Particle Flow Algorithm (PFA) for forward region jet reconstruction with the ATLAS Inner Traker (ITk) detector setup at the High Luminosity Large Hadron Collider (HL-LHC)

Thabo James Lepota[1], **Rachid Mazini**[1] **and Mukesh Kumar**[1]

[1]School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, 2050, South Africa

E-mail: 568571@students.wits.ac.za

Abstract. The ATLAS experiment at the High Luminosity Large Hadron Collider (HL-LHC) will require advanced reconstruction techniques, particularly in the forward region, to cope with increased pileup. This work presents a Particle Flow Algorithm (PFA) development for the ITk detector, focusing on tower clusters rather than traditional topological clusters in the ($0 \leqslant \eta > 1.5$) region. The forward region indicates ($2 \leqslant \eta > 4$). The strategy integrates tracker momentum measurements with calorimeter energy deposits through cell based subtraction, prioritising energy density layers to resolve overlaps between tracking and calorimetric data. By employing tower clusters, which aggregate calorimeter cells into fixed $\eta \times \phi$ grids, we aim to improve computational efficiency while maintaining spatial granularity critical for forward jet reconstruction. The framework processes Event Summary Data (ESD), containing raw detector level information (tracker hits, calorimeter clusters), and is processed into Analysis Object Data (AOD), a condensed format storing high level physics objects (jets, leptons) optimised for analysis. The algorithm refines energy subtraction and calibration by implementing Gaussian fitting of $\langle E/p \rangle$ distributions across calorimeter layers, mitigating pileup effects in the forward region. This approach addresses the high pileup HL-LHC environment, balancing precision in jet energy resolution with computational scalability for the ITk detector's upgraded granularity.

1 Introduction

Motivation: The HL-LHC will operate at an unprecedented instantaneous luminosity, leading to an average of $\mu = 200$ simultaneous proton-proton collisions per bunch crossing (pileup) [3]. This extreme environment poses significant challenges for particle reconstruction, particularly in the forward region ($\mid \eta \mid > 2.0$, where the calorimeter granularity is coarser and tracking information is traditionally limited. The ATLAS Phase II upgrades aim to address these challenges by introducing the Inner Traker (ITk) detector (see Figure 1), which extends tracking coverage up to ($\mid \eta \mid \approx 4$), and the HGTD, which provides $\sim$30 ps timing resolution to help distinguish between primary and pileup vertices [4]. The Particle Flow Algorithm (PFA) has demonstrated significant improvements in jet and missing transverse energy (MET) reconstruction in the central detector region by combining tracking and calorimeter information [5]. However, its application in the forward region has been limited due to the lack of high-granularity tracking and the dynamic clustering approach used in TopoClusters. With the improved capabilities of the ITk and HGTD, we propose an adapted PFA strategy that replaces TopoClusters with TopoTowers fixed size $\eta - \phi$ calorimeter towers to better match the forward detector geometry. This modification is expected

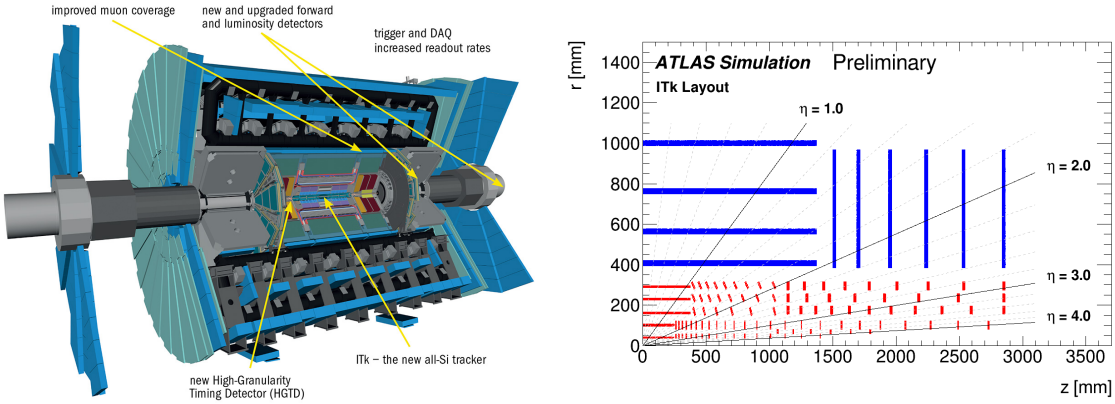

Figure 1: (*left*) Schematic of the important upgrades to the ATLAS detector for the HL-LHC era, showing the location of ITk detector [1]. (*right*) Schematic showing of the ITk layout for the HL-LHC phase of ATLAS. A zoomed view of pixel detector showing only one quadrant which is an active detector. Shown in blue are active elements of the strip detector and red shows the pixel detector. The horizontal axis lies along beam line with zero being the interaction point. The vertical axis is the radius measured from the interaction region [2].

to improve energy resolution and reduce algorithmic complexity in high pileup environments, ultimately improving physics performance in forward jet dominated processes such as vector boson fusion (VBF) and high energy Quantum Chromodynamics (QCD) studies [2]. Figure 2 shows how the PFA works, the key features includes data integration whereby it uses both calorimeter systems and tracking to optimise object reconstruction. Track energy (E_{track}) over momentum (p) matching to compute $^{E_{track}}/_{p}$ which accurately subtract calorimeter energy linked to charged tracks, avoiding double counting. Also, there is a split showering correction, to identify and merge split clusters to ensure correct energy association and prevent $^{E_{track}}/_{p}$ underestimation or misreconstruction.

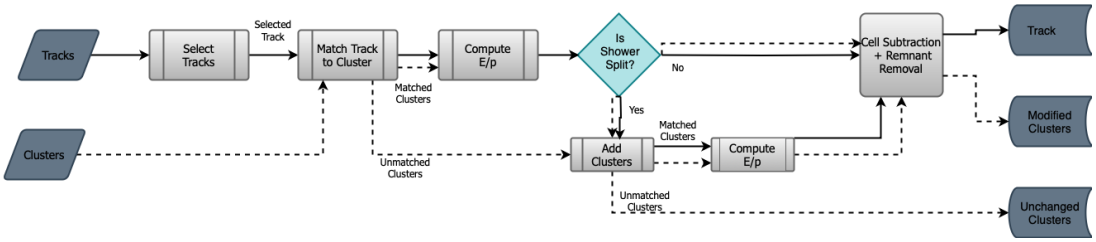

Figure 2: flowchart of the particle flow algorithm (PFA). Beginning with track selection and proceeding until the energy linked with the selected tracks has been removed from the calorimeter. At the end, charged particles, TopoClusters/TopoTowers which have not been modified by the algorithm, and remnants of TopoClusters/TopoTowers which have had part of their energy removed remain,in addition to clusters unchanged by the PFA.

2 Methodology

*Detector and data samples:*TopoClusters are dynamic, fine grained 3D clusters formed from calorimeter cells based on signal significance and spatial connectivity, optimised for precise shower reconstruction in central detector regions ($\mid \eta \mid < 2$). Whereas, TopoTowers are fixed size grids (0.1×0.1 in $\eta - \phi$) that combines energy deposits into coarse bins, prioritising computational efficiency and noise suppression in high-pileup forward regions ($\mid \eta \mid > 2$). While TopoClusters excel in granular tracking for isolated particles, TopoTowers simplify energy subtraction in coarse calorimeters [6]. Both are important to the PFA, in central regions precision is enabled by TopoClusters, while TopoTowers ensure robustness in forward reconstruction at the HL-LHC. Monte Carlo (MC) simulations of single pion events without pileup are used as input samples for calibration and validation studies. Roughly two thirds of visible energy is as a result of charged pions accounting for most of the jet's charged component [7, 8]. Reconstruction data are processed from Event Summary Data (ESD) which has full event details to Analysis

Object Data (AOD) which has physics optimised objects, where Topotowers are incorporated in forward η regions as shown in Figure 3.

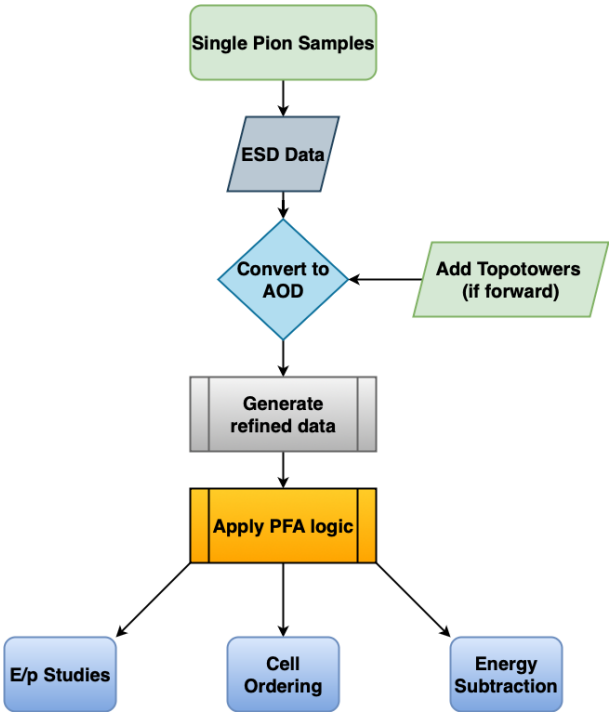

Figure 3: Framework workflow for the PFA analysis in the ATLAS forward region. The process begins with single pion samples which are processed as ESD data. This data is then converted to AOD format, where Topotowers are added for forward region reconstruction. Refined data are generated and the PFA logic is applied, building up to in three key analysis outputs: E/p Studies, Cell Ordering, and Energy Subtraction.

PFA adaptation for forward region: In the forward region TopoClusters are replaced by TopoTowers in the PFA, to better match the forward calorimeter segmentation and minimise computational complexity under extreme pileup. It is beneficial to focus on robustness then fine resolution when working under high pileup conditions ($\mu = 200$). E/p *Calibration Approach:* The E/p ratio, defined as calorimeter energy (E) within a $\Delta R < 0.2$ cone around a charged track normalised by track momentum p, is used to calibrate energy subtraction. Binning depends on: Track pseudorapidity, η in fine bins up to 4.0, p in bins from 0 to 100 GeV, Layer of Highest Energy Deposit (LHED) corresponding to calorimeter regions. Gaussian fits and histogram mode analyses of E/p distributions provide mean and resolution estimates for calibration. To calibrate the $^{E_{track}}/_{p}$ distributions, a fitting strategy is used, which involves using Gaussian fits within the predefined bin, where the fit range is set to surround the data between the mean plus or minus the root mean square (RMS) of the distribution. This approach aims to record the core behaviour of the E/p ratio while excluding the tails that may arise from background or pileup effects. To ensure reliability, the fit results undergo validation: if the absolute difference between the histogram mean and the Gaussian fit mean exceeds 0.1, an alternative metric is used for the calibration constant specifically, the histogram mode is selected if it is greater than 0.1; otherwise, the histogram mean is adopted. Fit quality is assessed using the chi squared per degree of freedom (χ^2/NDF), with acceptable thresholds set at less than 2.0 for central detector regions and less than 3.0 for the forward region, reflecting the different detector resolutions and granularities. Complementing this, energy density profiling is performed to guide shower subtraction by calculating the average energy density per concentric ring around the track axis, where each ring corresponds to a fixed $\Delta\eta \times \Delta\phi$ area of 0.05×0.05. The subtraction algorithm prioritises cells with the highest energy density first, ensuring that the core energy deposits belonging to charged particles are removed before peripheral and potentially pileup contaminated cells, thus improving subtraction accuracy in the challenging high pileup forward environment.

Cell ordering and energy subtraction strategy Energy subtraction is implemented cell by cell with priority given to cells of highest energy density within coaxial rings defined in $\eta - \phi$ space around the track. This ensures correct subtraction of core shower energy before peripheral deposits, particularly important in coarse forward calorimeters.

3 Results

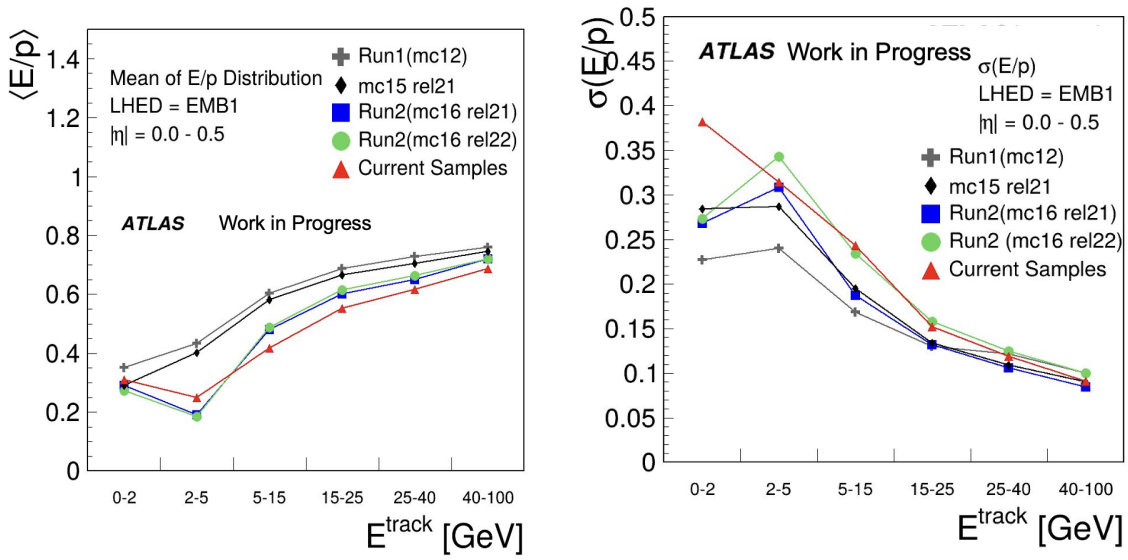

Figure 4: Central region results (LHED = EMB1) for $0 \leqslant |\ \eta\ | < 0.5$, showing (*left*) the mean E/p ratio rising and plateauing with increasing E_{track}. (*right*) the resolution $\sigma(E/p)$ improving with E_{track}. Data from Run 1, Run 2, and current MC samples are overlaid for comparison and validation of the particle flow calibration.

Central region validation: Figure 4 presents key central region results for the first electromagnetic barrel calorimeter layer (LHED = EMB1) within the pseudorapidity interval $0 \leqslant |\ \eta\ | < 0.5$. The left panel displays the mean E/p ratio, defined as the calorimeter energy measured in a cone of $\Delta R < 0.2$ around a charged particle track normalized by the track momentum (E_{track}). As observed, the mean E/p increases with rising track transverse momentum, reaching a plateau between roughly 0.6 and 0.7 at higher momenta (above $\sim$15 GeV). This behaviour reflects the calorimeter's intrinsic response to charged hadrons and accounts for energy losses in upstream material, validating the PFA's calibration approach in the central detector region. The right panel shows the resolution $\sigma(E/p)$, characterized by the standard deviation of the E/p distribution, which improves significantly with increasing E_{track}. At low transverse momenta (below 5 GeV), the resolution is comparatively poor (around 0.4–0.5), due to larger fluctuations in shower development and threshold effects. As E_{track} increases, the resolution stabilizes to values between approximately 0.15 and 0.3, indicative of better shower containment and detector response consistency. Overlaid data from Run 1, Run 2, and current MC samples exhibit close agreement, demonstrating robust detector performance and effective calibration of the PFA across multiple data taking periods. Figure 5 shows histograms of E/p distributions for charged tracks in different transverse momentum bins, with Gaussian fits applied over mean $\pm$ RMS. Fit parameters (mean, standard deviation, error) and quality (χ^2/NDF) are indicated. At low momenta (0–2 GeV), the Gaussian mean is around 0.3–0.4, reflecting energy loss and partial shower containment, increasing to about 0.6–0.7 at higher momenta, consistent with calorimeter response. Fit quality is generally good ($\chi^2/\text{NDF} < 2$), supporting the fit's robustness. In some low-momentum bins, discrepancies between Gaussian and histogram statistics trigger a validation step employing alternative metrics to ensure reliable calibration.

Cell ordering: The cell ordering plots shown in Figure 6 analyse the average energy density distribution in the ATLAS EMB1 calorimeter around charged particle tracks with $0 < |\eta_{\text{track}}| < 0.5$ and $5\,\text{GeV} < E_{\text{track}} < 15\,\text{GeV}$. By grouping calorimeter cells into concentric rings in η-ϕ space, the dependence of energy deposition on radial distance from the track axis is investigated. For a fixed ring radius of 1, the energy density distribution exhibits a broad spread (RMS = 0.635) with a mean of $\log_{10}(\langle E_{\text{density}}\rangle) = -1.44$, while at a larger radius of 10, the distribution narrows (RMS = 0.45) with a lower mean of -1.72, consistent with the expected shower profile. As you vary the ring radius there is a steep decline in energy density from $\sim 10^{-0.2}$ to $10^{-1.6}$ as the radius increases, well described by a two component fit.

Forward region: The performance of the forward region is characterised using distributions shown in Figure 7. Figure 7(*left*) shows a consistent response with mean $E/p = 0.54$) across η regions with 7790 forward entries, and a slight η- dependence with mean $\eta = -0.015$ The energy density profiles with 114223 simulated events

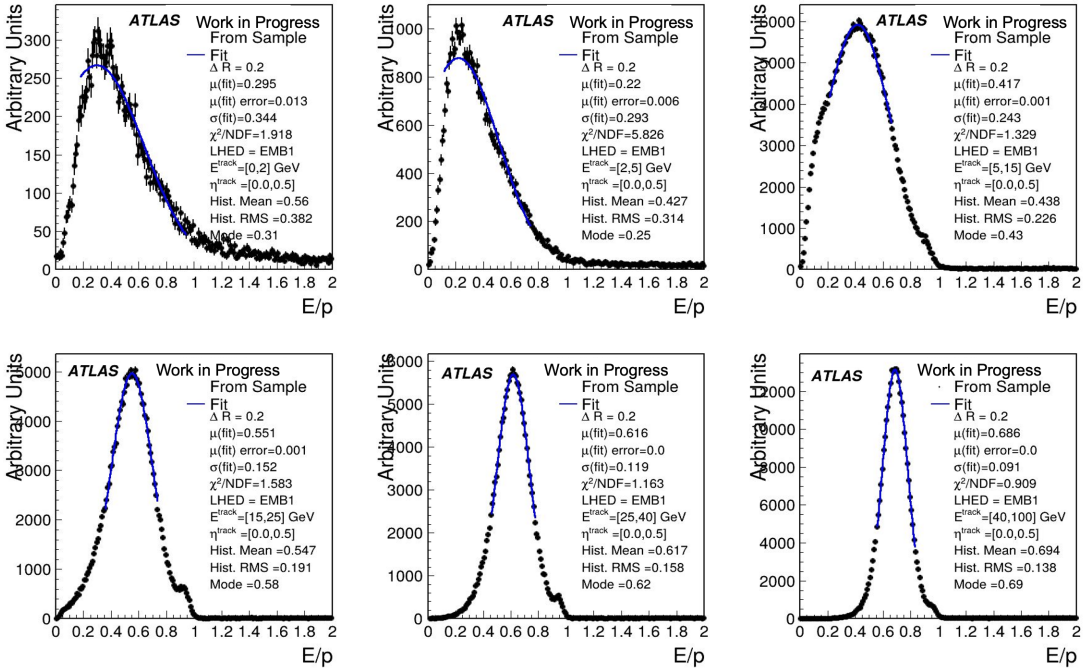

Figure 5: shows fits of Gaussian functions to the E/p distributions in several momentum bins for the EMB1 layer in the central region ($0 <\mid \eta_{track} \mid < 0.5$).

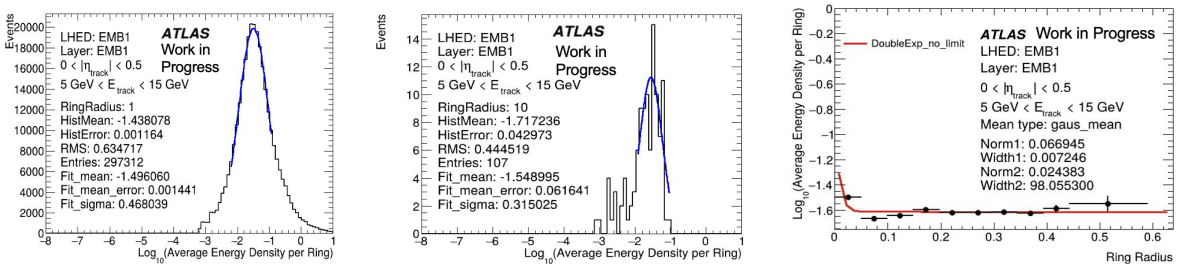

Figure 6: Distribution of average energy density per ring in the ATLAS EMB1 calorimeter for charged particle tracks with $0 <\mid \eta_{track} \mid < 0.5$ and 5 GeV$\mid \eta_{track} \mid$<15 GeV. (*left*) Energy density profile for a fixed ring radius of 1, showing a broad distribution with fit results overlaid. (*middle*) Distribution for a larger ring radius of 10, exhibiting a narrower spread. (*right*) Dependence of energy density on ring radius, fitted with a two component model. Larger radii sample the sparse,homogeneous lateral shower distribution, while smaller radii capture the dense shower core.

showing an asymmetric distribution with mean of $= 0.196 \pm 0.41\, GeV/mm^2$, where most deposits cluster at low energy while maintaining a significant high density tail (see Figure 7(*middle*)). Figure 7(*right*) shows a maps the Layer with Highest Energy Deposits (LHED) across the forward region layers showing the spatial distribution of the energy deposits in this region.

4 Summary and Outlook

Figures and references demonstrate the impact of coarse granularity: TopoTowers replacing spatial resolution for pileup robustness and computational simplicity, but calibration must compensate for the loss of detail. Detector layout and timing diagrams illustrate how ITk and HGTD create the first opportunity for forward region PFA by extending high precision tracking and timing into areas previously uninstrumented for such analyses. The PFA workflow in the forward region is now at an advanced stage: TopoTower clustering and PFA adaptation (cell/tower energy subtraction and E/p binning) are complete, and preliminary E/p distributions and energy density profiles

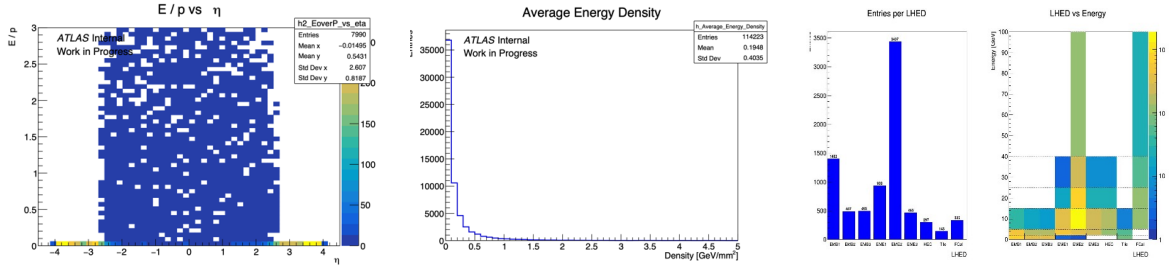

Figure 7: E/p distributions for the forward region ($\mid \eta \mid > 2$). (*left*) E/p as function η, with entries extended to 4, shows an average E/p of 0.54 with η-dependence across detector regions. (*middle*) Energy density distribution in forward calorimeter cells shows the asymmetric profile shows mainly low energy deposits with mean 0.19 GeV/mm^2)) with a high energy tail. (*right*) Distribution of Layer of Highest Energy Deposits (LHED) across forward calorimeter cells with entries showing that there is a presence of entries in the forward region. The mapping shows energy deposit patterns throughout the barrel electromagnetic calorimeter.

show stable, expected behaviour even under high pile-up conditions. Validation of the implemented methods using realistic physics samples (with jets and $Z \rightarrow ee$ samples), as well as ongoing tuning of E/p calibration and tower subtraction, are in progress. The remaining steps focus on refining subtraction parameters, broadening validation to complex event topologies, and exploring the use of machine learning for further improvement. This progress enables robust jet and MET reconstruction in the forward detector, a critical advancement for ATLAS operation at the HL-LHC.

References

[1] R. M. Bianchi and ATLAS Collaboration, "ATLAS experiment schematic or layout illustration," 2022. [Online]. Available: https://cds.cern.ch/record/2837191

[2] ATLAS Collaboration, "Expected Tracking Performance of the ATLAS Inner Tracker at the HL-LHC," CERN, Tech. Rep., 2019. [Online]. Available: https://cds.cern.ch/record/2669540

[3] A. Collaboration, "Technical Design Report for the ATLAS Phase-II Upgrade," CERN, Tech. Rep., 2017. [Online]. Available: https://cds.cern.ch/record/2285584

[4] ATLAS Collaboration, "Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade," CERN, Tech. Rep., 2020. [Online]. Available: https://cds.cern.ch/record/2719855

[5] A. Collaboration, "R=0.4 jets input comparison and monte carlo calibration with the atlas detector," CERN, Tech. Rep., 2022. [Online]. Available: https://cds.cern.ch/record/2824558

[6] G. Aad *et al.*, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J. C*, vol. 77, p. 490, 2017. [Online]. Available: https://cds.cern.ch/record/2138166

[7] I. G. Knowles and G. D. Lafferty, "Hadronization in decay," *J. Phys. G: Nucl. Part. Phys.*, vol. 23, no. 7, pp. 731–789, 1997. [Online]. Available: http://dx.doi.org/10.1088/0954-3899/23/7/003

[8] M. G. Green, *Electron-Positron Physics at the Z*, ser. Studies in High Energy Physics, Cosmology and Gravitation. IOP Publishing, 1998.