PROCEEDINGS OF SCIENCE

Constituent-based approaches to quark–gluon jet tagging for the High-Luminosity LHC

Florencia L. Castillo[a] **on behalf of the ATLAS collaboration**

[a] *Laboratoire d'Annecy de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France*

E-mail: fcastill@cern.ch

Jet constituents provide a finer description of the radiation pattern inside a jet than global observables. In simulations for ATLAS Run-2 data (2015-2018), transformer-based taggers trained on low-level inputs significantly outperformed traditional approaches using high-level variables with conventional neural networks for quark–gluon discrimination. As the High-Luminosity LHC era approaches, with increased luminosity and center-of-mass energy, the ATLAS detector will undergo thorough/major upgrades, including an extended inner tracker that enhances coverage in the forward region, previously uncovered by a tracking detector. This study evaluates how these improvements enhance the accuracy and robustness of jet taggers, which is essential for key analyses such as vector boson fusion, vector boson scattering, and supersymmetry searches, where precise jet identification is critical for suppressing backgrounds.

The European Physical Society Conference on High Energy Physics (EPS-HEP2025)
7-11 July 2025
Marseille, France

 ISSN 1824-8039 . Published by SISSA Medialab.

PoS(EPS-HEP2025)536

1. Introduction

Quark–gluon tagging is an important technique in high-energy physics, leveraging the distinct kinematic and substructure properties of jets initiated by quarks versus gluons. Quark-initiated jets are typically narrower, with higher transverse momentum (p_T) constituents and fewer particles due to the lower color factor of quarks ($C_F = 4/3$), resulting in reduced gluon radiation compared to gluons ($C_A = 3$). In contrast, gluon-initiated jets exhibit broader and softer radiation patterns with more diffuse substructures. This discrimination is critical for enhancing signal-to-background separation in numerous physics analyses at the Large Hadron Collider (LHC), including Vector Boson Fusion (VBF), Vector Boson Scattering (VBS), supersymmetry (SUSY) searches, and heavy resonance searches. The technique is particularly impactful in the forward region ($|y| > 2.5$), where jet identification faces challenges due to limited tracking coverage and higher particle densities, yet many new physics signatures produce high-rapidity jets [1].

The High-Luminosity LHC (HL-LHC), set to begin operations in 2030, will deliver an unprecedented integrated luminosity of up to 3000 fb^{-1} in proton–proton collisions. However, this increased luminosity introduces significant experimental challenges, with an average of 140 pile-up interactions per bunch crossing (peaking at 200), leading to dense event environments, higher detector occupancy, and increased background contamination. To address these conditions, the ATLAS experiment is undergoing major upgrades, notably the replacement of the current inner detector with an all-silicon Inner Tracker (ITk). The ITk, comprising pixel and strip silicon sensors, extends charged-particle tracking to pseudorapidities up to $|y| < 4.0$ from the current $|y| < 2.5$ [2].

This study investigates the impact of these HL-LHC upgrades on quark–gluon tagging performance when integrated with advanced machine learning architectures. We employ transformer-based models, the Particle Transformer (ParT) [3], which processes low-level jet information. We evaluate whether the quark–gluon discrimination efficiency observed in LHC Run-2 simulations (2015–2018) [4] persists under the challenging HL-LHC pile-up conditions, where increased event complexity may degrade jet substructure observables. Additionally, we quantify the benefits of incorporating forward-tracking information into the ParT model, comparing to two fully connected (FC) baselines: one that employs eight high-level jet variables optimized for jet characterization, and an FC-reduced version that emulates the ATLAS quark–gluon tagger using five high-level variables from Run 2 analyses [5].

2. Methodology

Taggers are trained on simulated VBF Higgs samples (Powheg[6]+Herwig7[7]) and dijet samples (Pythia8 [8]) under HL-LHC conditions. The simulations include an average pile-up of 140, introducing additional particles that complicate event reconstruction and increase background noise. Jets are reconstructed using the anti-k_t algorithm with a radius parameter $R = 0.4$. The inputs to the jets are Particle Flow Objects (PFOs), which combine calorimeter topo-clusters and matched tracks, these are referred to as the jet constituents. These elements provide a detailed, particle-level representation of the jet's internal structure, enabling advanced discrimination between different jet types.

To ensure balanced and effective training, the p_T spectrum is flattened during the training phase. This is achieved by reweighting events to produce a uniform distribution across p_T bins, preventing the model from being biased toward the more abundant low-p_T jets that dominate natural spectra due to QCD dynamics. For evaluation and performance assessment, uniform weights are applied to all events, ensuring unbiased metrics. Event selection focuses on the two leading (highest p_T) jets in each event, with each required to have $p_T > 20$ GeV to guarantee sufficient energy for reliable reconstruction and tagging. The tagger is defined in two regions: central ($|y| < 2.5$) and forward ($2.5 < |y| < 4.0$). The tagger definitions and input variables are summarized in Table 1.

Tagger	Description	Features
ParT	Processes up to 50 PFOs per jet, ordered by descending p_T. Concatenates topo-tower, track, and constituent inputs in the forward region for HL-LHC, extending ATLAS Run 2 configurations [4].	**Single-constituent**: Δy^a: Relative rapidity ($y^a - y^{\text{jet}}$) $\Delta\phi^a$: Azimuthal angle difference ($\phi^a - \phi^{\text{jet}}$) ΔR^a: $\sqrt{(\Delta y^a)^2 + (\Delta\phi^a)^2}$ Log terms: $\log p_T^a$, $\log E^a$, $\log(p_T^a/p_T^{\text{jet}})$, $\log(E^a/E^{\text{jet}})$ Mass: Constituent mass (m^a) **Pairwise**: ΔR_{ab}: Angular separation ($\sqrt{(y^a - y^b)^2 + (\phi^a - \phi^b)^2}$) m_{ab}^2: Invariant mass ($(p^{\mu,a} + p^{\mu,b})^2$) Lund variables: $k_T = \min(p_T^a, p_T^b) \cdot \Delta R_{ab}$, $z = \min(p_T^a, p_T^b)/(p_T^a + p_T^b)$
FC	Employs eight high-level jet variables for tagging, optimized for jet characterization.	p_T: Jet transverse momentum m: Jet mass EMFrac: Electromagnetic fraction Width: Jet width (from PFOs, charged PFOs, and tracks with $p_T > 1$ GeV) PFOs: Number of PFOs Charged PFOs: Number of charged PFOs ($p_T > 1$ GeV)
FC-reduced	Emulates ATLAS quark–gluon tagging with five high-level variables from Run 2 analyses [5].	p_T: Jet transverse momentum η: Pseudorapidity PFOs: Number of PFOs w^{PFO}: PFO width $C_1^{\beta=0.2}$: Two-point energy correlation

Table 1: Characteristics and features of taggers. Indices a, b denote different PFOs within a jet.

3. Results

The performance of the jet tagger is evaluated using the gluon-jet rejection, defined as $1/(1-\epsilon_g)$, where ϵ_g is the gluon efficiency, and measured at a fixed quark-jet efficiency of $\epsilon_q = 0.5$. This metric quantifies the tagger's ability to distinguish quark-initiated jets from gluon-initiated jets. The evaluation is shown as a function of jet rapidity ($|y|$), across low-p_T jets and high-p_T jets.

Additionally, the tagger's robustness against pile-up is assessed for pile-up levels of 60, 140, and 200 interactions.

The ParT tagger, which employs a transformer-based architecture to process detailed jet constituent-level inputs is compared against the FC tagger and a reduced version of the FC tagger (FC-reduced), which both of them rely on high-level jet features. In the central region, the ParT tagger outperforms the FC tagger across both low- and high-p_T ranges, as shown in Figure 1. At low p_T, the ParT tagger achieves approximately 10% higher gluon-jet rejection compared to the FC tagger, while at high p_T, the improvement reaches up to 25%. This enhanced performance is attributed to the ParT tagger's ability to capture complex, non-linear patterns in the jet substructure, which are more pronounced at higher transverse momenta due to the increased collimation of quark jets.

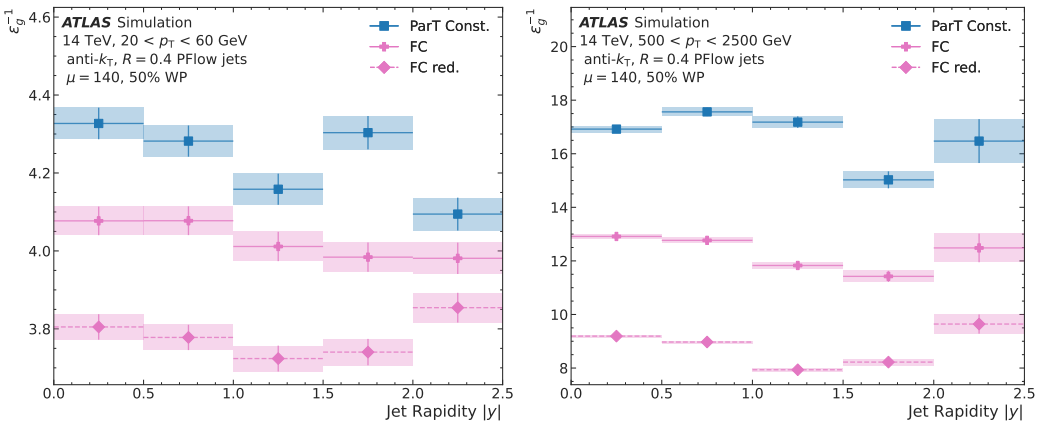

Figure 1: Gluon jet rejection (ϵ_g^{-1}) vs. jet rapidity ($|y|$) in the central region, across low- and high-p_T at $\epsilon_q = 0.5$. ParT (blue) and FC taggers (solid and dashed pink) are compared under HL-LHC conditions with pile-up 140. Error bars show statistical uncertainties [10].

In the forward region (Figures 2), the ParT tagger achieves 20–30% better gluon rejection than the FC and FC-reduced taggers by leveraging constituent, track, and topo-tower information. With only constituents (ParT Const.), performance is reduced because the drop in track efficiency decreases the number and quality of reconstructed tracks [9], limiting the effectiveness of track-dependent constituent features. Adding tracks and topo-tower inputs (ParT Const. + Tower + Track) improves performance, as the reconstructed tracks provide additional discriminating information not fully captured by the constituents, while the transformer architecture can handle missing track information. Including topo-tower inputs alone (ParT Const. + Tower) provides further complementary gains, enhancing overall discrimination in the forward region.

The robustness of the ParT tagger against pile-up is evaluated in the forward region under varying pile-up conditions of 60, 140, and 200 additional interactions per bunch crossing, as shown in Figure 3. Pile-up introduces additional particles from simultaneous proton-proton collisions, which can contaminate jet reconstruction and degrade tagger performance. The ParT tagger, using concatenated inputs from jet constituents, topo-towers, and tracks, demonstrates remarkable stability across these conditions. For both low- and high-p_T jets, the gluon-jet rejection exhibits minimal degradation, with performance losses of less than 5–10% at the highest pile-up level of 200.

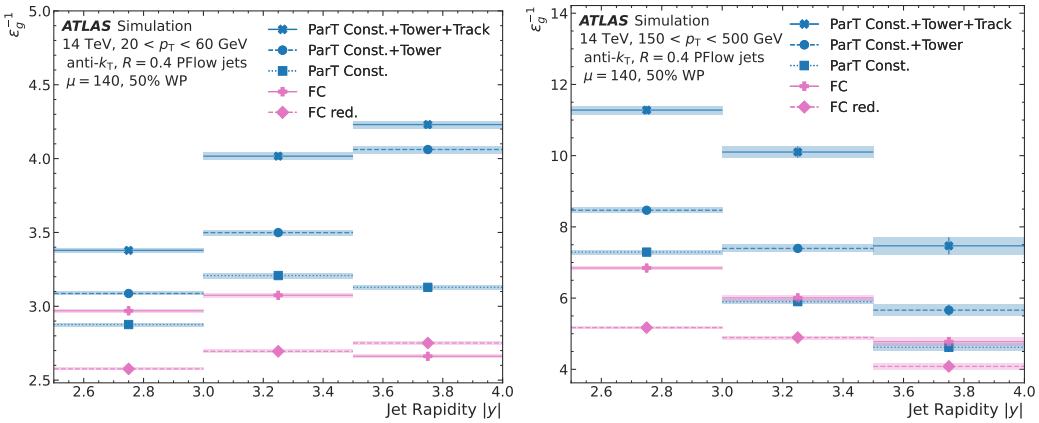

Figure 2: Gluon jet rejection (ϵ_g^{-1}) vs. jet rapidity ($|y|$) in the forward region, low-p_T and range, at $\epsilon_q = 0.5$. ParT (blue) and FC taggers (solid and dashed pink) are compared under HL-LHC conditions with pile-up 140. Error bars show statistical uncertainties [10]

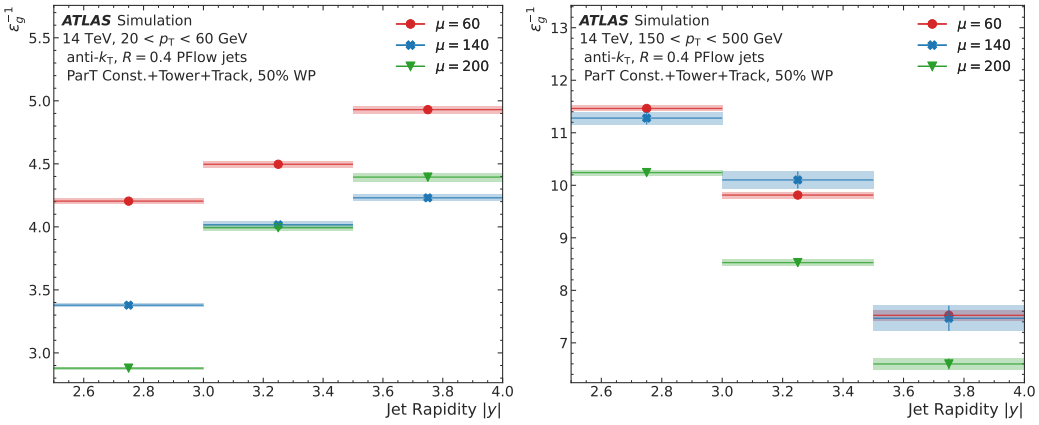

Figure 3: Gluon jet rejection (ϵ_g^{-1}) as a function of jet rapidity ($|y|$) in the central region for low- and high-p_T jets at $\epsilon_q = 0.5$. Results are shown for the ParT tagger using concatenated constituent, topo-tower, and tracking inputs, evaluated under pile-up conditions of 60 (red), 140 (blue), and 200 (green). Error bars indicate statistical uncertainties [10].

4. Conclusion

This study demonstrates that the transformer-based ParT tagger, leveraging low-level jet constituent information, substantially improves quark–gluon discrimination at the HL-LHC. It consistently outperforms the FC tagger, which relies on high-level jet features, across all detector regions and the full p_T spectrum. The inclusion of tracking information from the upgraded tracker further enhances performance in the forward region, achieving up to 30% higher gluon-jet rejection compared to FC taggers. The ParT tagger exhibits robust performance under high pile-up conditions, making it well-suited for the challenging HL-LHC environment.

Acknowledgements Supported by the Agence Nationale de la Recherche (ANR) under the program "Advanced Tracking Algorithms for Particle Physics (ATRAPP)" (ANR-21-CE31-0022) and CERN via the ATLAS Collaboration.

References

[1] ATLAS Collaboration, *Prospective study of vector boson scattering in WZ fully leptonic final state at HL-LHC*, ATLAS Note ATL-PHYS-PUB-2018-023 (2018), https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2018-023.

[2] ATLAS Collaboration, *Expected Tracking Performance of the ATLAS Inner Tracker at the High-Luminosity LHC*, JINST **20**, P02018 (2025), doi:10.1088/1748-0221/20/02/P02018.

[3] H. Qu, C. Li and S. Qian, *Particle Transformer for Jet Tagging*, (2022), arXiv:2202.03772 [hep-ph].

[4] ATLAS Collaboration, *Constituent-Based Quark Gluon Tagging using Transformers with the ATLAS detector*, ATLAS Note ATL-PHYS-PUB-2023-032 (2023), https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-032.

[5] ATLAS Collaboration, *Performance and calibration of quark/gluon-jet taggers using 140 fb^{-1} of pp collisions at 13 TeV with the ATLAS detector*, Chin. Phys. C **48**, 023001 (2024), doi:10.1088/1674-1137/acf701.

[6] P. Nason and C. Oleari, *NLO vector-boson fusion processes matched with shower in POWHEG*, JHEP **02**, 037 (2010), doi:10.1007/JHEP02(2010)037, arXiv:0911.5299 [hep-ph].

[7] J. Bellm *et al.*, *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76**, 196 (2016), doi:10.1140/epjc/s10052-016-4018-8, arXiv:1512.01178 [hep-ph].

[8] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015), doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012 [hep-ph].

[9] ATLAS Collaboration, *Expected tracking performance of the ATLAS Inner Tracker at the High-Luminosity LHC*, JINST **20**, P02018 (2025), doi:10.1088/1748-0221/20/02/P02018.

[10] ATLAS Collaboration, *JETM-2025-01: Quark-gluon tagging performance at the HL-LHC*, Public plots, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2025-01/, accessed 9 June 2025.

PoS(EPS-HEP2025)536