

Machine learning-based Particle Flow at the FCC-ee

Dolores Garcia¹, Gregor Kržmanc², Michele Selvaggi¹

¹CERN, Geneva, Switzerland.

²ETH Zürich, Zürich, Switzerland.

Contributing authors: dolores.garcia@cern.ch; gregor.krzmanec@cern.ch;
michele.selvaggi@cern.ch;

Accurate and adaptive particle-flow reconstruction is a key requirement to maximise the physics potential and to facilitate the design of the FCC-ee detectors. Classical algorithms for particle-flow reconstruction require a detailed understanding of the detector and manual tuning, these requirements could be circumvented by using machine learning. This work describes an end-to-end particle-flow algorithm based on graph neural networks. We demonstrate the effectiveness of the approach using different datasets: 10-15 collimated particles resembling a jet and $H \rightarrow s\bar{s}$ physics events. The jet-like dataset allows to train and test the algorithm on a controlled and balanced sample. We discuss how such an approach can be further applied to improve physics performance at the FCC-ee.

1 Introduction

The FCC-ee will push the intensity frontier by providing unprecedented instantaneous luminosity at the Z , WW , ZH and $t\bar{t}$ thresholds [1]. Designing optimal detectors and implementing reliable event reconstruction algorithms that maximise the physics potential of the FCC is crucial for the success of the program. The Higgs boson decays preferentially to hadrons, therefore the optimal identification and reconstruction of hadronic jets is crucial for the measurement of Higgs boson properties. In particular, requiring that the mass peaks from $W, Z \rightarrow jj$ hadronic decays should be resolvable at the level of 3σ , imposes that the W and Z invariant masses should be reconstructed with a resolution σ_m at the level of their natural width $\Gamma/m \approx 2.5\%$. Moreover the measurement of rare hadronic Higgs decay rates, such as $H \rightarrow c\bar{c}, s\bar{s}$, requires excellent hadronic mass resolution since the relative uncertainty on these rare modes scales as $\sqrt{\sigma_m}$.

The Particle-Flow (PF) algorithm was first introduced in the context of e^+e^- collisions at LEP by the ALEPH experiment [2], with the goal of improving the angular and energy resolution of hadronic final states. Other variants have later been introduced at DELPHI [3], H1 and ZEUS [4]. The CMS experiment has pioneered its use in hadronic collisions at the LHC [5, 6], and ATLAS has recently implemented its own version [7]. The PF approach has a proven record of effective improvement of the hadronic jet energy resolution in e^+e^- , e^-p , and pp collisions. Therefore, event reconstruction in future collider experiments such as FCC-ee [1], CEPC [8], ILC [9] and CLIC is based on PF. In addition, the implementation of an adaptive particle-flow and fully functional reconstruction framework is key requirement to drive the design and optimisation of future experiments.

The conventional approach to the measurement of hadronic signatures is to rely purely on calorimetric measurements. The calorimeter resolution can be written, neglecting the noise and constant terms as:

$$\frac{\sigma_{E(H)}(E)}{E} = \frac{S_{E(H)}}{\sqrt{E}}. \quad (1)$$

Typical values for the stochastic term S in electromagnetic and hadronic calorimeters are given in Table 1. The visible energy of hadronic jets or events is roughly distributed as $f_{\pm} = 60\%$ of charged hadrons ($\pi^{\pm}, K^{\pm}, \dots$), $f_{\gamma} = 30\%$ of photons from π^0 decays, and $f_n = 10\%$ neutral hadrons ($n, K_L, \dots$). The ideal visible energy resolution obtained by a pure calorimetric measurement can then be expressed as:

$$\sigma_{\text{CAL}}^2(E) = \sum_{i \in \gamma} \sigma_E^2(E_i) + \sum_{i \in \text{had}} \sigma_H^2(E_i) \quad (2)$$

approximately leading to:

$$\frac{\sigma_{\text{CAL}}(E)}{E} = \frac{S_{\text{CAL}}^{\text{eff}}}{\sqrt{E}} \approx \frac{S_H \sqrt{f_{\pm} + f_n}}{\sqrt{E}} \quad (3)$$

In contrast, the PF approach relies on the tracking system to measure the momenta of charged particles, while photons and neutral hadrons are reconstructed with the calorimeters:

$$\sigma_{\text{PF}}^2(E) = \sum_{i \in \text{tracks}} \sigma_{\text{track}}^2(E_i) + \sum_{i \in \gamma} \sigma_{\text{E}}^2(E_i) + \sum_{i \in \text{had}} \sigma_{\text{H}}^2(E_i) \quad (4)$$

The tracking resolution can be written as $\sigma_{\text{track}}(E) \sim (10^{-3} - 10^{-4})E^2$, which, despite the quadratic scaling with energy, can be safely neglected for particles with $E < 50$ GeV. This leads, in analogy with Eq. 3, to:

$$\frac{\sigma_{\text{PF}}(E)}{E} = \frac{S_{\text{PF}}^{\text{eff}}}{\sqrt{E}} \approx \frac{S_{\text{E}}\sqrt{f_{\gamma}} \oplus S_{\text{H}}\sqrt{f_n}}{\sqrt{E}} \quad (5)$$

The idealised effective stochastic terms $S_{\text{CAL}}^{\text{eff}}$ and $S_{\text{PF}}^{\text{eff}}$ are summarised in Table 1 and clearly show the dramatic increase in performance in the PF approach. Due to the complex nature of hadronic showers, both pure ideal calorimetric and ideal PF measurements of the visible energy are dominated by the hadronic resolution of the detector. From Equation 5 it is manifest that in order to achieve the highest visible energy and mass resolution it is crucial to efficiently identify and measure the neutral hadrons with the highest possible resolution. It should be noted that this derivation assumes that all the visible energy components, charged, photons and neutral hadrons are reconstructed efficiently. This also assumes no tracking nor calorimeter inefficiencies and no mistake in the track to calorimeter cluster association.

stochastic term	CMS	ATLAS/CLD	IDEA
nominal S_{E}	5%	10%	10%
nominal S_{H}	100%	50%	30%
ideal $S_{\text{CAL}}^{\text{eff}}$	83%	42%	25%
ideal $S_{\text{PF}}^{\text{eff}}$	32%	17%	11%

Table 1 Comparison of the stochastic term for different detector technologies and the effective stochastic term obtained with pure calorimetric and particle-flow reconstruction.

Therefore the key ingredients for an asymptotically optimal PF algorithm performance can be summarised as:

- Close to 100% tracking efficiency allowing for an optimal measurement of the charged visible energy component
- Fine transverse and longitudinal calorimeter segmentation allowing for an optimal geometrical matching of the extrapolated track trajectory to the electromagnetic and hadronic showers
- Efficient calorimeter clustering algorithm allowing to identify and reconstruct individual photon and hadron showers.
- Excellent electromagnetic and hadronic calorimeter resolutions allowing for an optimal comparison of the tracking versus calorimeter energy measurements after the geometrical matching

Hadronic signatures represent a sizable fraction of the final states at colliders, but are not exhaustive. Electrons and muons identification and momentum determination, whether isolated or non-isolated within a jet also benefit from a collective treatment of all the detector components. For example, the recovery of electron bremsstrahlung emissions within the tracker requires a procedure that combines tracking and ECAL. Similarly muons are identified by a combination of tracking, calorimetry and muon detectors information. A complete PF reconstruction should encompass a wide range of tasks, with the goal of providing the exhaustive list of stable final particles resulting from the collisions and interacting with the detector as well as a measurement of their associated momenta. In this paper we describe an end-to-end PF algorithm implementation using machine learning techniques (MLPF). Combining the information of all the various detector sub-systems, such as tracker, electromagnetic and hadronic calorimeters and muon detectors, the MLPF algorithm returns a list of reconstructed particles along with a measurement of the momentum $\vec{p}$ at vertex and type (electron, muon, charged hadron, neutral hadron or photon). Section 2 summarizes existing classical and machine-learning based approaches to PF reconstruction. In Section 3 we describe our generated dataset and the construction used as input for the training of the algorithm. Section 4 describes the various steps and implementation of MLPF and in Section 5 the performance of MLPF is discussed and compared to the Pandora algorithm.

2 Related Work

2.1 Classic approaches to particle flow

Existing implementations of particle-flow algorithms all share an iterative clustering stage, to generate a starting set of clusters, and a post-processing to match this clusters to tracks and find neutral clusters. They differ in the rules and cuts applied to make the different assignments which are usually detector specific. These algorithms have allowed for excellent physics results. However, they are largely parameterized and sometimes include steps that are specific to a detector geometry. These problems make their development cycle long and reduce their adaptability to new detector concepts.

[10] generates a set of tracks and clusters, based on proximity in the detector cells. Then, it iteratively extrapolates the tracks to find clusters and produces candidates for electrons and muons, which are then removed from the clusters and tracks lists. This is achieved making use of metrics that take into account the shape of the showers in the ECAL and muon chamber respectively. Then photons and π^0 are identified using different rules on the energy, distance to tracks and shower shapes. Finally, the charged hadrons and neutral hadrons are constructed. Neutral hadrons are reconstructed indirectly by identifying an excess energy in the calorimeters. Later approaches [11, 12] differ from [10] in that they reconstruct all particles in the event using more sophisticated clustering techniques. The PandoraPFA algorithm [12] performs multiple clustering steps including: hit ordering, seeding, topological merging of clusters and reclustering for higher energies. [12] also proposes photon identification and recovery algorithms.

2.2 ML approaches to particle flow

Novel end-to-end approaches to particle reconstruction have been recently proposed in the context of high-luminosity LHC [13–15] and HGCal [16, 17].

[13–15] consider a set of clusters and tracks in the detector and use Graph Neural Networks (GNNs) to reconstruct a list of particle objects. This is done by regressing a particle per object (cluster or track) where some will have empty labels as the number of particles is smaller than the number of clusters. For this, they use a regression loss to predict the momentum components and a classification loss to predict the PID. They evaluate the performance on a particle gun and physics events. On a single particle gun, the performance evaluation shows similar metrics for charged hadrons compared to PandoraPFA while lower efficiency for neutral hadrons and photons. On physics samples, they show improved jet p_T response width over PandoraPFA. They also consider starting from a set of raw calorimeter hits but in this case the performance is similar to PandoraPFA.

[16, 17] propose a different approach to particle reconstruction based on object condensation, described in more detail in section 4. They start from raw detector hits and also using a graph representation with a GNN architecture based on [18]. This approach gives a more natural representation of the reconstruction problem as it does not regress a particle per hit as in [13]. In the object condensation approach, the regression task is to encode the hits in an embedding space where hits from the same particle are clustered close together and far from hits from other showers. This is achieved through a loss function with repulsive and attractive potentials described in 4. Our approach is also based on the object condensation approach as starting from the raw data could in theory allow for larger gains by avoiding under optimal clustering.

3 Simulated dataset

We generate different datasets of e^+e^- collisions and propagate these particles in the detector using GEANT4. We use the CLD detector model [19], since it is the only FCC-ee detector for which the geometry and particle reconstruction are available at this stage.

The datasets include two datasets based on a custom particle gun and one based on physics events:

- The particle gun dataset generates particles of PID described in table 3 with balanced classes between $\{\text{CH}, \text{NH}, \gamma, \mu, e\}$. The gun generates an arbitrary number (between 10 and 15) of high-energy ($0.5 < E < 50$ GeV) stable final state hadrons.
- The second dataset uses the same particle gun but generates particles collimated in an arbitrary direction in the detector.
- Finally, in the third dataset we generate $H \rightarrow s\bar{s}$ events using Pythia.

The evaluation dataset has 10k events for all datasets.

An ML-ready dataset is created by storing the calorimeter hits and the reconstructed tracks, as well as the associations of each hit/track to the stable generator MC particles for evaluation purposes. Specifically, the inputs to our pipeline, $\mathcal{X}$ are a set of hits from the ECAL and HCAL calorimeters and the muon system, and a set of

Dataset	Number of particles	PIDs
Particle gun	10-15	$\pi^+, K_L, n, \gamma, e^+, \mu, K^+, p$
Jet-like particle gun	10-15	$\pi^+, K_L, n, \gamma, e^+, \mu, K^+, p$
$H \rightarrow s\bar{s}$		

Table 2 Details of the different datasets considered for training.

tracks $\mathcal{X} = \{\mathcal{X}_{\text{ECAL}}, \mathcal{X}_{\text{HCAL}}, \mathcal{X}_{\text{Muon}}, \mathcal{X}_{\text{T}}\}$. For every detector hit, the energy (E), position (x, y, z), and type in the detector are stored. The hit type is stored as a boolean and tracks whether the corresponding detector element is part of the ECAL (electromagnetic calorimeter), the HCAL (hadronic calorimeter), or muon subdetectors, respectively.

$$h_i = [E, x, y, z, \text{type}] \quad (6)$$

$$\text{type} \in [\text{ECAL}, \text{HCAL}, \text{MUON}, \text{track}] \quad (7)$$

If the element is a track, the position is the track state at the entrance of the calorimeter. In each event, there are $\mathcal{O}(600)$ hits per particle.

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 (E, p_x, p_y, p_z) and particle type (neutron or proton) (id) of the corresponding ground-truth particles is stored.

These events are reconstructed with the CLD pipeline [19], which contains an implementation of PandoraPFA, to obtain a baseline reconstruction.

3.1 Particularities of the data

When training an ML-model it is important that the data is balanced, e.g in an image classification problem the same number of images for each class would be available for training. There are certain particularities of particle reconstruction that make the dataset imbalances and that need to be taken into account when designing an ML-based particle flow. These are listed below:

- The distribution of types of particles in a jet is imbalanced: about 62% of jet energy is carried by charged particles, 27% by photons and 10% by long-lived neutral hadrons (e.g. $n, \bar{n}, K_L$) and around 1.5% by neutrinos. Therefore the number of particles of each class is also not roughly equal, for this reason we generate a particle gun for training the ML-based particle flow. Otherwise, the model will decrease the loss by becoming better at classes with higher representation. Particle-flow aims to reconstruct five particle classes, $\{\text{CH}, \text{NH}, \gamma, \mu, e\}$, therefore we design the dataset to be balanced according to these.
- In a jet 72% of the energy is measured in the HCAL, but the majority of hits are recorded in the ECAL due to higher granularity. It is therefore important to account for the imbalance of HCAL, ECAL and track hits in the loss function, as we describe in the next section. The distribution of energy of the particles also affects the imbalance between ECAL and HCAL hits as higher energy hadrons will deposit more hits further into the HCAL.

4 ML-based Particle Flow

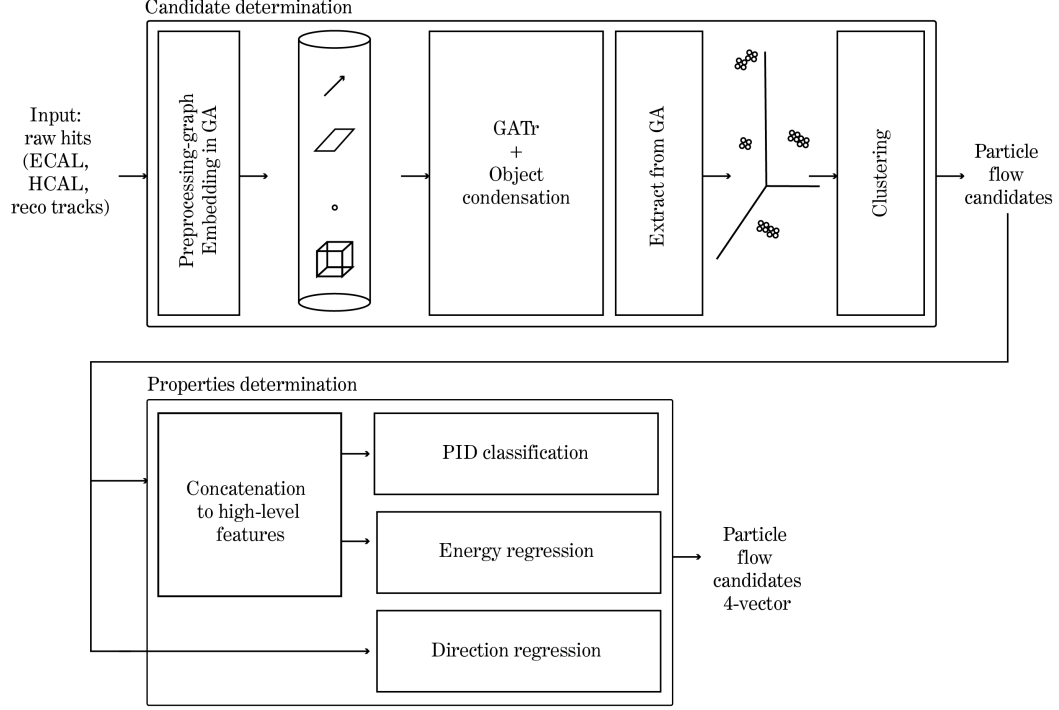


Fig. 1 Overview of the MLPF end-to-end approach for Particle Flow. The raw hit are represented as multivectors in the geometric algebra. The multivectors are embedded into a low-dimensional space using the Geometric Algebra Transformer (GATr) architecture and an object condensation approach, finally a clustering step determines the particle flow candidates. The four-vector determination is the second step of the pipeline, including, PID classification, as well as energy and direction regression.

Our pipeline is composed of multiple steps: First, the particle flow candidates are detected and second, the properties (PID, four-vector) are determined for each candidate.

4.1 Particle-flow candidates determination

This step is a clustering that defines the particle-flow candidates. First, the hits/tracks are represented as objects in a geometric algebra. Then, the GATr network [20] is used together with an object condensation loss to make the clusters of hits separable in an embedding space. Finally, a clustering algorithm is applied on this space to obtain particle flow candidates.

Graph representation

As an initial step, the set of hits for each event is converted into a graph. However, unlike traditional approaches, the GATr architecture allows to represent the raw hits as different geometric objects and transformations in the projective geometric algebra $\mathcal{G}_{3,0,1}$ using multivectors. This method allows to have as an input planes, lines, point or scalars. We choose to embed the hits as points with two associated scalars, one corresponding to the energy of the hit, E and the other one to the type of hit. Then, all multivectors in the event are connected, resulting in a fully connected graph.

Multivector processing

In order to obtain particle-flow candidates we perform an end-to-end clustering step. The task is to embed all multivectors (hits/track) belonging to the same particle close together in some embedding space. 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 [6, 9] aim to improve the separability of nearby showers, to minimise the error on the measured momenta and the mis-identification probability of the ancestor particle species by optimally combining all the available sub-detector information.

We start from the hits from the different systems (ECAL, HCAL, MUON, reconstructed tracks), unified in the multivector format, which represents the different geometrical objects in 16 dimensions. Then, we use the GATr architecture [21]. GATr consists of multiple transformer blocks. Each block has a LayerNorm, an equivariant multivector self-attention layer, a second LayerNorm, an equivariant multivector MLP with geometric bilinear interactions, and another residual connection. This architecture is $E(3)$ equivariant. The $E(3)$ -equivariant linear layers map multivectors using grade projections and multiplication with a homogenous basis vector. The geometric bilinears allow to build more expressive networks by allowing grade mixing. The network outputs multivectors which are used by the object condensation loss, defined below.

Object condensation loss and particle flow candidates

The object condensation approach for multi-object detection [16] allows to reconstruct an unknown number of objects through a dedicated loss function. The intuition behind this approach is as follows: the loss function requires the output of hits belonging to the same track to be close in a low dimensional embedding space and those belonging to different tracks to be far away. In this way, in the embedding space the tracks are easily separable.

The network is trained to predict a scalar $0 < \beta_i < 1$ and a coordinate x per node. The scalar is used to define a charge q_i with an increasing gradient:

$$q_i = \arctan^2 \beta_i + q_{min}. \quad (8)$$

This charge for every node is used to define a potential $V_{i,k}(x)$. The potential affecting nodes belonging to an object is approximated by the potential of the node with the highest charge in the object, α .

The full loss function is:

$$L = L_V + L_\beta \quad (9)$$

$$L_V = \frac{1}{N} \sum_{j=1}^N q_j \sum_{k=1}^K \left(M_{jk} \hat{V}_k(x_j) + (1 - M_{jk}) \check{V}_k(x_j) \right) \quad (10)$$

Each node j contributes with an attractive potential $\hat{V}_k(x_j)$ for the object they belong to, and with a repulsive potential $\check{V}_k(x_j)$ for the other $K - 1$ objects. The mask M_{jk} is 1 if j belongs to k and 0 otherwise. This potential is multiplied by the charge. L_β is defined to enforce one condensation point per object.

$$L_\beta = \frac{1}{K} \sum_k (1 - \beta_{\alpha k}). \quad (11)$$

Starting from the output multivectors of the network we define the coordinates of each node as the point component extracted from the multivector and the β for each node as the scalar component.

Clustering step

Finally, we cluster in the 3D embedding space using the HDBSCAN [22] clustering algorithm. Other clusterings are available, such as the one presented in [16], which uses the β in the embedding space to select cluster centers, or DBSCAN. However, HDBSCAN is chosen because it allows to find clusters of different densities and does not need a threshold on β . However, the clustering step is key to detecting fakes and an improved clustering could help reduce the fake rate.

4.2 Properties of the particle-flow objects

4.2.1 PID and energy correction

After clustering, new graphs are obtained by fully connecting the obtained clusters. Particle-level properties (PID and energy E) are predicted from concatenated *high-level features* and the outputs from a GATr head. The high-level features used describe key parameters of the clusters:

- The distribution of the energies among calorimeters (sum of the energies of the hits, fractions stored in ECAL and HCAL, dispersion of the energies in the ECAL and HCAL)
- The position of the shower and its generic properties (number of hits, average x , y , z and η , ϕ)
- Information about the track: track momentum and χ^2

Intuitively, the energy of the particle is mostly determined by the sum of the energies of the hits in the cluster. Similarly, PID can in most cases be identified by

looking at the presence of a track, and whether the energy is stored in the ECAL or also in HCAL. However, by applying a geometric machine learning model, more complex information about the shower geometry can be extracted to complement these features and further improve the prediction.

The particle mass is estimated from a classifier with 4 classes (e^- , γ , charged hadrons and neutral hadrons). For the charged hadrons, the mass is set to m_{π^+} and for the neutral hadrons to m_n .

4.2.2 Direction estimation

The momentum direction $\frac{\vec{p}}{|\vec{p}|}$ can be predicted from a separate GATr head. However, we find empirically that regressing the momentum direction does not provide a clear advantage over Pandora. Therefore we use a classical approach for charged and neutral candidates:

Charged candidates

We take as direction the normalized track momentum from the track state at IP.

Neutral candidates

Similarly to Pandora, we estimate $\frac{\vec{p}}{|\vec{p}|} \sim \frac{\vec{w}}{|\vec{w}|}$ where $w = \frac{\sum_i E_i \vec{r}_i}{\sum_i E_i}$ is a vector pointing from the interaction point $(0, 0, 0)$ to the energy-weighted average of the hit coordinates $\vec{r}_i$.

5 Results

5.1 Metrics

The following features allowing for comparison of MLPF with Pandora are computed.

- Energy resolution is estimated by fitting a Gaussian PDF $\mathcal{N}(\mu, \sigma)$ onto the histogram of E_{model}/E_{true} and reporting the variance divided by mean σ/μ .
- Angular resolution is estimated by looking at the difference in angles Φ and θ , defined according to the CMS coordinate system [23]. In case the difference in angles is greater than π , the smaller angle $\phi - \pi$ is taken for the difference.
- 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.
- Fake energy rate: the fraction of the energies of the particles that are reconstructed without matching to a true shower with respect to the energy of the true showers.

5.2 Particle gun

The model is evaluated per parts considering 1) the ground truth clustering (perfect clusters) with the ML-based properties determination and 2) the full MLPF end-to-end approach as described in Section 4.

Energy and angular resolution

Figure 2 shows the energy and angular resolution plots as a function of energy and separated by particle ID. The first column on the left shows the resolution of raw energy; therefore, for the true clustering, it is always zero. The MLPF also performs a better clustering than pandora for neutrons, K_L , π and protons. This improved clustering results in improved energy resolution for the neutral hadrons. For the charged hadrons, the energy is set to the momentum of the track so no differences are observed. The MLPF model also shows equivalent angular resolution to pandora, with improved resolution for neutral hadrons.

PID classification

Figure 3 displays confusion matrices for both models and the pandora baseline. We highlight that the muons in the range of 1-10 GeV are correctly classified by the MLPF 95% of the times (38% for pandora). This is due to the lower energy muons not always depositing hits on the muon detectors, therefore they are classified as charged hadrons by the pandora algorithm. However, the ECAL and HCAL deposits of these muons are characteristic and can be distinguished from the charged hadrons by the MLPF-PID block.

Efficiency

Figure 4 shows the efficiency metrics described in Section 5.1. For the ground truth clustering the efficiency is always one and the fake rate and fake energy rates are zero as the matching is always perfect. Overall, the efficiency at low energies is higher for MLPF with a similar fake rate and fake energy rate.

Mass resolution

Figure 5 shows the mass resolution of the different approaches. The highest resolution is achieved with the ML-based particle property determination with perfect clustering. This shows that the clustering can still be improved to achieve even higher performance. The full end-to-end MLPF approach outperforms pandora showing also higher energy resolution (Figure 5 right). This improvement comes from the improved performance in neutral hadrons, as we show next in an ablation study.

To determine the impact of each component on the overall performance we perform an ablation study fixing certain parts of the particle-flow both for the baseline and MLPF. We present the results considering only the neutral hadrons as here is where the main differences are observed.

The baseline for the ablation study is pandora shown in Figure 6 top left.

- The effect of the energy correction and direction estimation is shown in Figure 6 top center and left. MLPF’s improved energy correction for hadrons also results in improved mass and energy resolutions. For these plots we only consider particles that are matched by both pandora and MLPF in order to remove the effect of the efficiency.
- The effect of increased efficiency (together with energy correction and direction estimation) is shown in Figure 6 bottom left, which improves the mass resolution further.

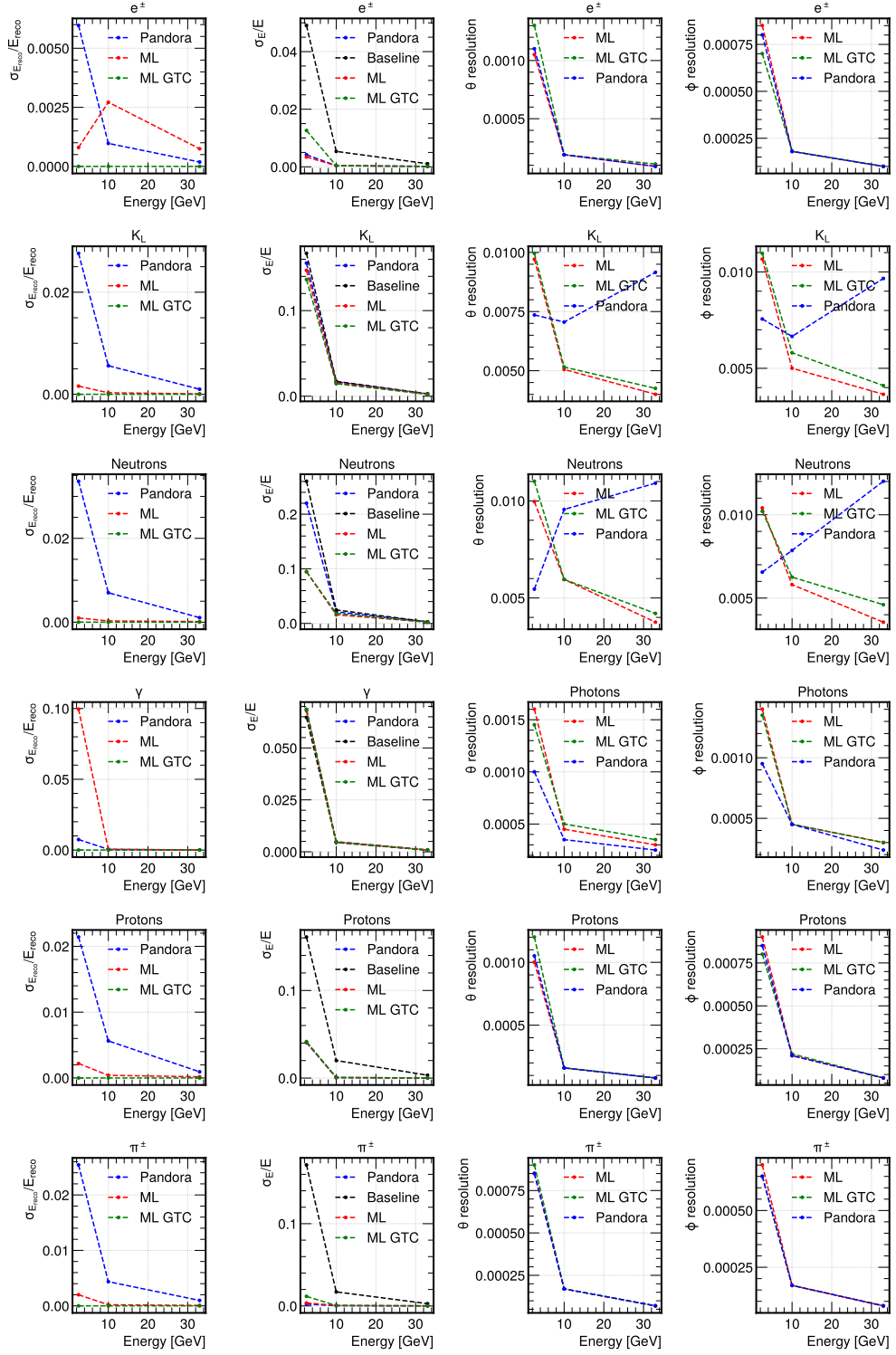


Fig. 2 Energy and angular resolution plots per energy for the main particle types.

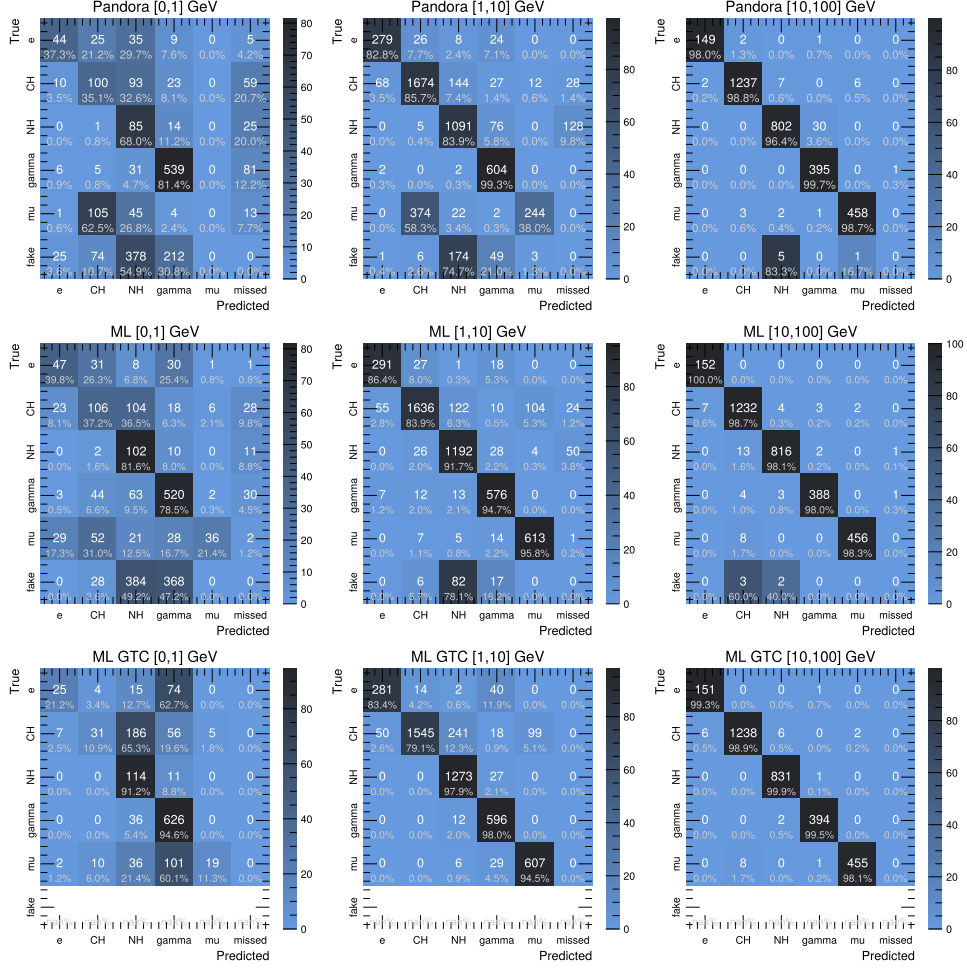


Fig. 3 Confusion matrix for different models, split into energy bins, for the model trained on the particle gun dataset. The fake particles are binned by the sum of the energies of the hits.

- The PID also shows improved performance for MLPF with 6% misidentification rate of pandora of neutral hadrons into photons.

Therefore, we can conclude that it is not only one block of MLPF that results in higher performance but that the different aspects of the reconstruction help, with the most significant improvement being increased efficiency.

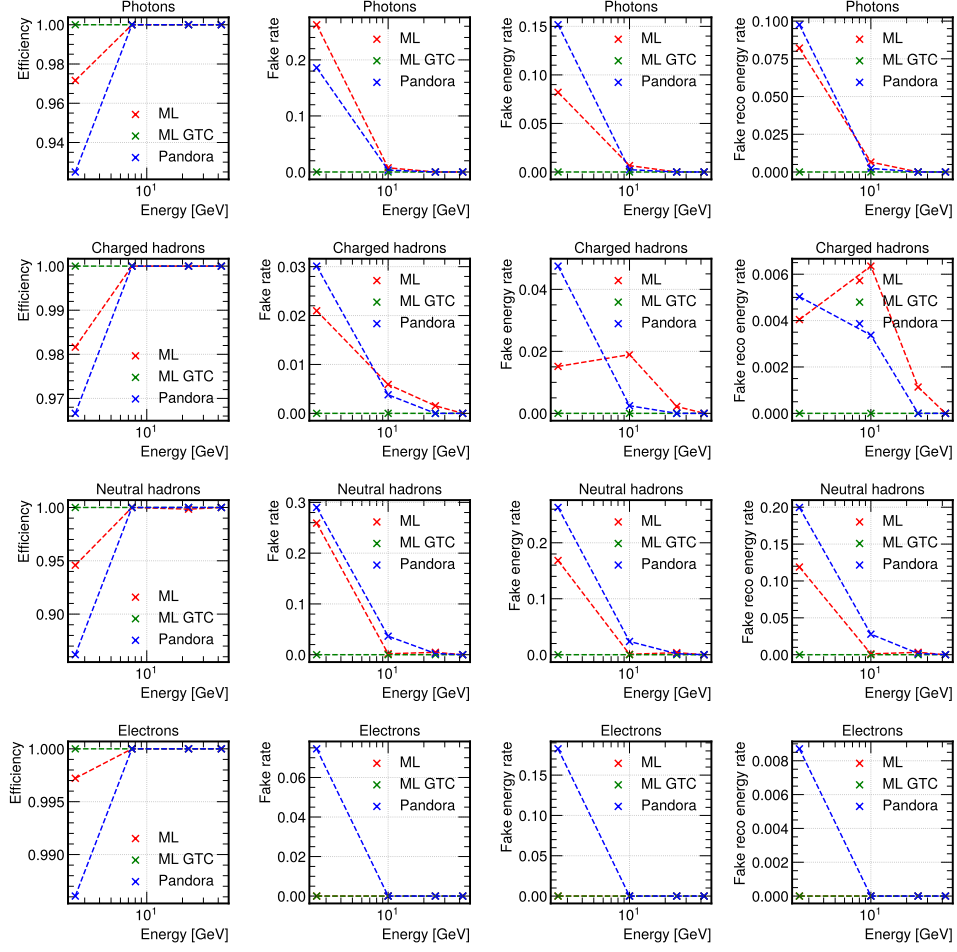


Fig. 4 Plots of reconstruction performance metrics for the particle gun dataset.

5.3 Jet-like particle gun

Energy and angular resolution

Figure 7 shows the energy and angular resolution plots as a function of energy and separated by particle ID. The first column on the left shows the resolution of raw energy. The MLPF shows improved energy resolution for the neutral hadrons. For the charged hadrons, the energy is set to the momentum of the track so no differences are observed. The MLPF model also shows equivalent angular resolution to pandora, with improved resolution for neutral hadrons.

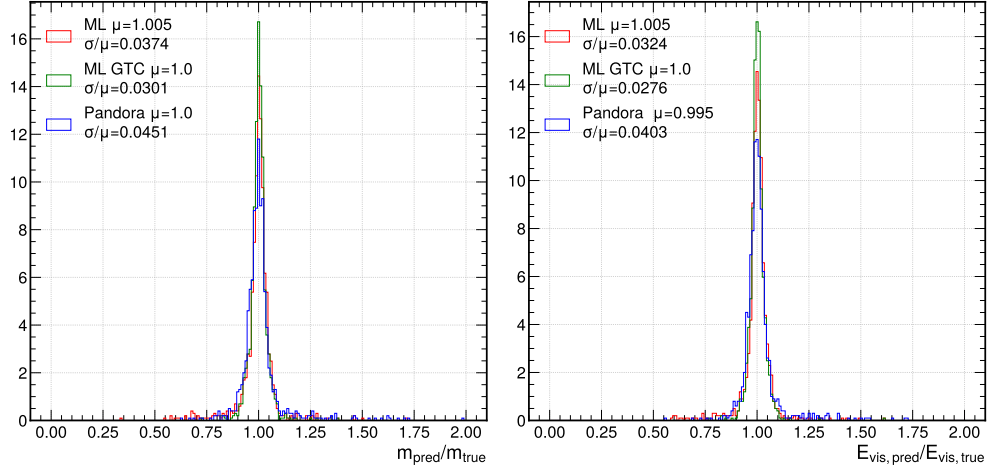


Fig. 5 Mass (left) and visible energy (right) histogram for the particle gun dataset with ideal clustering (green) and with ML trained clustering (red).

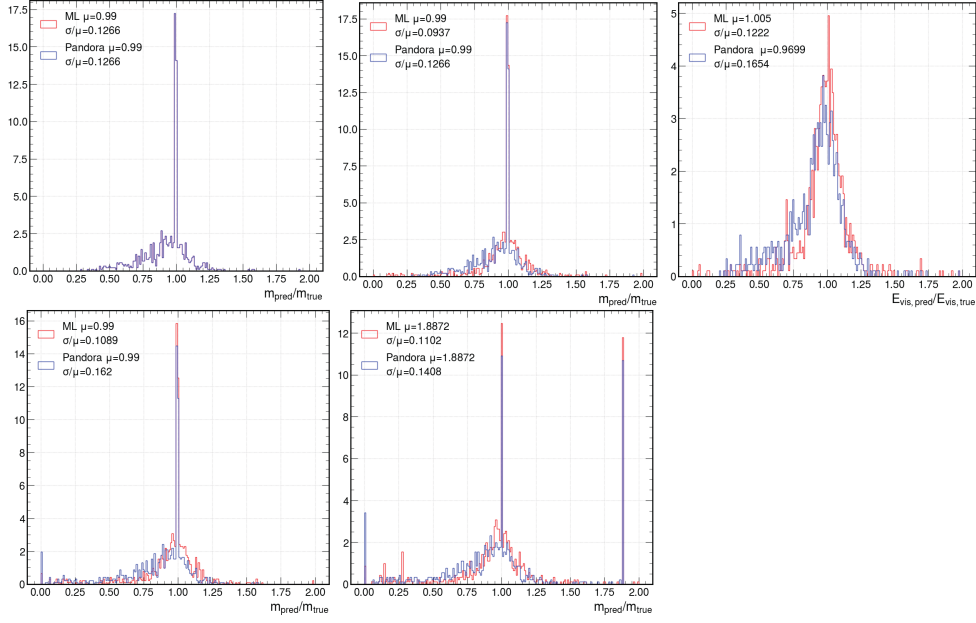


Fig. 6 Ablation study of the impact of different parts of the MLPF pipeline. Top left: neutral hadrons pandora baseline. Top center/right: effect of improved energy correction and direction estimation on mass and energy response distributions. Bottom left: effect of efficiency. Bottom center: effect of PID determination.

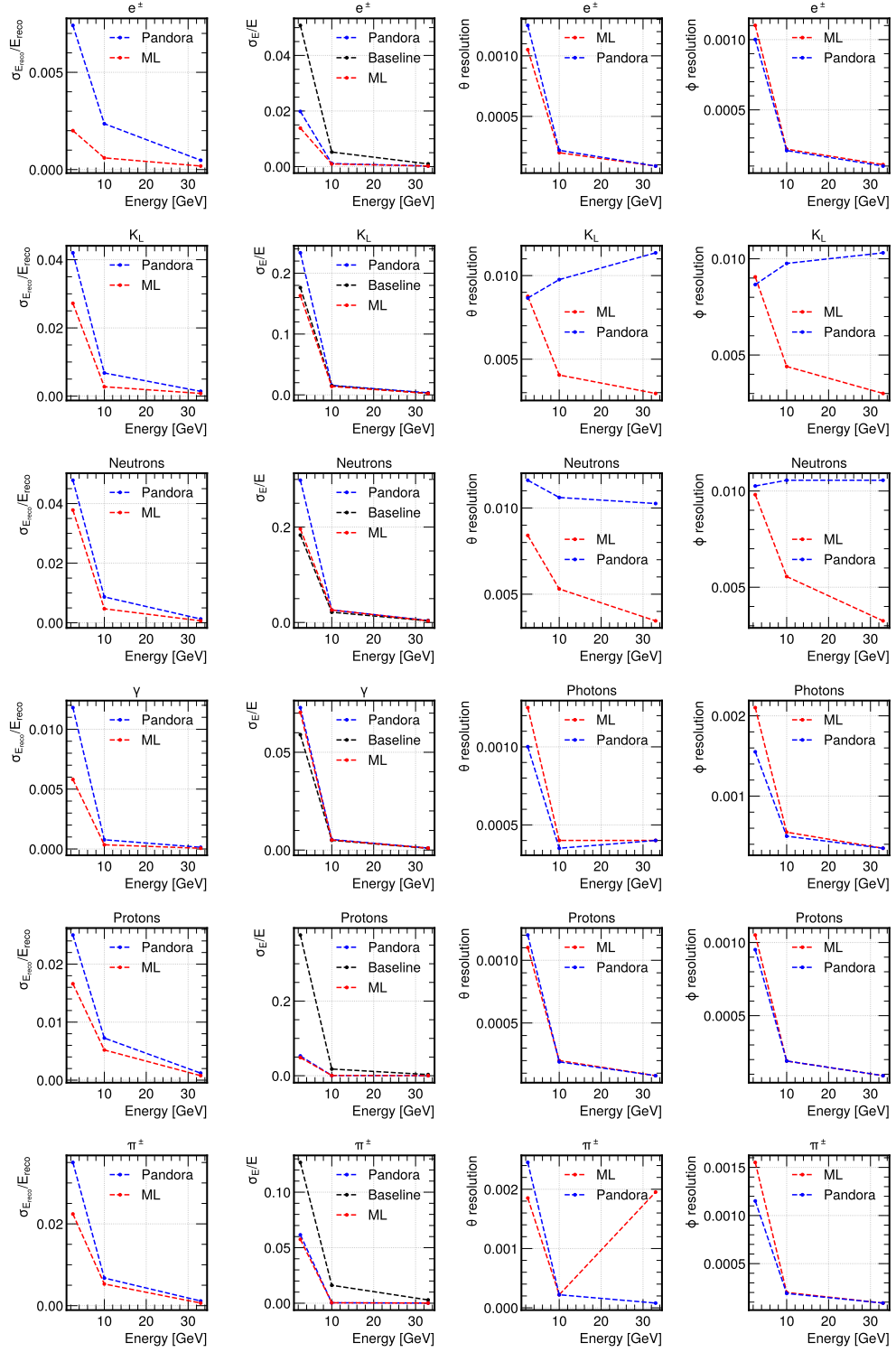


Fig. 7 Energy and angular resolution plots per energy for the main particle types for the jet-like particle gun dataset.

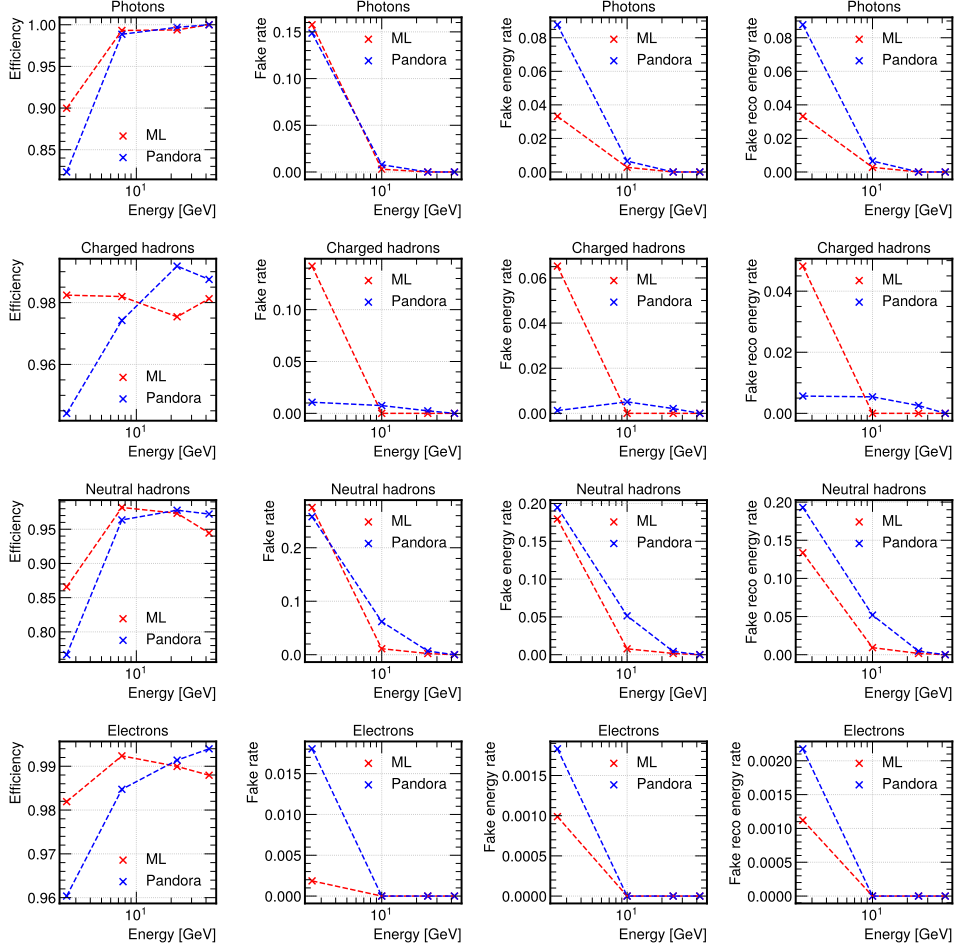


Fig. 8 Plots of reconstruction performance metrics for the jet-like particle gun dataset.

PID classification

Efficiency

Figure 8 shows the efficiency metrics described in Section 5.1. Overall, the efficiency at low energies is higher for MLPF with a similar fake rate and fake energy rate. However, the efficiency at higher energies is worse, this is due to close by showers that are hard to cluster for the algorithm.

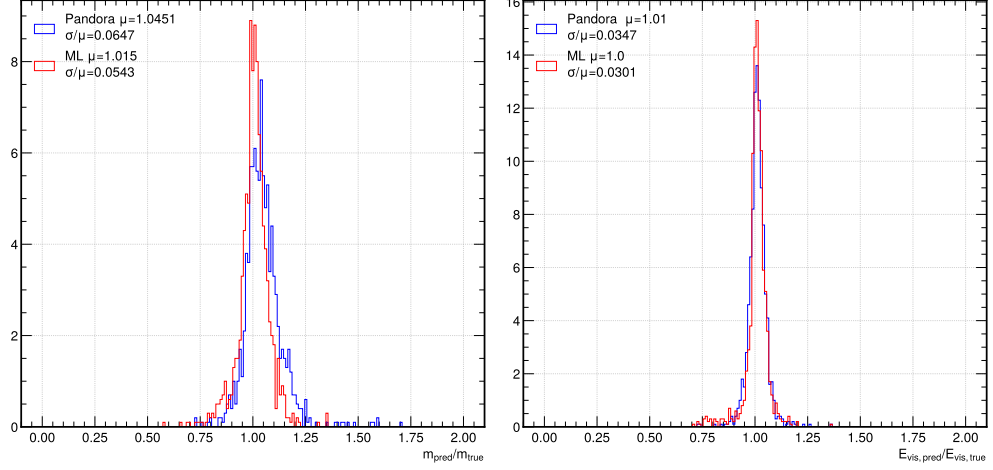


Fig. 9 Mass (left) and visible energy (right) histogram for the jet-like particle gun dataset with pandora (blue) and with ML trained clustering (red).

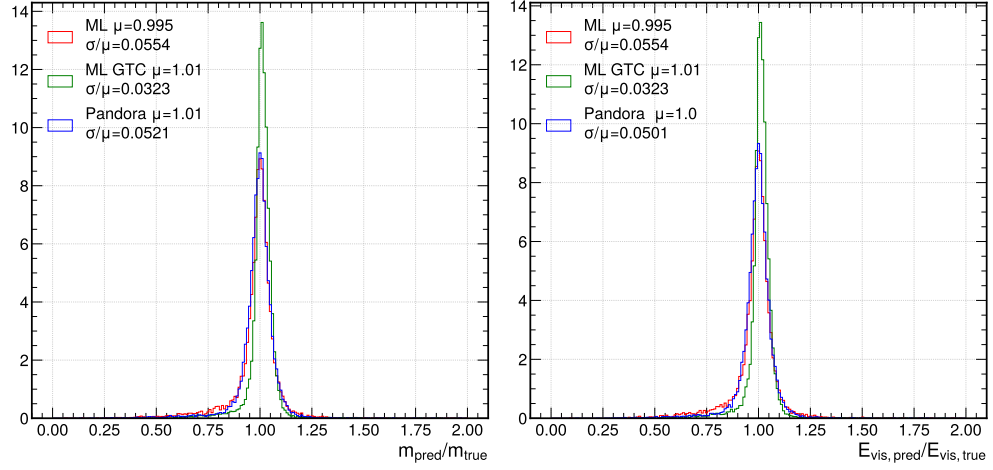


Fig. 10 Mass (left) and visible energy (right) histogram for the $H \rightarrow s\bar{s}$ dataset with ideal clustering (green) and with ML trained clustering (red).

Mass resolution

5.4 $H \rightarrow s\bar{s}$

The current performance of MLPF on the dataset of $H \rightarrow s\bar{s}$ physics events is reported in Figure 10. We highlight that the energy correction head used for ML was trained on the Hss dataset, while the energy correction used for ML with ground truth clustering was trained on the particle gun dataset.

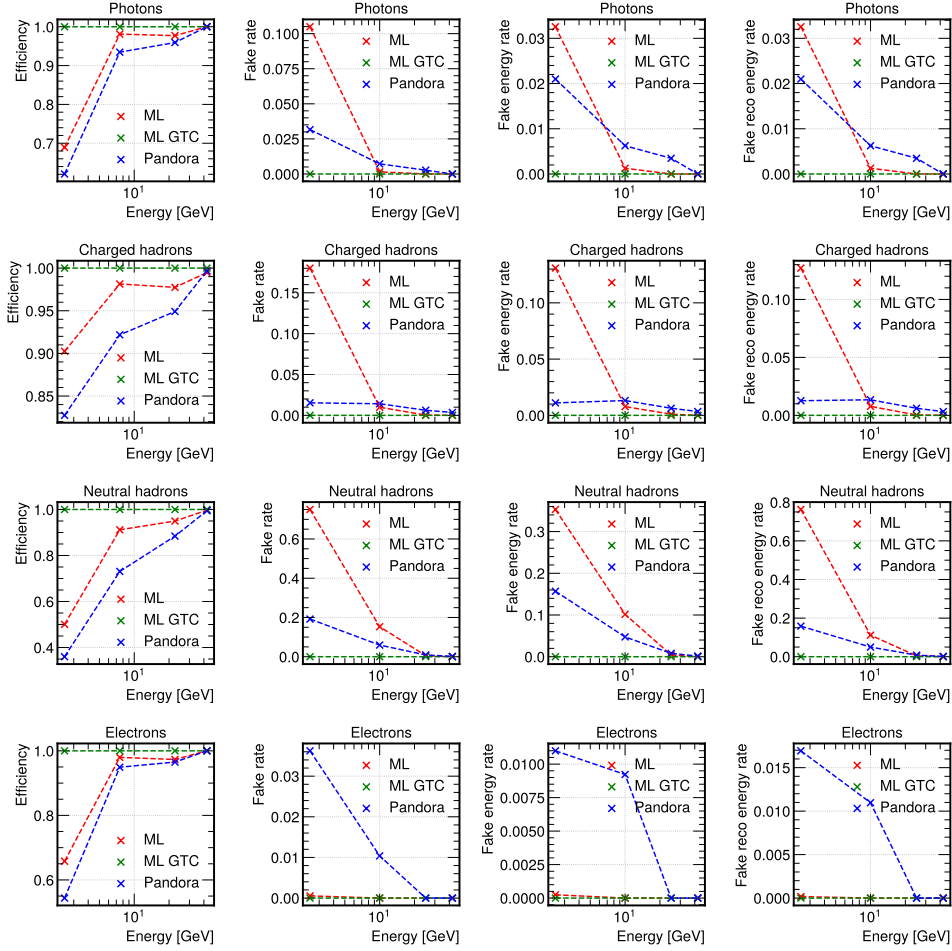


Fig. 11 Clustering metrics for the $H \rightarrow s\bar{s}$ dataset.

The high fake rate (Figure 11) needs to be addressed in the future. Many fakes stem from the clusters without a well-defined condensation point: Figure 12 shows that we can remove more fakes by removing the clusters where the highest beta falls in the range $[0.03, 0.3]$. However, as seen from Figure 13, the applied filter doesn't contribute significantly to the overall mass resolution, so other strategies to use the information stored in the fake showers need to be addressed.

6 Conclusion

We present a graph neural network-based model trained for particle flow in the proposed CLD detector at the FCC-ee.

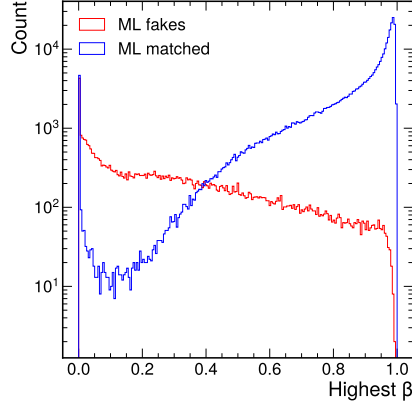


Fig. 12 Distributions of the highest β values in the cluster for the fake and matched particles for the $H \rightarrow s\bar{s}$ dataset.

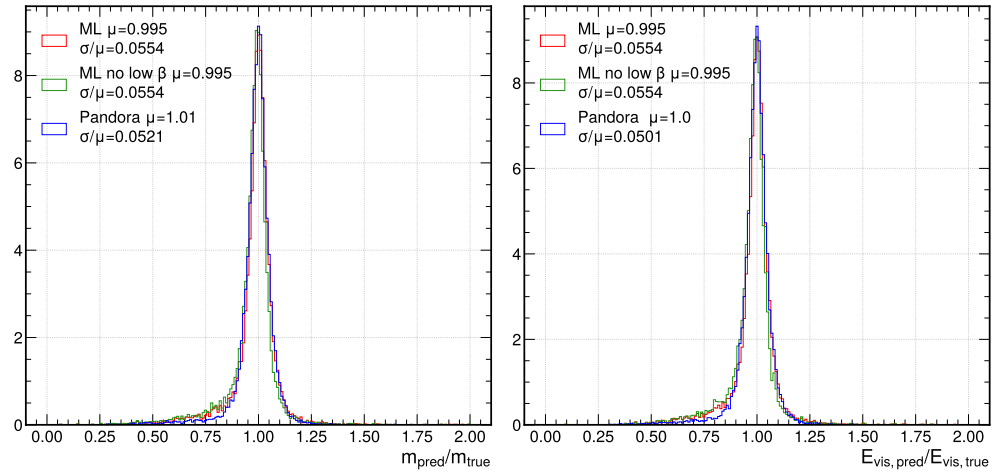


Fig. 13 Mass (left) and visible energy (right) histograms for the ML approach, as well as the ML approach with removed particles with highest $\beta \in [0.03, 0.3]$. **GK - It's not really a super relevant plot, it can also be removed or put in the appendix**

We demonstrate that state-of-the-art geometric graph neural network architectures can be used to improve particle flow performance of the classical baseline Pandora. Our work serves as a successful implementation of machine learning for particle-flow reconstruction. A key advantage of our approach compared to Pandora is that there is no need for detector-specific parameter tuning.

However, there are still various aspects of the model that need more detailed assessment. Additional training strategies to reduce the fake rate at low energies while

maintaining high efficiency, could be explored. hat extent our model trained on a 10-15 particle gun dataset works well on physics events. Additionally, training strategies to enable reconstruction at different detectors with the same model could be explored. That could enable even faster detector optimization.

7 Perspectives and outlook

Predicting an unknown number of elements and their properties from diverse inputs is a broad task that extends beyond the confines of the particle physics community. While the problem has been researched extensively in the context of computer vision, how to apply it effectively to unstructured data such as graphs or point clouds is still an emerging area of research.

Acknowledgements

This work has been partly supported by the Future Circular Collider Innovation Study (FCCIS) project, which has received funding from the European Union’s Horizon 2020 research and innovation programme under grant No 951754. Gregor is also supported by the Ad futura scholarship of the Public Scholarship Fund of Slovenia. We acknowledge the support of the ETH Zürich Institute for Particle Physics and Astrophysics for funding Gregor’s stays at CERN.

References

- [1] A. Abada, et al., FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2. *Eur. Phys. J. ST* **228**(2), 261–623 (2019). <https://doi.org/10.1140/epjst/e2019-900045-4>
- [2] D. Buskulic, et al., Performance of the ALEPH detector at LEP. *Nucl. Instrum. Meth. A* **360**, 481–506 (1995). [https://doi.org/10.1016/0168-9002\(95\)00138-7](https://doi.org/10.1016/0168-9002(95)00138-7)
- [3] P. Abreu, et al., Performance of the DELPHI detector. *Nucl. Instrum. Meth. A* **378**, 57–100 (1996). [https://doi.org/10.1016/0168-9002\(96\)00463-9](https://doi.org/10.1016/0168-9002(96)00463-9)
- [4] M. Wing, Precise measurement of jet energies with the ZEUS detector. *Frascati Phys. Ser.* **21**, 617–623 (2001). [arXiv:hep-ex/0011046](https://arxiv.org/abs/hep-ex/0011046)
- [5] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET (2009)
- [6] A.M. Sirunyan, et al., Particle-flow reconstruction and global event description with the CMS detector. *JINST* **12**(10), P10003 (2017). <https://doi.org/10.1088/1748-0221/12/10/P10003>. [arXiv:1706.04965](https://arxiv.org/abs/1706.04965) [physics.ins-det]
- [7] M. Aaboud, et al., Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C* **77**(7), 466 (2017). <https://doi.org/10.1140/epjc/s10052-017-5031-2>. [arXiv:1703.10485](https://arxiv.org/abs/1703.10485) [hep-ex]
- [8] M. Ruan, H. Videau, in *International Conference on Calorimetry for the High Energy Frontier* (2013), pp. 316–324
- [9] M.A. Thomson, Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth. A* **611**, 25–40 (2009). <https://doi.org/10.1016/j.nima.2009.09.009>. [arXiv:0907.3577](https://arxiv.org/abs/0907.3577) [physics.ins-det]
- [10] D. Buskulic, D. Casper, I. De Bonis, D. Decamp, P. Ghez, C. Goy, J.P. Lees, M.N. Minard, P. Odier, B. Pietrzyk, et al., Performance of the aleph detector at lep. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **360**(3), 481–506 (1995)
- [11] L. Gray, in *38th International Conference on High Energy Physics (ICHEP 2016)*, (Link) (2016)
- [12] M. Thomson, Particle flow calorimetry and the pandorapfa algorithm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **611**(1), 25–40 (2009)
- [13] J. Pata, E. Wulff, F. Mokhtar, D. Southwick, M. Zhang, M. Girone, J. Duarte, Improved particle-flow event reconstruction with scalable neural networks for current and future particle detectors. *Communications Physics* **7**(1), 124 (2024)

- [14] J. Pata, J. Duarte, J.R. Vlimant, M. Pierini, M. Spiropulu, Mlpf: efficient machine-learned particle-flow reconstruction using graph neural networks. *The European Physical Journal C* **81**, 1–14 (2021)
- [15] F. Mokhtar, J. Pata, J. Duarte, E. Wulff, M. Pierini, J.R. Vlimant, Progress towards an improved particle flow algorithm at cms with machine learning. *arXiv preprint arXiv:2303.17657* (2023)
- [16] J. Kieseler, Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data. *The European Physical Journal C* **80**, 1–12 (2020)
- [17] S.R. Qasim, K. Long, J. Kieseler, M. Pierini, R. Nawaz, in *EPJ Web of Conferences*, vol. 251 (EDP Sciences, 2021), p. 03072
- [18] S.R. Qasim, J. Kieseler, Y. Iiyama, M. Pierini, Learning representations of irregular particle-detector geometry with distance-weighted graph networks. *The European Physical Journal C* **79**(7), 1–11 (2019)
- [19] N. Bacchetta, J.J. Blaising, E. Brondolin, M. Dam, D. Dannheim, K. Elsener, D. Hynds, P. Janot, A. Kolano, E. Leogrande, et al., Cld—a detector concept for the fcc-ee. *arXiv preprint arXiv:1911.12230* (2019)
- [20] J. Brehmer, P. de Haan, S. Behrends, T. Cohen, in *Advances in Neural Information Processing Systems*, vol. 37 (2023). URL <https://arxiv.org/abs/2305.18415>
- [21] J. Brehmer, P. De Haan, S. Behrends, T.S. Cohen, Geometric algebra transformer. *Advances in Neural Information Processing Systems* **36** (2024)
- [22] L. McInnes, J. Healy, S. Astels, et al., hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2**(11), 205 (2017)
- [23] G. Flucke, Alignment of the cms silicon tracker. *Journal of Physics: Conference Series* (2012)