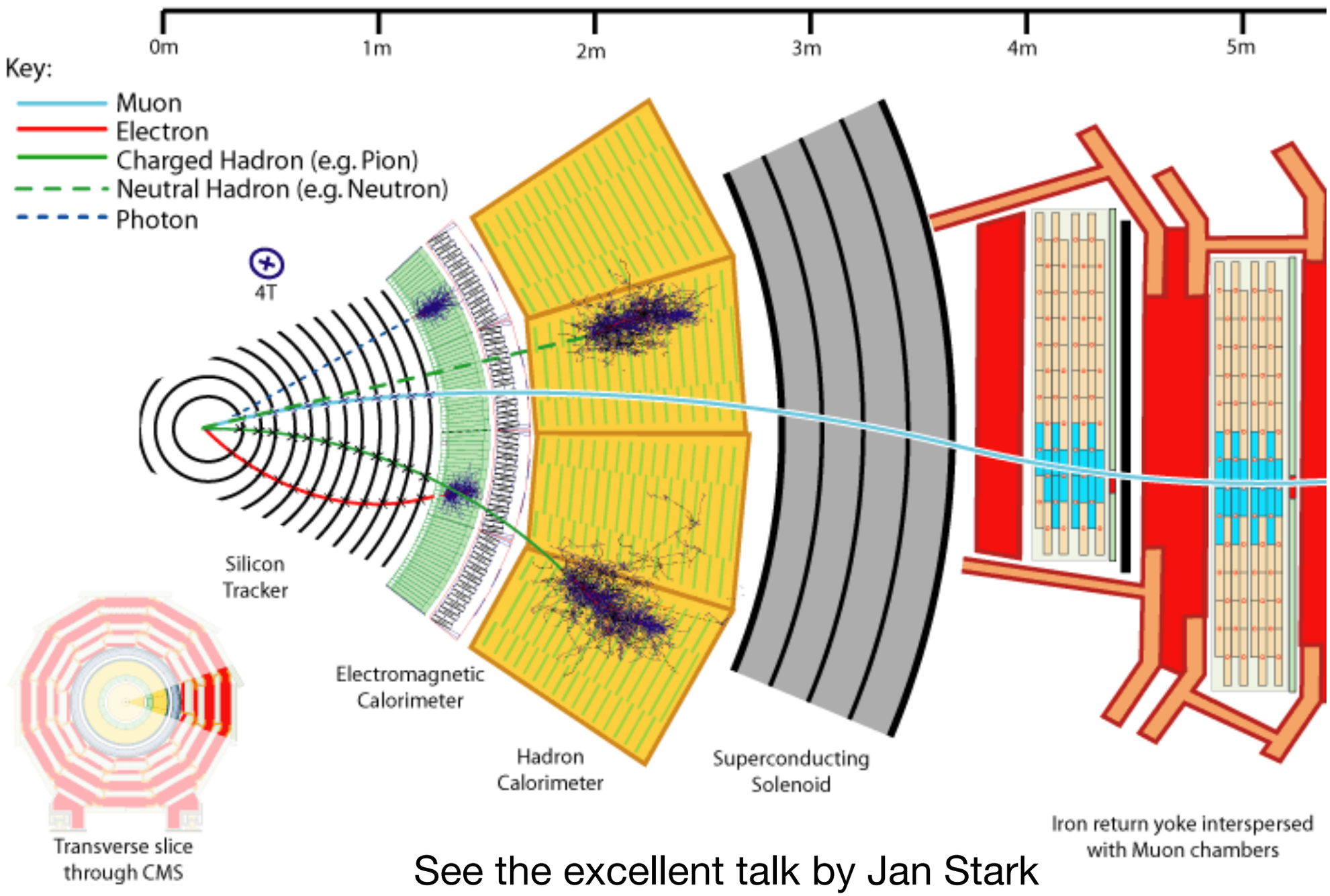
Overview of Machine Learning for Calorimeter and Particle Flow

Joosep Pata (NICPB, Estonia)

April 19, 2022
Learning To Discover
Institut Pascal, Université-Paris-Saclay

Multilayered detectors



See the excellent talk by Jan Stark earlier today.

Algorithmic reconstruction

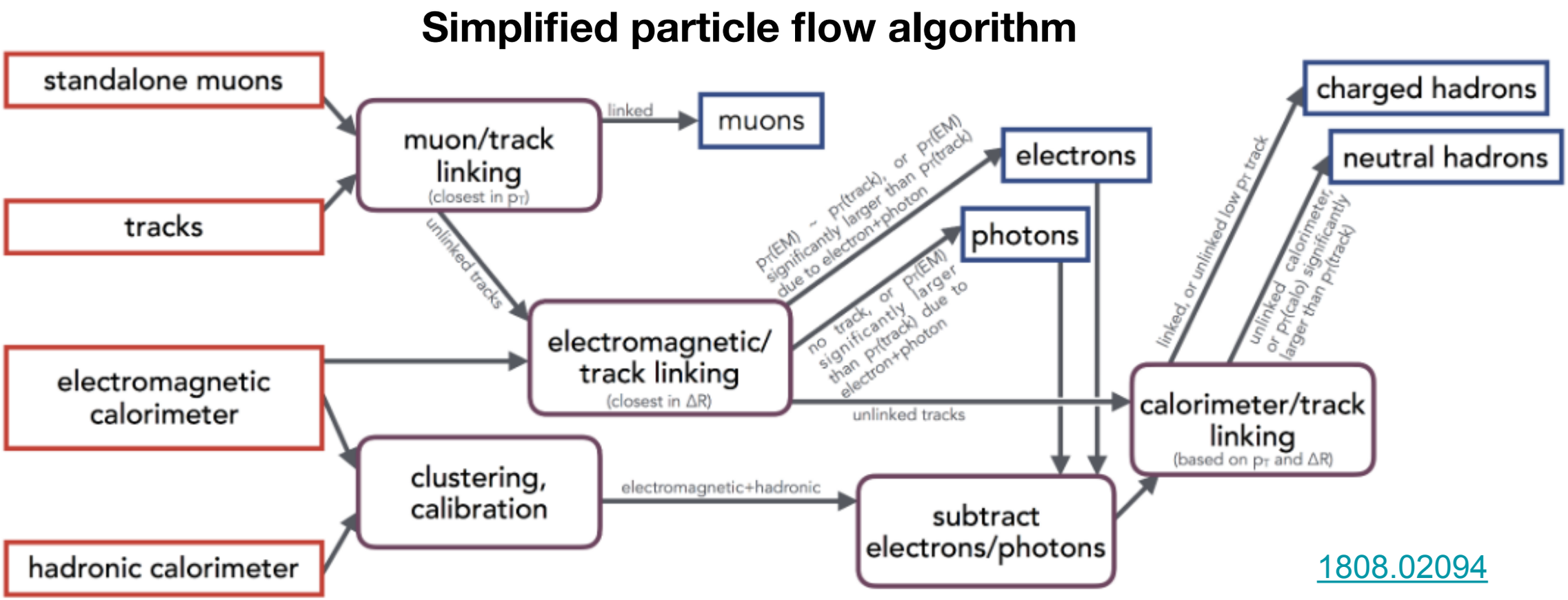
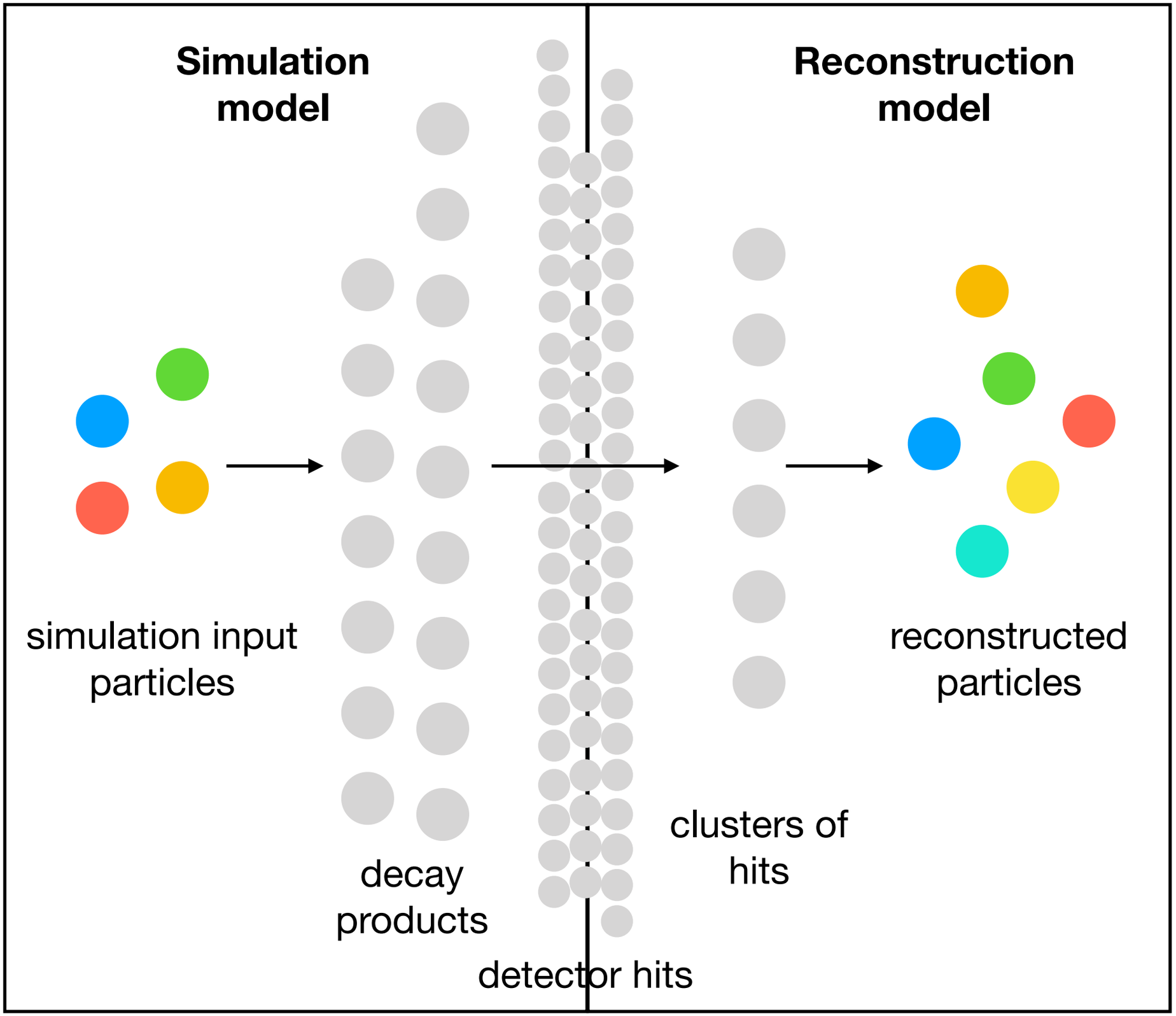


Figure 2. Schematic of particle flow algorithm for CMS Level-1 trigger correlator.

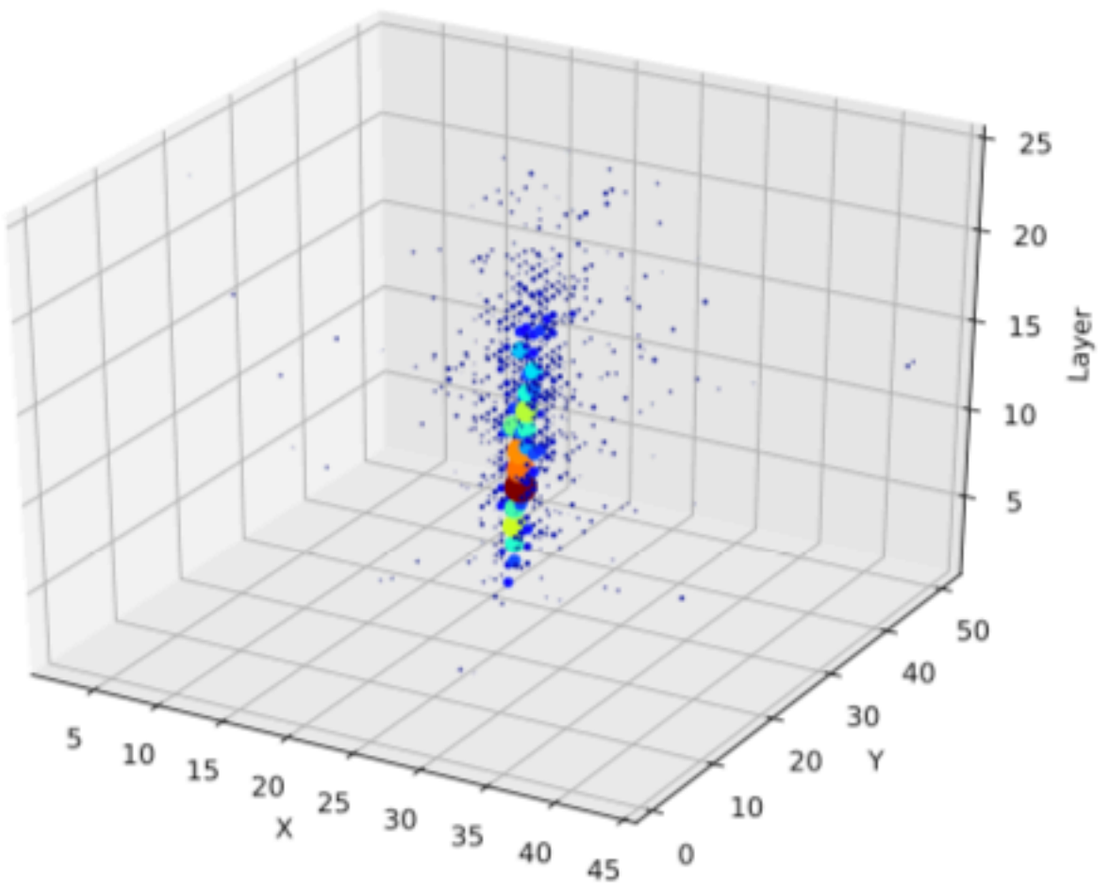
Simulation to reconstruction



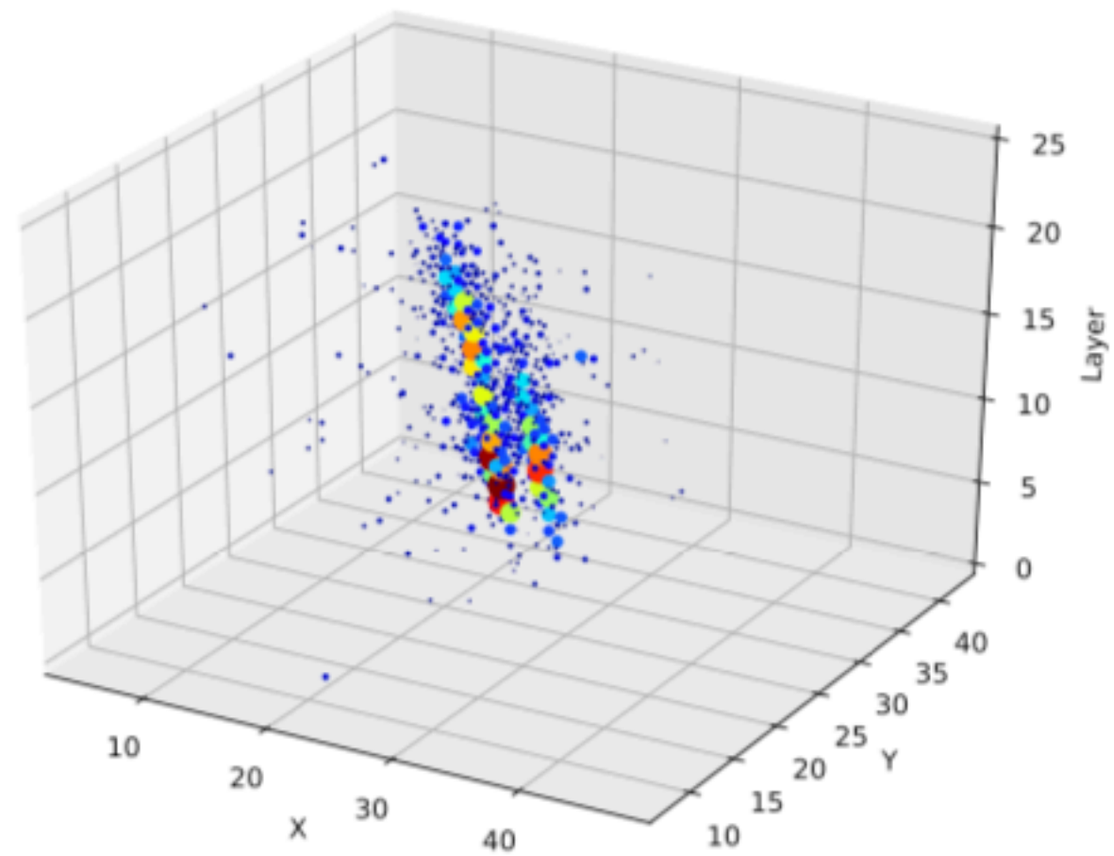
Single particle showers

Data are very sparse!

photon

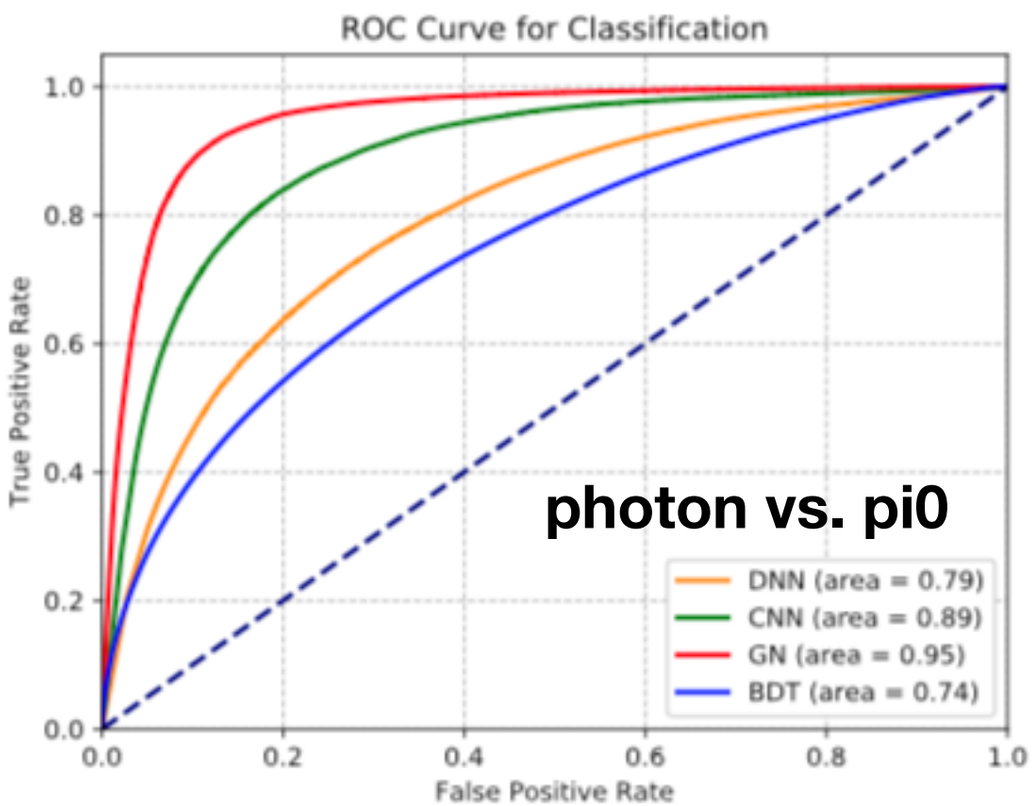


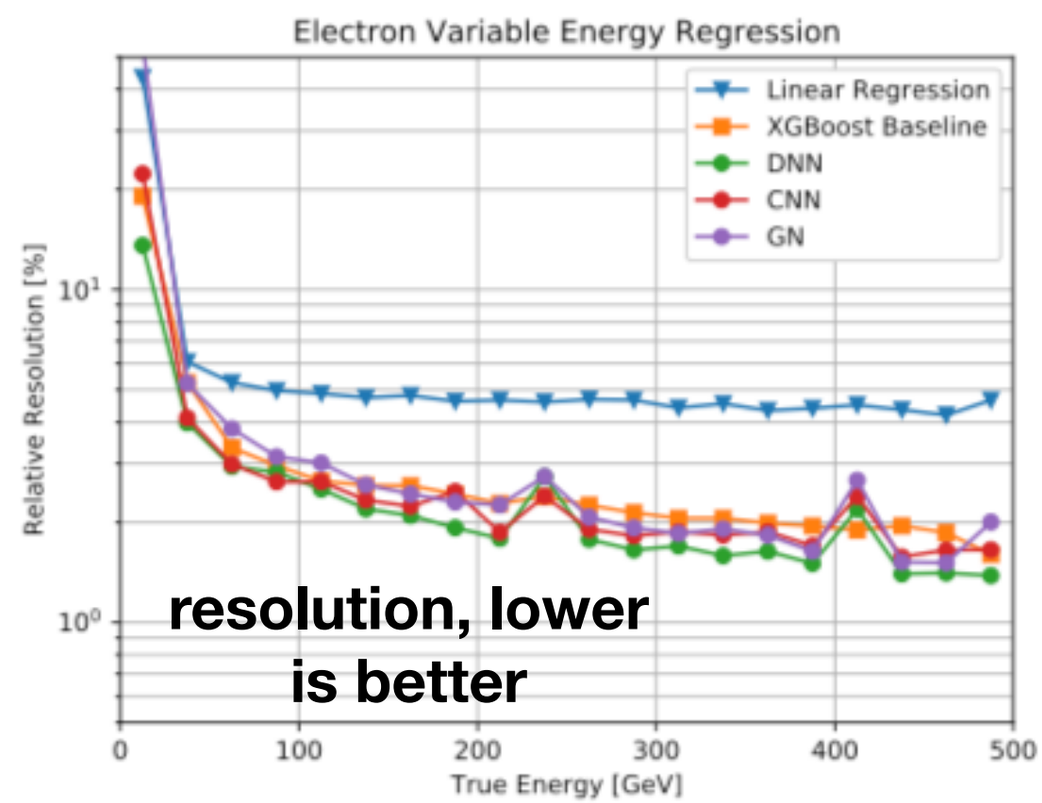
pi0 -> two photons



Belayneh, D., Carminati, F., Farbin, A. *et al.* Calorimetry with deep learning: particle simulation and reconstruction for collider physics. *Eur. Phys. J. C* **80,** 688 (2020). https://doi.org/10.1140/epjc/s10052-020-8251-9

Multi-task learning



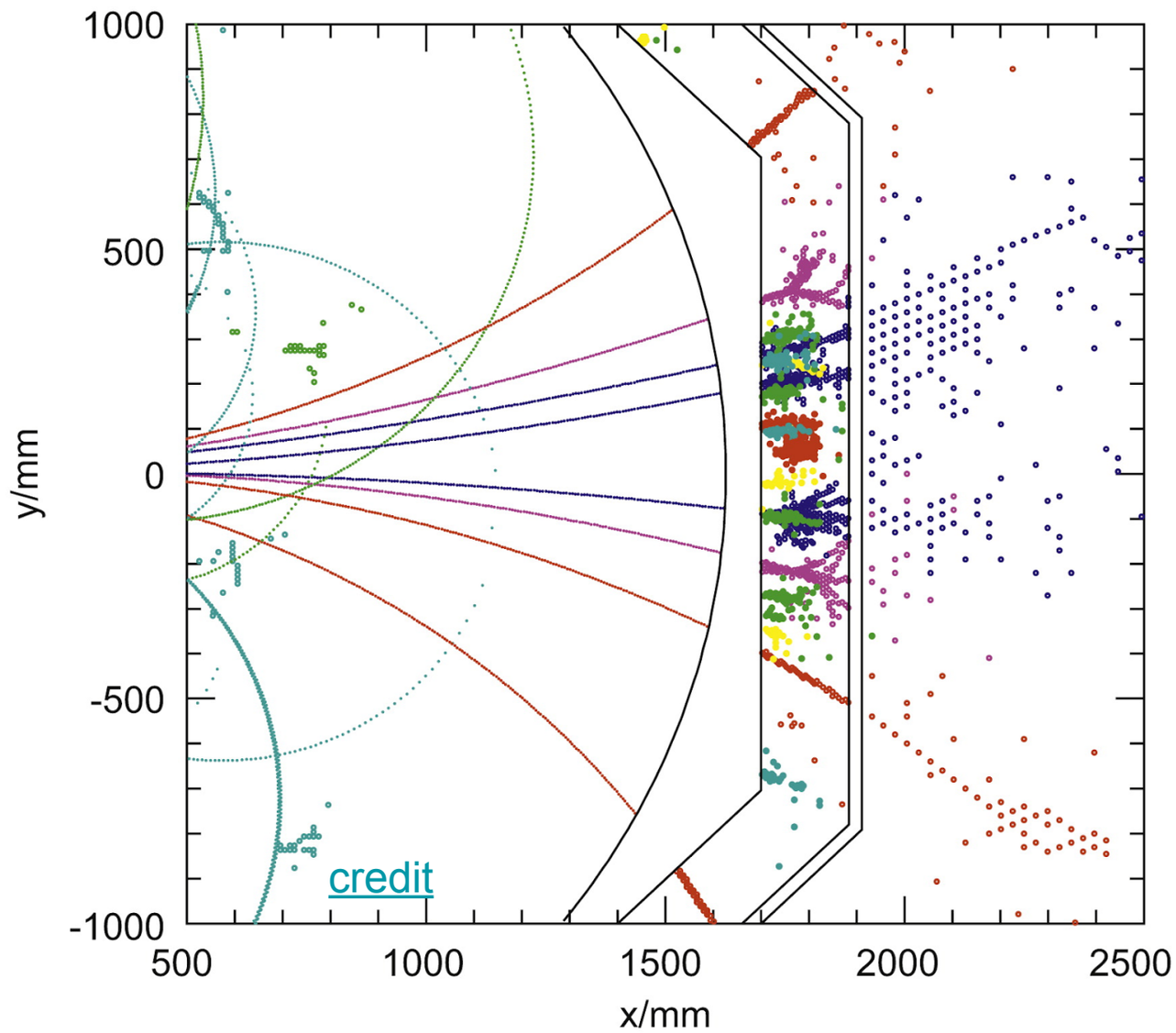


DNNs/CNNs on granular detetors are performant for shower identification and energy regression.

Belayneh, D., Carminati, F., Farbin, A. *et al.* Calorimetry with deep learning: particle simulation and reconstruction for collider physics. *Eur. Phys. J. C* **80,** 688 (2020). https://doi.org/10.1140/epjc/s10052-020-8251-9

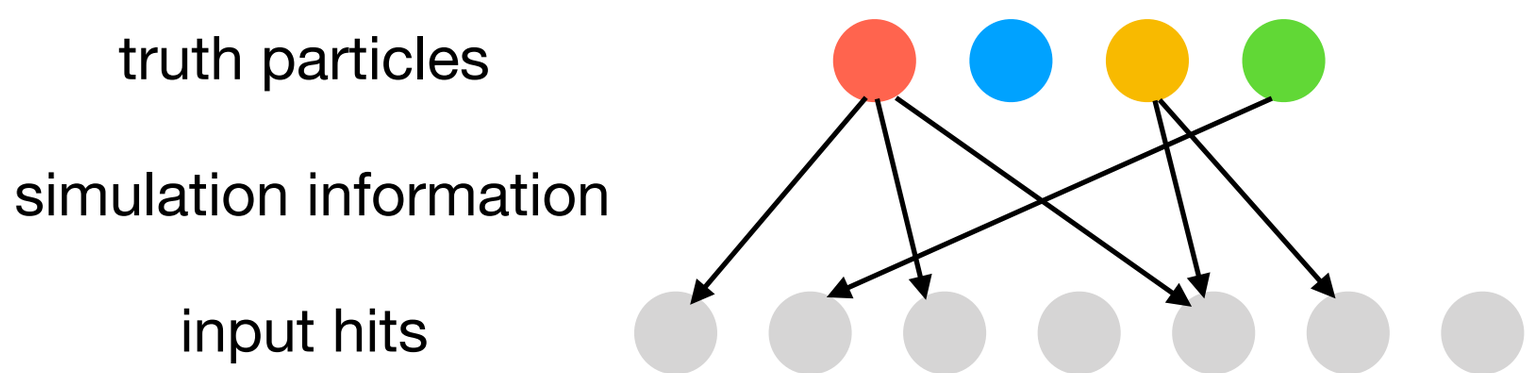
Calorimeter clustering

- Segment the energy deposits (hits) according to the originator particles
- The hits are embedded in a complicated feature space (Cartesian position, energy, signal significance, timing, layer information, ...)
- Showers from different particles may overlap spatially
- **Standard heuristic approaches** based on seeding & collecting neighbors, typically iterative



Particle representation

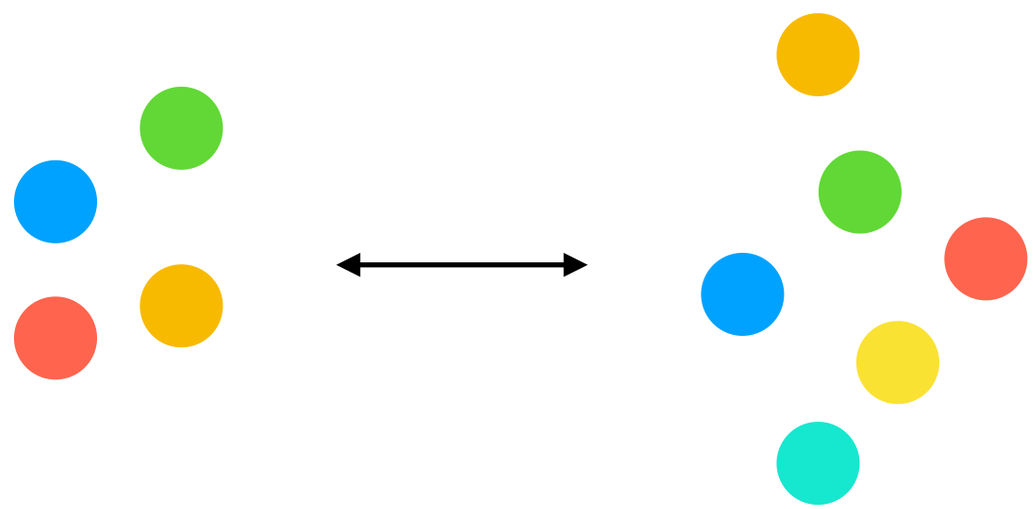
- The ground truth is a set of simulation particles (p4, ID)
- The input is the set of all calorimeter hits (energy, location)



An unknown number of different truth particles (segmentation labels).

Set-to-set problem

Each particle is described by a multi-class label, and is embedded in a complex, problem-dependent feature space.



ground truth
simulation input particles

prediction
reconstructed particles

How to compare two sets of arbitrary size with complex features?
How to do it differentiably, in a performant way?

Energy flow

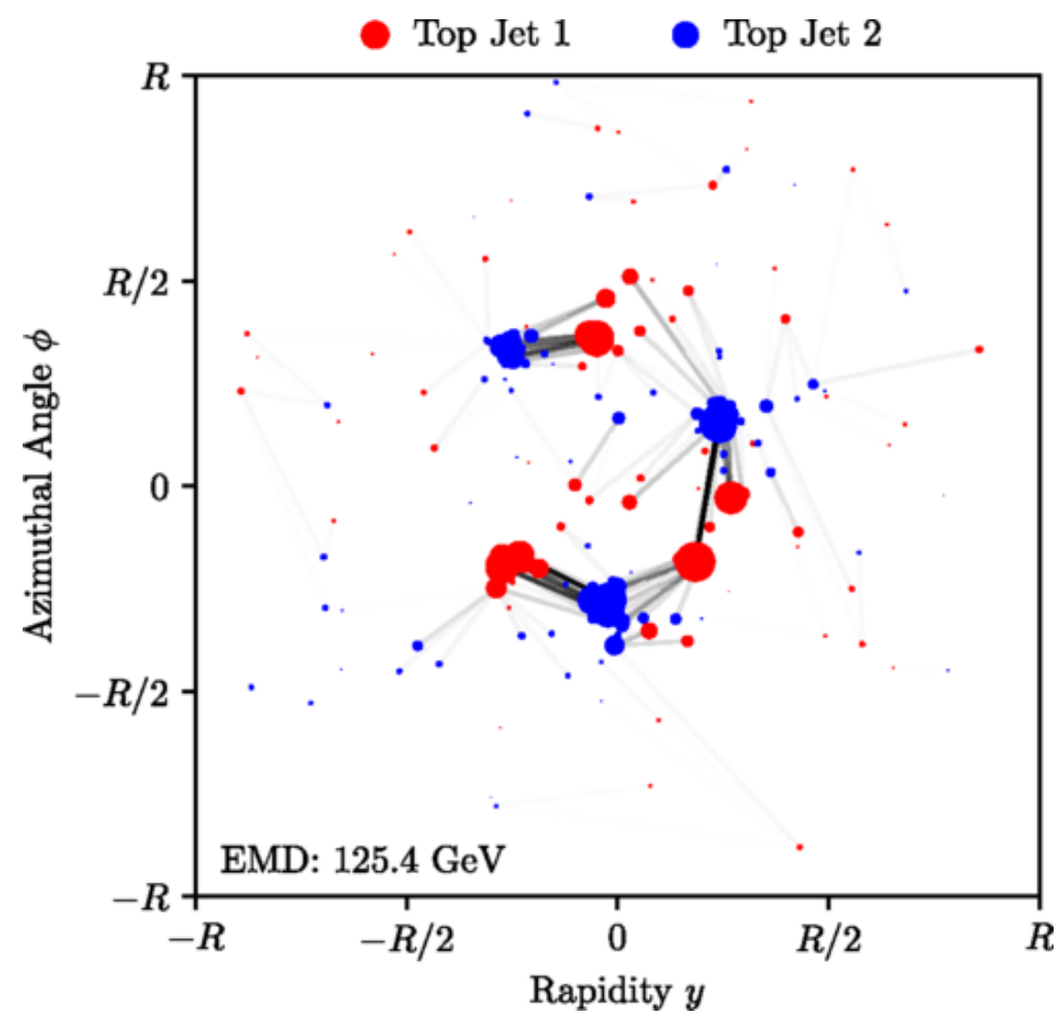
- Use Earth Mover's Distance to define a differentiable loss between two sets of particles described by (E, eta, phi)

- Good theoretical properties, not sensitive to soft particles / collinear radiation

- Optimal Transport is challenging to practically compute on large sets

$$\mathrm{EMD}(\mathcal{E}, \mathcal{E}') = \min_{\{f_{ij} \geq 0\}} \sum_{ij} f_{ij} \frac{\theta_{ij}}{R} + \left| \sum_i E_i - \sum_j E'_j \right|,$$

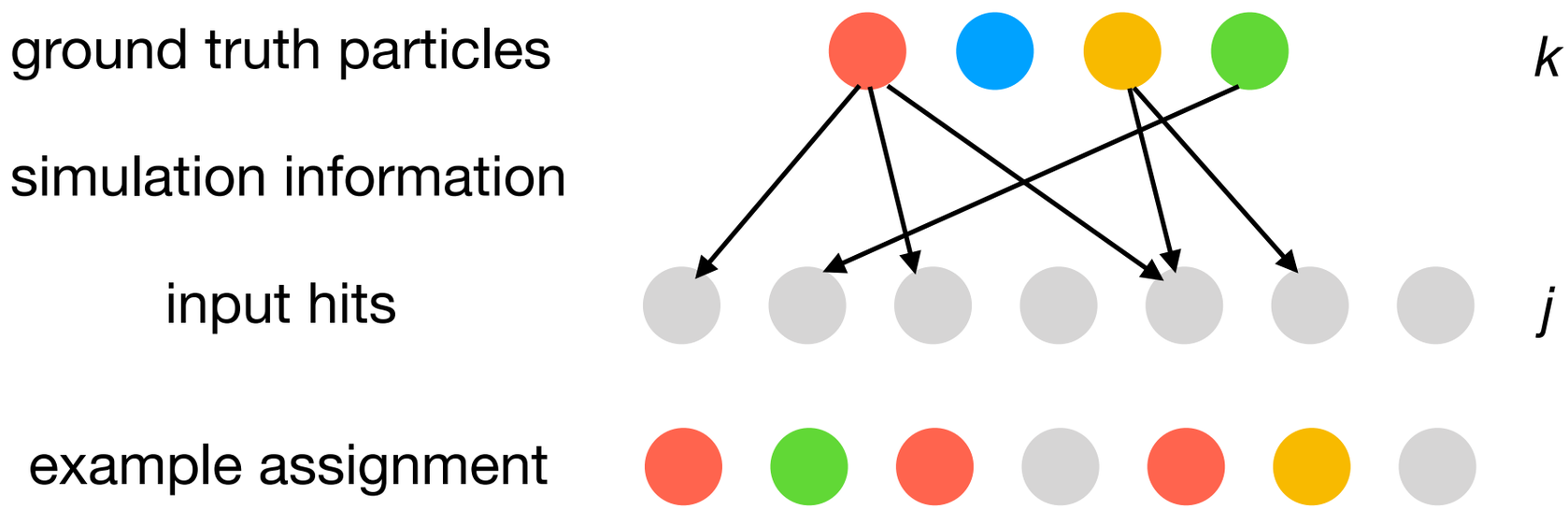
$$\sum_j f_{ij} \leq E_i, \qquad \sum_i f_{ij} \leq E'_j, \qquad \sum_{ij} f_{ij} = E_{\min},$$



Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler; Phys. Rev. Lett. **123**, 041801; https://doi.org/10.1103/PhysRevLett.123.041801

Object condensation

Boundedness: the number of truth particles usually cannot be larger than the number of inputs (typically it's much smaller).



Each input represents exactly one truth particle, with attractive/repulsive potentials in a learned space x_j between correct/incorrect assignments.

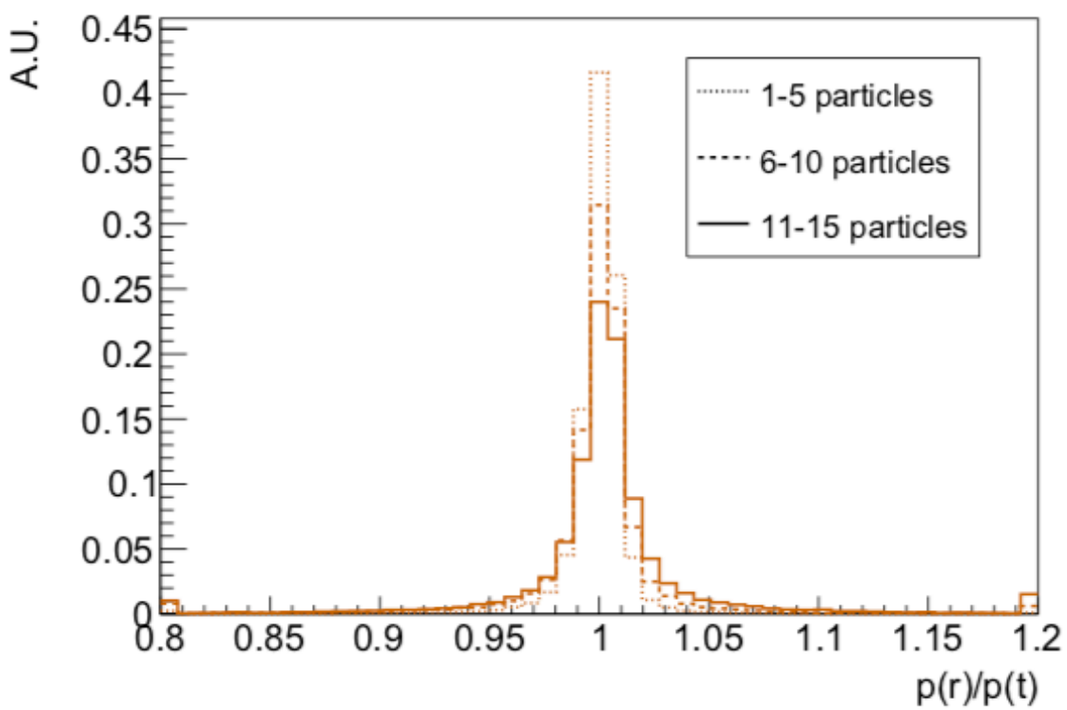
$$L_V = \frac{1}{N}\sum_{j=1}^{N} q_j \sum_{k=1}^{K} \Big(M_{jk}\check{V}_k(x_j) + (1 - M_{jk})\hat{V}_k(x_j) \Big) .$$

attractive repulsive

Kieseler, J. Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *Eur. Phys. J. C* **80,** 886 (2020). https://doi.org/10.1140/epjc/s10052-020-08461-2

A simplified set-to-set loss

This approximation is fairly model-independent (e.g. not tied to GNNs). The exact form of the potentials is a hyperparameter.



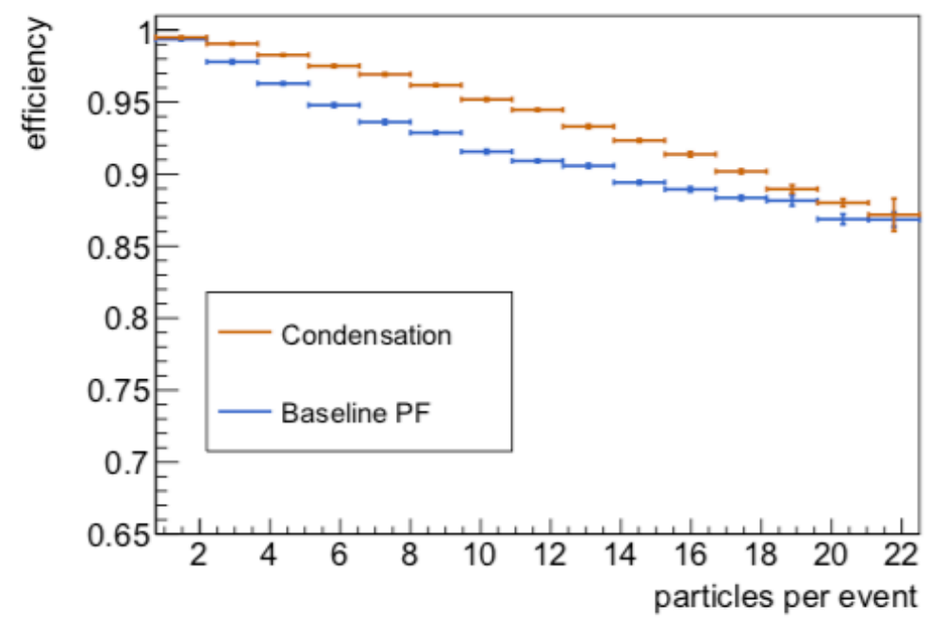


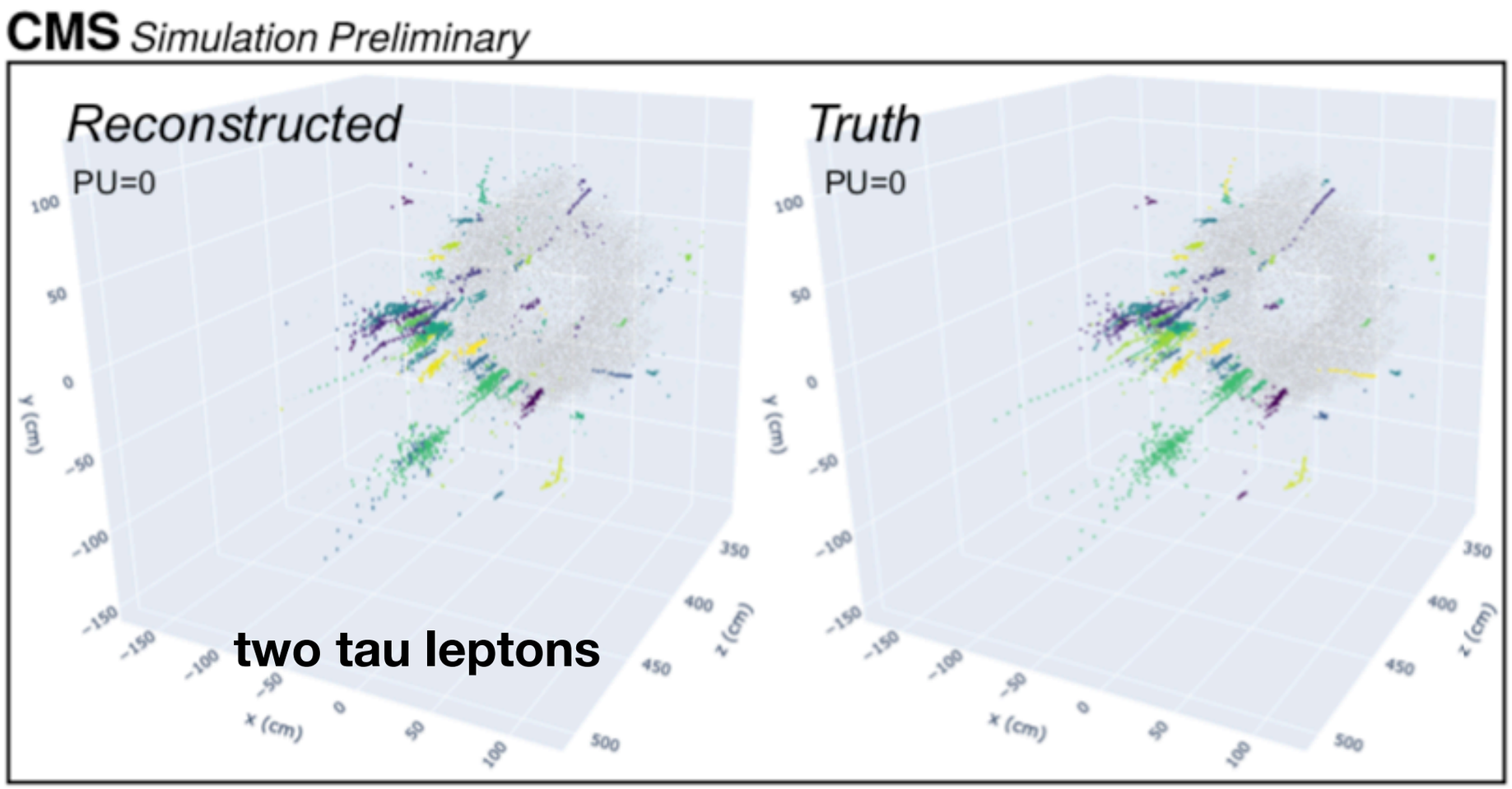
Fig. 6 Reconstruction efficiency as a function of the particle multiplicity in the event

Can be used for constructing particle reconstruction models across a varied number of inputs.

Kieseler, J. Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *Eur. Phys. J. C* **80,** 886 (2020). https://doi.org/10.1140/epjc/s10052-020-08461-2

Realistic clustering with ML

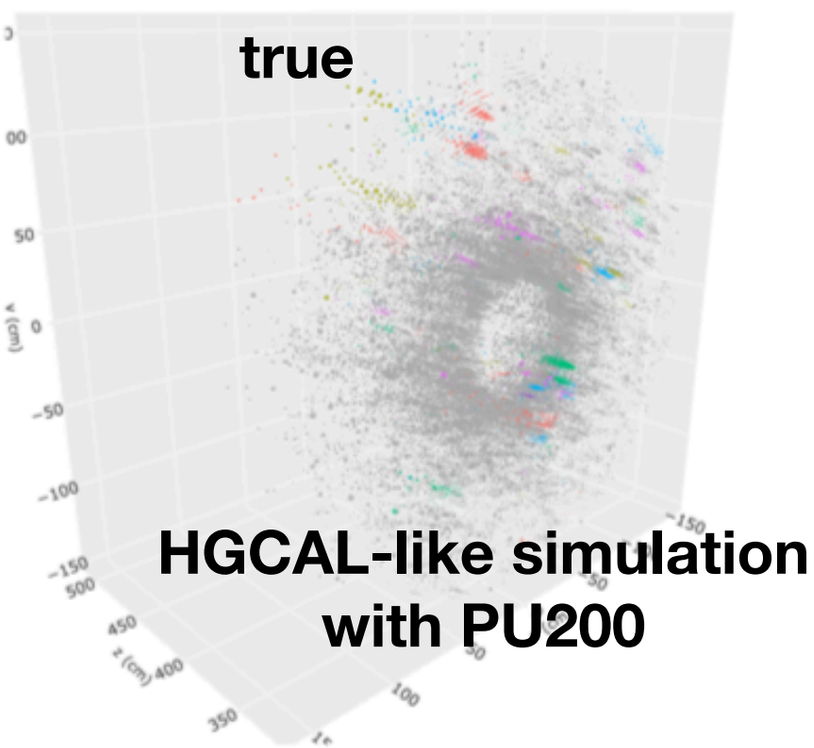
Simulation-level particles → simulation energy deposits → reconstructed energy deposits → predict the cluster label (or noise) for each hit.

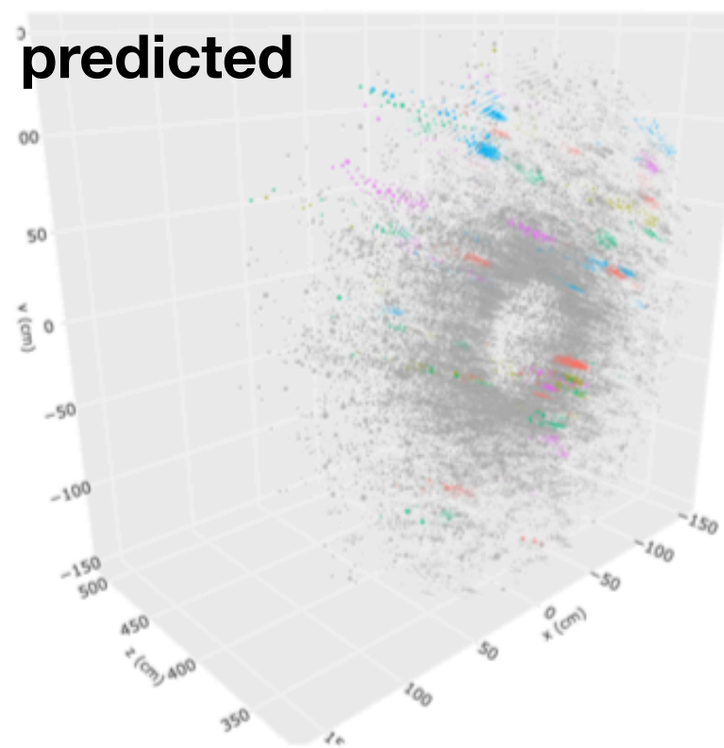


Saptaparna Bhattacharya, Nadezda Chernyavskaya, Saranya Ghosh, Lindsey Gray, Jan Kieseler et al. GNN-based end-to-end reconstruction in the CMS Phase 2 High-Granularity Calorimeter. ACAT 2021. https://doi.org/10.48550/arXiv.2203.01189

Clustering to particle reconstruction

The physics task involves in using the clustering to predict the energy of the particle initiating the clustered shower.





Loss function over cluster, regressing the energy

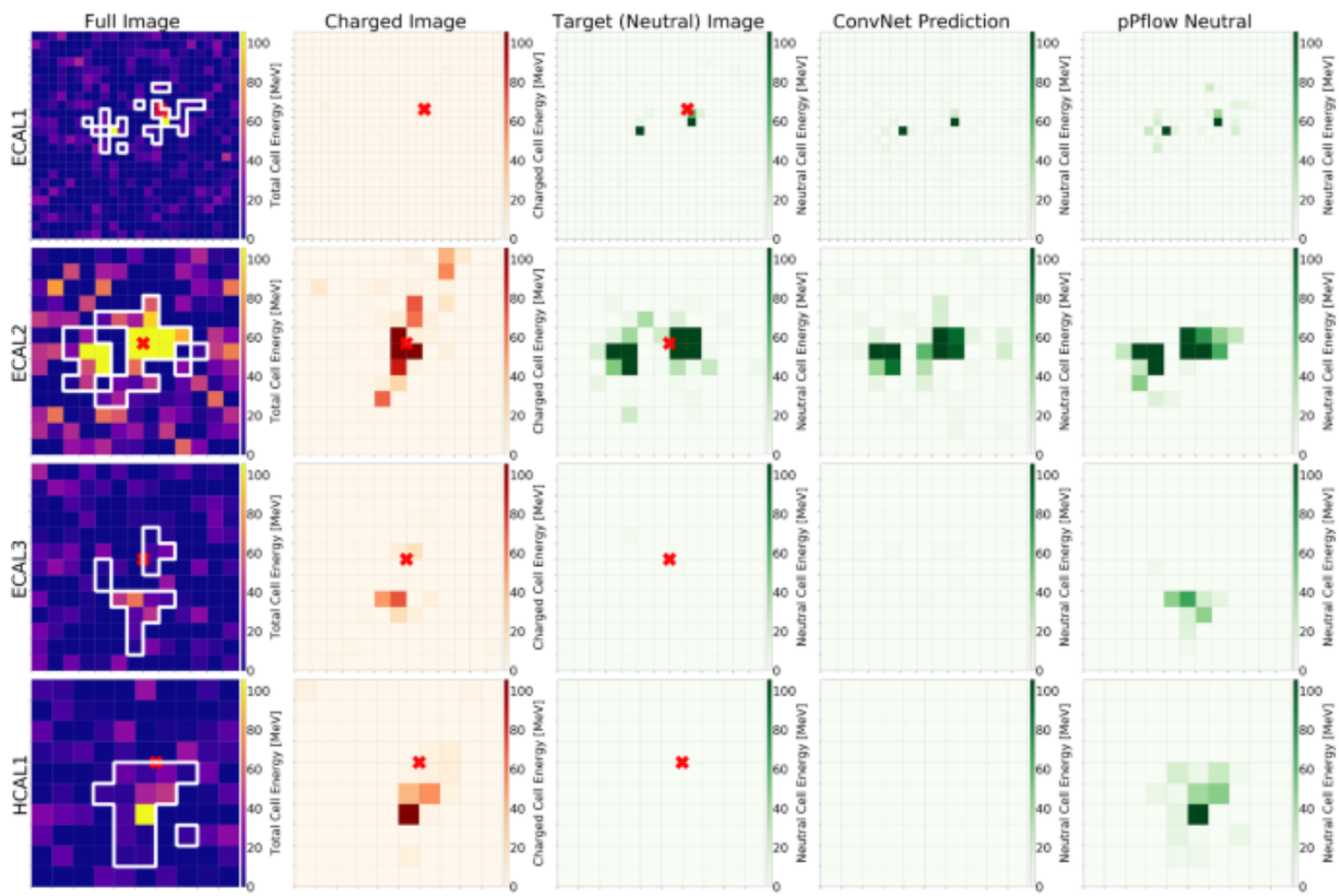
$$L_P = \sum_{t \in T} \frac{w_t}{\sum_{h \in H_t} \xi(h)} \sum_{h \in H_t} \xi(h) L_E,$$

$$L_E = \log\left(\left(\frac{E_{\text{true},t} - \psi_h E_{\text{dep},t}}{\sqrt{E_{\text{true},t}} + 0.003}\right)^2 + 1\right),$$

Shah Rukh Qasim, Nadezda Chernyavskaya, Jan Kieseler, Kenneth Long, Oleksandr Viazlo et al. End-to-end multi-particle reconstruction in high occupancy imaging calorimeters with graph neural networks. https://doi.org/10.48550/arXiv.2204.01681

Particle Flow with ML

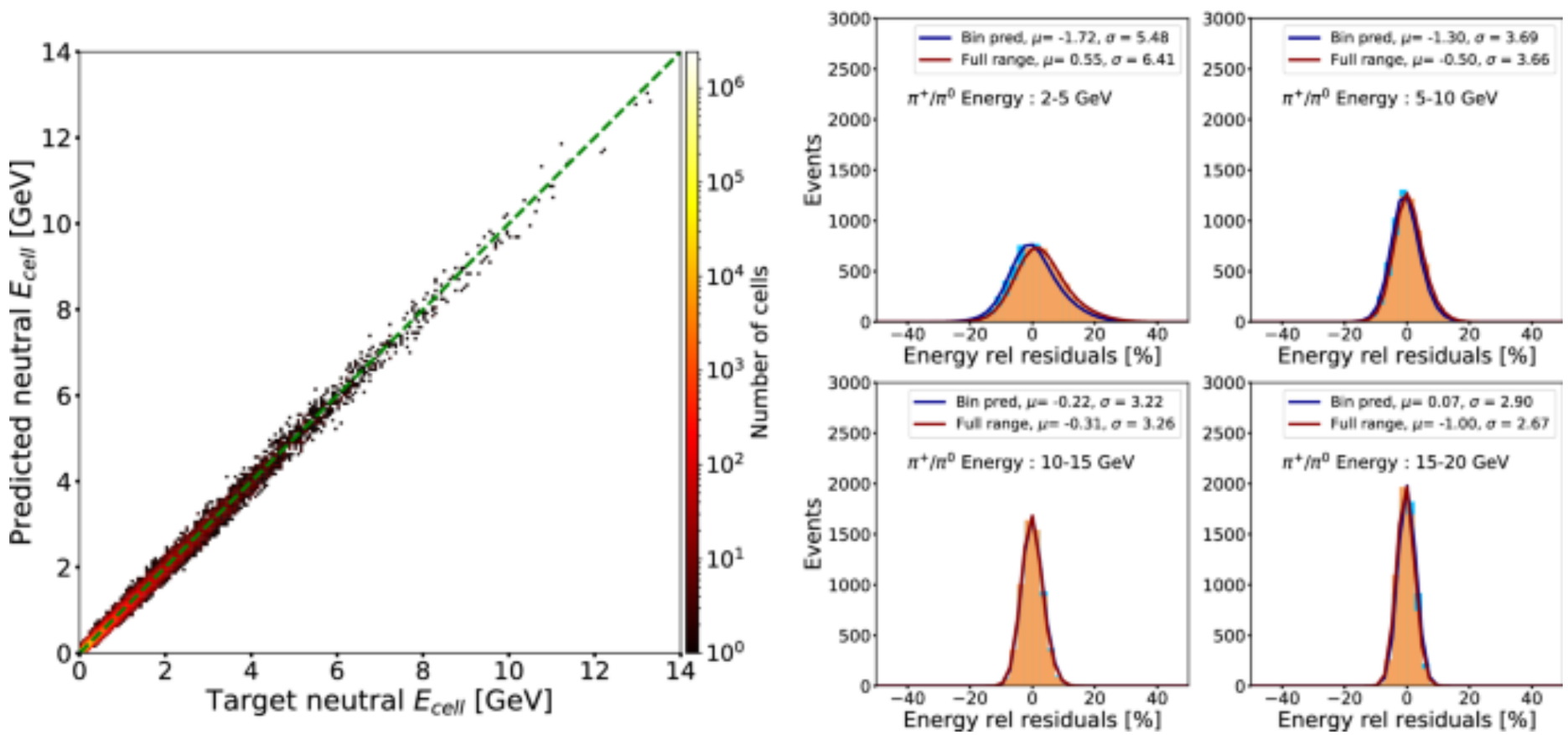
The full event: a multilayered calorimetric image + tracks.
Predict the neutral energy deposits.



Di Bello, F.A., Ganguly, S., Gross, E. et al. Towards a computer vision particle flow. Eur. Phys. J. C 81, 107 (2021). https://doi.org/10.1140/epjc/s10052-021-08897-0

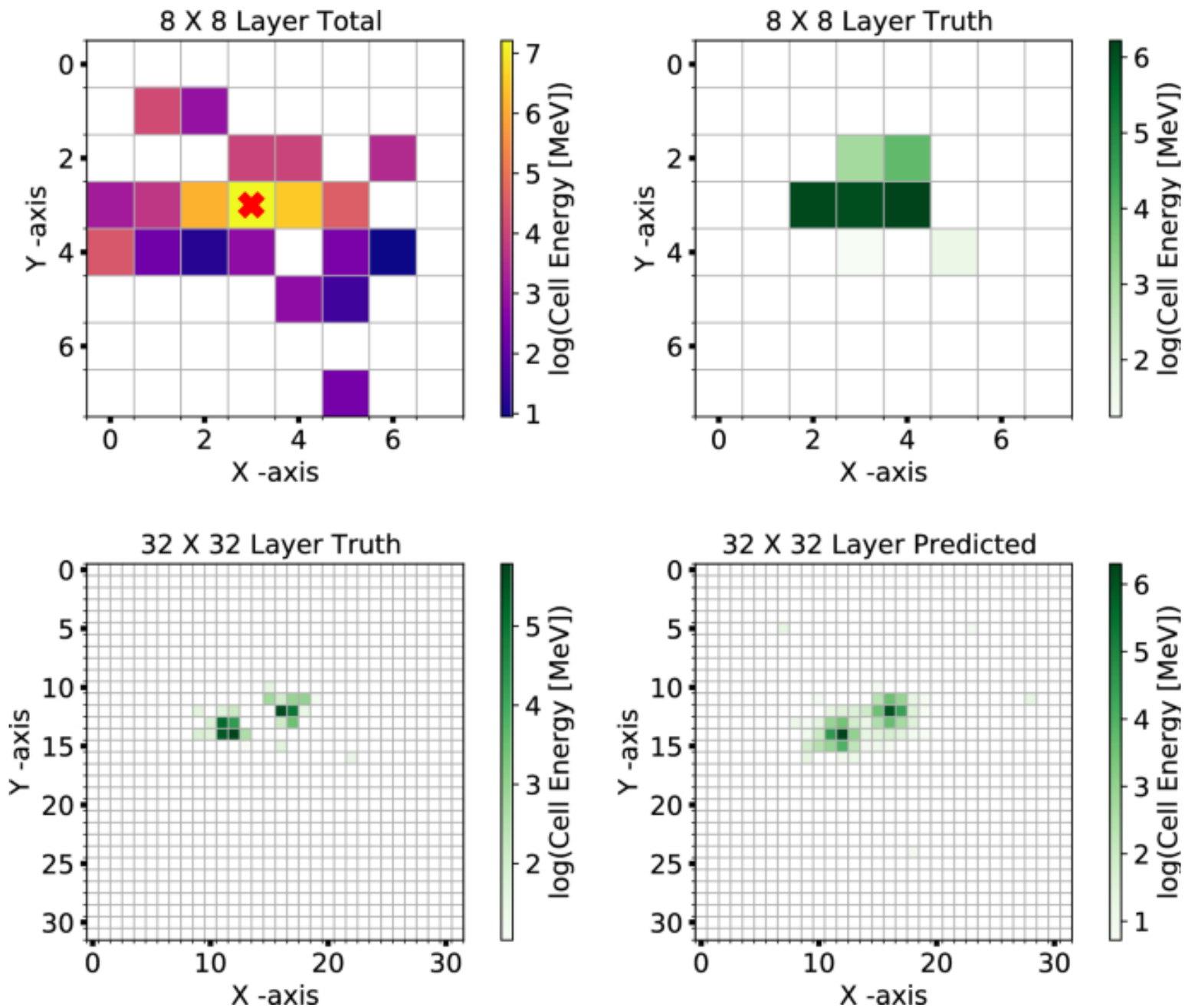
Neutral energy regression

The image-based approach is competitive for the cell neutral energy prediction compared to the algorithmic baseline.



Di Bello, F.A., Ganguly, S., Gross, E. et al. Towards a computer vision particle flow. Eur. Phys. J. C 81, 107 (2021). https://doi.org/10.1140/epjc/s10052-021-08897-0

Super-resolution

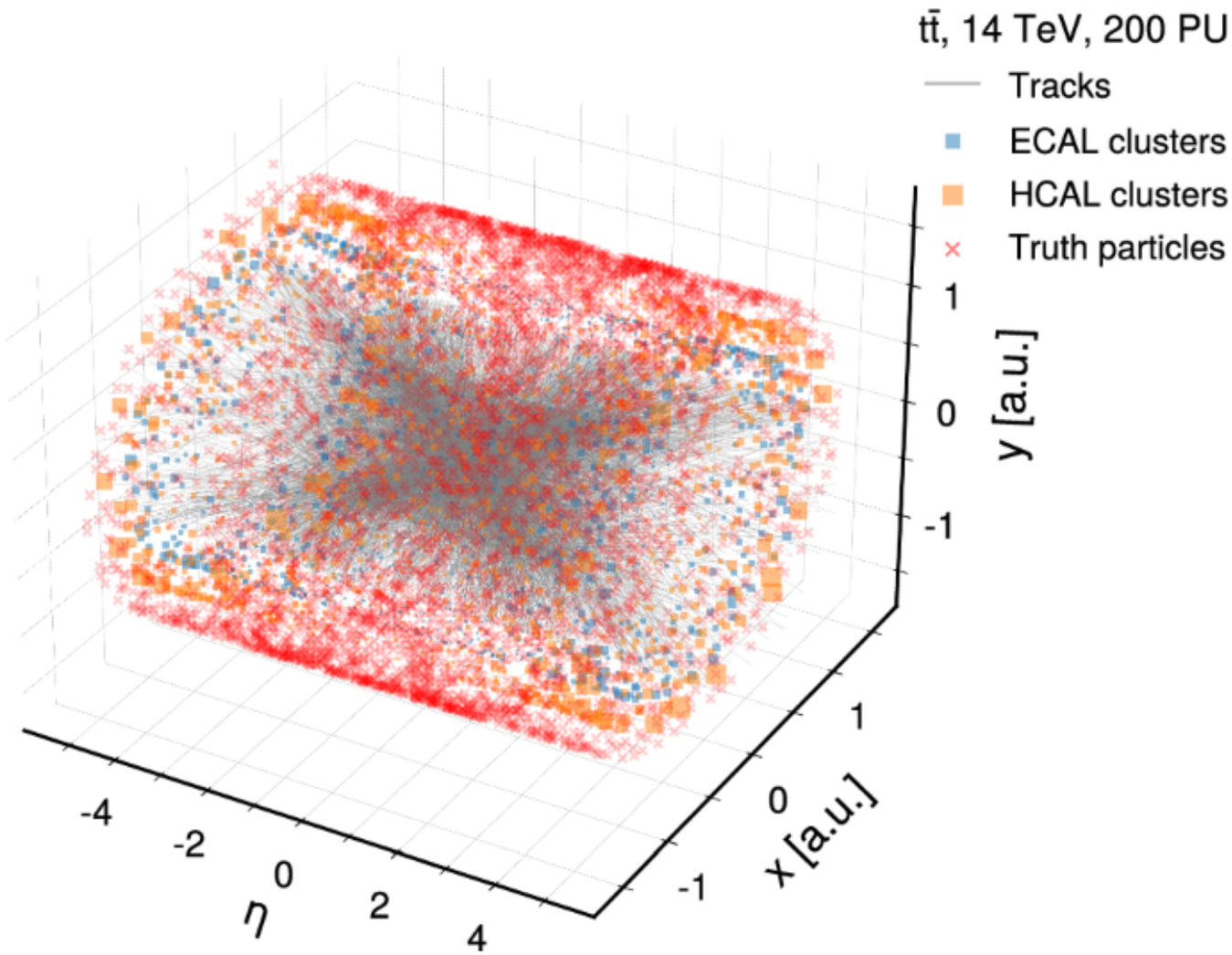


Di Bello, F.A., Ganguly, S., Gross, E. et al. Towards a computer vision particle flow. Eur. Phys. J. C 81, 107 (2021). https://doi.org/10.1140/epjc/s10052-021-08897-0

Sparse representations

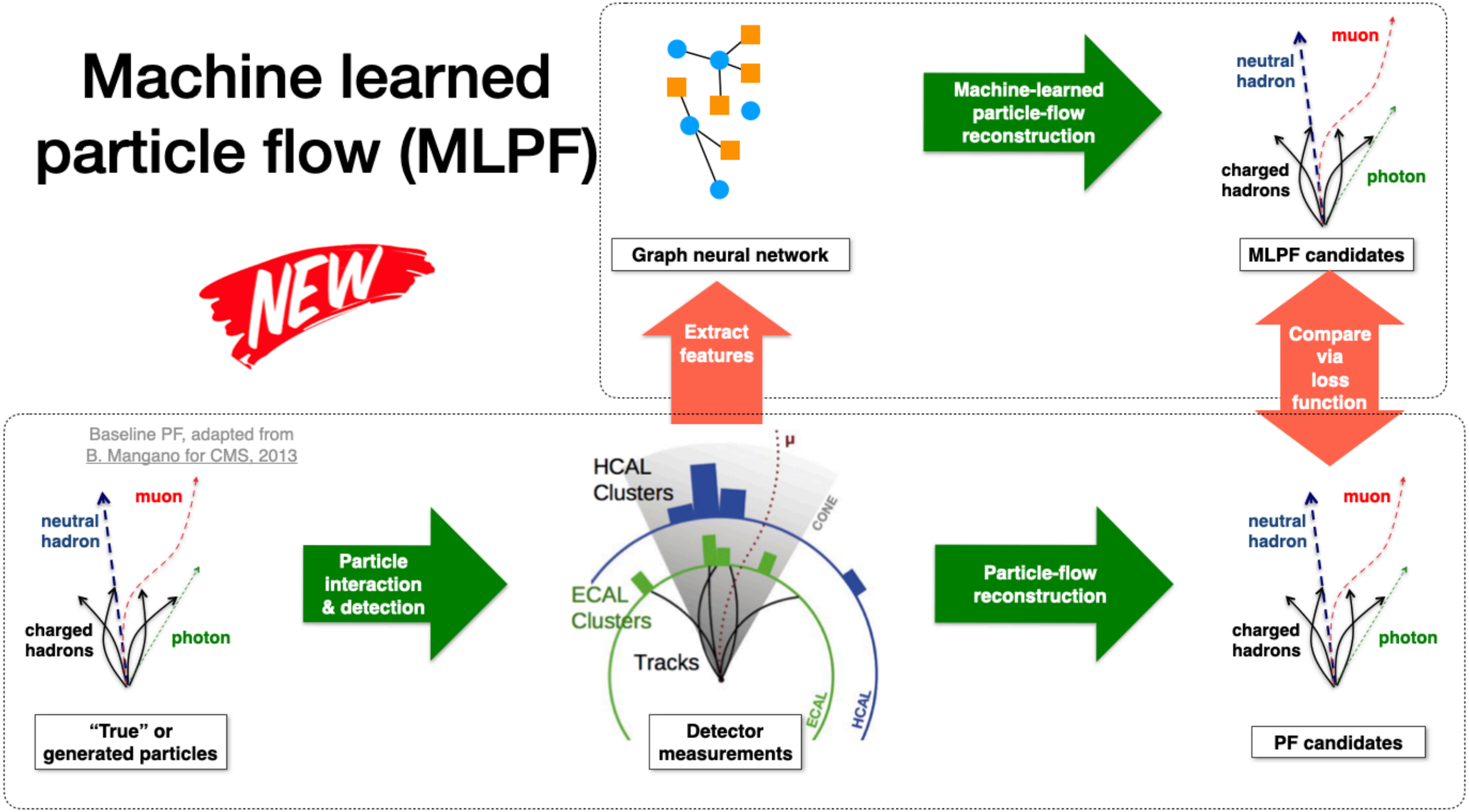
Starting from tracks and calorimeter clusters, aim to reconstruct the full set of input particles.

Inputs are heterogeneous, no natural underlying topology or associations.



Pata, J., Duarte, J., Vlimant, JR. et al. MLPF: efficient machine-learned particle-flow reconstruction using graph neural networks. Eur. Phys. J. C 81, 381 (2021). https://doi.org/10.1140/epjc/s10052-021-09158-w

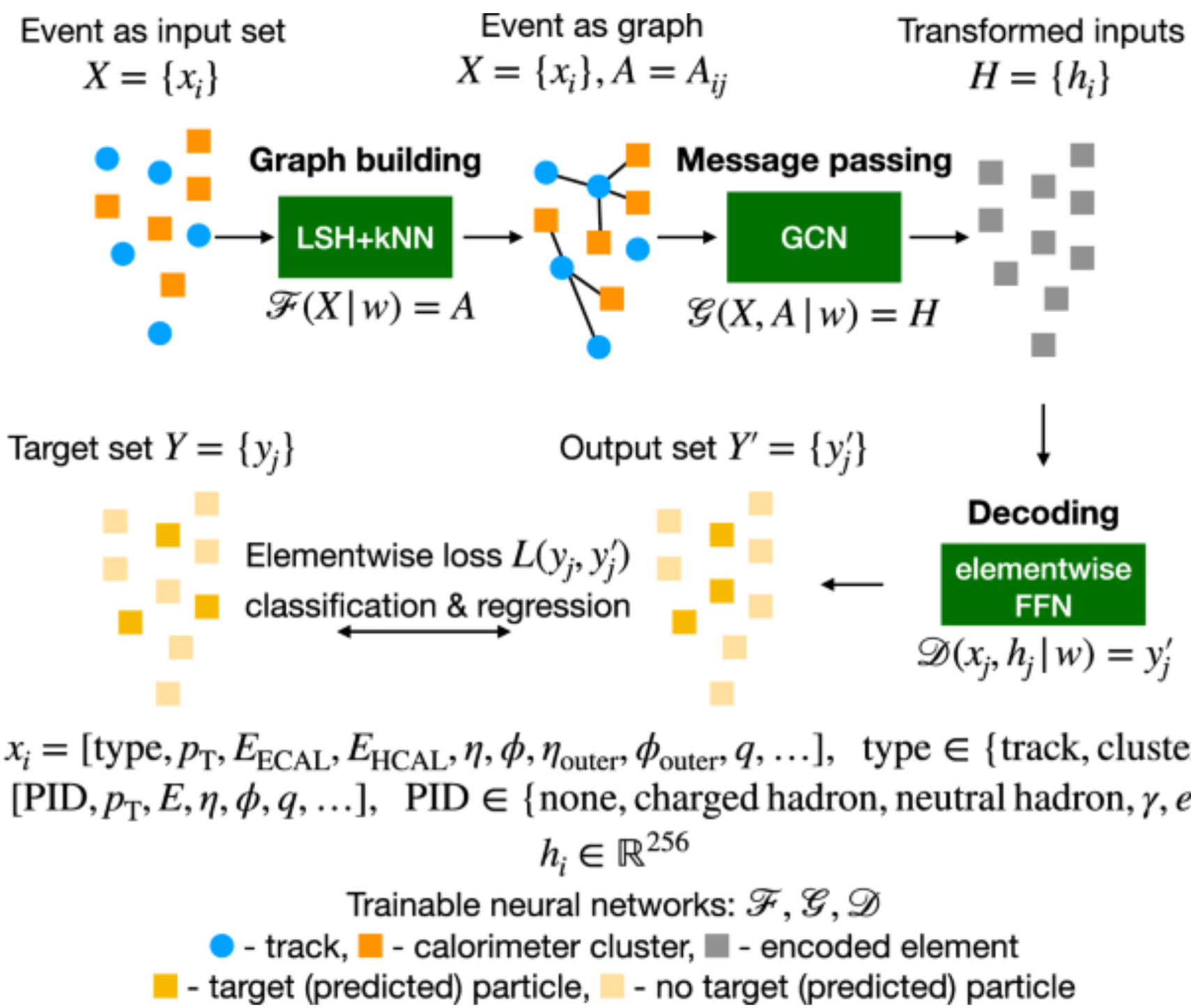
Machine learned particle flow (MLPF)



Pata, J., Duarte, J., Vlimant, JR. et al. MLPF: efficient machine-learned particle-flow reconstruction using graph neural networks. Eur. Phys. J. C 81, 381 (2021). https://doi.org/10.1140/epjc/s10052-021-09158-w

Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

A simplification: treat the inputs as a homogenous set.



$$x_i = [\text{type}, p_T, E_{\text{ECAL}}, E_{\text{HCAL}}, \eta, \phi, \eta_{\text{outer}}, \phi_{\text{outer}}, q, \ldots], \quad \text{type} \in \{\text{track}, \text{cluster}\}$$

$$y_j = [\text{PID}, p_T, E, \eta, \phi, q, \ldots], \quad \text{PID} \in \{\text{none}, \text{charged hadron}, \text{neutral hadron}, \gamma, e^{\pm}, \mu^{\pm}\}$$

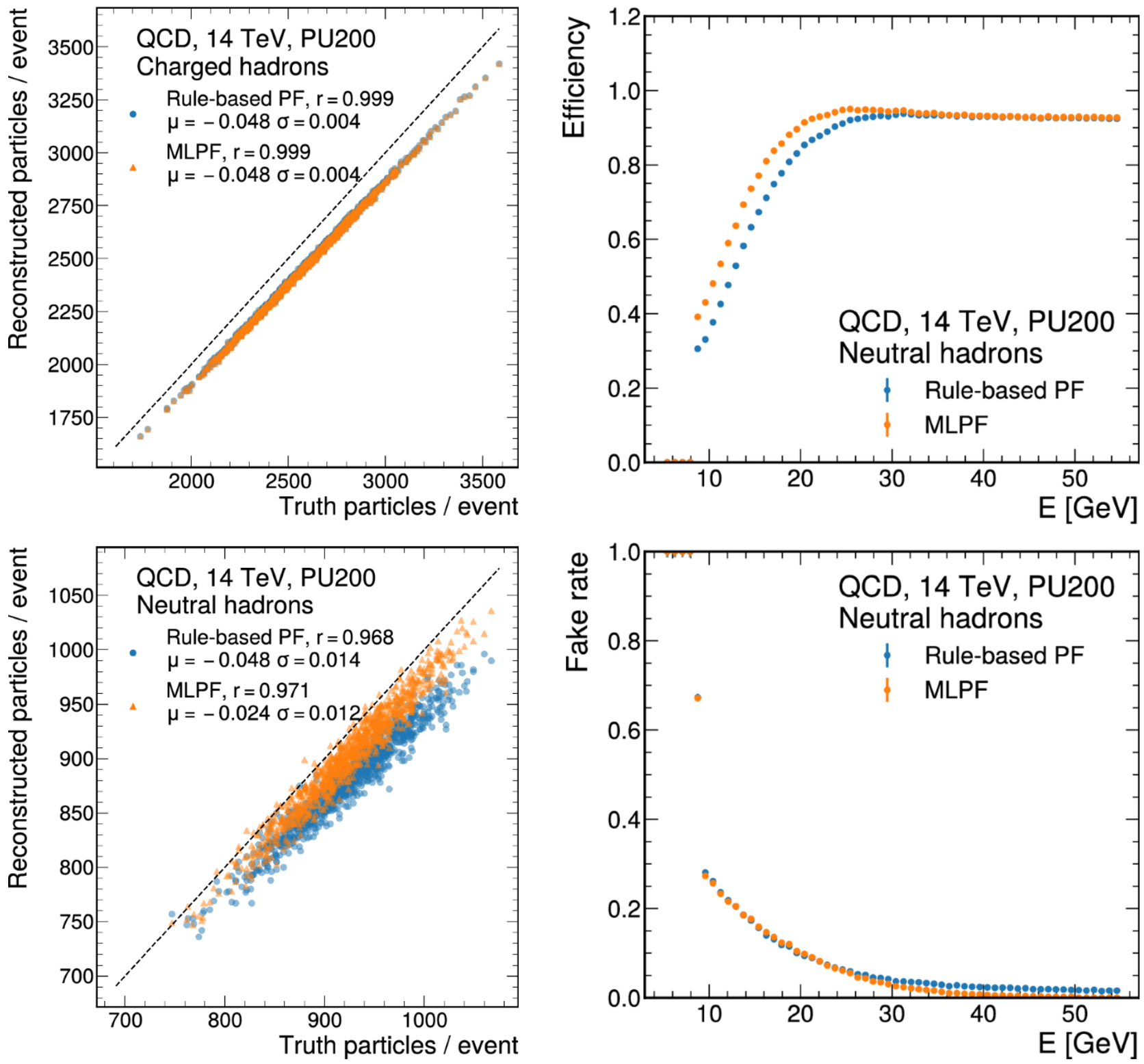
$$h_i \in \mathbb{R}^{256}$$

Trainable neural networks: $\mathscr{F}, \mathscr{G}, \mathscr{D}$

● - track, ■ - calorimeter cluster, ■ - encoded element

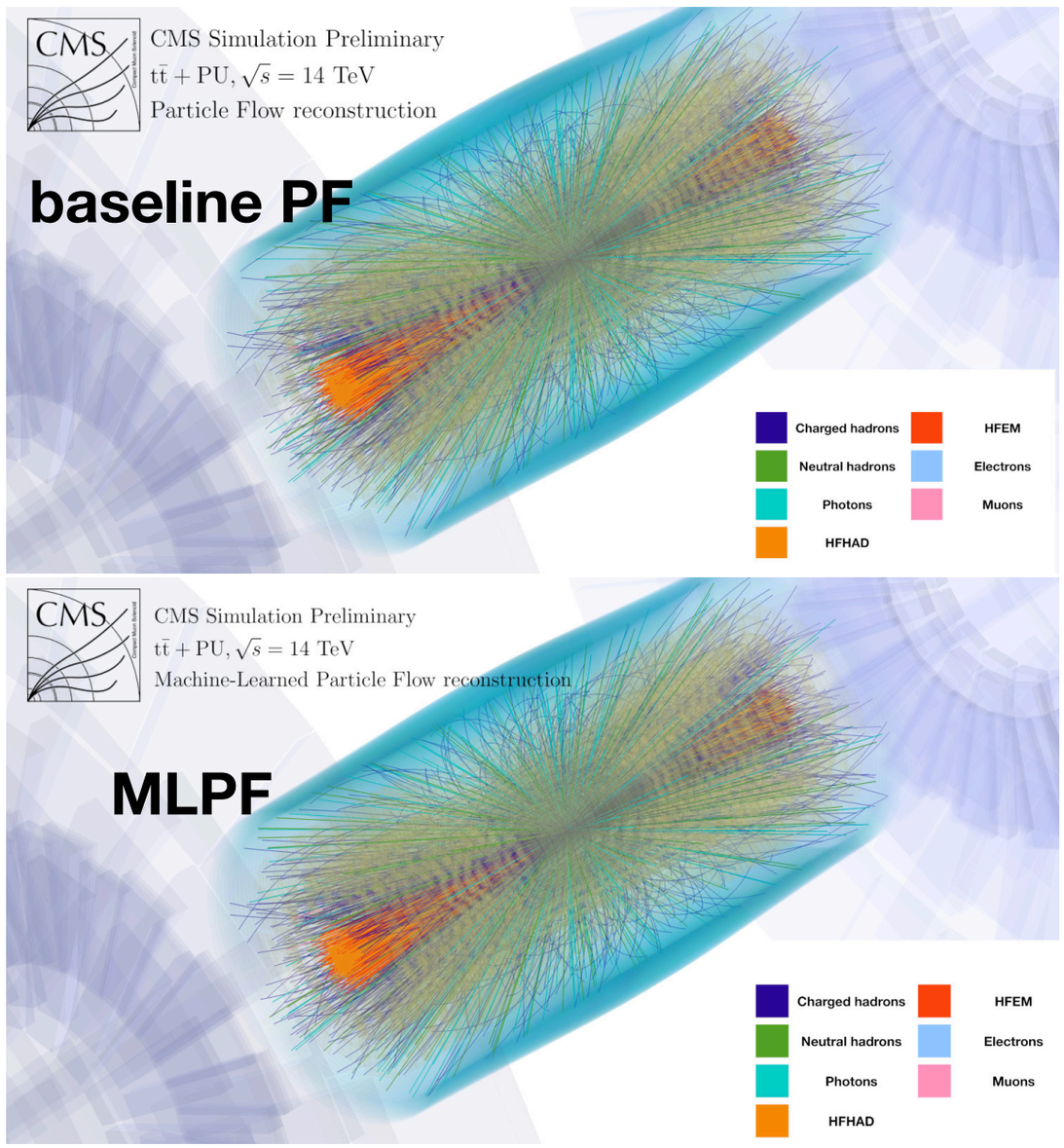
■ - target (predicted) particle, ■ - no target (predicted) particle

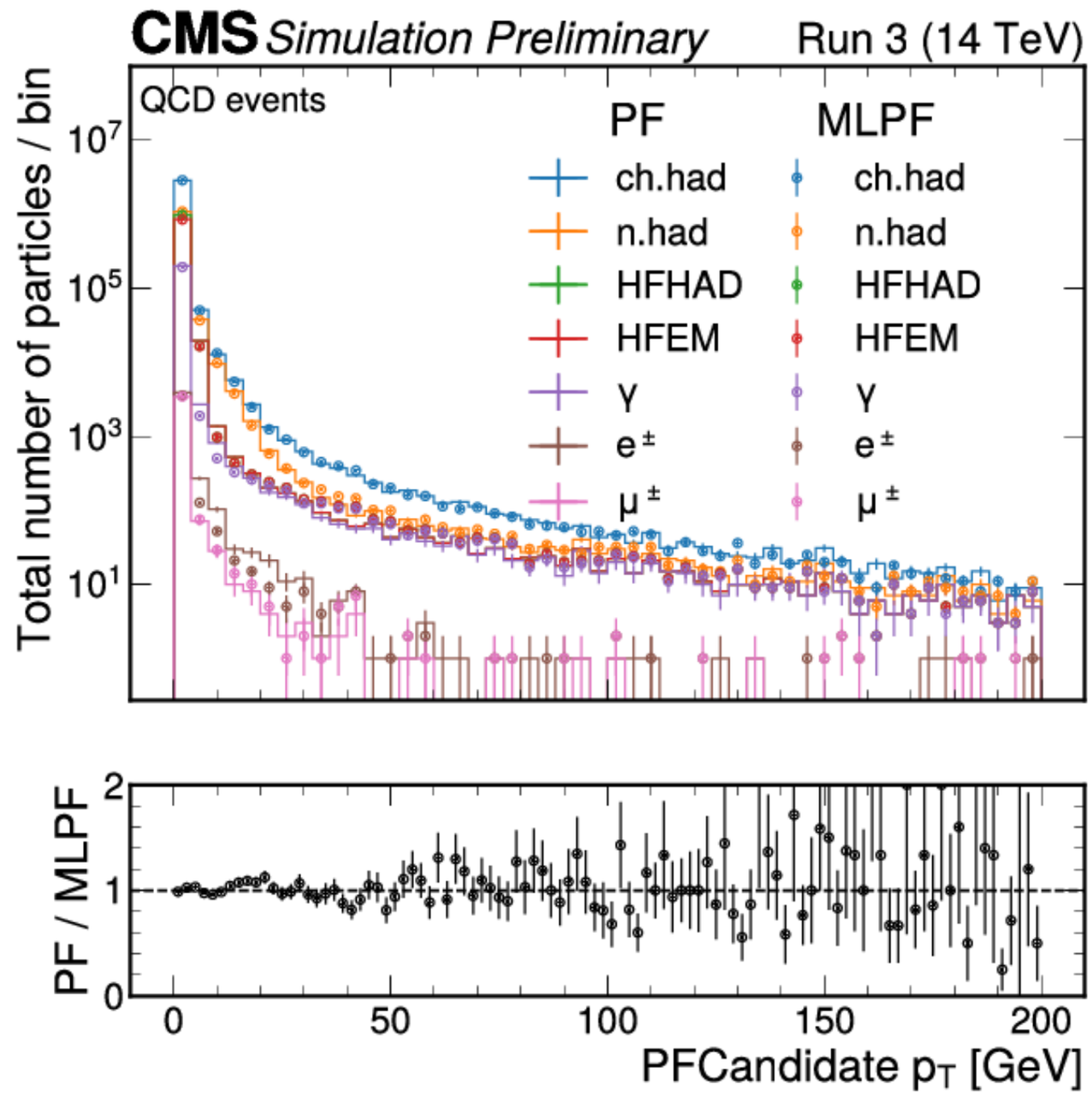
Pata, J., Duarte, J., Vlimant, JR. et al. MLPF: efficient machine-learned particle-flow reconstruction using graph neural networks. Eur. Phys. J. C 81, 381 (2021). https://doi.org/10.1140/epjc/s10052-021-09158-w



Pata, J., Duarte, J., Vlimant, JR. et al. MLPF: efficient machine-learned particle-flow reconstruction using graph neural networks. Eur. Phys. J. C 81, 381 (2021). https://doi.org/10.1140/epjc/s10052-021-09158-w

In a realistic environment

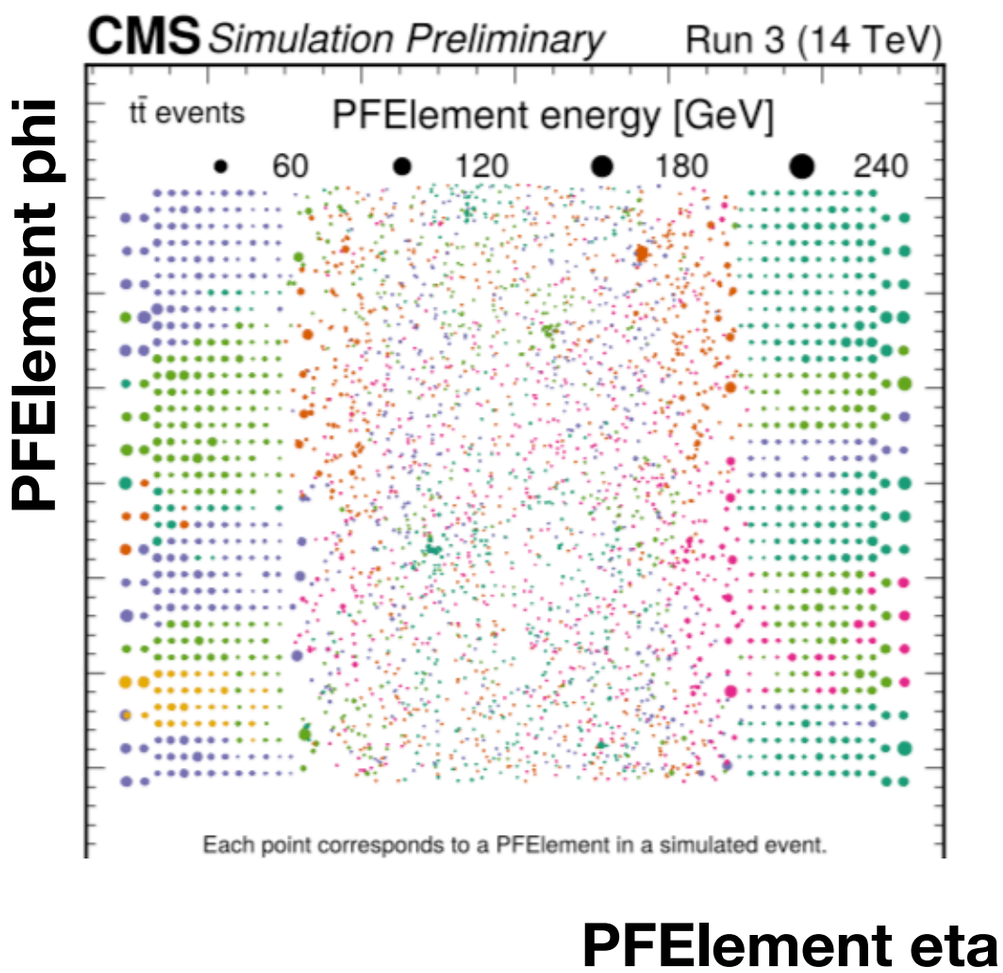


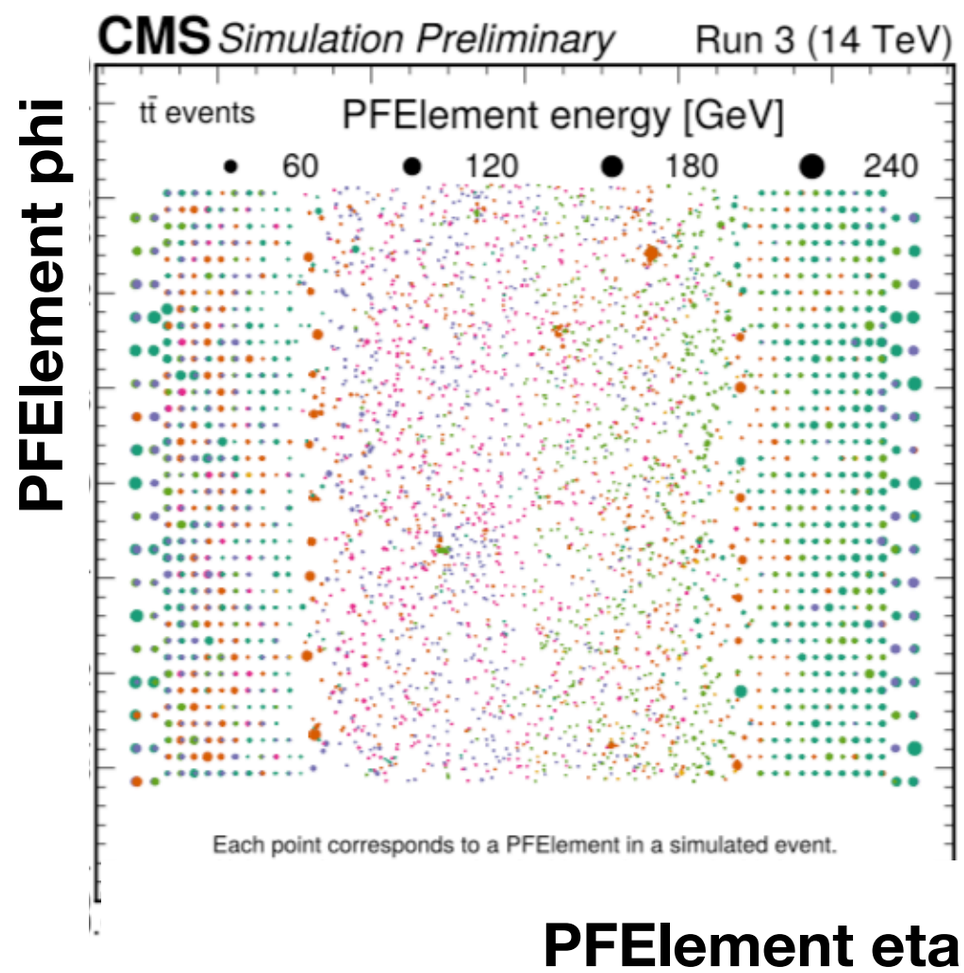


Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

Clustering to reconstruction

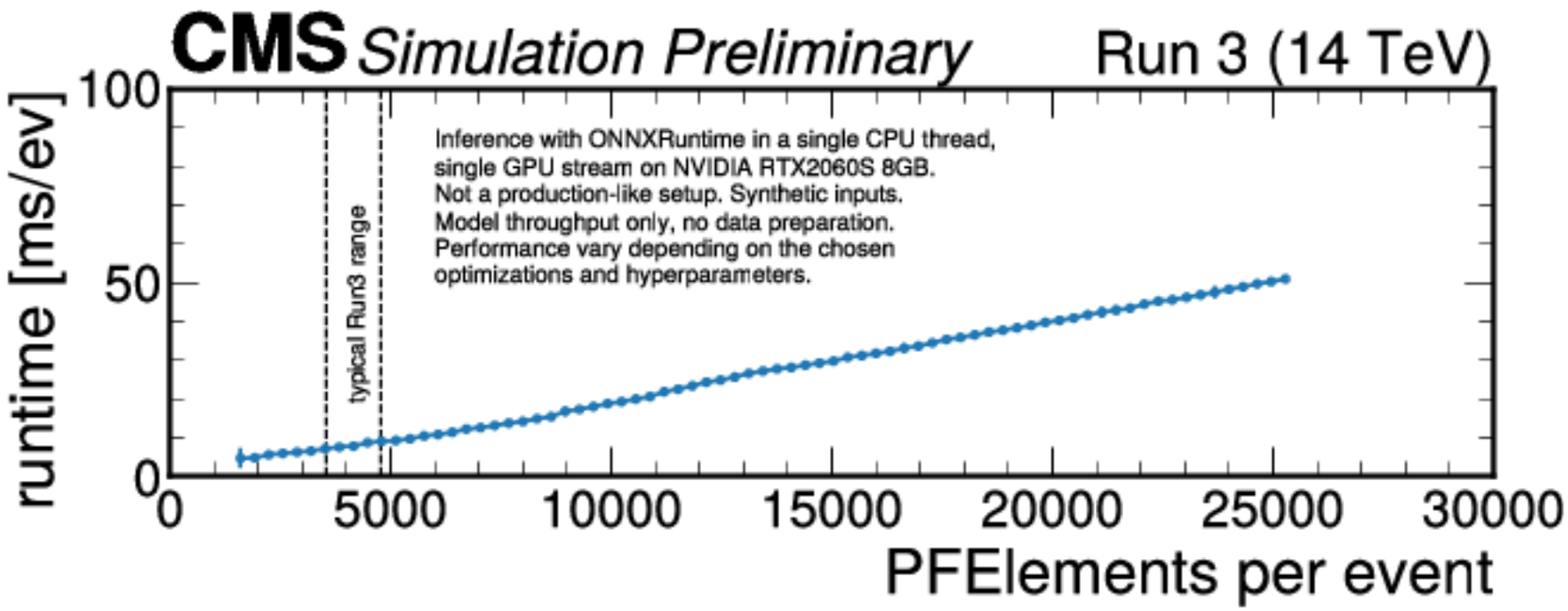
Clustering (graph building) is an internal model parameter, not a model target.

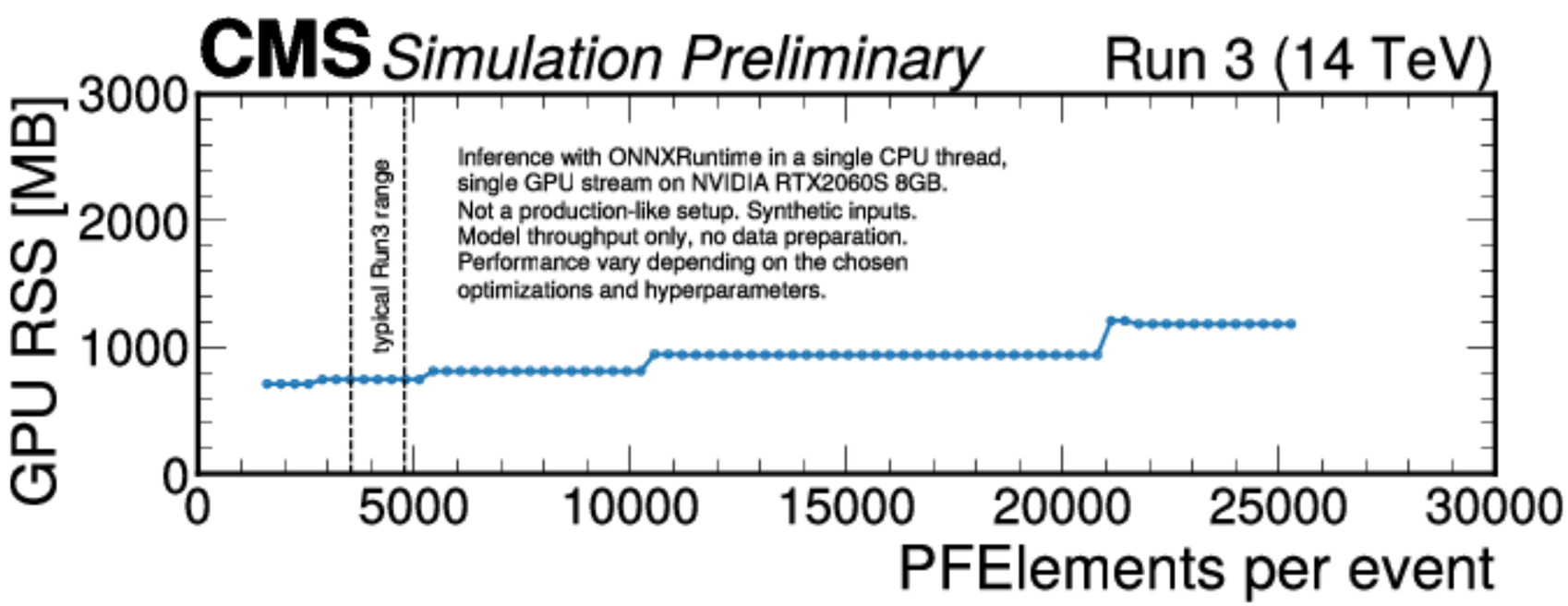




Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

Computational scalability

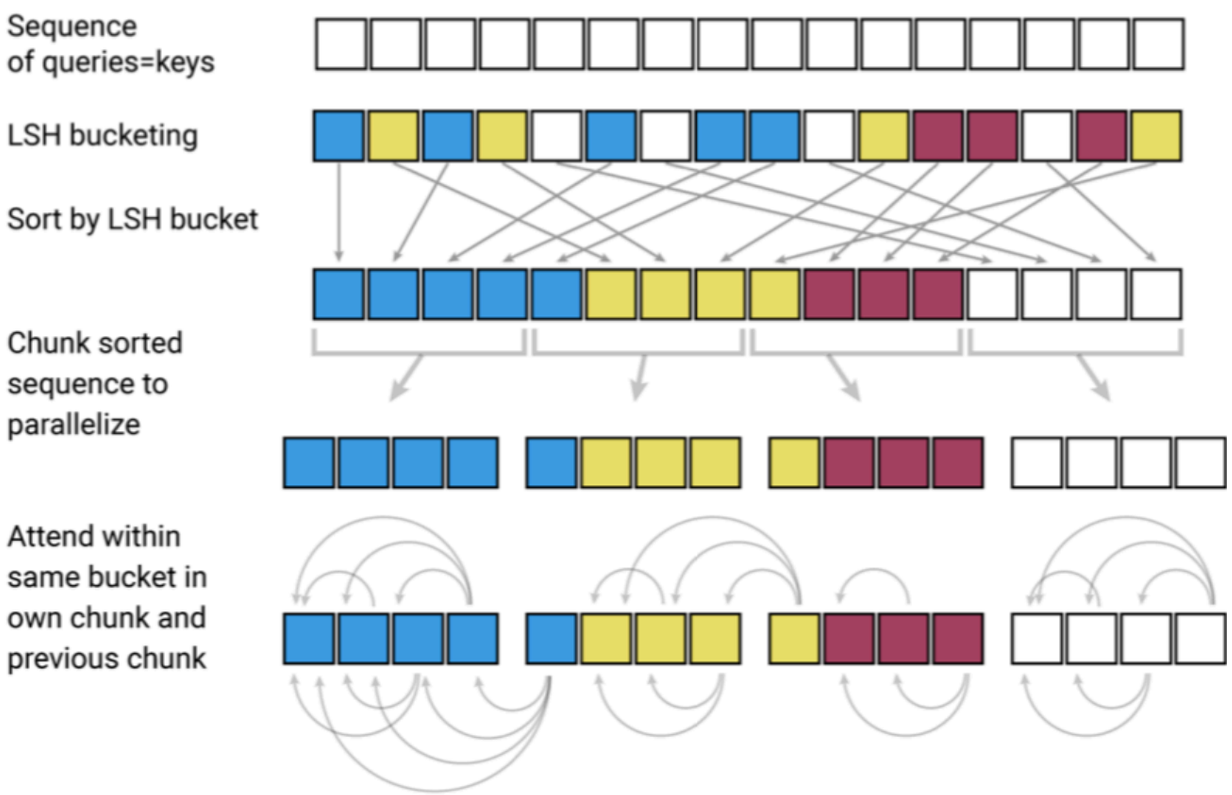


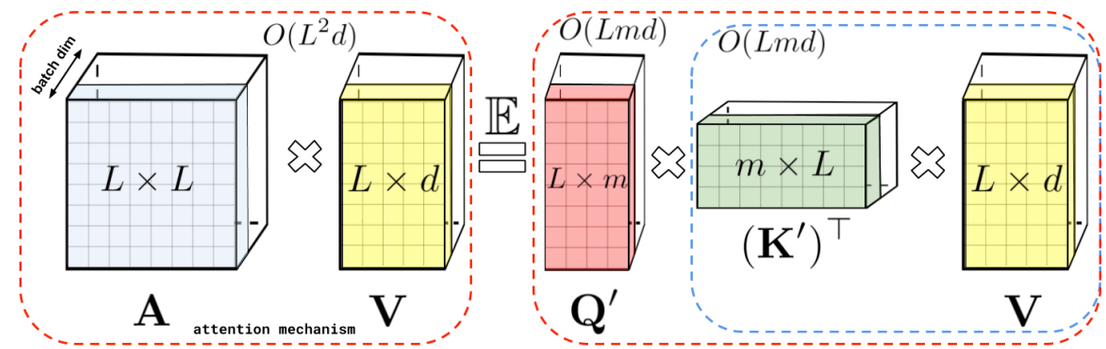


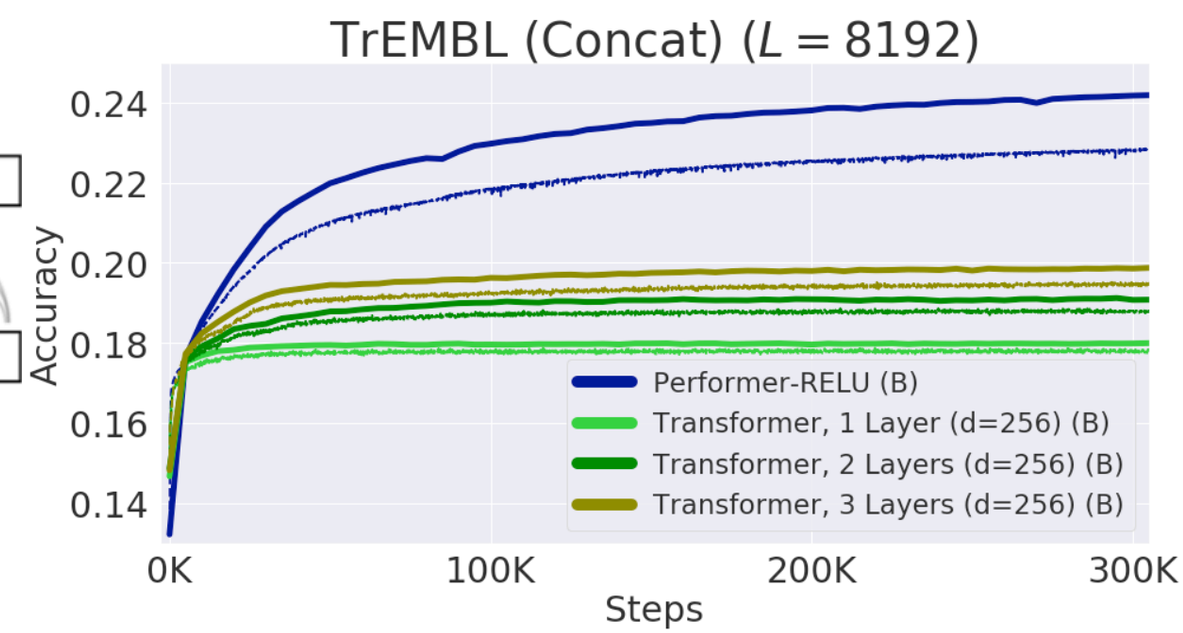
Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

Scalable models

The computational scaling of models on large sets/ sequences is an active topic.



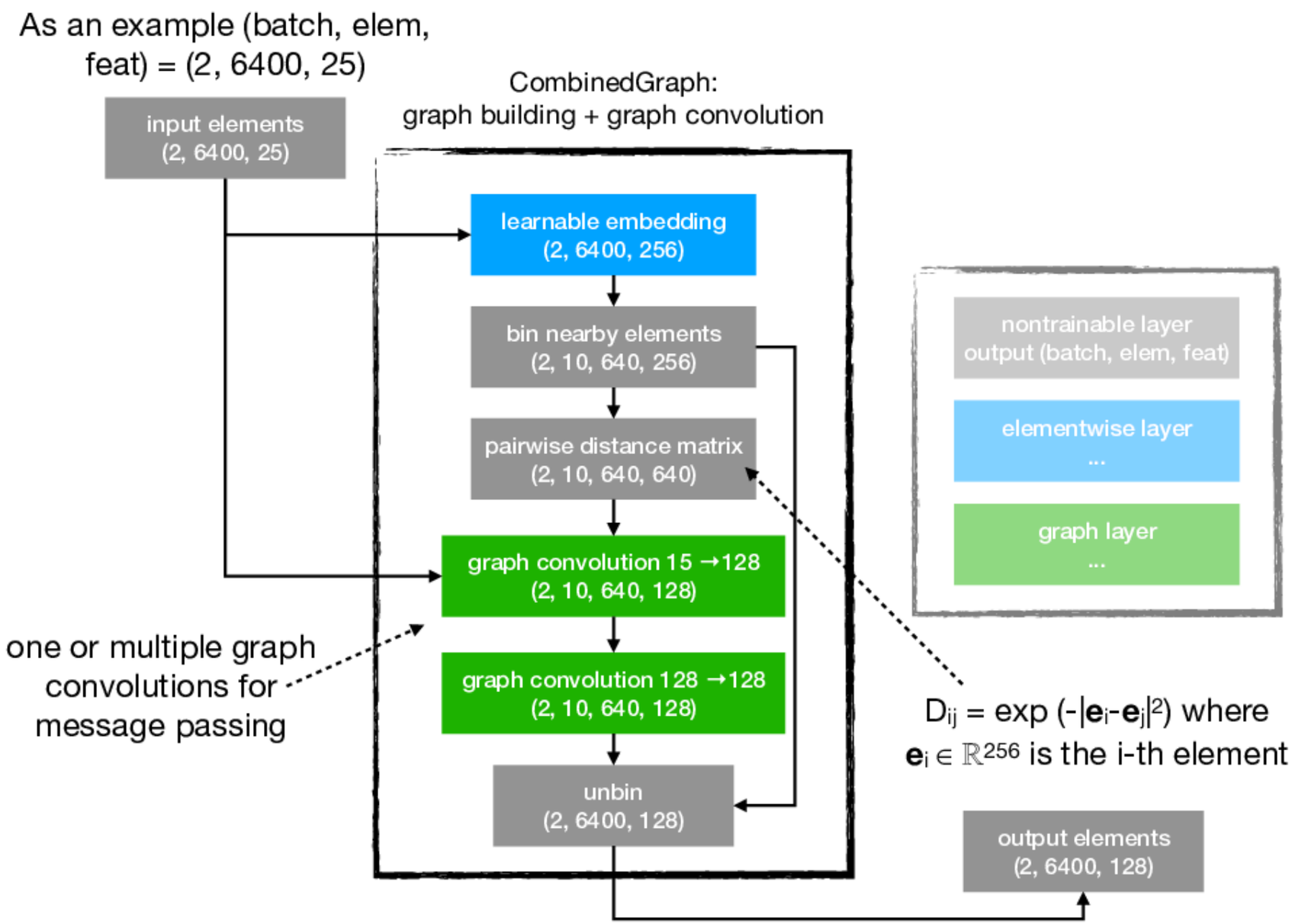




https://arxiv.org/pdf/2001.04451.pdf

https://arxiv.org/abs/2001.04451

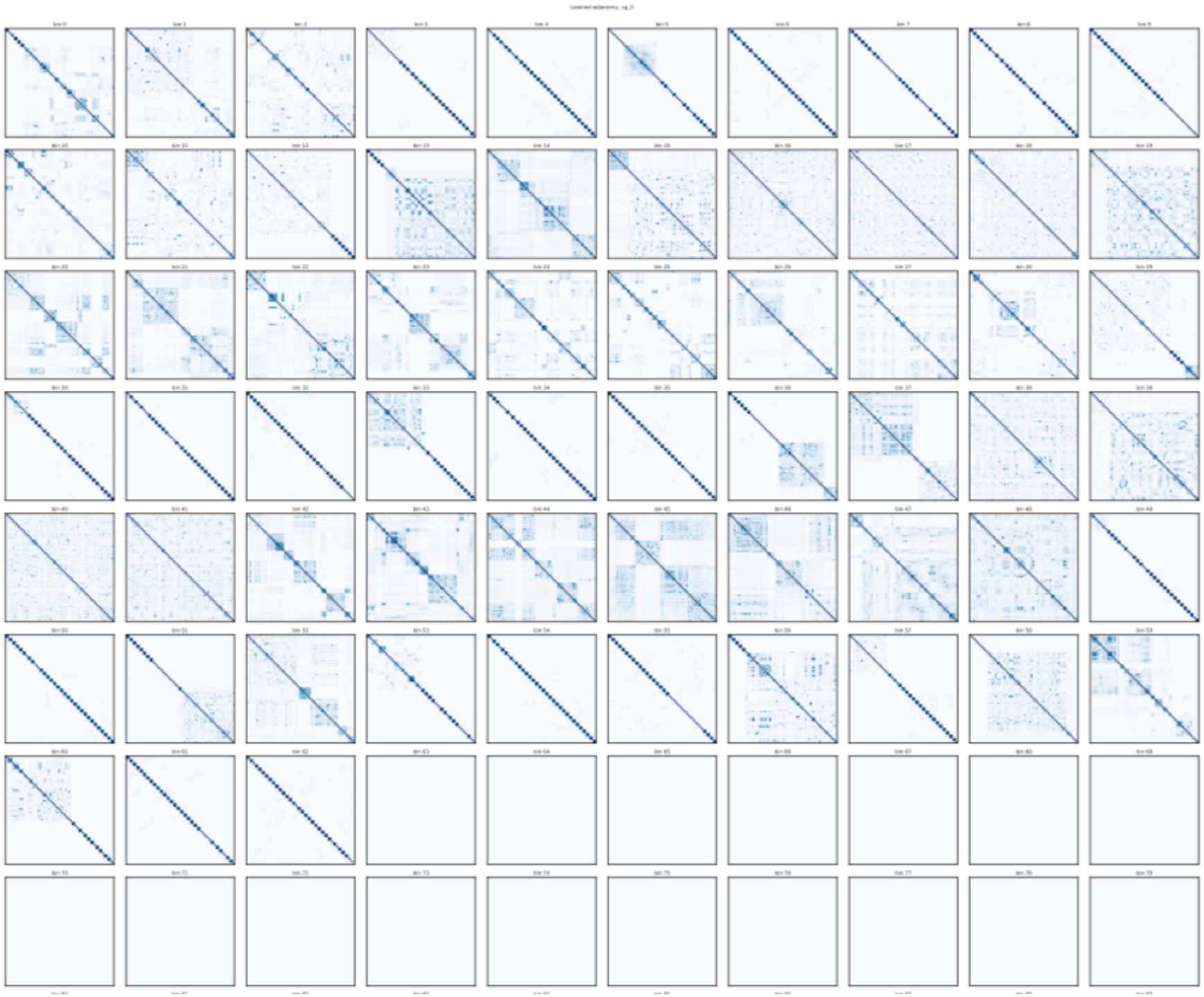
Implementation for sets



Uses built-in dense matrix, reshape and scatter/gather operations in TF.
Requires batch-mode graphs. No N^2 allocation or computation needed.

Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

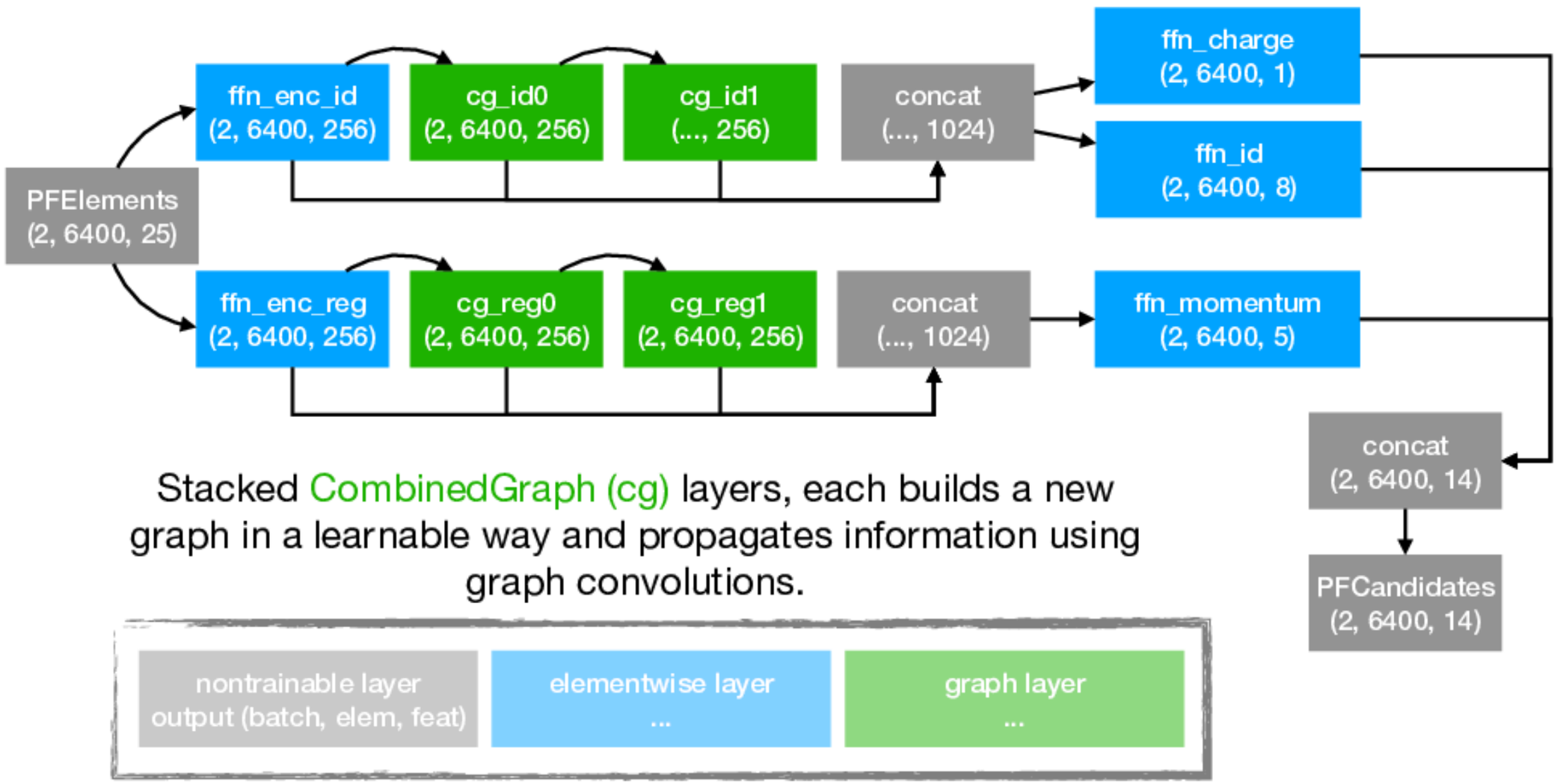
Disjoint event graphs



Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

Stacked models

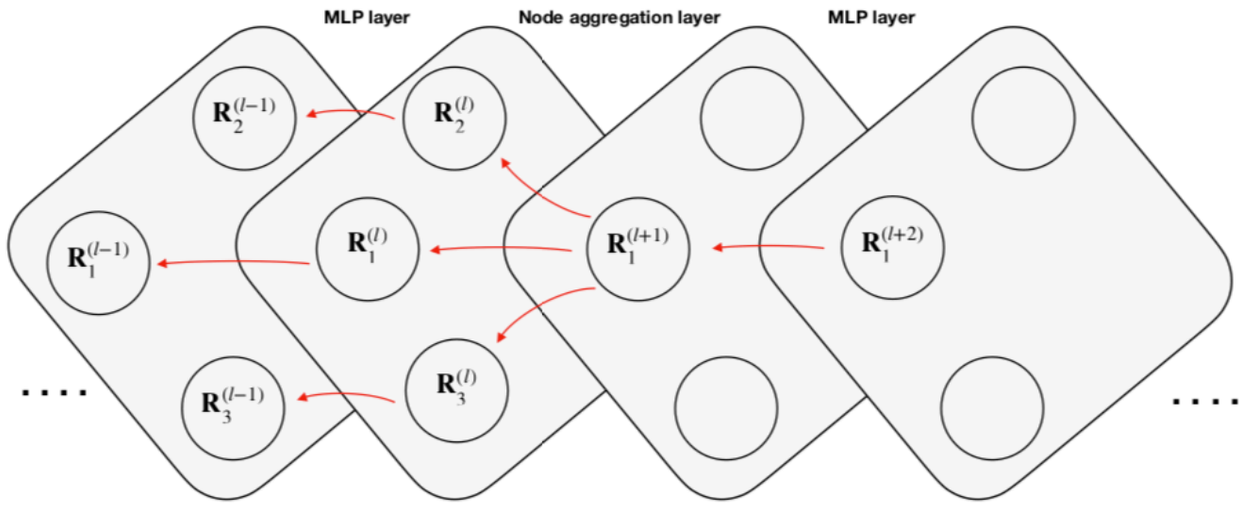
As an example (batch, elem, feat) = (2, 6400, 25)



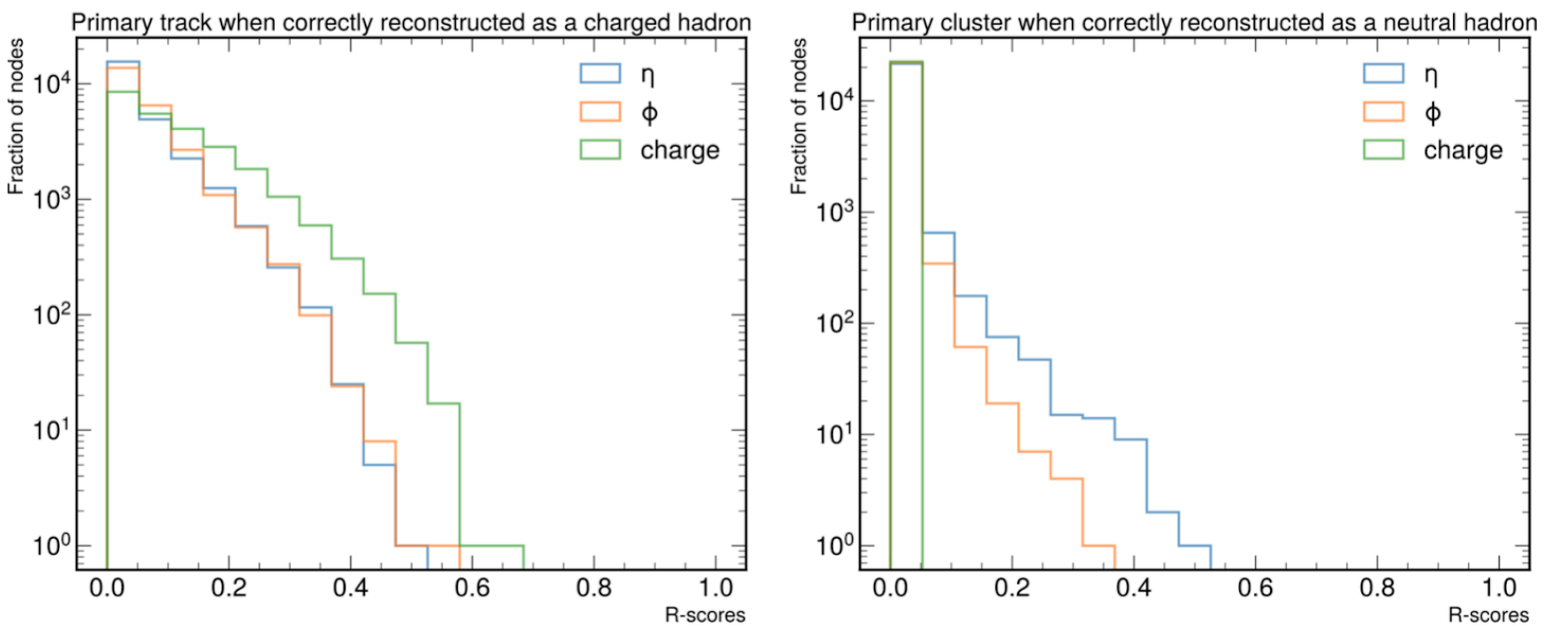
Pata, J. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330

Interpretability

- What inputs are relevant for a particular model output?
- Compute layerwise relevance scores **R**
- Aggregate along the graph structure



$$\mathbf{R}_j^{(l)} = \sum_k \frac{x_j A_{jk}}{\sum_m x_m A_{mk}} \mathbf{R}_k^{(l+1)}$$



"Explaining machine-learned particle-flow reconstruction"; Farouk Mokhtar, Raghav Kansal, Daniel C Diaz Javier Duarte, **JP**, Maurizio Pierini, Jean-Roch Vlimant. NeurIPS 2021, Machine Learning and the Physical Sciences, https://doi.org/10.48550/arXiv.2111.12840

Community datasets

Datasets for realistic particle reconstruction are involved to set up and can be expensive to generate, but crucial for development and comparisons.
An opportunity to build a community-based reconstruction data challenge?

As an example: a simplified Delphes dataset for PF

February 24, 2021 (v1.1) | Dataset | Open Access

View

Simulated particle-level events of ttbar and QCD with PU200 using Pythia8+Delphes3 for machine learned particle flow (MLPF)

Pata, Joosep; Duarte, Javier Mauricio; Vlimant, Jean-Roch; Pierini, Maurizio; Spiropulu, Maria;

Dataset of 50,000 top quark-antiquark (ttbar) and 5,000 QCD events produced in proton-proton collisions at 14 TeV, overlaid with minimum bias events corresponding to a pileup of 200 on average. The dataset consists of detector hits as the input, generator particles as the ground truth and reconstruc

Uploaded on February 24, 2021

1 more version(s) exist for this record

https://zenodo.org/record/4559324#.YImnSy0RoWo

Discussion

- **Realistic, open benchmark datasets and baseline algorithms!**

- **What are the goals of reconstruction?** Unique clustering/ segmentation, particle-level physics reconstruction, or event-level (jets, MET) physics reconstruction?

- **Useful loss functions for optimizing reconstruction.** To what extent (and based on what criteria?) do the hyperparameters in the loss function need to be tuned to the task? Can we rely on the model to learn what is not reconstructable?

- **To what extent can ML for simulation and ML for reco approaches inform each other?** Are they the inverse of each other? Can one construct an invertible model (e.g. normalizing flow) that does both?

- So far, synthetic data (simulation) is driving the efforts. What is the role of data-driven approaches, e.g. **learning representations from data**, fine-tuning on specific tasks?

References

- Belayneh, D., Carminati, F., Farbin, A. *et al.* Calorimetry with deep learning: particle simulation and reconstruction for collider physics. *Eur. Phys. J. C* **80,** 688 (2020). https://doi.org/10.1140/epjc/s10052-020-8251-9

- Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler; Phys. Rev. Lett. 123, 041801; https://doi.org/10.1103/PhysRevLett.123.041801

- Kieseler, J. Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. Eur. Phys. J. C 80, 886 (2020). https://doi.org/10.1140/epjc/s10052-020-08461-2

- Saptaparna Bhattacharya, Nadezda Chernyavskaya, Saranya Ghosh, Lindsey Gray, Jan Kieseler et al. GNN-based end-to-end reconstruction in the CMS Phase 2 High-Granularity Calorimeter. ACAT 2021. https://doi.org/10.48550/arXiv.2203.01189

- Shah Rukh Qasim, Nadezda Chernyavskaya, Jan Kieseler, Kenneth Long, Oleksandr Viazlo et al. End-to-end multi-particle reconstruction in high occupancy imaging calorimeters with graph neural networks. https://doi.org/10.48550/arXiv.2204.01681

- Di Bello, F.A., Ganguly, S., Gross, E. et al. Towards a computer vision particle flow. Eur. Phys. J. C 81, 107 (2021). https://doi.org/10.1140/epjc/s10052-021-08897-0

- JP., Duarte, J., Vlimant, JR. et al. MLPF: efficient machine-learned particle-flow reconstruction using graph neural networks. Eur. Phys. J. C 81, 381 (2021). https://doi.org/10.1140/epjc/s10052-021-09158-w

- JP. et al. Machine Learning for Particle Flow Reconstruction at CMS. ACAT 2021. https://doi.org/10.48550/arXiv.2203.00330, http://cds.cern.ch/record/2792320/files/DP2021_030.pdf?version=1