Improvements in the measurements of Higgs boson's Yukawa coupling to b- & c- quarks in ATLAS(*)

Giovanni Besio on behalf of the ATLAS Collaboration
INFN, Sezione di Genova - Genova, Italy
Dipartimento di Fisica, Università di Genova - Genova, Italy

Summary. — This proceeding details improvements for measuring the Higgs boson's coupling to b- and c-quarks at the ATLAS experiment, focusing on the associated production with a vector boson. The analysis relies largerly on the performance of the algorithms devoted to the identification of b- and c-jets, referred to as flavor tagging algorithms. A new transfomer model algorithm,GN2, has recently been employed by the ATLAS collaboration, significantly improving flavor tagging performance compared to previous approaches. Preliminary studies show a 11% to 16% improvement in the analysis sensitivity for the measurment of the $H \to b\bar{b}$ coupling and 45% to 56% for the search of the $H \to c\bar{c}$ coupling.

1. – Introduction

Since the Higgs boson's discovery in 2012, a primary goal at the Large Hadron Collider (LHC) has been to precisely measure its properties, especially the Yukawa couplings that generate fermion masses. Measuring these couplings is a critical test of the Standard Model (SM). While couplings to third-generation fermions have been observed, interactions with lighter quarks remain challenging. This work focuses on improving the measurements of the Higgs boson coupling to the bottom (b) and charm (c) quarks within the ATLAS experiment [2]. The analysis targets the Higgs boson production in association with a vector boson (W or Z). By selecting leptonic decays of W/Z boson, a clean signature can be achieved, significantly suppressing the large Quantum Chromodynamics (QCD) multi-jet background and enhancing the analysis sensitivity. The Higgs boson's primary decay is to a pair of b-quarks ($H \to b\bar{b}$, BR $\approx$ 58%), while the decay to c-quarks ($H \to c\bar{c}$, BR $\approx$ 3%) is more elusive. Recent full Run-2 results [1] already improved the $H \to b\bar{b}$ measurement and set the most stringent limits on $H \to c\bar{c}$ to date (Fig.1). This proceeding outlines strategies to further enhance these measurements for the ongoing Run 3 data-taking period.

(*) IFAE 2025 - "Poster" session

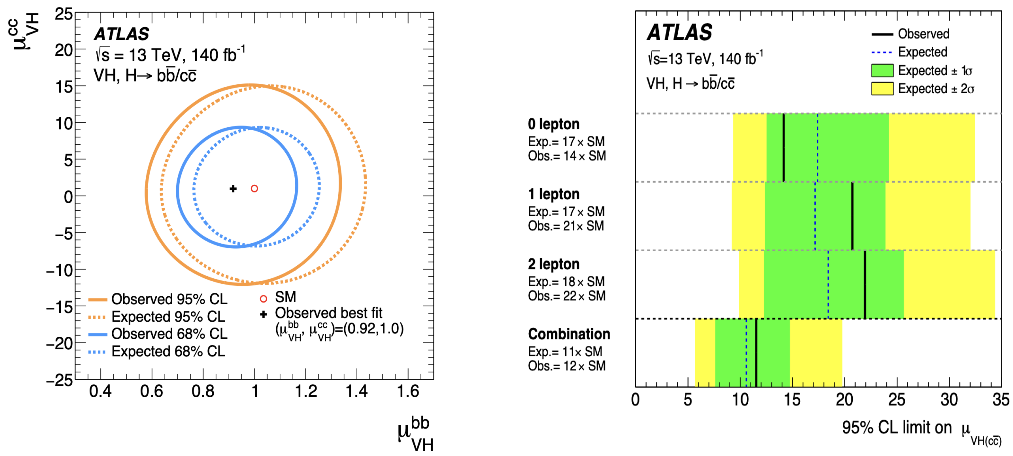

Fig. 1.: 68 and 95% confidence level of the signal strength μ_{VH} (ratio of the observed signal cross-section to the predicted signal cross-section) for $H \to b\bar{b}$ and $H \to c\bar{c}$(left), Exclusion limits on μ_{VH}^{cc}(right) [1].

2. – Advance flavour tagging with GN2

A critical technique for identifying Higgs boson decays to quarks is flavour tagging—the ability to determine the original flavour of the quark (b-quark, c-quark, or a light-flavour quark/gluon) that produced a jet. The performance of flavour tagging algorithms is paramount to the sensitivity of the $H \to b\bar{b}$ and $H \to c\bar{c}$ analyses.
For Run 3, ATLAS is deploying a new, state-of-the-art flavour tagging algorithm named GN2 [3]. This algorithm represents a significant leap forward, built upon a transformer-based architecture. Transformers, which have revolutionized natural language processing, are now being applied to particle physics problems due to their powerful ability to model complex relationships within data, such as the tracks inside a jet.

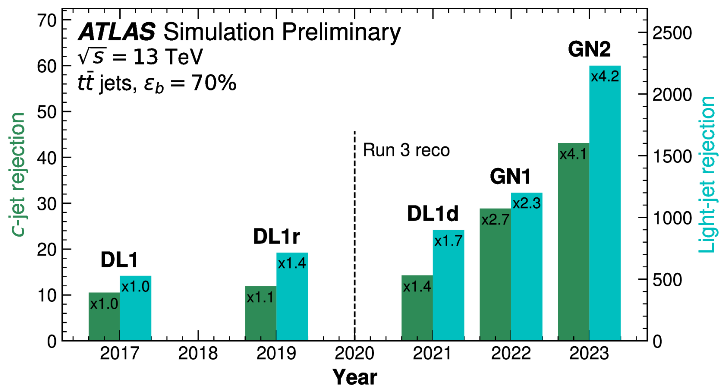

Fig. 2.: Rejection performance comparison between GN2 (transformer-based architecture) and previous taggers [4].

The development of GN2 follows the success of previous Graph Neural Network (GNN) based taggers (GN1), which had already improved flavour tagging performance. GN2 is expected to provide substantial gains over its predecessor, DL1r. As shown in Fig. 3, preliminary simulation studies predict an improvement of approximately 30% in c-jet and light-jet rejection for a fixed b-jet identification efficiency of 70%. Where rejection is defined as the inverse of mis-tagging efficiency of light or c-jets

3. – Analysis overview and Event topology

The analysis is structured into three main channels based on the decay mode of the associated vector boson (V), as illustrated by the Feynman diagrams in Fig. 3 and discussed in Table 1:

Category	Description
0-Lepton Channel	The Z boson decays into a pair of neutrinos ($Z \to \nu\nu$). These events are identified by a large amount of missing transverse energy (MET) in the detector, corresponding to the undetected neutrinos.
1-Lepton Channel	The W boson decays into a charged lepton (electron or muon) and a neutrino ($W \to l\nu$). These events are characterized by one isolated lepton, MET, and jets.
2-Lepton Channel	The Z boson decays into a pair of charged leptons ($Z \to ll$). These events feature two isolated, same-flavour, opposite-sign leptons.

TABLE I.: Description of the 3 channels.

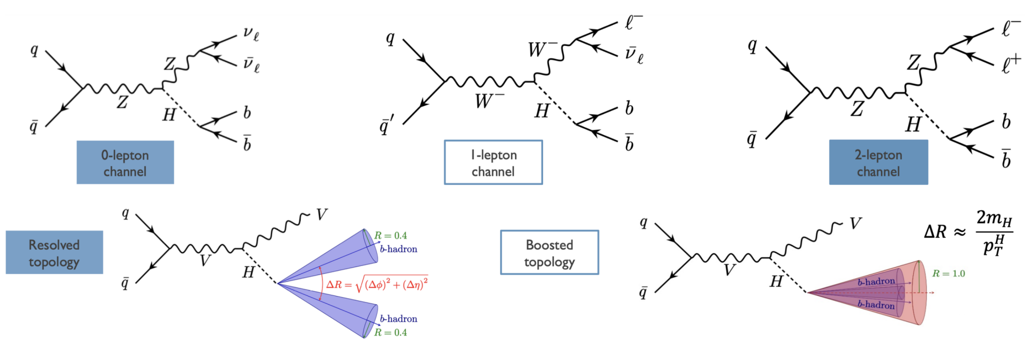

Fig. 3.: VH Analysis overview.

Furthermore, the analysis considers two distinct event topologies depending on the transverse momentum (p_T^V) of the Higgs boson candidate, outlined in Table 2:

Topology	Description
Resolved Topology	At low p_T^V, the two b-quarks from its decay are well-separated and are reconstructed as two distinct, small-radius (R=0.4) jets.
Boosted Topology	At high p_T^V, the decay products are highly collimated. The two b-quarks are merged into a single, large-radius (R=1.0) "fat jet".

TABLE II.: Description of the 2 topologies.

The high p_T region of the analysis is used to study the differential distirbution of the Higgs boson transverse momentum, while the resolved analysis dominates the inclusive measurement on μ_{VH}, and it will be the subject of this paper.

4. – Methodology: sensitivity and categorization

To fully exploit the enhanced capabilities of the new flavour tagging algorithms, the analysis employs an advanced event categorization scheme known as Pseudo-Continuous 2D Tagging. This method moves beyond a simple binary classification of jets (e.g., "b-tagged" or "not b-tagged"), instead, it uses the continuous output scores from the GN2 tagger for both b-jet and c-jet likelihoods to assign each jet to one of several exclusive categories. This 2D approach allows for a more granular separation of signal and background processes. For example, a category can be defined to be rich in c-jets while having low contamination from b-jets and light-flavour jets, which is essential for the $H \to c\bar{c}$ search.

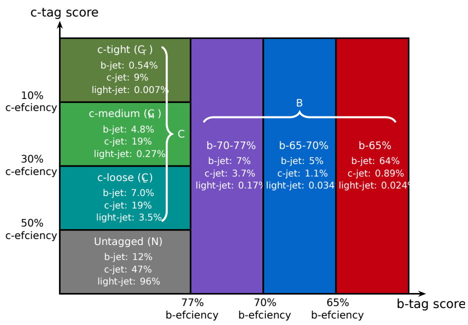

Fig. 4.: Pseudo-Continuous 2D tagging scheme [1].

Figure 4 illustrates this categorization scheme. The space is divided based on the b-tag and c-tag score. Working points (WPs) are defined at different b(c)-tagging efficiencies (e.g., 77%, 70%, 65%). This defines a multi-bin analysis where the signal-to-background ratio varies significantly between categories, enhancing the overall statistical power of the measurement.

5. – Preliminary results and conclusions

To quantify the expected benefit of these improvements, preliminary sensitivity studies are conducted. These studies focus on evaluating the increase in statistical significance of the analysis using the GN2 tagger, with respect to the previous analysis, that used the DL1r tagger [5].

The statistical-only significance is calculated using the Asimov formula.

$$S = \sqrt{\sum_i 2[(s_i + b_i)\ln(1 + \frac{s_i}{b_i}) - s_i]} \quad (1)$$

This formula provides an estimate of the expected median significance by comparing the expected number of signal events (s_i) and background events (b_i) across all bins (i) of a given distribution.

The work done on the VH, $H \to b\bar{b}, c\bar{c}$ analysis, show promising results. The improved performance of the new flavour tagger leads to a relevant increase of significance.

The key findings from these studies are:

- 0-Lepton channel: for $H \to b\bar{b}$ an estimated 16% overall improvement in significance is observed. The improvement is even more pronounced—up to 28% —in the regions of highest transverse momentum of the vector boson ($p_T^V > 400 GeV$). Seemingly for $H \to c\bar{c}$ an overall improvement of 56% is observed.

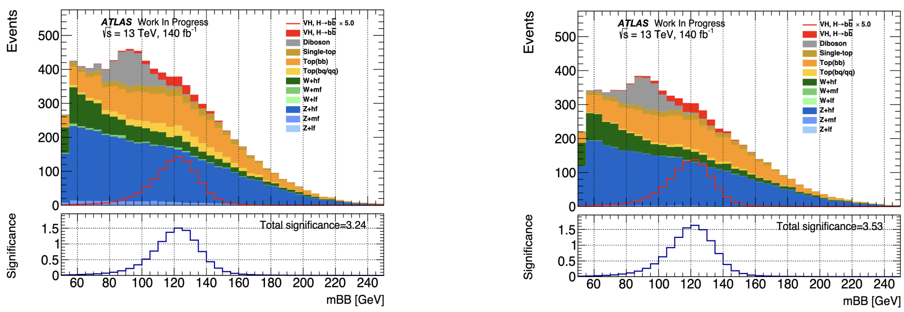

Fig. 5.: Binned significance as a function of the invariant mass of the Higgs candidate, the inclusive significance is also shown in the Figures. Left) The results of the published analysis [1], Right) the new version of the analysis.

- 2-Lepton channel: for $H \to b\bar{b}$ a significant improvement of 11% overall in significance is observed. The improvement is even more pronounced—up to 25% —in the regions of highest transverse momentum of the vector boson ($p_T^V > 400 GeV$). Seemingly for $H \to c\bar{c}$ an overall improvement of 45% is observed.

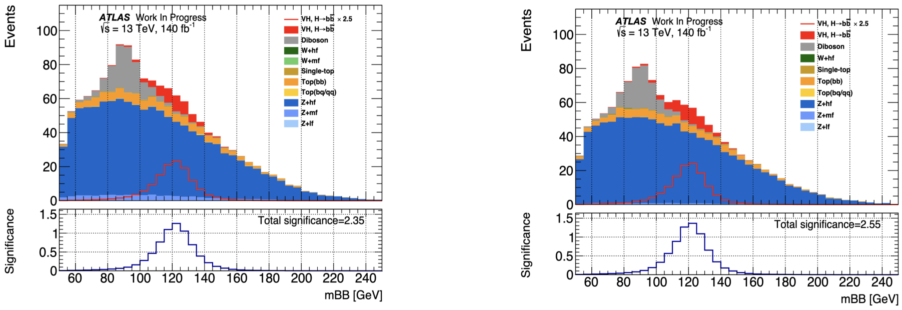

Fig. 6.: Binned significance as a function of the invariant mass of the Higgs candidate, the inclusive significance is also shown in the Figures. Left) The results of the published analysis [1], Right) the new version of the analysis.

Figure 5 and 6 presents the projected di-b-jet invariant mass (mBB) distributions respectively for the 0-lepton and 2-lepton channels. The significance plots show the expected composition of the selected events, with the $H \to b\bar{b}$ signal (red) visible on top of various background sources. The bottom panels quantify the significance per bin, leading to the total expected significance values that underpin the improvement estimates.

6. – Conclusions

The measurement of the Higgs boson's couplings to b- and c-quarks is a vital component of the LHC physics program. The developments presented here demonstrate a clear path toward improving these fundamental measurements at the ATLAS experiment.
The new transformer-based flavour tagger, GN2, provides a substantial boost in performance, enabling more effective rejection of c-jet and light-flavour jet backgrounds. This, combined with advanced analysis strategies like 2D pseudo-continuous tagging, allows for a more powerful separation of signal from background. Preliminary studies for the VH, $H \to b\bar{b}$ analysis, predict significant gains in statistical significance of 11% to 16%, for the $H \to c\bar{c}$ analysis show gains in significance of 45% to 56%.
These improvements will be crucial for achieving a more precise measurement of the Higgs-bottom Yukawa coupling and for pushing the sensitivity of the search for the rare Higgs-charm decay with the full Run 3 dataset and beyond.

REFERENCES

[1] The ATLAS Collaboration, *Measurements of WH and ZH production with Higgs boson decays into bottom quarks and direct constraints on the charm Yukawa coupling in 13 TeV p p collisions with the ATLAS detector*, *JHEP*, **04** (2025) 075;
[2] The ATLAS Collaboration, *JINST*, **3** (2008) S08003
[3] The ATLAS Collaboration, *Transforming jet flavour tagging at ATLAS*, *CERN*, **EP** (2025) 103;
[4] The ATLAS Collaboration, *FTAG-2023-01*, **https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2023-01/**;
[5] The ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, *Eur. Phys. J.* , **CC** (2023) 83,681