Towards Particle Flow Event Reconstruction at the Future Circular Collider with GNNs

Dolores Garcia
CERN
Switzerland
dolores.garcia@cern.ch

Gregor Kržmanc
University of Ljubljana
Slovenia
gregor.krzmanc@cern.ch

Philipp Zehetner
University of Munich
CERN
philipp.zehetner@cern.ch

Jan Kieseler
Karlsruhe Institute of Technology
ETP, Germany
jan.kieseler@kit.edu

Michele Selvaggi
CERN
Switzerland
michele.selvaggi@cern.ch

Abstract

Reconstructing particle properties from raw signals measured in particle physics detectors is a challenging task due to the complex shapes of the showers, variety in density and sparsity. Classical particle reconstruction algorithms in current detectors use a multi-step pipeline, but the increase in data complexity of future detectors will reduce their performance. We consider a geometric graph representation due to the sparsity and difference in density of particle showers. We introduce a dataset for particle level reconstruction at the Future Circular Collider and benchmark the performance of state-of-the-art GNN architectures on this dataset. We show that our pipeline performs with high efficiency and response and discuss how this type of data can further drive the development of novel geometric GNN approaches.

1 Introduction

The Future Circular Collider (FCC) (2) will push the energy and intensity frontiers in the search for new physics. For each collision, the final state consists of a number of stable particles which interact with the detector sensors in their path producing detectable hits. The pattern left by each particle is unique and has general properties depending on its momentum and its species. Figure 1 shows the produced hits in an event. In order to analyse the event generated by a given particle collision and have access to higher information data, the final state particles properties - relativistic 4-momentum, production vertex and type - need to be reconstructed from the observed hits. All physics studies then, take the produced list of output particles to perform a higher level processing. Therefore, having highly accurate reconstruction algorithms that properly map the detector outputs to a list of particles is of high relevance to maximise the the physics potential of the new generation of colliders.

Current reconstruction algorithms often start by identifying a set of seeds, and then proceed to add the remaining hits using hypothesis and clustering techniques. Although proficient, these algorithms have begun to show their limitations, in particular as the scale and complexity of data grow exponentially. In this context the lessons and advancements offered from point cloud object detection and GNNs offer fresh perspectives. While there is a foundational similarity in the data to point clouds, the nuances of particle physics call for more specialized tools. For example, the measured hits include energy and the information about the type of detector subsystem sensor as well as the coordinates. Pioneering developments, such as GravNet (17) and Object Condensation (13), heed this call. GravNet offers an innovative approach to node clustering and particle classification, by learning a rewiring of the graph during training based on a learnt embedding space, and considering a potential function on this space

NeurIPS 2023 New Frontiers in Graph Learning Workshop (NeurIPS GLFrontiers 2023).

as weights for message passing. Furthermore, it scales to complex events with up to 300 000 hits with low resource requirements on a single GPU (16). Concurrently, Object Condensation proposes an object reconstruction pipeline that is independent of object size or density by making use of attractive and repulsive potentials. These approaches largely differ from classical GNN approaches due to the nature of the task of particle reconstruction, and represent a bridge between approaches for geometric graphs and point clouds.

However, the tools and the sequence needed to simulate collision (matrix element generation (5), parton shower and hadronisation (20), detector simulation (11; 4), digitisation and post-processing (10)) events and making them AI ready need enhanced domain understanding, increasing the barrier for the AI community to meaningfully contribute to these problems and limiting the application of GNNs to scientific research.

This paper presents a dataset created by simulating events on future colliders using the CLD detector concept (6), one of the proposed detectors for the FCC-ee. Additionally, we present the SOTA end-to-end methodology for reconstructing PF candidates. Further, we discuss how this task may lead to new applications of GNNs in large scale geometric graphs.

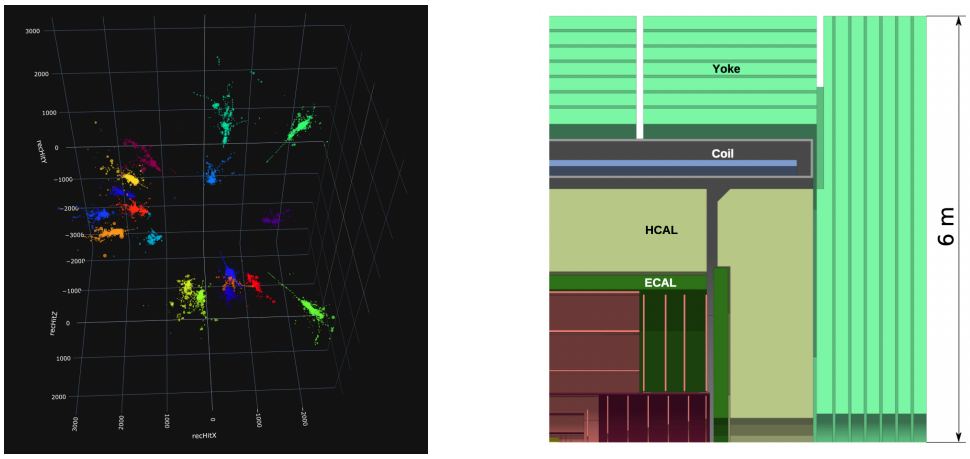


Figure 1: Left: example event from the training set. The plot contains the hits in detector coordinates, the size represents the energy deposit in the calorimeter and the color of each shower the particle id. Right: vertical cross-section of the CLD detector showing the different types of sensor (HCAL, ECAL) as in (6).

2 Related work

GNNs has been extensively applied to solve computationally challenging experimental particle physics-related problems such as jet tagging (18; 8; 1), track reconstruction (14), and data generation (12). Approaches such as machine-learned particle-flow reconstruction (MLPF) (13; 15) have been used to combine calorimeter clusters and tracks to generate a set of particles. Despite the effectiveness of the approach when applied to preprocessed calorimeter clusters as input, the approach in (15) cannot be directly applied to more granular shower shapes produced by the higher-granularity calorimeters.

For the use in the CMS High-Granularity Calorimeter (HGCAL), object condensation was proposed for particle reconstruction (13; 16; 9). Drawing from a physics-based learning objective, it has been successfully applied to generic problems and demonstrates potential for application to more diverse graph segmentation or similar tasks (13).

3 Dataset

The dataset consists of particle collision events with an arbitrary number (between 10 and 15) of high-energy ($0.5 < E < 50$ GeV) stable final state hadrons (protons, neutrons, K_L, $\pi^{\pm}$). These hadrons interact with the CLD detector calorimeter (7), giving rise to detectable hits. When a hadron

enters the calorimeter, it initiates a hadronic shower through nuclear interactions with the absorber material. The secondary particles produced in the shower are detected by the active elements, and their energies are measured to determine the energy of the incoming hadron. Events are simulated using the KEY4HEP turnkey software stack (22) and GEANT4 detector simulation (4).

Nodes In each event, there are $\mathcal{O}(600)$ hits per particle (Figure 3). For every detector hit, the energy (e), position (x, y, z), and type in the detector are stored. The hit type is stored in the form of a boolean (is) and tracks whether the corresponding detector element is part of the ECAL (electromagnetic calorimeter) or the HCAL (hadronic calorimeter) subdetector, respectively. As the ground truth information, for every hit, we keep track of the corresponding truth particle(s) responsible for generating the energy deposit. This information is obtained by using the history of the particle that generated the hit to link it back to the stable particle that produced it. The 4-momenta (p_x, p_y, p_z, e) and particle type (neutron or proton) (id) or the corresponding ground-truth particles is stored. The hadronic calorimeter hits store, on average, more energy than the electromagnetic hits, highlighting the heterogeneousness of the systems, as shown in Figure 2.

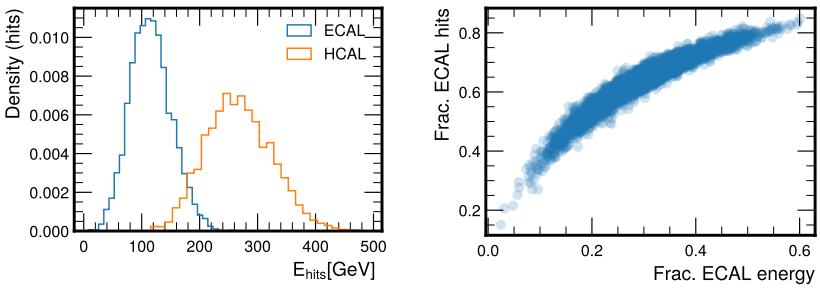


Figure 2: Distributions of the per event energies deposited in the ECAL and HCAL sub-detectors (left). Fraction of the number of ECAL hits vs. the fraction of the energy stored in the ECAL hits per event (right).

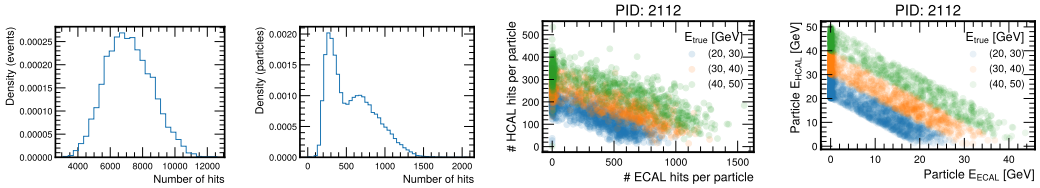


Figure 3: From left to right: Distribution of the numbers of hits (nodes) per event, number of hits (nodes) per particle, numbers of HCAL and ECAL hits for neutrons for different energy ranges, energies left by the HCAL and ECAL hits for neutrons for different energy ranges.

Graph The set of hits is then converted into a graph. The hits are the nodes in the graph and the edges are obtained by running a custom scalable KNN on the detector coordinates of the hits. The graph is dynamically updated by the network as discussed in the next section.

4 Particle reconstruction with GNNs

Task The learning task consists of reconstructing an unknown number of particles using the hits in a given event. The task is supervised and each hit has a link to one particle. This task is similar to a point cloud object instantiation. Although showers (collection of hits generated by one particle) could be well separated in space, *particle flow* algorithms (19; 21) aim to improve the separability of nearby showers, to minimise the error on the measured momenta and the mis-indentification probability of the ancestor particle species by optimally combining all the available sub-detector information.

Baseline We present the SOTA framework for this problem: a GravNet architecture wrapped in Object condensation framework (13; 16), which defines the the types of outputs and the loss function. Object condensation (13; 16) has been recently proposed to reconstruct multiple objects and their properties from a variable number of nodes (hits). The approach trains a model first to predict coordinates in an abstract space termed *clustering space*.[1] In the clustering space, 'central' hits of each cluster termed *condensation points* are additionally predicted, and the model is trained such that there is one condensation point (hit) per particle and the hits belonging to that particle are close to it. In addition to the clustering space coordinates, a scalar value is outputted by the model that constitutes a measure of how likely the given point is to be a condensation point (*beta value*; β_i). The β_i of the condensation point (α) is enforced to be close to 1 while being close to 0 for the other points. Additionally, the β values additionally provide for weights q_i of the nodes $q_i = \operatorname{arctanh}^2 \beta_i + q_{min}$. Such structure of clustering space is achieved by jointly optimizing multiple losses, (1), (2), and (3). We use $q_{min} = 3$ unless otherwise specified. The set of condensation points is termed α and the set of hits that are not alphas is termed $\cancel{\alpha}$. The losses used to optimize the model are defined in Equations (1), (2) and (3). The first term in (1) is added to keep the beta values of the hits belonging to $\cancel{\alpha}$ close to zero, (2) and (3) represent the repulsive and attractive potential to the condensation points. The losses are standardized per truth particle: α_j belongs to event j and hits belonging to event j that are not alphas are termed $\cancel{\alpha_j}$.

$$\mathcal{L}_\beta = \frac{1}{N_{part.}} \left(\sum_{\alpha_j \in \alpha} \frac{1}{N_{hits,j}} (\beta_{\alpha_j} + \sum_{i \in \cancel{\alpha_j}} (1 - \beta_i)) \right) \tag{1}$$

$$\mathcal{L}_{V,\text{repulsive}} = \frac{1}{N_{part.}} \sum_{\alpha_j \in \alpha} \frac{1}{N_{hits,j}} \sum_{i \notin \cancel{\alpha_j}} \left(q_i q_{\alpha_j} \exp(-4 \left| x_i - x_{\alpha_j} \right|) \right) \tag{2}$$

$$\mathcal{L}_{V,\text{attractive}} = \frac{1}{N_{part.}} \sum_{\alpha_j \in \alpha} \frac{1}{N_{hits,j}} \sum_{i \in \cancel{\alpha_j}} \left(q_i q_{\alpha_j} \left| x_i - x_{\alpha_j} \right|^2 \right) \tag{3}$$

GravNet (17) is a message passing MP graph neural network that dynamically updates the graph structure at each layer using a KNN over a set of coordinates obtained from the features. The message passing functions use the distances in the embedding space as a weight to aggregate the features of neighbouring nodes. We consider the architecture presented in (17) with 4 layers and $k = [16, 64, 16, 64]$ as the number of neighbours for each layer.

Figure 4 shows the outputs for a single event at different steps of training, where the coordinates are the outputted coordinates of the model in the embedding space, the size of the points represents β and each colour represents a true particle id. It can be observed in this figure that as training advances, the model pulls the hits belonging to a single particle closer together and in some cases, like the red and green shower it can separate the showers but leaves some residual hits that are pulled by both condensation points.

For the inference, the Object Condensation framework selects a threshold for the β outputs per node to be considered a condensation point, and proceeds to assign to each condensation points all hits with predicted coordinates inside a radius r_c.

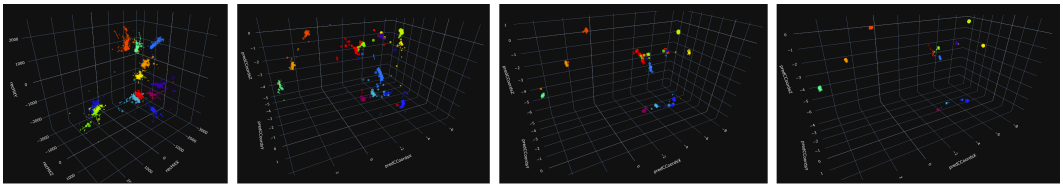
In order to thoroughly evaluate results, each reconstructed shower needs to be matched to a truth shower or discarded as a fake. We quantify the similarity between two showers as an energy-weighted intersection over union (EIOU) as proposed in (16). The predicted showers are matched to the truth showers by maximizing the sum of EIOU as in (16).

The architectures and models described in this section are implemented in TensorFlow (3).

5 Results

The performance of the reconstruction algorithm is evaluated on 5000 events that were generated with the same number of particles and energy range as the training set. In order to measure the

[1] Additionally, other features such as particle energy and momentum can be predicted.



Ground truth Epoch 1 Epoch 3 Epoch 12

Figure 4: Clustering space evolution with training for a sample event shown on the left. The size of the points in clustering space represents the β value predicted by the model, whereas the size of the ground truth hits is proportional to their energy.

performance of our algorithms we define multiple metrics that are used in the High Energy Physics community to evaluate the performance of a particle flow algorithm:

Efficiency: the fraction of particles that are reconstructed and matched to one of the true showers with respect to the total number of true showers.
Fake rate: the fraction of particles that are reconstructed without matching to a true shower with respect to the total number of true showers.
Energy response quantities: the mean, termed response, and the mean-corrected standard deviation, termed resolution, of the distribution of reconstructed predicted energy divided by the energy of the particle for matched particles. This metric can be considered for both the true and reconstructed energy of the particle. The response would peak at 1 in a perfect reconstruction for the reconstructed energy. The reconstructed energy is the energy that can be reconstructed from the hits of a particle. It can be smaller than the true energy.
Per shower metrics: for each matched shower the containment is the percentage of energy the predicted shower contains of the matched true particle, calculated as $\frac{\sum_n M_{n,i} E_n}{E_{\text{reco},i}}$, where $M_{n,i}$ is a mask that is 1 if hit n belongs to particle i and zero otherwise. The purity of the shower is the percentage energy from the true particle that belongs to the reconstructed particle, which is calculated as $\frac{\sum_n M_{n,i} E_n}{E_{\text{pred},i}}$.

The results are presented in Figure 5. The efficiency calculated for each energy bin is shown in Figure 5a. The efficiency increases with the increase in p_T and reaches values close to 1 for $p_T > 30$. We hope that by adding track information to the dataset the efficiency at lower p_T increases, since only one track would correspond to each shower (for charged particles). The fake rate as a function of the reconstructed particles energies is presented in Figure 5b. There is a large number of fake particles produced around 1 GeV and the percentage of fake particles sharply decreases for higher p_T. Fake particles appear as a result of learning a bad distribution for the β, as it assigns a high β to hits that already have a condensation point. We aim to reduce this effect by improving the β term in the loss function to more accurately reflect the optimal distribution. For the matched true showers we evaluate the response and resolution as function of the reco p_T and show the results for each energy bin in Figures 5c and 5d. The response shows that the clustering correctly assigns a high fraction of energy to the matched reconstructed particles and the resolution shows that the distribution is peaked. In both, we observe that the performance is worse at lower energies. As a next step, the model should also regress the true energy of the particle, which as explained in the metrics is different to the reco energy (that can be observed in the event). This could be achieved by training a sub-block of the model that considers each sub-graph resulting from the condensation and regresses particle level features, such as mass, energy and coordinates. Finally, containment and purity metrics as a function of the reconstructed energy for matched showers are shown in Figures 5e and 5f. The error vars represent the standard deviation of the metrics for the showers inside the energy bin. These are excellent metrics to evaluate if the model is mixing hits from different showers. Both containing and purity are high and there is a increasing trend in both for increasing p_T. We can also observe this in the clustering space in Figure 4, where most of the clusters are well separated but some remain close to each other and have particles that are pulled by multiple condensation points.

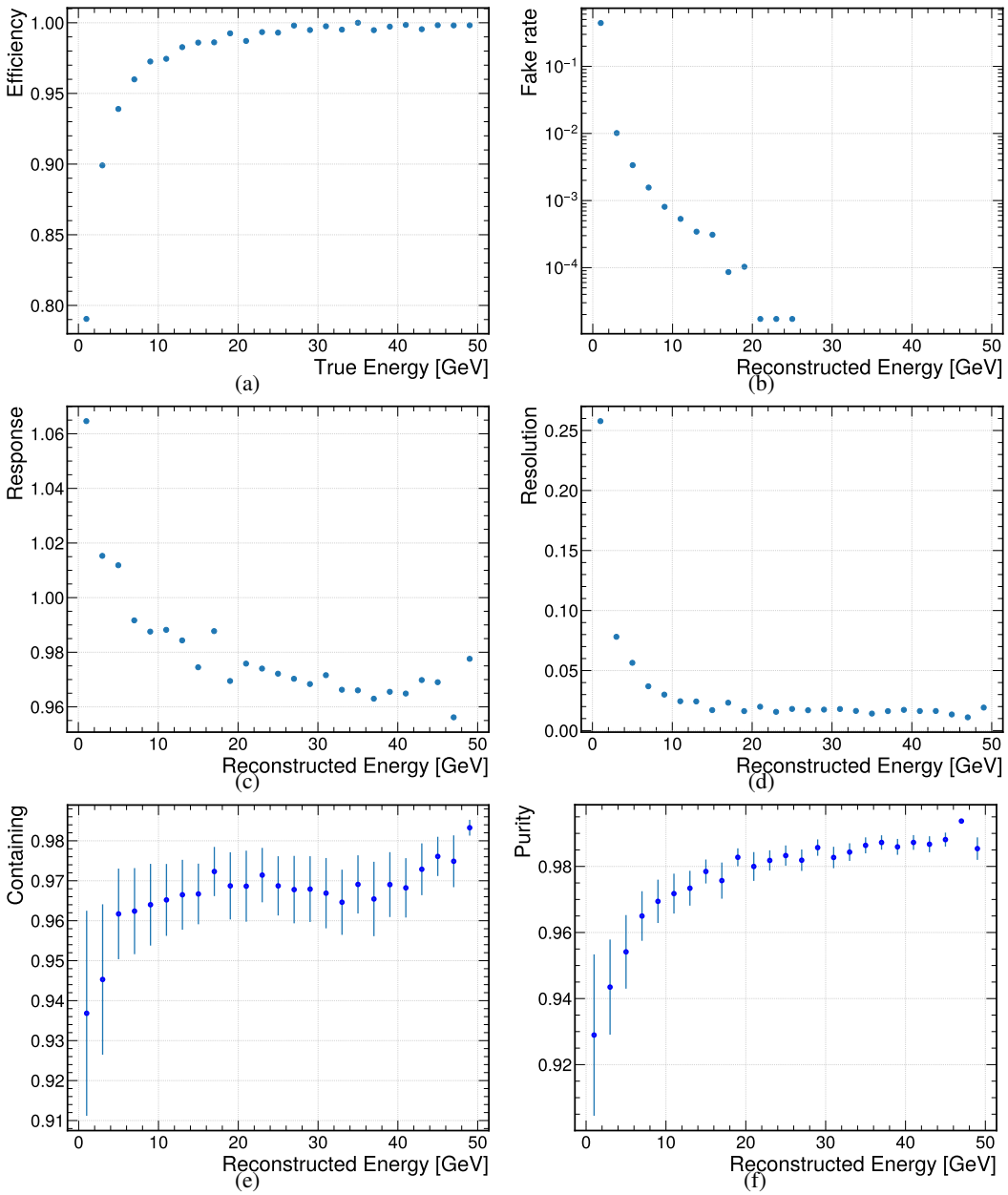


Figure 5: Reconstruction performance on evaluation dataset. (a) Efficiency as function of true energy. (b) Fake rates as a function of the reco particle's energy.(c) Response as a function of reco energy of matched showers. (d) Resolution as a function of reco energy of matched showers.

6 Conclusion

We introduce a new dataset with a high energy physics driven task, particle level reconstruction at future colliders. We show how SOTA GNN architectures perform and present metrics relevant for the community. However, in order to bring GNNs to the performance required for the FCC-ee studies, several items are missing. The clustering performance has to be improved at low energies $E < 2$ GeV where particle multiplicities are the largest. In addition tracking information has to be included, and tracks have to be part of the input for the clustering. Finally the identification of the particle species, in conjunction with all the various sub-detector components (not only the calorimeters) has to be demonstrated. We hope that this complex problem is of interest to the GNN

community. With this public dataset, we open the door to new challenges and research directions that could benefit both the high energy physics and the ML communities.

Acknowledgments and Disclosure of Funding

We would like to thank CERN OpenLab for the support and resources over the course of this work, in particular, Maria Girone and Eric Wulff. We would also like to thank Joosep Pata for the helpful discussions at the start of this work. We thank Michelangelo Mangano and the Future Collider Unit at CERN for funding Gregor's stay at CERN during the preparation of this paper. We acknowledge EuroHPC Joint Undertaking for awarding us access to Karolina at IT4Innovations, Czech Republic.

References

[1] Graph Neural Network Jet Flavour Tagging with the ATLAS Detector, 2022. Place: Geneva. URL: https://cds.cern.ch/record/2811135.

[2] AEA Abada, M Abbrescia, SS AbdusSalam, I Abdyukhanov, J Abelleira Fernandez, A Abramov, M Aburaia, AO Acar, PR Adzic, Prateek Agrawal, et al. Fcc-ee: the lepton collider: future circular collider conceptual design report volume 2. *The European Physical Journal Special Topics*, 228:261–623, 2019.

[3] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.

[4] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.

[5] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. arXiv:1405.0301, doi:10.1007/JHEP07(2014)079.

[6] N. Bacchetta, J.-J. Blaising, E. Brondolin, M. Dam, D. Dannheim, K. Elsener, D. Hynds, P. Janot, A. M. Kolano, E. Leogrande, L. Linssen, A. Nürnberg, E. F. Perez, M. Petrič, P. Roloff, A. Sailer, N. Siegrist, O. Viazlo, G. G. Voutsinas, and M. A. Weber. CLD – A Detector Concept for the FCC-ee, December 2019. arXiv:1911.12230 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/1911.12230.

[7] Nicola Bacchetta, J-J Blaising, E Brondolin, M Dam, D Dannheim, K Elsener, D Hynds, P Janot, AM Kolano, E Leogrande, et al. Cld–a detector concept for the fcc-ee. *arXiv preprint arXiv:1911.12230*, 2019.

[8] Franco Bedeschi, Loukas Gouskos, and Michele Selvaggi. Jet flavour tagging for future colliders with fast simulation. *Eur. Phys. J. C*, 82(7):646, 2022. arXiv:2202.03285, doi:10.1140/epjc/s10052-022-10609-1.

[9] S Bhattacharya, N Chernyavskaya, S Ghosh, L Gray, J Kieseler, T Klijnsma, K Long, R Nawaz, K Pedro, M Pierini, G Pradhan, S R Qasim, O Viazlo, P Zehetner, and on behalf of the CMS Collaboration. Gnn-based end-to-end reconstruction in the cms phase 2 high-granularity calorimeter. *Journal of Physics: Conference Series*, 2438(1):012090, feb 2023. URL: https://dx.doi.org/10.1088/1742-6596/2438/1/012090, doi:10.1088/1742-6596/2438/1/012090.

[10] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, 389:81–86, 1997. doi:10.1016/S0168-9002(97)00048-X.

[11] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014. arXiv:1307.6346, doi:10.1007/JHEP02(2014)057.

[12] Raghav Kansal, Javier Duarte, Hao Su, Breno Orzari, Thiago Tomei, Maurizio Pierini, Mary Touranakou, jean-roch vlimant, and Dimitrios Gunopulos. Particle Cloud Generation with Message Passing Generative Adversarial Networks. In *Advances in Neural Information Processing Systems*, volume 34, pages 23858–23871. Curran Associates, Inc., 2021. URL: https://papers.nips.cc/paper/2021/hash/c8512d142a2d849725f31a9a7a361ab9-Abstract.html.

[13] Jan Kieseler. Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph and image data. *The European Physical Journal C*, 80(9):886, September 2020. arXiv:2002.03605 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/2002.03605, doi:10.1140/epjc/s10052-020-08461-2.

[14] Daniel Thomas Murnane, Alexis Vallier, Charline Rougier, Paolo Calafiura, Jan Stark, Xiangyang Ju, Steven Andrew Farrell, Sylvain Caillou, Mark Neubauer, and Markus Julian Atkinson. Graph Neural Network Track Reconstruction for the ATLAS ITk Detector, 2022. URL: https://cds.cern.ch/record/2809518.

[15] Joosep Pata, Javier Duarte, Jean-Roch Vlimant, Maurizio Pierini, and Maria Spiropulu. MLPF: Efficient machine-learned particle-flow reconstruction using graph neural networks. *The European Physical Journal C*, 81(5):381, May 2021. arXiv:2101.08578 [hep-ex, physics:physics, stat]. URL: http://arxiv.org/abs/2101.08578, doi:10.1140/epjc/s10052-021-09158-w.

[16] Shah Rukh Qasim, Nadezda Chernyavskaya, Jan Kieseler, Kenneth Long, Oleksandr Viazlo, Maurizio Pierini, and Raheel Nawaz. End-to-end multi-particle reconstruction in high occupancy imaging calorimeters with graph neural networks. *The European Physical Journal C*, 82(8):753, August 2022. arXiv:2204.01681 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/2204.01681, doi:10.1140/epjc/s10052-022-10665-7.

[17] Shah Rukh Qasim, Jan Kieseler, Yutaro Iiyama, and Maurizio Pierini. Learning representations of irregular particle-detector geometry with distance-weighted graph networks. *The European Physical Journal C*, 79(7):1–11, 2019.

[18] Huilin Qu and Loukas Gouskos. ParticleNet: Jet Tagging via Particle Clouds. *Phys. Rev. D*, 101(5):056019, 2020. arXiv:1902.08570, doi:10.1103/PhysRevD.101.056019.

[19] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017. arXiv:1706.04965, doi:10.1088/1748-0221/12/10/P10003.

[20] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. arXiv:1410.3012, doi:10.1016/j.cpc.2015.01.024.

[21] M. A. Thomson. Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth. A*, 611:25–40, 2009. arXiv:0907.3577, doi:10.1016/j.nima.2009.09.009.

[22] Valentin Völkl, Placido Declara Fernandez, Gerardo Ganis, Benedikt Hegner, Clement Helsens, Andre Sailer, Erica Brondolin, Graeme A Stewart, Juraj Smiesko, Frank Gaede, Thomas Madlener, Wenxing Fang, Tao Lin, Xiaomei Zhang, Jiaheng Zou, Xingtao Huang, Teng Li, Sang Hyun Ko, Joseph Wang, Wouter Deconinck, and Sylvester Joosten. The Key4hep turnkey software stack. *PoS*, ICHEP2022:234, 2022. URL: https://cds.cern.ch/record/2869547, doi:10.22323/1.414.0234.