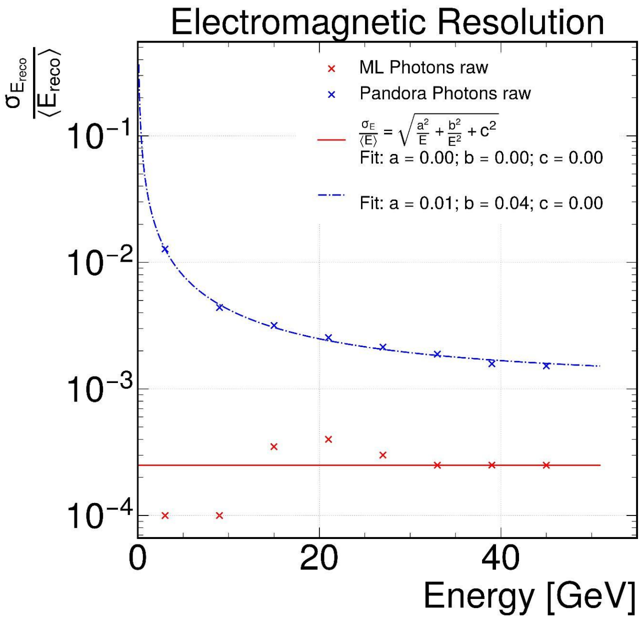


Dataset with jet-like 40-50 particles



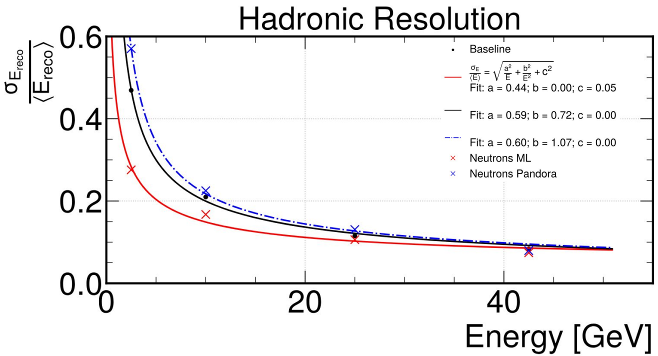
Results:

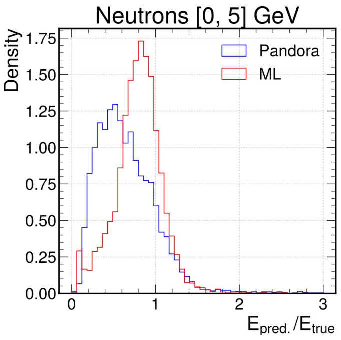


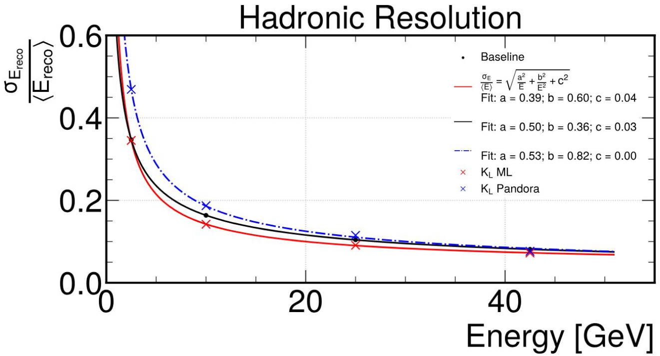


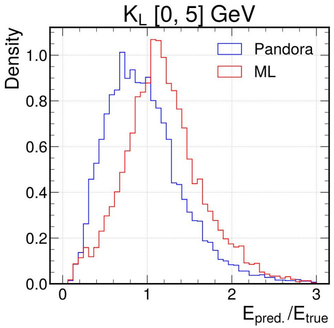
- The ML model has no energy correction applied
- Baseline is the sum of the hits energy

Results





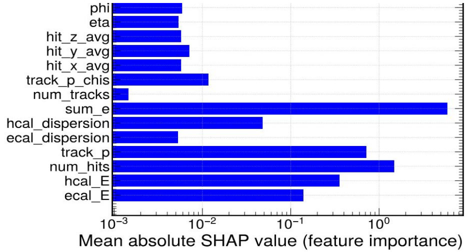




- **Model:** ML clustering + ML energy correction
- **Baseline:** Pandora
- Improved resolution for neutral hadrons
  - SHAP point to sum of the hits as relevant feature
- Jet resolution is driven by neutral hadron (HCAL) resolution

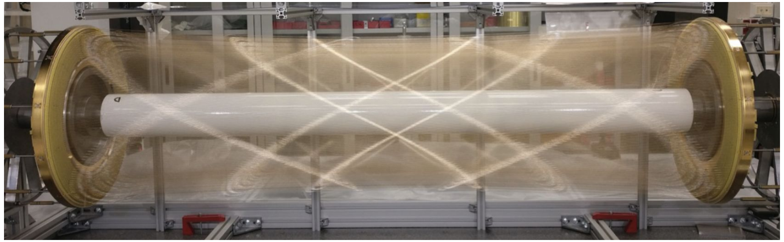
$$\sigma^2(E_{\mathrm{vis}}) = \sum_{i\in \mathrm{tr}} \sigma^2_{\mathrm{tr}}(E^{(i)}_{\mathrm{tr}}) + \sum_{i\in\gamma} \sigma^2_{\mathrm{ecal}}(E^{(i)}_{\gamma}) + \sum_{i\in \mathrm{nh}} \sigma^2_{\mathrm{hcal}}(E^{(i)}_{\mathrm{nh}})$$

- Charged have similar performance
  - Based on track measurement
- **Next steps:** p reconstruction, physics events, impact on visible mass



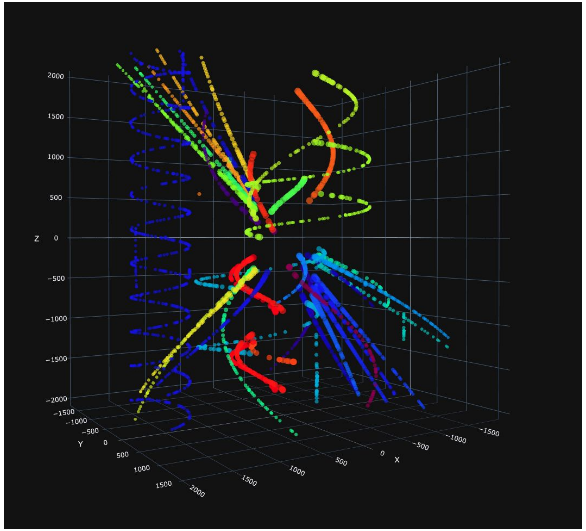
Pattern recognition for the IDEA Drift Chamber

- Investigating Machine Learning based tracking for the IDEA detector (in addition to the conventional tracking)
- Simulating events with Pythia + ddsim + digi
- No beam background included
- Dataset inputs of DCH:
  - Wire geometry:
    - Layer, superlayer
    - Stereo angle
    - → Coordinates
  - Hit:
    - Distance along the wire
    - Distance to the wire
- Hits from the vertex detector



Top: Drift Chamber [1]

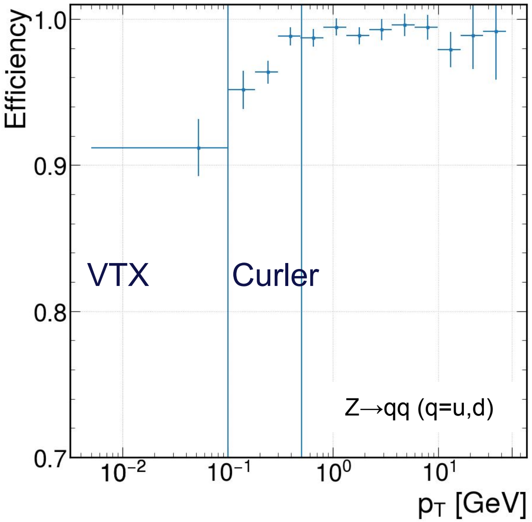
Bottom: example event from full sim

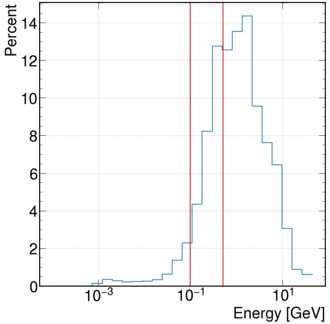


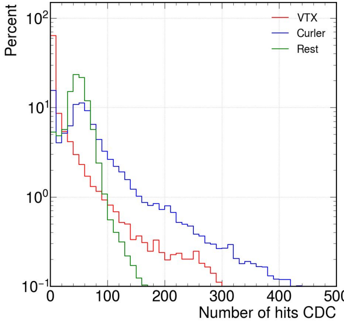
[1] Tassielli, G. F., Baldini, A. M., Cavoto, G., Cei, F., Chiappini, M., Chiarello, G., ... & Voena, C. (2020). The drift chamber of the MEG II experiment. *Journal of Instrumentation*, *15*(09), C09051.

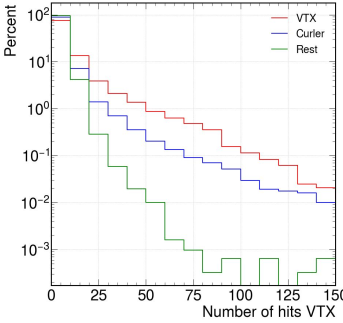
Pattern recognition for the IDEA Drift Chamber

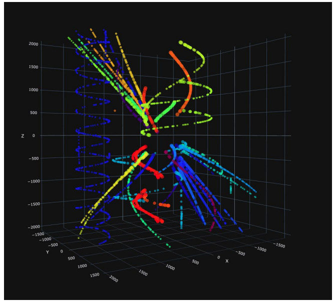
- Efficiency: track is matched if 50% of hits belong to the corresponding MC particle (Belle II defines it with 5% hit efficiency [1])
- Tracking efficiency shows expected results with the dependence on transverse momentum and has to be further evaluated
  - Two imbalance factors: Tracks on different regions (with different geometries) are not equally accounted for and this needs to be addressed in the next iteration. Number of hits in each region
- Pattern recognition is implemented as a Gaudi algorithm in key4hep
  - The model is ONNX exported
- Model can be easily adapted to changes in the geometry





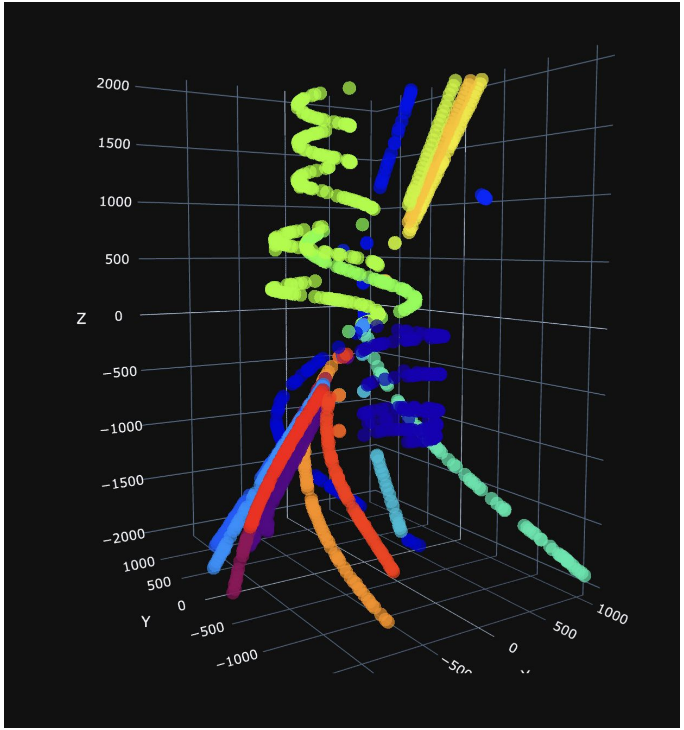


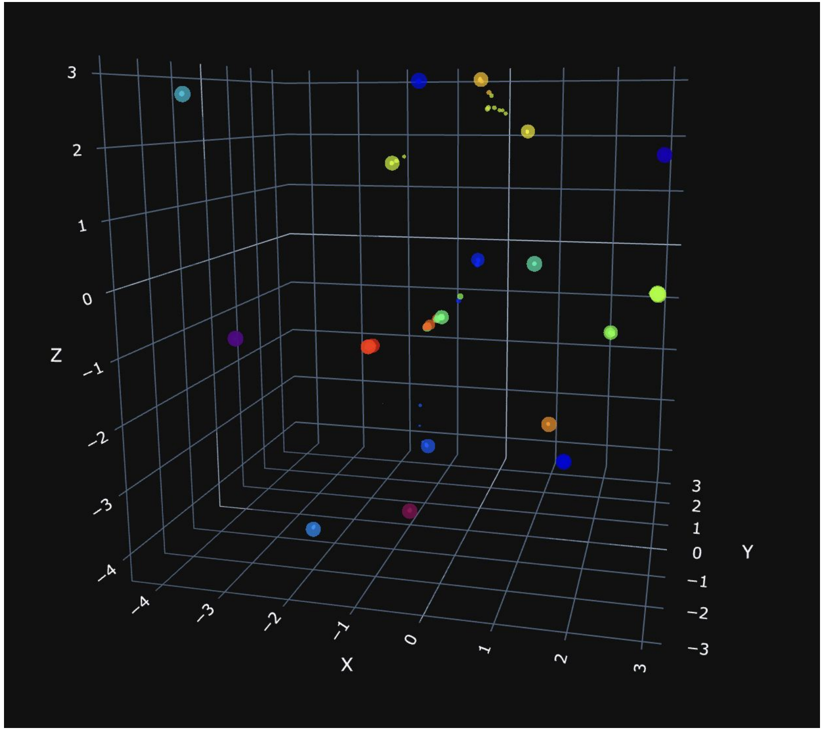




[1] Bertacchi, V., Bilka, T., Braun, N., Casarosa, G., Corona, L., Cunliffe, S., ... & Zani, L. (2020). Track finding at Belle II. *arXiv preprint arXiv:2003.12466*.

Pattern recognition for the IDEA Drift Chamber





Summary and next steps

MLPF

MLPF allows for a detector agnostic calorimeter clustering with similar performance to Pandora

Next steps:

- Tackling full PF, adding 4-vector reconstruction
- Evaluation on more complex datasets with physics events and jet metrics

IDEA wire chamber pattern recognition

First pattern recognition version available in key4hep

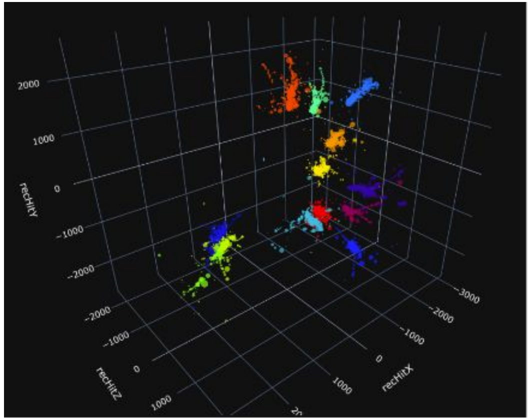
Next steps:

- Improve the tracking by generating more balanced datasets
- Resolve left right ambiguity of the DC hits
- Fitting of tracks and parameter estimation for TrackStates

Thanks to the Key4hep team!

Thank you

End-to-end approach



CP

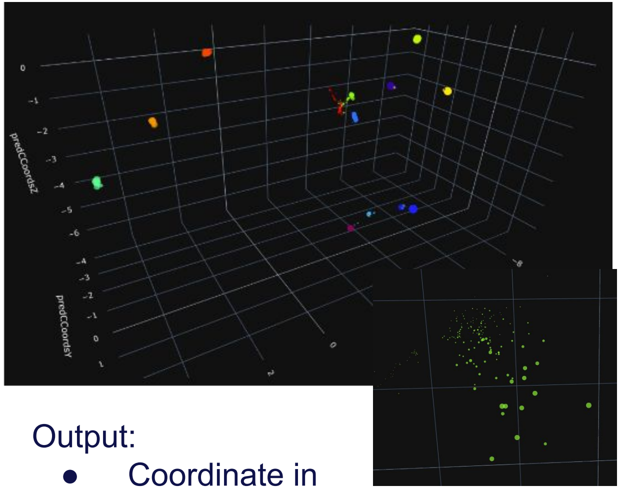
+ GNN

$$q_{\alpha k} = \max_i q_i M_{ik}.$$

$$\breve{V}_k(x) = \|x - x_\alpha\|^2 q_{\alpha k}, \text{ and}$$

$$\hat{V}_k(x) = \max(0, 1 - \|x - x_\alpha\|) q_{\alpha k}.$$

$$L_V = \frac{1}{N} \sum_{j=1}^{N} q_j \sum_{k=1}^{K} \left(M_{jk} \breve{V}_k(x_j) + (1 - M_{jk}) \hat{V}_k(x_j) \right).$$



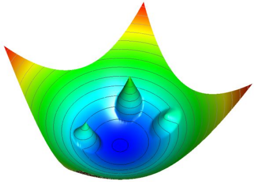
Input:

- A set of hits from different sensors (coordinates, type of hit, energy, A)
- Each one node in the graph O(600) per particle

- **Each object 1 condensation point (CP)**
- **Repulsive +Attractive potentials for each CP**

Output:

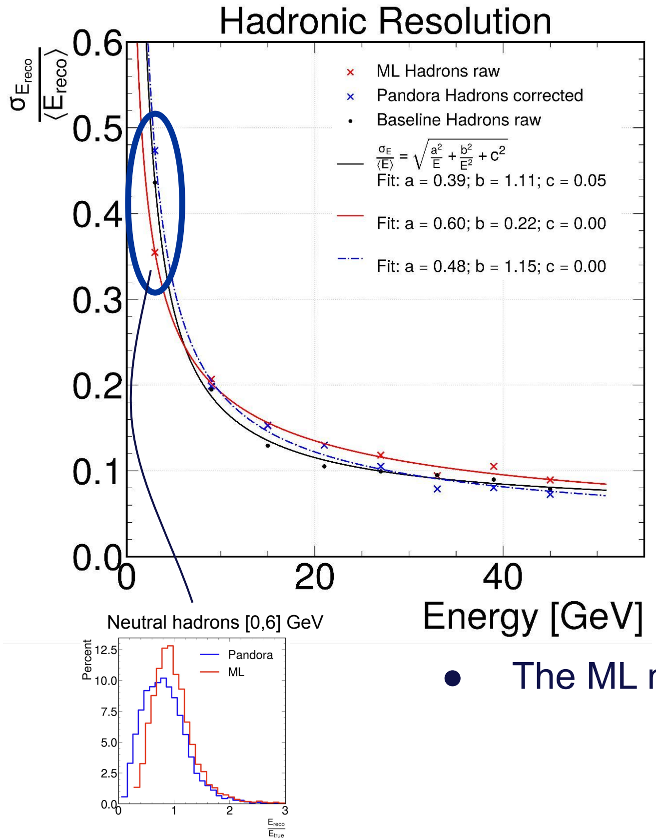
- Coordinate in embedding space (3D>)
- Beta (q)
- Use clustering space to build showers

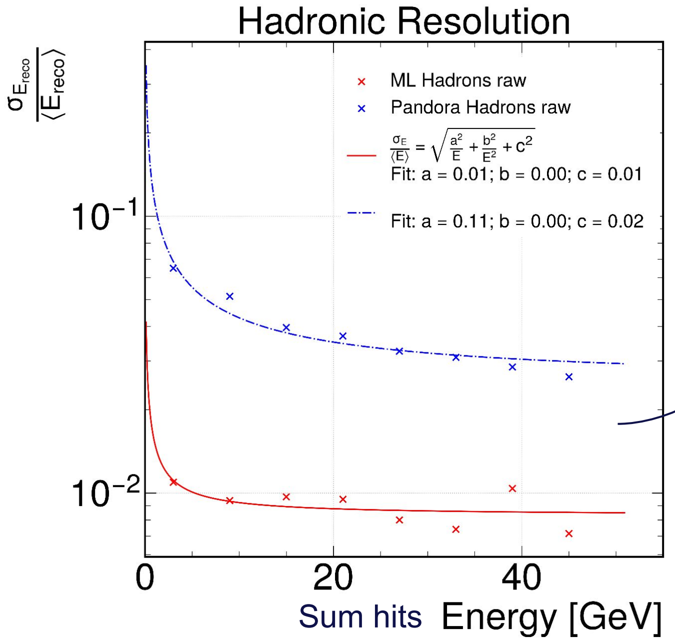


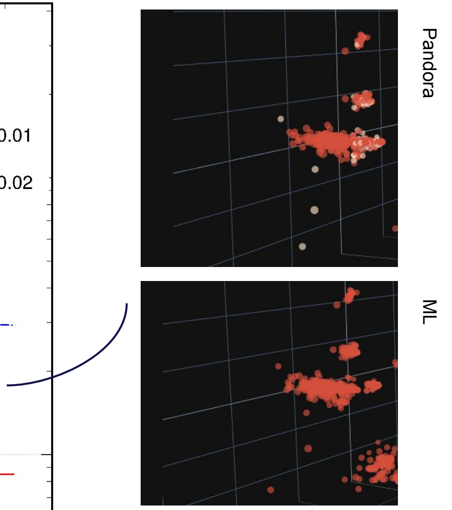
[1] ACKS : Michele Selvaggi, Gregor Krzmanc, Jan Kieseler, Philipp Zehetner

[1] Kieseler, J. (2020). Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *The European Physical Journal C*, *80*, 1-12.
[2] Qasim, Shah Rukh, et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks." *The European Physical Journal C* 79.7 (2019): 1-11.

Results:



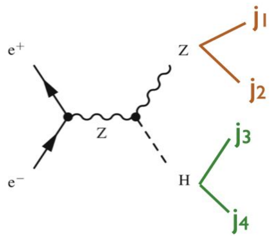


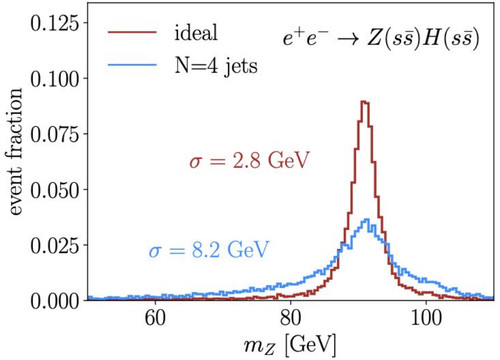


- The ML model has no energy correction applied

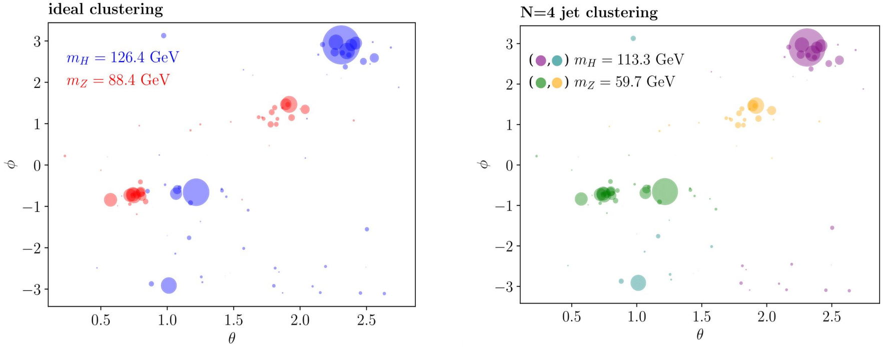
Clustering Color Singlets

- FCC-ee would serve as a Higgs factory, electroweak and top at highest luminosities
  - Measure Higgs particle properties and interactions in challenging decay modes
- Identification of color-neutral resonances relies on clustering final state into jets
- Calorimetry is expected to be much improved at future e+e− colliders, so that the 2-jet invariant mass resolution will be dominated not by detector resolution but rather by mis-clustering [1] (A)
- Jets are not well defined but color connection is physical, this may help **improve the mass estimation for color singlets** (H,Z,W) **and remove more background**





A Comparison of clustering performance vs ideal reconstruction



B Example of miss clustering

[1] ACKS : Michele Selvaggi

[1]Fujii, K., Grojean, C., Peskin, M. E., Barklow, T., Gao, Y., Kanemura, S., ... & Murayama, H. (2020). ILC study questions for snowmass 2021. *arXiv preprint arXiv:2007.03650*..
[2] Gallicchio, J., & Schwartz, M. D. (2010). Seeing in color: jet superstructure. *Physical review letters*, *105*(2), 022001.

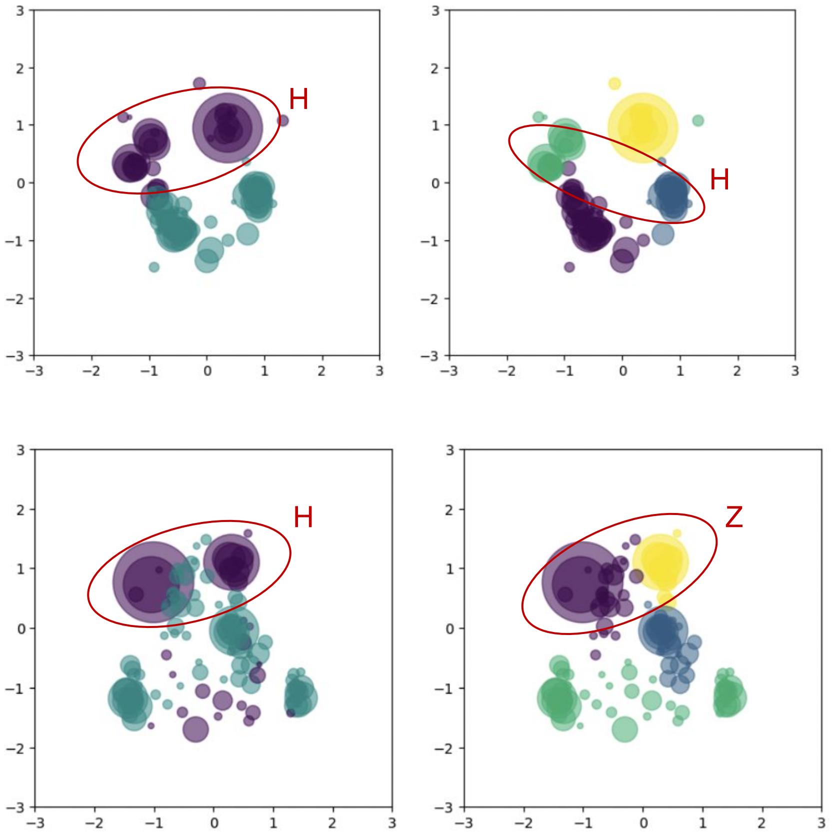
Clustering Color Singlets

Loss in performance can be due to:

- Miss matching of jets pairs
- Miss clustering of soft particles leading to degraded resolution

Possible solutions:

- Parameter tuning (BAO)
- Optimize distance metrics?: piecewise continuous function, hard optimization problem
- **End-to-end approach**



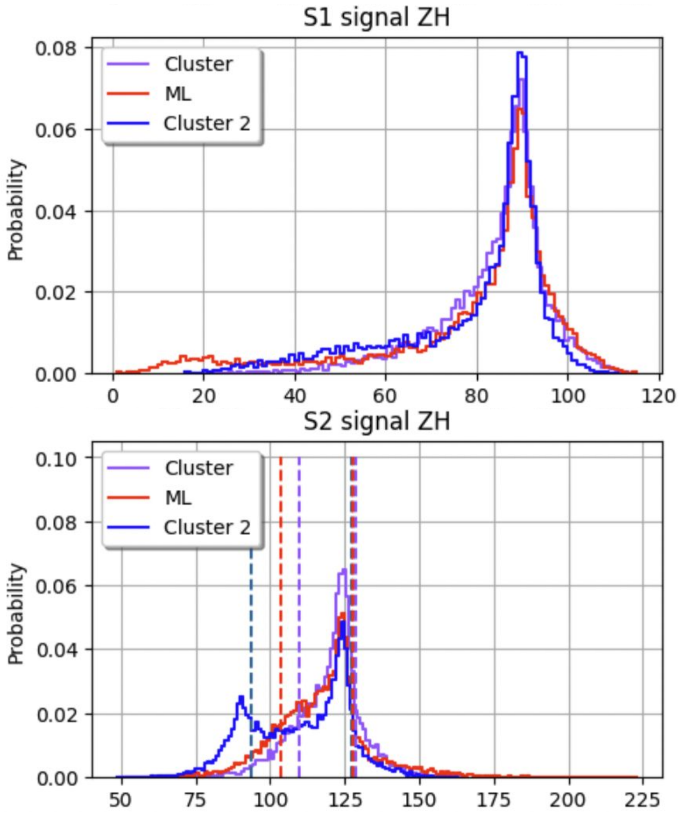
A Mismatching of jets pairs

[1] ACKS : Michele Selvaggi

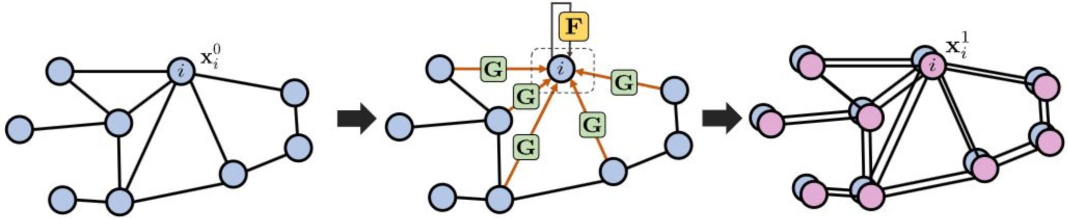
[1]Fujii, K., Grojean, C., Peskin, M. E., Barklow, T., Gao, Y., Kanemura, S., ... & Murayama, H. (2020). ILC study questions for snowmass 2021. *arXiv preprint arXiv:2007.03650*..
[2] Gallicchio, J., & Schwartz, M. D. (2010). Seeing in color: jet superstructure. *Physical review letters*, *105*(2), 022001.

Clustering Color Singlets

- **GNN** - Node classification (instantiation) problem, permutation invariant and equivariant
- Arch: FC - Graph **Transformer [1]**
- **Results:**
  - Similar performance to classical approach
  - **Baselines:**
    - **Chi-squared**
      $\chi^2 = 1/\sigma_H(M_{1/2}-M_H)^2 + 1/\sigma_z (M_{2/1}-M_Z)^2$
    - **Z only**
      $\chi^2 = (M_{1/2}-M_Z)^2$
- Wiring is important, using information about the ordering (tree structure) performance can be improved



A. Mass distributions of signal



[1] ACKS : Michele Selvaggi

[1]Dwivedi, V. P., & Bresson, X. (2020). A generalization of transformer networks to graphs. *arXiv preprint arXiv:2012.09699*..