

# Machine Learning for the Cluster Reconstruction in the CALIFA Calorimeter at R3B

T. Jenegger<sup>a</sup>, N. Hartman<sup>a</sup>, R. Gernhäuser<sup>a</sup>, L. Fabbietti<sup>a</sup>, L. Heinrich<sup>a</sup>

<sup>a</sup>*TUM School of Natural Sciences, Technical University of Munich, Germany*

---

## Abstract

The R3B experiment at FAIR studies nuclear reactions using high-energy radioactive beams. One key detector in R3B is the CALIFA calorimeter consisting of 2544 CsI(Tl) scintillator crystals designed to detect light charged particles and gamma rays with an energy resolution in the per cent range after Doppler correction. Precise cluster reconstruction from sparse hit patterns is a crucial requirement. Standard algorithms typically use fixed cluster sizes or geometric thresholds. To enhance performance, advanced machine learning techniques such as agglomerative clustering were implemented to use the full multi-dimensional parameter space including geometry, energy and time of individual interactions. An Edge Detection Neural Network exhibited significant differences. This study, based on Geant4 simulations, demonstrates improvements in cluster reconstruction efficiency of more than 30%, showcasing the potential of machine learning in nuclear physics experiments.

*Keywords:* R3B Experiment, CALIFA Calorimeter, Cluster Reconstruction, Machine Learning, Simulation

---

## 1. Introduction

With the advancements in facilities dedicated to the production of radioactive beams at relativistic energies, such as the Facility for Antiproton and Ion Research (FAIR) at GSI, significant progress is expected for our understanding of exotic nuclei far from stability [1]. FAIR will provide high-intensity relativistic radioactive beams of rare isotopes with energies in the range of one GeV per

14 nucleon, enabling investigations with full kinematic reconstruction [2]. A key  
 15 experimental setup designed for this purpose is the **R**eactions with **R**elativistic  
 16 **R**adioactive **B**eams (R3B) setup, providing access to high-resolution spectro-  
 17 scopic data. This setup serves as a unique tool for unveiling the structure of  
 18 nuclei and their reaction dynamics with unprecedented precision.  
 19 At the core of the R3B Setup is the CALIFA calorimeter (Calorimeter for the  
 20 In-Flight Detection of Gamma Rays and Light Charged Particles), a highly  
 21 segmented detection system composed of 2544 CsI(Tl) scintillator crystals that  
 22 hermetically enclose the target area in the polar angular range of  $7^\circ < \theta < 140^\circ$   
 23 (see Fig. 1). This design enables the simultaneous measurement of gamma  
 24 rays down to  $E_\gamma \approx 100$  keV and light charged particles, such as protons and  
 25 deuterons, up to several hundred A MeV [3]. To ensure optimal performance,  
 26 extensive research has been conducted to refine the geometric design, minimize  
 27 scattering and energy loss due to the mechanical structure [4], and develop a  
 28 dead-time-free data acquisition system capable of handling high-rate experi-  
 29 ments [5]. Furthermore, a seamless integration within the R3BRoot framework  
 30 [6] has been achieved, enabling offline data analysis from the raw to the cali-  
 31 brated data level and ultimately to the cluster level, where individual hits are  
 32 recombined for the final energy reconstruction.  
 33 This study presents the results of a hierarchical machine learning model to en-  
 34 hance the energy reconstruction of gamma rays in CALIFA. Using simulated  
 35 Geant4 data, the performance of the geometrical R3B clustering algorithm is  
 36 compared to an agglomerative clustering model [7] and a multi-layer perceptron  
 37 architecture [8], demonstrating the potential of machine learning techniques in  
 38 improving reconstruction efficiency and accuracy.

39

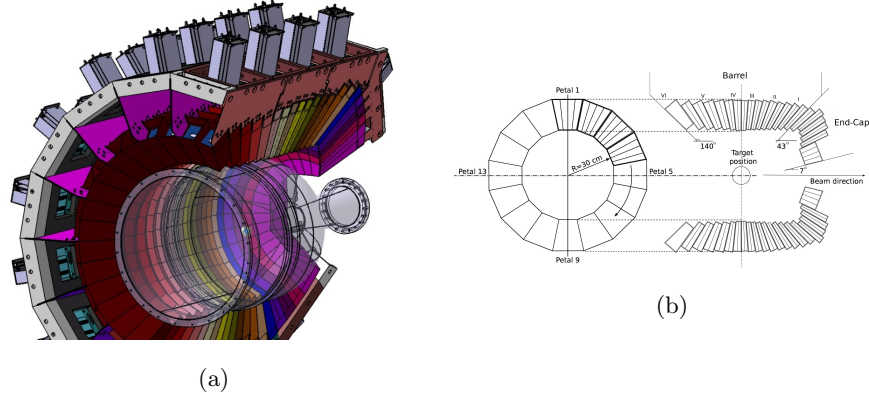


Figure 1: (a) Graphical representation of the CALIFA detector. Carbon fiber alveoli and aluminum holders fix the 15 to 22 cm long CsI(Tl) crystals. The gray boxes surrounding the holding structure represent the preamplifiers. (b) Cross profile and longitudinal section of the detector. The azimuthal angular coverage of the crystals vary between  $1.5^\circ$  (End-Cap) to  $3^\circ$  (Barrel). Figures taken from Ref. [9].

## 2. Methodology

### 2.1. Challenges in Relativistic Gamma Spectroscopy

While the detection of light charged particles such as protons typically yields well-localized energy deposits in segmented detector arrays, the detection of gamma rays which emerge from the reaction vertex presents significant challenges. These primarily arise from the inherently sparse and spatially distributed energy deposits resulting from the interaction mechanisms of photons with the scintillator material (see Fig. 2) [10].

At photon energies below approximately 300 keV, the photoelectric effect dominates the interaction cross-section in the CALIFA detector material (CsI(Tl)). As the photon energy increases, Compton scattering becomes the predominant process. For photon energies exceeding the pair production threshold ( $E_\gamma > 2m_e c^2 \approx 1.022 \text{ MeV}$ ), electron-positron pair creation becomes possible and is the dominant interaction mechanism above  $E_\gamma \approx 6 \text{ MeV}$ . Compton scattering broadens the clustering by the deflection of the incident gamma ray. According to the Klein-Nishina formula, the scattering is predom-

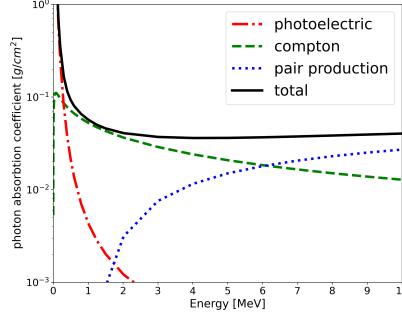


Figure 2: Photon absorption coefficients in CsI in the range from 100 keV to 10 MeV with data from XCOM database [11].

inantly forward-focused for moderate to high photon energies [12], leading to additional clusters in neighboring crystals.

At high photon energies, the dominant interaction mechanism in the detector material is pair production (see Fig. 2), in which the incident photon converts into an electron-positron pair during the initial interaction. The subsequent annihilation of the positron results in the emission of two additional gamma photons, each with an energy of 511 keV. These secondary photons often escape the initial interaction site, leading to a significant fraction of the incident photons energy being deposited in multiple detector elements.

For gamma rays emitted by nuclei at rest, this behavior gives rise to well-defined single- and double-escape peaks in the recorded energy spectra – corresponding to the escape of one or both 511 keV photons, respectively – if these photons exit the cluster volume without interaction.

In experiments involving relativistic ion beams, such as those exploited at R3B, Doppler broadening significantly affects the observed spectral features, including the single- and double-escape peaks. Moreover, for primary gamma rays with energies well above the pair production threshold ( $E_\gamma > 2m_e c^2$ ), both the electron and the positron produced in the initial interaction are subject to substantial energy loss via Bremsstrahlung. These effects contribute to a complex and highly non-trivial interaction pattern of gamma rays within the segmented

76 detector system.

77

## 78 *2.2. Data Structure and geometrical R3B Clustering Algorithm*

79 The fundamental data entity in the analysis is a hit, defined as a discrete  
80 signal recorded by an individual detector segment at a specific time. To suppress  
81 contributions from low-energy background, only signals exceeding a predefined  
82 energy threshold are registered. In the present analysis, this threshold was set  
83 to 100 keV.

84 In the standard data acquisition (DAQ) configuration, all CALIFA detector hits  
85 occurring within a  $\pm 4\mu\text{s}$  time window are grouped into a single event. Each  
86 individual hit  $i$  in one of the detector crystals is represented by a data structure  
87 containing the calibrated energy deposit  $E_i$ , the polar angle  $\theta_i$ , the azimuthal  
88 angle  $\phi_i$ , and a time stamp  $t_i$ , which is synchronized using the White Rabbit  
89 Precision Time Protocol [13].

90 In the geometrical R3B clustering approach, the time information  $t_i$  is not uti-  
91 lized during the spatial reconstruction of clusters.

92

93 For the cluster reconstruction all detected hits are sorted in a list by descend-  
94 ing energy. A user-defined cluster shape, typically a cone with an aperture of  
95 0.25 rad, is chosen. This value represents an optimal balance between the com-  
96 pact, high-energy clusters characteristic of light charged particles and the more  
97 diffuse showers produced by gamma rays.

98 The clustering process begins by assigning the hit with the highest energy as  
99 the center of the first cluster. The algorithm iterates through the remaining hits  
100 in the sorted list. A hit is added to the current cluster if its angle relative to  
101 the cluster's central hit is within the defined aperture. After all hits in the list  
102 have been processed, the assigned hits are removed. The hit with the highest  
103 remaining energy is then selected as the center of a new cluster, and the process  
104 is repeated until no unassigned hits remain.

105

### 106 2.3. Simulation Setup

107 Simulated datasets are used to evaluate and compare the performance of  
108 the clustering algorithms presented in this work. A geometrical model of the  
109 detector, closely matching the experimental setup, was implemented within  
110 the R3BROOT framework. The simulation employs a GEANT4-based Monte  
111 Carlo [14] back-end, which accounts for all relevant secondary interaction pro-  
112 cesses. This approach enables realistic modeling of energy deposition and pro-  
113 vides access to ground-truth labels for each individual interaction.

114  
115 The CALIFA detector geometry used in the simulation corresponds to the  
116 configuration implemented in early 2024. At that time, the iPhos region (polar  
117 angles  $19^\circ - 43^\circ$ ) was fully instrumented, while only the forward half of the  
118 Barrel region ( $43^\circ - 87^\circ$ ) was active. The forward-most CEPA region ( $7^\circ - 19^\circ$ )  
119 was not yet equipped.

120 Gamma-ray energies were sampled from a uniform distribution between 0.3 MeV  
121 and 10 MeV. The interaction of the primary gamma rays with the CsI(Tl) scin-  
122 tillation material was modeled using Geant4<sup>1</sup>.

123 To emulate realistic event topologies of signal and background, three gamma  
124 rays were generated per event, resulting in multiple detector hits. Timing infor-  
125 mation was coarsely simulated by assigning to each primary gamma a random  
126 emission time within the  $\pm 4 \mu\text{s}$  event window. The corresponding hit times were  
127 then Gaussian-smeared with a standard deviation of 200 ns to reflect the timing  
128 spread of the electronic signal of slow CsI(Tl) scintillator crystals.

129 Event selection is limited to cases in which all three gamma rays are emitted  
130 within the geometrical acceptance of the CALIFA detector, which only partially  
131 encloses the target region. For gamma rays that deposit only a fraction of their  
132 energy in the detector volume – such as in cases where the incident gamma ray  
133 undergoes Compton scattering, deposits part of its energy in the calorimeter,  
134 and subsequently escapes the active volume – the corresponding true energy is

---

<sup>1</sup>Version: geant4-11-02; used physics list: QGSP\_BERT\_HP

adjusted to reflect only the energy actually deposited in CALIFA.  
The resulting dataset was split into training and test subsets, comprising 13,000  
and 7,000 events, respectively.

#### 2.4. Performance Metrics

To quantitatively assess the performance of the clustering algorithms presented in this work, a set of four custom metrics was defined. Three of these are event-based, while an optional fourth metric evaluates clustering quality on a per-cluster basis:

- **True Positive (TP)**: All hits in an event are correctly assigned to their respective clusters.
- **False Positive (FP)**: At least one hit in an event is incorrectly merged into a cluster it does not belong to.
- **False Negative (FN)**: At least one hit is not merged into its true cluster and instead forms a spurious cluster.
- **False Mixed (FM)**: An event is classified as false mixed if it contains both FP and FN characteristics – i.e., at least one hit is incorrectly merged, and at least one true cluster is partially reconstructed.

In addition, a cluster-based metric is defined:

- **Well Reconstructed (WR)**: The ratio of correctly reconstructed clusters to the total number of true clusters in the dataset.

These metrics allow a comprehensive evaluation of clustering accuracy, robustness, and failure modes.

Special attention must be given to the false negative rate, which is closely associated with the complex interaction pattern in the segmented detector. These processes produce widely spread hits that cannot be merged using the geometrical R3B clustering method, thereby motivating the development of a multi-layer perceptron architecture to improve clustering performance at the boundaries

162 (see Subsection 2.6).

163

## 164 2.5. Agglomerative Clustering

165 To incorporate temporal information into the clustering process—unlike the  
166 geometrical R3B algorithm, which omits it—a generic, well-established method  
167 was adopted: agglomerative clustering [7] as implemented in the `SciPy` library  
168 [15]. This unsupervised learning algorithm enables flat clustering based on hi-  
169 erarchical linkage with a user-defined threshold.

170 Each hit was mapped into spherical coordinates  $(\theta, \phi, r)$ , where the radial com-  
171 ponent  $r$  encodes time information. To ensure non-negative radii, the acquisition  
172 time window of  $\pm 4 \mu\text{s}$  was shifted by  $+4.5 \mu\text{s}$ . The Ward linkage criterion [16],  
173 which minimizes intra-cluster variance, was employed as the distance metric.  
174 The threshold parameter was optimized to yield the best performance according  
175 to the custom-defined *true positive* (TP) and *well reconstructed* (WR) metrics.  
176 As shown in Table 1, the agglomerative clustering algorithm demonstrates im-  
177 proved performance both on an event level (true positive rate) and on a cluster  
178 level (correctly reconstructed clusters) compared to the geometrical R3B cluster-  
179 ing. However, this improvement is accompanied by an increased false negative  
180 rate, indicating that the algorithm tends to under-merge hits near the edges of  
181 clusters. This limitation motivated the development and application of an edge  
182 detection neural network, which is introduced in the following subsection.

## 183 2.6. Edge Detection Neural Network

184 To enhance the clustering performance, particularly at the boundaries of hit  
185 distributions, a multi-layer perceptron architecture was developed using the Py-  
186 torch library [17] to perform pairwise classification of detector hits. This model  
187 is applied either to individual raw hits or to hits pre-clustered via agglomerative  
188 clustering, on an event-by-event basis.

189 The model takes 12 input features for each hit pair  $(i, j)$ : absolute values of  
190 energy  $(E_i, E_j)$ , polar angle  $(\theta_i, \theta_j)$ , azimuthal angle  $(\phi_i, \phi_j)$ , and time  $(t_i, t_j)$ .



191 Additionally, four differential features are computed:  $\Delta E = |E_i - E_j|$ ,  $\Delta\theta =$   
192  $|\theta_i - \theta_j|$ ,  $\Delta\phi = |\phi_i - \phi_j|$ , and  $\Delta t = |t_i - t_j|$ . These differential inputs are  
193 helpful for training stability and convergence with our limited model sizes tested.  
194 In particular,  $\Delta\phi$  resolves the discontinuities caused by the periodicity of the  
195 azimuthal angle (e.g., distinguishing between  $\phi = 355^\circ$  and  $\phi = 5^\circ$ ), which  
196 would otherwise introduce large erroneous differences in angular comparisons.

197 Of the 12 features, only the hit time is normalized to the  $[0, 1]$  interval; all  
198 other values are used in their native physical units. The neural network ar-  
199 chitecture takes the 12-dimensional input vector and passes it through a fully  
200 connected feed-forward network with one hidden layer of  $10^3$  nodes, followed by  
201 a rectifier linear unit activation function (ReLU) [18]. Two additional hidden  
202 layers, each with  $10^2$  nodes, are applied sequentially. The output layer consists  
203 of a single node with a sigmoid activation, yielding a score in the interval  $[0, 1]$ ,  
204 where values close to 1 indicate that the hits (or clusters) are likely to originate  
205 from the same event cluster.

206

207 Training is performed using the binary cross-entropy loss function [19, 20]  
208 and stochastic gradient descent (SGD) [21] with a fixed learning rate of  $5 \times 10^{-3}$ .  
209 Given the moderate size of the training dataset, full-batch training is employed  
210 without mini-batching. The model is trained for  $8 \times 10^4$  epochs. After training, a  
211 threshold is applied to the prediction scores to classify hit pairs. This threshold  
212 is tuned to optimize the performance across all defined metrics, as described in  
213 Subsection 2.4. Final clusters are then formed by grouping all connected hit  
214 pairs based on the predicted associations.

215 The edge detection NN was implemented and tested in three configurations:

- 216 • **Plain Edge NN:** The model is applied directly to individual hits without  
217 any pre-clustering. All clustering is performed based solely on the NN  
218 predictions.
- 219 • **R3B + Edge NN:** The data are first clustered using the geometrical R3B  
220 clustering algorithm as an initial clean-up step. For each resulting cluster,

an energy-weighted center of mass is calculated, replacing individual hits. The NN is then trained exclusively on false negative cases, i.e., events where reconstructed clusters exhibit detached hits. In application, the geometrical R3B clustering is first applied to the test data, followed by the NN to refine cluster boundaries and reduce the false negative rate as clean-up step.

- **Agglo + Edge NN:** This strategy mirrors the R3B+Edge approach, with the key difference that time information is incorporated. As in the R3B+Edge model, the NN is trained on false negative cases to perform a final clean-up step after pre-clustering the hits using the agglomerative clustering algorithm described in the previous subsection. The significant reduction of the false negative rate achieved by the clean-up step in the Agglo+Edge implementation is demonstrated in Fig. 3, which compares the reconstructed energy spectra from simulations of mono-energetic 2.1 MeV gamma events using the geometrical R3B clustering and the Agglo+Edge method.

### 3. Discussion

Clustering Model	TP(†)	FP(↓)	FN(↓)	FM(↓)	WR(†)
Geometrical R3B Clustering	60.6	5.3	25.2	8.9	80.4
Agglomerative Clustering	62.8	<b>3.3</b>	32.0	1.9	84.1
Edge Clustering (no time)	63.4±0.3	7.2±0.3	24.8±0.7	4.6±0.1	82.4±0.1
Edge Clustering (with time)	74.7±0.5	3.4±0.6	20.5±1.3	1.4±0.1	89.2±0.1
R3B + Edge (no time)	67.4±0.3	8.5±0.3	16.0±0.4	8.0±0.3	82.2±0.1
Agglo + Edge (with time)	<b>81.3±0.3</b>	5.1±0.0	<b>12.2±0.3</b>	<b>1.5±0.1</b>	<b>91.0±0.1</b>

Table 1: Summary of performance metrics as defined in Subsection 2.4, evaluated for the different clustering algorithms. The models *Geometrical R3B Clustering*, *Edge Clustering (no time)*, and *R3B + Edge (no time)* utilize only angular and energy information on a per-hit basis for cluster reconstruction. In contrast, *Agglomerative Clustering*, *Edge Clustering (with time)*, and *Agglo + Edge (with time)* additionally incorporate time-of-hit information into the clustering process. Uncertainties reported for the four edge detection neural network variants correspond to the standard deviation of the results obtained from ten independent training runs.

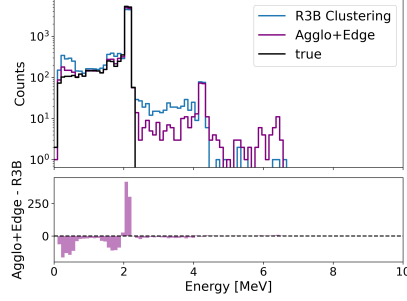


Figure 3: Reconstructed gamma cluster energy spectrum from simulated events, each consisting of three 2.1 MeV photons emitted from the target point. This showcase can be regarded as a “*worst-case*” scenario, since at this energy the event topology is strongly dominated by Compton scattering, leading to a comparatively broad spatial distribution of the energy deposits, as depicted in Fig. 2. The upper histogram shows the reconstructed cluster energy distribution using the geometrical R3B clustering (blue), the Agglo+Edge (pink) method accordingly, and in black the true energy cluster distribution. The lower panel displays the bin-by-bin count difference between the two approaches. The Agglo+Edge model demonstrates a significant improvement by successfully reattaching escaped hits, notably in cases where sparse energy deposits around 1.6 MeV and 0.5 MeV result from pair production and subsequent annihilation processes of the original gamma photons. This clean-up step leads to a marked reduction in false negatives (i.e. reduction of bin counts at 0.5 and 1.6 MeV) compared to the geometrical R3B clustering and an enhancement of 2.1 MeV peak.

238 The results of this study are summarized in Table 1, organized according to  
 239 increasing levels of reconstruction complexity. For completeness, the previously  
 240 obtained results from the comparison between the “baseline” geometrical R3B  
 241 clustering algorithm and the agglomerative model are also included.  
 242 The agglomerative model shows improved performance over the R3B baseline  
 243 in terms of both event-level true positives (TP) and cluster-level (WR) val-  
 244 ues. However, it exhibits inferior performance with respect to the false negative  
 245 (FN) rate, indicating a tendency to miss relevant hits during reconstruction.  
 246 This limitation motivated the development of an Edge Detection Neural Net-  
 247 work, initially evaluated as a standalone clustering algorithm and subsequently  
 248 integrated into the agglomerative framework, yielding the combined model de-  
 249 noted as *Agglo + Edge*.

250 The *Agglo + Edge* model demonstrates superior performance across all evalu-  
 251 ated metrics, achieving an overall correct reconstruction rate of 81.3%, signifi-  
 252 cantly outperforming the *Geometrical R3B Clustering* algorithm, which reaches  
 253 60.6%.

254 A visual representation of an example event, contrasting the incorrectly merged  
 255 hits from the geometrical R3B clustering with the correctly reconstructed clus-  
 256 tering using the Agglo+Edge model, is shown in Fig. 4.

257

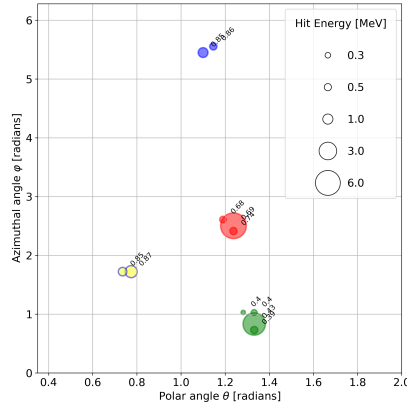


Figure 4: Example of a simulated event involving three primary photons, illustrating the performance difference between the Agglo+Edge clustering method and the geometrical R3B clustering approach. Each marker represents a detected hit, plotted as a function of the polar angle  $\theta$  and the azimuthal angle  $\varphi$ . The edge color of each circle indicates the true cluster assignment (ground truth), while the fill color denotes the cluster assignment according to the geometrical R3B clustering. The size of each circle reflects the energy deposited in the detector segment. Numbers adjacent to the hits represent the normalized hit times. In this event, the geometrical R3B clustering incorrectly assigns the two hits at  $(\theta \approx 0.8 \text{ rad}, \varphi \approx 1.8 \text{ rad})$ , with normalized times of 0.85 and 0.87 (blue edge, yellow fill) respectively, to a separate cluster, resulting in a *False Negative* (FN). In contrast, the Agglo+Edge method correctly assigns all hits to their respective clusters.

258 To further explore the capabilities of neural network-based clustering ap-  
 259 proaches, two additional models were evaluated: a standalone *Edge Detection*  
 260 *Neural Network* and a hybrid approach combining *Geometrical R3B Clustering*  
 261 with edge-based postprocessing (*R3B + Edge*). Notably, both of these mod-

262 els operate without incorporating time-of-hit information, similar to the R3B  
 263 baseline. Nonetheless, both outperform the *Geometrical R3B Clustering*, under-  
 264 scoring the potential of edge-based neural network models for improving cluster  
 265 reconstruction in high-granularity detector systems.  
 266 The incorporation of pre-clustering step in both Agglo+Edge and R3B+Edge  
 267 modes acts as a clean-up stage, reducing false negatives and enabling the Edge  
 268 model to specialize more effectively in merging decisions.  
 269 The edge detection NNs presented here represent a special case of Graph Neural  
 270 Networks (GNNs) [22], which, along with the more sophisticated transformer  
 271 models [23, 24], have seen widespread adoption in particle physics over the past  
 272 five years [25–27]. Interestingly, for this application, using an unsupervised  
 273 learning algorithm (agglomerative clustering) to first define a graph structure  
 274 presented a powerful inductive bias for our application which much improved  
 275 our results over the standalone edge-NN.

#### 276 4. Outlook

277 The results presented in the previous section clearly demonstrate that high-  
 278 level machine learning approaches, such as the Edge Detection NN, can sig-  
 279 nificantly enhance the accuracy of cluster reconstruction. These models not  
 280 only reduce distortions in the measurement process but also exhibit increased  
 281 sensitivity to low-statistics reactions – an important feature for experiments  
 282 targeting rare processes.

283 It is noteworthy that even the models, which do not utilize time-of-hit infor-  
 284 mation (similarly to the *Geometrical R3B Clustering*), outperform the baseline  
 285 method. This underscores the general effectiveness of neural network-based  
 286 methods in extracting structural features from detector data.

287 The inclusion of time-of-hit information proves to be a critical factor for enhanc-  
 288 ing clustering performance. As this observable is typically available for CALIFA  
 289 at R3B, the results of this study support the recommendation to incorporate it  
 290 into the reconstruction pipeline wherever possible.

Furthermore, these findings are intended to encourage broader adoption of advanced machine learning techniques by experimental groups, particularly in setups involving highly granular detectors. Such tools offer substantial performance benefits and can support more precise event reconstruction. One inherent limitation of the applied approach is its inability to correct for overly aggressive pre-clustering. In particular, false positive assignments introduced during the initial stage cannot be mitigated during the subsequent clean-up step by the edge-NN. This limitation is visible in Fig. 3, where a slight excess of reconstructed counts at  $E_{reco} \approx 6.3$  MeV is observed, likely indicating erroneous merging of unrelated hits due to excessive clustering. Despite this artifact, the high false negative rate – exceeding the false positive rate by more than a factor of five in the baseline R3B clustering (see Table 1) – motivated the development of a clustering strategy that prioritizes the recombination of hits to form complete clusters. To further validate the cluster reconstruction models presented here, conceptual methods are envisioned for application to source calibration data. Subsequent work could consider also adding a subsequent cluster splitting step in an end-to-end optimizable algorithm. Although, in principle, transformers could learn the graph structure directly from hit distributions, initial tests showed limited performance, highlighting an opportunity for the community to further develop combined machine learning-based reconstruction methods. From a computational standpoint, both the geometrical R3B clustering and the agglomerative clustering algorithms scale quadratically with the number of input hits, exhibiting a time complexity of  $\mathcal{O}(N^2)$ , where  $N$  denotes the number of detector hits per event. The combined methods –R3B + Edge and Agglo + Edge – induce additional computational overhead due to the Edge Detection Neural Network (NN) employed in the second stage. The current network architecture comprises three fully connected hidden layers with up to  $10^3$  neurons each, resulting in large matrix operations that dominate the runtime for typical events with  $N \sim \mathcal{O}(10^2)$ . Consequently, future work will focus on optimizing the Edge Detection NN by significantly reducing the model size to enable faster

322 execution while improving performance compared to the conventional geomet-  
323 rical R3B clustering.

324 Additionally, transformer-based models [23] – capable of analyzing full event  
325 topologies – may offer further improvements in clustering accuracy by captur-  
326 ing complex, global features.

327 The methods developed in this work can be directly integrated into the R3B  
328 data analysis chain as analysis tasks within the R3BRoot framework [6]. Their  
329 application is of particular relevance for heavy-ion experiments where the signal  
330 reconstruction is challenged by large background contributions, the production  
331 of  $\delta$ -electrons, and the simultaneous emission of a large number of neutrons and  
332 gamma-rays. In such scenarios, the improved reconstruction performance is ex-  
333 pected to enhance the precision of invariant-mass spectroscopy and kinematical  
334 reconstruction, thereby contributing directly to the scientific output of the R3B  
335 program.

### 336 **Acknowledgements**

337 The work was supported by BMBF 05P24WO2 and Excellence Cluster ORI-  
338 GINS from the DFG (Excellence Strategy EXC-2094-390783311). It was made  
339 possible through the close collaboration of experts from different disciplines  
340 within the Cluster of Excellence ORIGINS [28].

### 341 **Declaration of generative AI and AI-assisted technologies in the writ-** 342 **ing process.**

343 During the preparation of this work the authors used AI-assisted tools, in-  
344 cluding ChatGPT (OpenAI) and Gemini (Google), in order to improve the  
345 readability and language of the article. After using this tools, the authors re-  
346 viewed and edited the content as needed and take full responsibility for the  
347 content of the published article.

- 348 [1] N. Kalantar-Nayestanaki and C. Scheidenberger, Experiments at the In-  
349 terface of Nuclear, Atomic, and Hadron Physics with FRS at GSI and  
350 Super-FRS at FAIR, *Nuclear Physics News* 34 (2024) 21–26.
- 351 [2] Y. Leifels, Status and physics perspectives of FAIR, *Il Nuovo Cimento* 100  
352 (2025) 48.
- 353 [3] D. Cortina-Gil, H. Alvarez-Pol, T. Aumann et al., CALIFA, a Dedicated  
354 Calorimeter for the R3B/FAIR, *Nuclear Data Sheets* 120 (2014) 99–101.
- 355 [4] H. Alvarez-Pol, N. Ashwood, T. Aumann et al., Performance analysis for  
356 the CALIFA Barrel calorimeter of the R3B experiment, *Nuclear Instru-*  
357 *ments and Methods in Physics Research Section A: Accelerators, Spec-*  
358 *trometers, Detectors and Associated Equipment* 767 (2014) 453–466.
- 359 [5] T. Le Bleis, M. Bendel, R. Gernhäuser et al., A Digital Readout for CAL-  
360 IFA, [https://www.mll-muenchen.de/forschung/instrumentierung/](https://www.mll-muenchen.de/forschung/instrumentierung/califa_14.pdf)  
361 [califa\\_14.pdf](https://www.mll-muenchen.de/forschung/instrumentierung/califa_14.pdf), 2014. Accessed: 2025-04-06.
- 362 [6] D. Bertini, R3BRoot, simulation and analysis framework for the R3B  
363 experiment at FAIR, in: *Journal of Physics: Conference Series*, volume  
364 331, IOP Publishing, 2011, p. 032036.
- 365 [7] F. Nielsen, *Hierarchical Clustering*, Springer International Publishing,  
366 Cham, 2016, pp. 195–211.
- 367 [8] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu et al., Multilayer percep-  
368 tron and neural networks, *WSEAS Transactions on Circuits and Systems*  
369 8 (2009) 579–588.
- 370 [9] P. Cabanelas, H. Alvarez-Pol, J. Boillos et al., Commissioning of the califa  
371 barrel calorimeter of the r3b experiment at fair, in: *Journal of Physics:*  
372 *Conference Series*, volume 1667, IOP Publishing, 2020, p. 012006.
- 373 [10] H. Kolanoski and N. Wermes, *Teilchendetektoren*, Springer, 2016.



- [11] S. Seltzer, XCOM-Photon Cross Sections Database, NIST Standard Reference Database 8, <http://www.nist.gov/pml/data/xcom/index.cfm>, 2010. Accessed: 2025-04-10.
- [12] O. Klein and Y. Nishina, Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac, *Zeitschrift für Physik* 52 (1929) 853–868.
- [13] M. Lipiński, T. Włostowski, J. Serrano et al., White rabbit: A PTP application for robust sub-nanosecond synchronization, in: 2011 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication, IEEE, 2011, pp. 25–30.
- [14] S. Agostinelli, J. Allison, K. a. Amako et al., GEANT4a simulation toolkit, *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506 (2003) 250–303.
- [15] P. Virtanen, R. Gommers, T. E. Oliphant et al., SciPy 1.0: fundamental algorithms for scientific computing in Python, *Nature methods* 17 (2020) 261–272.
- [16] F. Nielsen and F. Nielsen, Hierarchical clustering, *Introduction to HPC with MPI for Data Science* (2016) 195–211.
- [17] S. Imambi, K. B. Prakash and G. Kanagachidambaresan, Pytorch, *Programming with TensorFlow: solution for edge computing applications* (2021) 87–104.
- [18] A. F. Agarap, Deep learning using rectified linear units (relu), *arXiv preprint arXiv:1803.08375* (2018).
- [19] S. Mannor, D. Peleg and R. Rubinstein, The cross entropy method for classification, in: *Proceedings of the 22nd international conference on Machine learning*, 2005, pp. 561–568.

- 401 [20] P.-T. De Boer, D. P. Kroese, S. Mannor et al., A tutorial on the cross-  
402 entropy method, *Annals of operations research* 134 (2005) 19–67.
- 403 [21] D. Newton, R. Pasupathy and F. Yousefian, Recent trends in stochastic  
404 gradient descent for machine learning and Big Data, in: 2018 Winter  
405 Simulation Conference (WSC), IEEE, 2018, pp. 366–380.
- 406 [22] P. W. Battaglia, J. B. Hamrick, V. Bapst et al., Relational inductive biases,  
407 deep learning, and graph networks, *arXiv preprint arXiv:1806.01261* (2018).
- 408 [23] A. Vaswani, N. Shazeer, N. Parmar et al., Attention is all you need, *Ad-  
409 vances in neural information processing systems* 30 (2017).
- 410 [24] X. Amatriain, A. Sankar, J. Bing et al., Transformer models: an introduc-  
411 tion and catalog, *arXiv preprint arXiv:2302.07730* (2023).
- 412 [25] G. DeZoort, S. Thais, J. Duarte et al., Charged particle tracking via edge-  
413 classifying interaction networks, *Computing and Software for Big Science*  
414 5 (2021) 1–13.
- 415 [26] X. Ju, D. Murnane, P. Calafiura et al., Performance of a geometric deep  
416 learning pipeline for HL-LHC particle tracking, *The European Physical  
417 Journal C* 81 (2021) 1–14.
- 418 [27] S. Van Stroud, P. Duckett, M. Hart et al., Transformers for Charged  
419 Particle Track Reconstruction in High Energy Physics, *arXiv preprint  
420 arXiv:2411.07149* (2024).
- 421 [28] Cluster of Excellence ORIGINS, Cluster of excellence origins, [https://](https://www.origins-cluster.de)  
422 [www.origins-cluster.de](https://www.origins-cluster.de), 2025. Accessed: 2025-05-03.