PROCEEDINGS OF SCIENCE

Searches for new phenomena using Anomaly Detection at the ATLAS experiment

---

Antonio D'Avanzo, on behalf of the ATLAS collaboration[a,*]

[a] *INFN Naples, University of Naples Federico II,*
Via Cintia, Complesso universitario di Monte S. Angelo ed. 6, Naples, Italy

E-mail: antonio.davanzo@cern.ch

After the discovery of the Higgs boson at the Large Hadron Collider at CERN, a new time period characterized by a lack of discoveries of Beyond Standard Model physics in particles accelerators is ongoing. Anomaly detection is a novel machine learning approach that can represent a new possible approach to searches, as it allows a very general method with the signatures of interest without losing sensibility to possible signals. ATLAS analyses are taking the first steps in this direction, following the results obtained from Classification Without Labels, often referred to as CWoLa, based resonant searches. These proceedings show the results obtained with anomaly detection approaches in ATLAS, where events are selected solely because of their incompatibility with a learned background-only model. In particular, the focus is on the search for a heavy resonance Y decaying into a Standard Model Higgs boson H and a new particle X in a fully hadronic final state, which represents the first application of fully unsupervised machine learning in an ATLAS analysis.

ATL-PHYS-PROC-2024-070
23 September 2024

42nd International Conference on High Energy Physics (ICHEP2024)
18-24 July 2024
Prague, Czech Republic

---

[*] Speaker

1. Introduction

The first search in ATLAS [1] to be conducted by means of anomaly detection (AD) [2] is the one in which heavy resonances A are searched for when they decay into two unknown bosons B and C in fully hadronic final states using a weakly-supervised Classification Without Labels (CWoLa) machine learning (ML) technique to define the signal region (SR) of interesting events [3]. Although no significant deviation from background was observed, the exclusion limits on the production cross section as a function of several mass hypothesis (Figure 1) highlighted that this approach performed better for generic mass values away from the ones compatible with the Z and W vector bosons, as expected by the weakly-supervised nature of the method.

2. Fully unsupervised search for new resonances in ATLAS

Following the results from this first weakly-supervised ML analysis, a new search was carried out based on a fully unsupervised anomaly detection approach. Data collected by the ATLAS detector during the full Run II (2015-2018) of the LHC, corresponding to an integrated luminosity of 139 fb^{-1}, was analysed to look for new particles Y and X produced in a resonant decay $pp \rightarrow Y \rightarrow XH$. The Higgs boson H decays in $b\bar{b}$, while the X is an unknown boson decaying fully hadronically [4].

The AD technique is employed for the identification of the X boson, along with a different model-dependent approach based on a heavy vector triplet (HVT) signal hypothesis, targeting the two-prong topology $X \rightarrow q\bar{q}$ in case no evidences are found in the anomalous signal region. The phase space explored is $1.5 < m_Y < 6$ TeV and $50 < m_X < 3000$ GeV, where m_Y and m_X are respectively the Y and X boson masses, and this allows H to be always produced with a significant Lorentz boost. For the X boson this happens only if m_X/m_Y is lower than $\simeq 0.3$. In the case of the two-prong hypothesis, two kinematic regimes can thus be distinguished: one with collimated quarks so that the X boson can be reconstructed with a single large-R jet (called merged) and one with two small-R jets with a significant angular opening (called resolved).

The Higgs boson is identified by means of a deep neural network (DNN) $H \rightarrow b\bar{b}$ tagger [5], the output of which is used to compute a D_{Hbb} score per jet used to select the Higgs boson at a 60% working point (WP) $D_{Hbb} > 2.44$. Before that though, the ambiguity in the H and X identification between the two most p_{T}-leading large-R jets for each event is resolved using this same tagger. The H candidate is chosen as the jet with the highest value of the D_{Hbb} score. An additional selection on the Higgs candidate mass, to constrain it in the Higgs mass window [75, 145] GeV, is used with the tagger WP to define the preliminary SR and five control regions (CRs). The X tagging finally completes the SR definition, making in total three separate SRs based on the X tagging procedure applied. The QCD multijet background estimation in each SR is achieved with a fully data-driven technique.The distributions of the data are reweighted from a control region 0 (CR0) to the SR with a function $w(x)$, obtained from a DNN in the training CR between 145 GeV and 175 GeV, validated in the CR between 65 GeV and 75 GeV and finally extrapolated to the Higgs boson mass window.

The AD algorithm for the X tagging consists on the prediction of a variational recurrent neural network (VRNN) [6]. The VRNN is trained exclusively on data, mostly composed of background events, using as input a sequence of ATLAS reconstructed constituents four-vectors of jets with

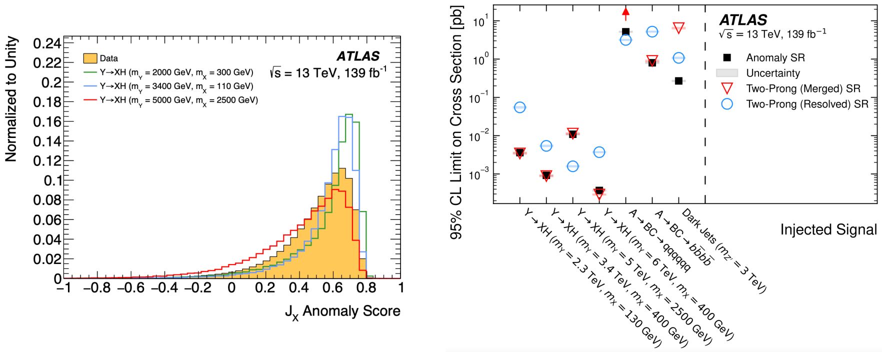

Figure 1: Anomaly Score distribution for the X candidate in data and MC signals (left); upper limits on production cross section at 95% CL for the process $pp \to Y \to XH$ (right), with comparisons between the AD procedures applied and model-dependent approaches, evaluated over several signal hypotheses.[4]

$p_{\mathrm{T}} > 1.2$ TeV. Once the training is complete, an anomaly score (AS) per jet is computed from the output. Large values (AS > 0.5) correspond to an X candidate produced in anomalous interactions, while low values correspond to SM processes compatibility. Moreover, this discriminant turned out to be not only sensitive to the two-prong decay of the X boson, but also to more exotic signal hypothesis.

The results are extracted based on the final state di-jet invariant mass distribution m_{JJ} in the SRs as defined above, separately for each X candidate mass value. As a consequence, many signal regions are defined for each (m_Y, m_X) interval. Model-independent discovery p-values are obtained by testing the background-only hypothesis in the anomaly SR, along with exclusion limits for the two-prong signal plus background hypothesis in case of no new physics evidence. No significant deviation from the expected SM background are observed, with a maximum deviation equal to $1.47\ \sigma$ of global significance in the anomaly SR. For this reason, upper limits at 95% C.L. on the production cross section $\sigma(pp \to Y \to XH \to q\bar{q}b\bar{b})$ are determined for the whole $m_Y\ vs\ m_X$ HVT signal grid considered, and in particular for other signal models. A comparison between SRs shows that the anomaly approach performs as good as the model-dependent exclusion regions when tested on the two-prong channel, while it performs better when tested on more exotic processes due to its general approach (Figure 1).

References

[1] ATLAS Collaboration, JINST 3 (2008) S08003.

[2] V. Belis, P. Odagiu and T. K. Aarrestad, Rev. in Phys. 12 (2024) 100091.

[3] ATLAS Collaboration, Phys. Rev. Lett. 125 (2020) 131801.

[4] ATLAS Collaboration, Phys. Rev. D 108 (2023) 052009.

[5] ATLAS Collaboration, ATL-PHYS-PUB-2020-019, url: https://cds.cern.ch/record/2724739.

[6] A. Kahn, J. Gonski, I. Ochoa, D. Williams and G. Brooijmans, JINST 16 (2021) P08012.