Anomaly detection approach for the search of high-mass diboson resonances in fully hadronic final states, using *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Fabrizio Norcia, on behalf of the ATLAS Collaboration
fabrizio.norcia@cern.ch, Università degli Studi di Napoli, Federico II - Napoli, Italy

Summary. — After the discovery of the Higgs boson at CERN LHC in Geneva, significant attention has been devoted to the research of new high-mass diboson resonances. Machine Learning (ML) techniques, particularly those using anomaly detection, offer innovative and model-independent approaches to these searches. Graph Neural Networks (GNNs) have shown promise in anomaly detection for high-energy physics. This work presents first results obtained using the LHC Olympics dataset and preliminary results on high-mass diboson resonances in fully hadronic final states from *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.

1. – Introduction

Currently, no evidence of Beyond Standard Model (BSM) physics has been observed at the Large Hadron Collider (LHC) [1]. BSM research can essentially be pursued in one of two ways: *Model-dependent*, targeting specific signal models; or *Model-independent*, which make minimal assumptions about the signal, offering generality at the cost of specificity. The research described in this paper is conducted using an anomaly detection approach to data collected with the ATLAS detector [2]. Anomaly detection is a Machine Learning (ML) technique that provides a systematic method for identifying anomalous patterns within a distribution of standard events. It typically relies on unsupervised training to achieve this. In the context of High Energy Physics (HEP), the application of this technique involves identifying event features that are not consistent with the predicted background.

2. – Target Processes

Surely, many hypothesized BSM processes involve massive particles that decay to quarks or gluons. Due to colour confinement and asymptotic freedom, this leads to hadronization, resulting in jets as the observable final state. This study targets scenarios where a potentially new resonance decays to two, possibly new, bosons, which in

turn decay hadronically. If the mass of the resonance is much greater than that of the bosons, then, by energy conservation, the bosons will carry large momentum, causing their decay products—multiple jets per boson—to be highly collimated. Subsequently, the hadronic showers produced by each boson, *i.e.* the jets, detected as energy deposits in the calorimeter cells, *i.e.* the constituents, will be spatially concentrated. Thus, it's more efficient to reconstruct (starting from these constituents and other detector information) each boson's decay products as a single large-radius jet, a "fat jet", rather than as multiple separate jets. The reconstruction performed here is the anti-k_T algorithm with $R = 1$ [3]. More than one fat jet could emerge from this method, but only the two p_T-leading ones are kept.

3. – Anomaly Detection in ATLAS

The first ATLAS application of anomaly detection for BSM searches was published in 2023 [4]. In this study, a search for a heavy resonance Y decaying to an Higgs boson H and a new particle X was presented. The Higgs boson was selected through its decay to bottom quarks, while the X candidate decays hadronically, thus leading to a fully hadronic final state $b\bar{b}q\bar{q}$. The results were interpreted in the Heavy Vector Triplet (HVT) theoretical framework model [5]. Due to large Y mass considered in the model ($m_Y = 1 \div 6$ TeV) Higgs candidates are always reconstructed as fat jets, while, X candidates (with mass ranges $m_X = 65 \div 3000$ GeV) can be reconstructed as large-R (*Merged* regime) or small-R (*Resolved* regime) jets, reflecting the value of the mass ratio m_X/m_Y.

Event Selection. The research employed a supervised Deep Neural Network (DNN) trained to identify bottom anti-bottom resonances, effectively serving as a Higgs tagger. For the X boson, two strategies were adopted: a model-dependent approach and, a model-independent one based on anomaly detection. The latter, in particular, applied only to the *Merged* regime, using a Variational Recurrent Neural Network (VRNN) [6] trained without supervision on the first 20 p_T-ordered constituents ([1]), assigning anomaly scores based on deviation from the learned distribution.

Results. No significant deviation from the background was observed. As a result, 95% confidence level (CL) upper limits on the cross section $\sigma(pp \to Y \to XH \to q\bar{q}b\bar{b})$ were derived in the two-dimensional (m_X, m_Y) mass plane, using a simultaneous fit to both the *Merged* and *Resolved* signal regions, incorporating all statistical and systematic uncertainties (see fig. 1). Moreover, as shown in fig. 2, the anomaly detection approach yields an observed 95% CL upper limit on the cross section that is competitive with the model-dependent method across four signal benchmarks. Notably, for three additional models beyond the target scenario, it even surpasses the performance of the VRNN trained specifically for HVT.

4. – Graph Neural Network for Anomaly Detection

The current analysis is founded on the idea that jets have a sparse internal structure, which makes their constituents ideal for being represented as nodes in a graph. Specifically, graphs are a mathematical structure consisting of a set of elements, called nodes, which can be linked together by "edges", representing relationships between them. A

([1]) A preselection of $p_T > 1.2$ TeV has been applied to the fat jet.

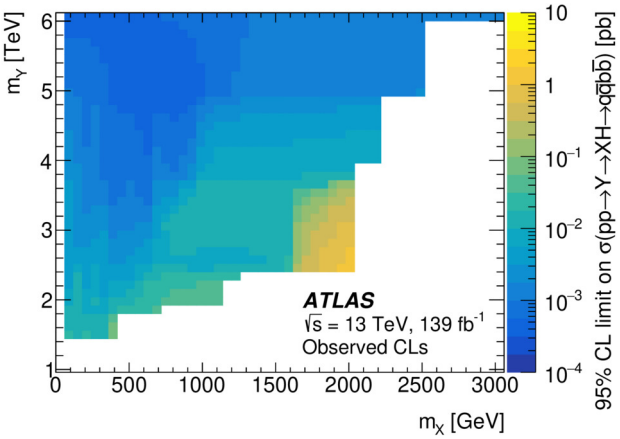

Figure 1. – 95% confidence level (CL) upper limits on the $Y \to XH \to q\bar{q}b\bar{b}$ cross section obtained using a simultaneous fit to both the *Merged* and *Resolved* signal regions, incorporating all statistical and systematic uncertainties. [4]

graph definition for jets is adopted, in which the constituents four-momentum vector are used as node features and the inverse of the distance in the $\eta - \phi$ plane serves as the only edge feature, present only if this distance is less than 0.2. This representation allows us to apply Graph Neural Networks (GNNs), a class of machine learning models specifically designed to operate on graph-structured data and effectively capture relational patterns. This GNN is trained and validated only on background graphs and subsequently tested on both the background and signal samples. During training, the embeddings of every node are updated at each layer by aggregating information passed to the node from

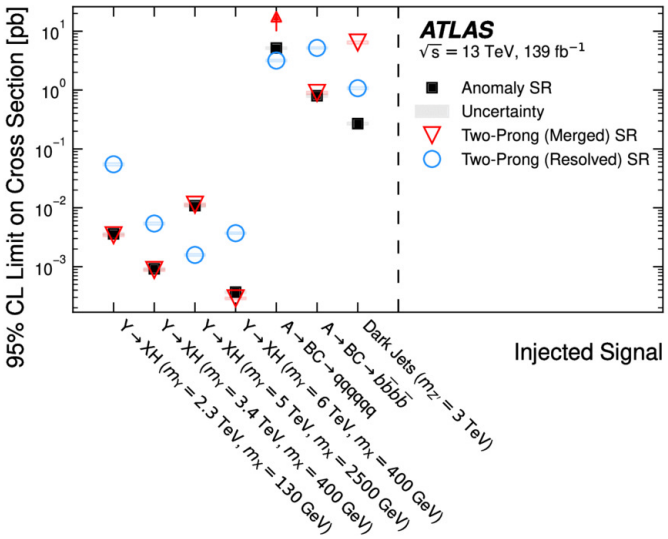

Figure 2. – 95% confidence level (CL) upper limits comparison between model-dependent and model-independent approaches. [4]

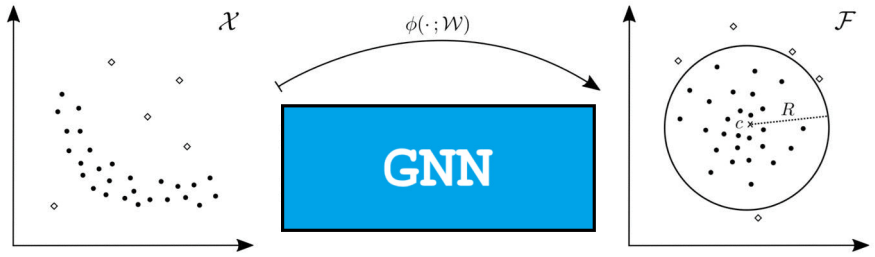

Figure 3. – Schematic of GNN training using the SVDD objective. • = background, ◊ = signal.

its neighbours. This is message passing. Two GNN architectures were tested for this analysis:

- *EGAT*, Edge Graph Attention Network, extends GAT [7] by incorporating, in parallel, layers that process edge features to layers that process node features. Like GAT it uses an attention mechanism to achieve this [8];
- *GIN*, Graph Isomorphism Network, employs Multi Layer Perceptron (MLP) on the standard message passing embedding [9].

In both cases, the GNN maps the graph features from the input space X to the output feature space F by Deep Support Vector Data Description (SVDD) objective, as depicted in fig. 3. The loss functions used for optimization are given in (1), with eq. (1a) for the GIN architecture and (1b) for the EGAT:

$$\text{Loss}_{\text{GIN}} = \frac{1}{N}\sum_{i=1}^{n}\|\phi(G_i;W)-c\|^2 + \frac{\lambda}{2}\sum_{l=1}^{L}\left\|W^l\right\|_F^2 \tag{1a}$$

$$\text{Loss}_{\text{EGAT}} = \frac{1}{N}\sum_{i=1}^{n}\|\phi(G_i;W)-c\|^2 \tag{1b}$$

After training, an anomaly score (AS) is assigned to every graph-jet:

$$\text{AS} = \|\phi(G_i;W^*)-c\|^2 \tag{2}$$

where W^* is the learned weights matrix and c is the average initial feature vector, *i.e.* the centre of the minimal enclosing hypersphere in F. Lastly, the anomaly score of an event is generally obtainable by combining the anomaly scores of the two fat-jets.

5. – LHC Olympics

An initial GNN optimization was conducted on the public LHC Olympics dataset [10], which includes simulated QCD dijet background and an HVT signal $Z' \to XY \to q\bar{q}q\bar{q}$ with $m_{Z'} = 3.5$ TeV, $m_X = 500$ GeV and $m_Y = 100$ GeV. Events are preselected with $p_T > 1.2$ TeV and $|\eta| < 2.5$, and reconstructed as two fat jets. The anomaly score for an

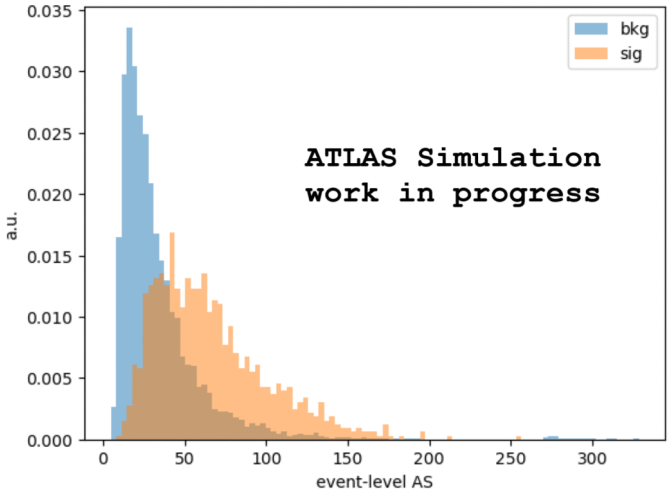

Figure 4. – Test-set results for the event anomaly score on the LHC Olympics dataset. Background events are shown in blue, and signal events in orange.

event, here, is defined as the average of the anomaly scores of the two jets. Fig. 4 shows the test-set results for *EGAT*, which slightly outperforms *GIN* as the area under the curve (AUC) of the receiving operating characteristic (ROC) is: 81.8% *vs.* 79.6% coming from *GIN*. Due to this, *EGAT* was chosen as the main architecture for the continuation of the study.

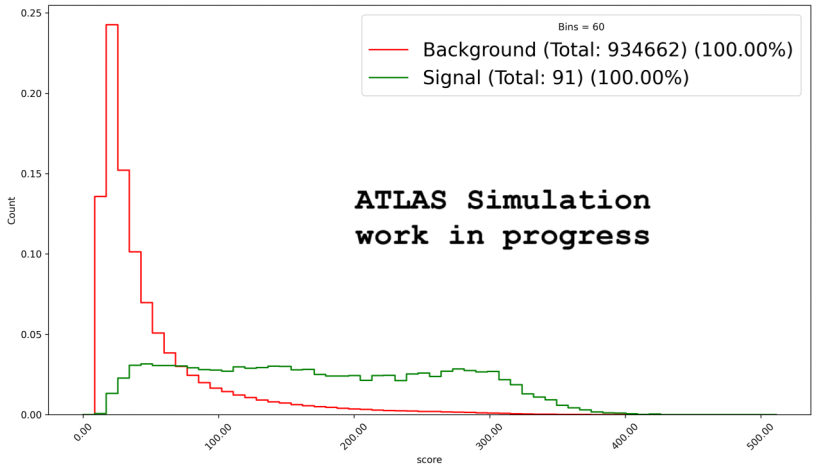

Figure 5. – Test-set results for the event anomaly score on the $Y \to XH$ with $m_Y = 3.4$ TeV and $m_X = 200$ GeV test signal. Background events are shown in red, and signal events in green.

6. – Preliminary results in ATLAS

The study is transitioning to Run 3 ATLAS data, currently validating on Monte Carlo (MC) samples such as HVT $Y \to XH$, $W' \to WW$, DarkJets and others. The

analysis adopts a fully data-driven strategy, defining a signal region via a cut on the anomaly score. Pre-selection cuts are applied, and new trigger items are being studied to potentially recover events below the 1.3 TeV mass threshold. As for the ML component of the study, graph nodes now include more features, and the edge feature is redefined as the negative exponential of the distance in the η-ϕ plane. Besides, the score for the event is now taken as the product of the anomaly scores of the two jets. More significantly though, in addition to the jet-level graph definition, an event-level graph has been constructed; this includes the constituents of both jets. Training has been performed for both approaches. Testing on the $Y \to XH$ sample with $m_Y = 3.4$ TeV and $m_X = 200$ GeV yields an AUC of 89.63% for the event-level graph versus 81.25% for the jet-level. Results for the event-level graphs are displayed in figs. 5 and 6.

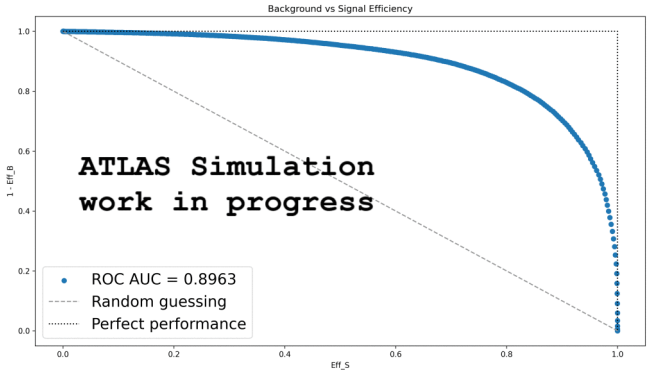

Figure 6. – ROC obtained from event anomaly score cuts on the $Y \to XH$ with $m_Y = 3.4$ TeV and $m_X = 200$ GeV test signal.

7. – Conclusions

The 2023 anomaly detection result seems very promising [4] and the GNN has also delivered a good performance so far. The study will continue validation on other benchmark signals, followed by the unblinding of ATLAS Run 3 data.

REFERENCES

[1] Evans L. and Bryant P. (editors), *JINST*, **3** (2008) S08001.
[2] ATLAS Collaboration, *JINST*, **3** (2008) S08003.
[3] Cacciari M., Salam G. P. and Soyez G., *J. High Energy Phys.*, **04** (2008) 063.
[4] ATLAS Collaboration, *Phys. Rev. D*, **108** (2023) 052009.
[5] Pappadopulo D., Thamm A., Torre R. and Wulzer A., *J. High Energy Phys.*, **09** (2014) 060.
[6] Kahn A. *et al.*, *JINST*, **16** (2021) P08012.
[7] Veličković P. *et al.*, 2018, arXiv:1710.10903v3 [cs.LG].
[8] Chen J. and Chen H., 2021, arXiv:2101.07671v1 [cs.LG].
[9] Xu K., Hu W., Leskovec J. and Jegelka S., 2019, arXiv:1810.00826v3 [cs.LG].
[10] https://lhco2020.github.io/homepage/