



Exploring the Potential of New Silicon Sensor Technologies for ITk Layers Replacement Beyond Run 4

Samantha Khavai

*Supervisors: Pierfrancesco Butti, Valentina Cairo, Markus Elsing
CERN Summer Student 2025*

Abstract

The High-Luminosity Large Hadron Collider (HL-LHC) will operate under extreme conditions of event pile-up beyond Run 4 phase, presenting significant challenges for the ATLAS detector and necessitating the development of advanced trigger-level techniques for early b-jet identification. This work investigates a hit-based b-tagging approach that assumes new technology of silicon sensor technologies in the upgraded Inner Tracker (ITk) can provide timing information at early trigger stage. Using only hit data from the first two ITk pixel layers, the study analyses spatial and angular hit distributions, applying vertex reconstruction techniques to enhance alignment with the jet axis. The prepared datasets will serve as input for training a Graph Neural Network (GNN), aimed at delivering a machine-learning solution capable of operating within the demanding constraints of real-time data processing beyond Run 4.

Contents

1	Introduction	2
2	Hit Extraction and Jet Matching	2
3	Analysis of ΔR Distributions for b-jets and Light Jets	3
3.1	Reconstructing the True Origin of the Jet	4
3.2	Hit Position Calculation	5
4	Corrected Distributions After Vertex Reconstruction	5
5	Training the Graph Neural Network	7
5.1	Training Setup	7
5.2	Monitoring and Results	7
6	Conclusion	8
7	Recommendations and Future Work	8
8	Acknowledgments	8

1 Introduction

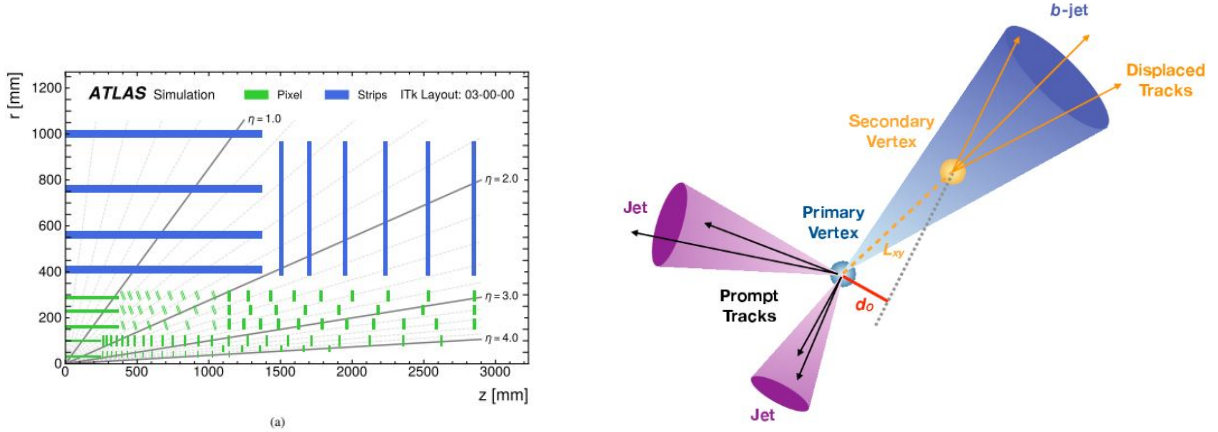
The High-Luminosity Large Hadron Collider (HL-LHC) expects approximately 200 proton–proton collisions per event in the forthcoming phases. This study takes two assumptions. One that with new technology advancements in silicon sensor technology, the ATLAS Inner Tracker (ITk) might be able to provide precise timing information directly at the trigger level and the second assumption that the ITk provides improved readout capability at rates above 1 MHz. With these assumptions, the study explores the possibility of incorporating hit-level timing data from the first two ITk pixel layers into the trigger system, allowing early b -tagging before full track reconstruction.

Building on concepts explored in ATL-PHYS-PUB-2023-023 and the Next Generation Trigger (NGT) study, this work focuses on “trackless” b -tagging using only hit information from ITk Layer 0 and Layer 1 at the earliest possible stage. This analysis investigates η – ϕ spatial patterns and ΔR angular distributions of hits, applying geometric corrections to recover the true jet origin along the beamline and improve hit alignment with the jet axis.

The end goal is to prepare corrected hit-level datasets for training and evaluating a Graph Neural Network (GNN), aiming to deliver a machine-learning-based b -tagging solution that is fast, robust, and suitable for integration into the ATLAS trigger system beyond Run 4.

2 Hit Extraction and Jet Matching

A jet is defined as a collimated spray of particles originating from a high-energy quark or gluon. In the context of the ATLAS Inner Tracker (ITk), hit-level information from the innermost pixel layers (Layer 0 and Layer 1) can provide the earliest signature of jet activity.



(a) ATLAS Inner Tracker layout in the r – z plane highlighting Layer 0 and Layer 1.

(b) Jet structure illustrating differences between light jets and b -jets.

Figure 1: Subfigure (a) shows the ITk layout in the r – z plane, while subfigure (b) illustrates the structure of light jets compared to b -jets.

Hit extraction was implemented from the first two ITk pixel layers, these layers can be clearly shown in figure 1(a), with Layer 0 located at a radius of about 34 mm from the

beamline and Layer 1 at about 100 mm. These layers provide the finest granularity and closest proximity to the interaction point, making them ideal for studying jet signatures before full track reconstruction.

In this study, jets were reconstructed using the anti- k_t algorithm with radius parameter $R = 0.4$ on calorimeter cells, and their directions were obtained by projecting back to the reference frame origin $(0, 0, 0)$. For each jet, hits were matched based on their proximity in the η - ϕ plane using the distance measure ΔR , defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (1)$$

where $\Delta\eta$ is the difference in pseudorapidity and $\Delta\phi$ is the difference in azimuthal angle between the hit and the jet axis. The pseudorapidity η is defined in terms of the polar angle θ (measured from the beamline) as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (2)$$

The quantity ΔR thus represents the angular separation in detector coordinates.

A matching requirement of $\Delta R < 0.4$ was imposed, which corresponds to the typical jet cone size used in ATLAS analyses. This criterion ensures that only hits lying within a narrow cone around the jet axis are considered, thereby suppressing contributions from the underlying event and pile-up.

The pseudorapidity range was restricted to $|\eta| < 1$, to understand hit-to-jet behavior. A schematic representation of the ΔR matching cone is shown in Figure 2 below.

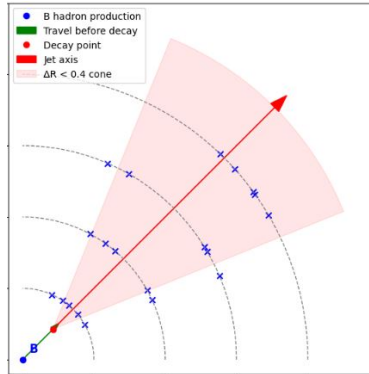


Figure 2: A schematic representation of the ΔR matching cone

3 Analysis of ΔR Distributions for b -jets and Light Jets

The initiating b -quark for b -jets hadronizes into a heavy b -hadron that carries a large fraction of the quark momentum. When the b -hadron decays, it produces displaced particles originating from a secondary vertex, but the decay products remain mostly aligned with the jet axis. In contrast, light jets, initiated by light quarks (u, d, s), decay promptly from the beamline resulting in wider angular spread of the hits on the innermost layers.

As ΔR increases, the distance from the jet axis grows, so the number of hits associated with the jet is expected to drop. Most jet constituents are concentrated near the core, and only a few particles extend outward. Therefore, both b -jets and light jets are expected to show a decreasing trend in their ΔR distributions as we move towards larger ΔR values.

However, the plotted ΔR distributions for both b -jets and light jets (Figure 3) show an unexpected rising trend at larger ΔR values, deviating from the expected falloff.

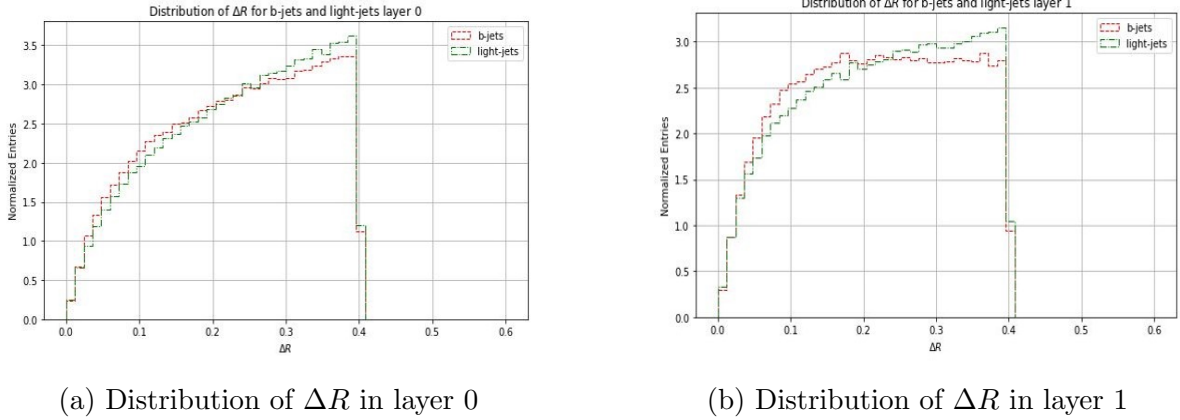


Figure 3: Subfigure (a) shows the Distribution of ΔR in layer 0, while subfigure (b) illustrates Distribution of ΔR in layer 1.

This anomaly indicates that something went wrong during data preparation: the analysis initially assumed that all jets originated from the reference frame origin $(0, 0, 0)$ in the center of the detector, which is not generally true. To correct this, the origin of each jet along the beamline was determined, and hit association was recomputed relative to this vertex. Applying this correction is expected to restore the physically expected ΔR behavior, with the number of associated hits decreasing as the distance from the jet axis increases.

3.1 Reconstructing the True Origin of the Jet

In the initial analysis, the pseudorapidity (η) of hits was computed assuming that all jets originated from the same nominal vertex along the beamline. This assumption led to systematic misalignment in the η values of the hits because, in reality, each jet originates from a different position along the z -axis.

The correction method involves finding the point of closest approach between the jet axis and the beamline. Specifically, two 3D points along the jet axis are used:

- $P_1 = (x_1, y_1, z_1)$, the mean position of hits in Layer 1 associated to the jet.
- $P_2 = (x_2, y_2, z_2)$, the mean position of calorimeter clusters along the jet direction.

The line connecting P_1 and P_2 defines the jet axis in 3D space. The beamline is assumed to be the z -axis passing through the origin $(0, 0, 0)$. The direction vector of the line is

$$\mathbf{d} = (dx, dy, dz) = (x_2 - x_1, y_2 - y_1, z_2 - z_1),$$

and the scalar parameter t_{min} at which the line comes closest to the z -axis is

$$t_{min} = -\frac{x_1 dx + y_1 dy}{dx^2 + dy^2}.$$

The corresponding z -coordinate of the closest point, which is taken as the reconstructed jet vertex, is then

$$z_{vtx} = z_1 + t_{min} dz.$$

Once the true z -vertex is determined, the hit positions are recalculated relative to this point.

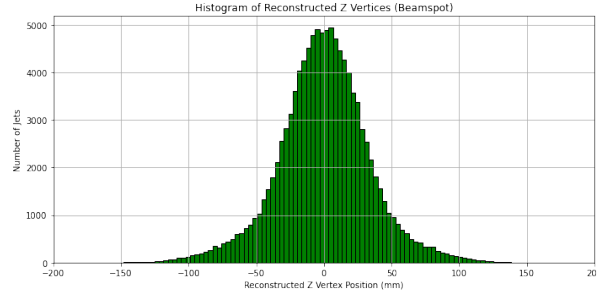


Figure 4: constructed z -vertex distributions

Figure 4 shows the reconstructed z -vertices distribution for a sample of jets. It is evident that jets originate from different points along the beamline. It also closely resembles the luminous region, which exhibit a longitudinal spread of approximately 50 mm.

3.2 Hit Position Calculation

To illustrate how hit positions are recalculated relative to the reconstructed z -vertex, consider a hit in Layer 0 with coordinates $(x_{hit}, y_{hit}, z_{hit})$ and a jet originating from a reconstructed vertex z_{vtx} .

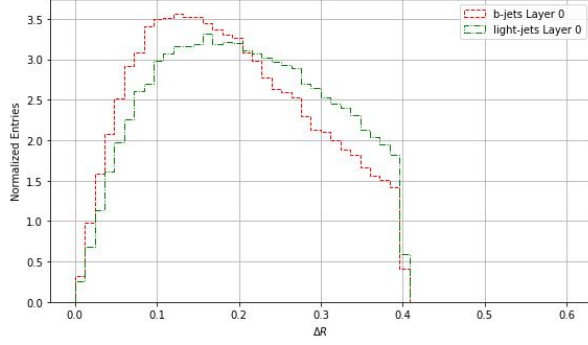
After vertex correction: The hit's z -coordinate is recalculated relative to the true jet vertex:

$$z_{hit}^{corr} = z_{hit} - z_{vtx}, \quad \theta_{hit}^{corr} = \arctan \frac{\sqrt{x_{hit}^2 + y_{hit}^2}}{z_{hit}^{corr}}, \quad \eta_{hit}^{corr} = -\ln \left(\tan \frac{\theta_{hit}^{corr}}{2} \right). \quad (3)$$

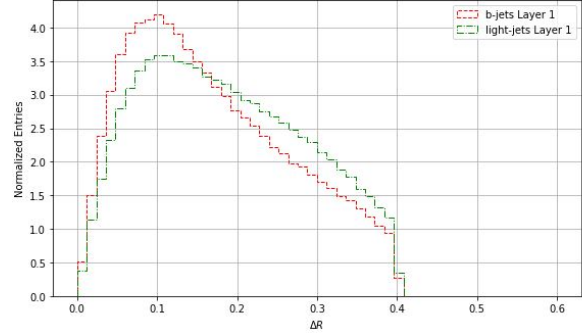
This shows how the pseudorapidity changes significantly when the hit position is corrected, which in turn affects the ΔR calculation and aligns the hits properly with the jet axis.

4 Corrected Distributions After Vertex Reconstruction

After applying this correction, the angular coordinates and ΔR distributions are consistent with physical expectations, as shown in Figure 5, thereby resolving the anomalous rising



(a) Distribution of ΔR in layer 0



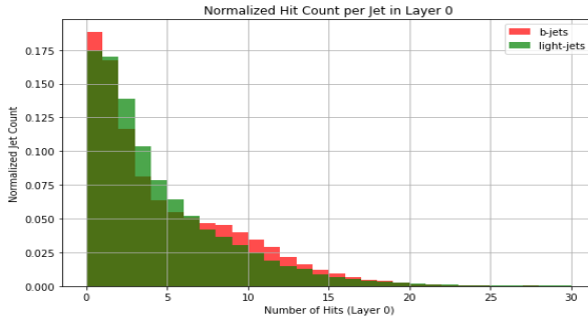
(b) Distribution of ΔR in layer 1

Figure 5: Subfigure (a) shows the Distribution of ΔR in layer 0, while subfigure (b) illustrates Distribution of ΔR in layer 1 after z- vertex correction

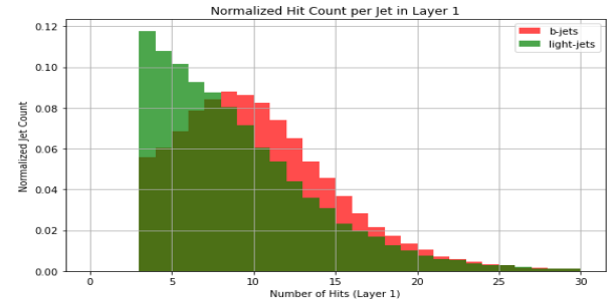
trends observed previously. Since the correction depends on the reconstructed vertex position, which itself is refined by the corrected hit information, the procedure needs to be iterated until convergence is achieved.

After applying this correction, the angular coordinates and ΔR distributions are consistent with physical expectations as shown in figure 5, resolving the anomalous rising trends observed previously.

The corrected hit distributions reveal the detector response across different layers. In layer 0, the innermost layer, a noticeable number of jets register zero hits. This occurs because some hits were lost during data preparation, reducing the apparent occupancy in this layer. In contrast, layer 1 exhibits higher hit counts for many b-jets, reflecting the increased interception of decay products in the second layer.



(a) Distribution of hits in layer 0



(b) Distribution of hits in layer 1

Figure 6: Subfigure (a) shows the Distribution of ΔR in layer 0, while subfigure (b) illustrates Distribution of hits in layer 1 after z- vertex correction

These distributions align with the expected detector geometry and particle propagation from the jet, confirming that the vertex correction improves both the angular alignment with the jet axis and the physical realism of hit patterns across layers.

5 Training the Graph Neural Network

The training of the GNN model was carried out on the Next Generation Trigger system, find documentation here ngt.docs.cern.ch/getting-started, which is equipped with an NVIDIA cluster comprising 100 GPUs with the end goal of studying possibility of early b-jet tagging by distinguishing between light-flavor, charm, and bottom jets at the hardware trigger level (L0), prior to the event filter. The model used per-hit spatial information from the tracking detectors (global x, y, z) combined with jet-level kinematic properties (p_T, η) to learn patterns associated with different jet flavors. Each jet in the dataset was associated with its flavor label, kinematic properties, and per-hit coordinates. The dataset was split into training, validation, and test subsets, with features normalized using a pre-computed dictionary to stabilize training.

5.1 Training Setup

Training was performed using PyTorch Lightning on a single GPU. Mixed-precision (16-bit) training was used to improve efficiency. The AdamW optimizer was employed with a one-cycle learning rate schedule, starting at 10^{-4} , peaking at 10^{-3} , and decreasing to 10^{-7} . A batch size of 128 was used, and the maximum number of epochs was set to 10. Categorical cross-entropy loss was used for classification, with equal weighting for all three classes. The preprocessing code is available at github.com/umami-hep/umami-preprocessing.

5.2 Monitoring and Results

The experiment was tracked using Comet ML. Training and validation losses were monitored at regular intervals, and loss curves were produced to assess convergence and model behavior. Due to the exploratory nature of this study, it was not possible to produce ROC curves at this stage. Initial attempts to save losses for post-training evaluation produced zero values, likely due to debugging and experimental setup issues. Therefore, the loss curves represent the main metric available for training assessment.

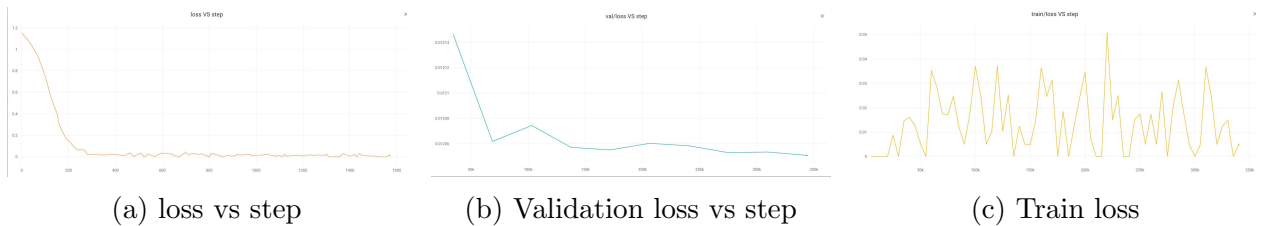


Figure 7: (a) Train loss vs step, (b) Validation loss vs step, (c) additional loss plot or comparison.

The flat training loss indicates a potential issue with the dataset or the model, likely related to the ΔR correction. The lack of convergence suggests that the network is not yet extracting the correct weight parameters. The validation loss, shown in Figure 7(b), decreases initially but flattens rapidly, which is an unexpected behaviour. The fluctuations observed in the training loss, Figure 7(c), arise from mini-batching, whereas the validation

loss is computed over the full dataset and therefore exhibits less noise. Figure 7(a) shows the loss after debugging by including the truth labels; the loss now decreases smoothly from the start, providing a positive indication of proper network training.

6 Conclusion

This study represents possibility of using only per hit information of the two innermost tracker layers for trackless b-tagging. The end focus was to establish a stable training workflow and understanding model behavior. The loss curves serve as the primary diagnostic tool for model convergence and highlight the potential of this approach. Future work will include debugging checkpoint saving to enable ROC-based evaluation and more extensive hyperparameter tuning.

7 Recommendations and Future Work

Based on the current results, the following recommendations are proposed to improve model performance and dataset quality:

- **Iterative ΔR correction:** Since the angular coordinates depend on the reconstructed vertex, the ΔR correction procedure should be iterated until convergence to ensure accurate hit positioning and reliable jet reconstruction.
- **Dataset refinement:** Investigate and clean the dataset to address any inconsistencies or missing information that may be affecting training, such as incomplete hit data or preprocessing errors.

8 Acknowledgments

I would like to express my sincere gratitude to CERN for offering me the opportunity to participate in the Summer Student Programme and for providing an inspiring environment to learn and grow as a young researcher.

I am deeply thankful to my supervisors, Pierfrancesco Butti, Valentina Cairo, and Markus Elsing, for their continuous guidance, support, and valuable insights throughout this project.

Above all, I thank God for granting me strength and perseverance during this journey, and I am profoundly grateful to my parents for their unconditional love, encouragement, and sacrifices that have made my academic pursuit possible.

References

- [1] ATLAS Collaboration, *Prospects for b-jet identification at the HL-LHC and beyond*, ATL-PHYS-PUB-2023-023, CERN, 2023. Available at: <https://cds.cern.ch/record/2851329>.

- [2] ATLAS Collaboration, *Next Generation Trigger: Exploring novel hardware and machine learning approaches for the HL-LHC*, ATL-DAQ-PUB-2024-001, CERN, 2024. Available at: <https://cds.cern.ch/record/2868152>.