EPJ Web of Conferences **295**, 09040 (2024) https://doi.org/10.1051/epjconf/202429509040
CHEP 2023

Use of Anomaly Detection algorithms to unveil new physics in Vector Boson Scattering

Giulia Lavizzari[1,2,*], *Giacomo* Boldrini[1,2], *Simone* Gennai[2], and *Pietro* Govoni[1,2]

[1]Università degli Studi di Milano Bicocca, Piazza della Scienza, 3, 20126 Milano MI, Italy
[2]INFN Milano-Bicocca, Piazza della Scienza, 3, 20126 Milano MI, Italy

Abstract. A new methodology to improve the sensitivity to new physics contributions to the Standard Model processes at LHC is presented.
A Variational AutoEncoder trained on Standard Model processes is used to identify Effective Field Theory contributions as anomalies. While the output of the model is supposed to be very similar to the inputs for Standard Model events, it is expected to deviate significantly for events generated through new physics processes. The reconstruction loss can then be used to select a signal enriched region which is by construction independent of the nature of the chosen new physics process. In order to improve further the discrimination power, an adversarial layer is introduced with a cross entropy term added to the loss function, optimizing at the same time the reconstruction of the input variables of the Standard Model and classification of new physics processes. This procedure ensures that the model is optimized for discrimination, with a small price in terms of model dependency to physics process.
In this work I will discuss in detail the above-mentioned method using generator level Vector Boson Scattering events produced at LHC assuming an integrated luminosity of 350/fb.

1 Introduction

In 2012, a significant milestone in particle physics was reached when the ATLAS and CMS collaborations confirmed the existence of the Higgs boson, marking a key moment in the validation of the Standard Model (SM) [1, 2]. The SM encodes our most comprehensive account of the fundamental fabric of matter, by describing all the known elementary particles, and three of the four fundamental forces that govern the interactions between them. Over time it proved capable of offering precise experimental predictions and theoretical explanations for numerous phenomena, but the persistent unresolved questions indicate the need for its expansion to include new physics contributions: for example, the presence of dark matter [3], the relative abundance of matter over antimatter [4] and the so-called hierarchy problem [5].

Many Beyond Standard Model (BSM) theories have been proposed, attempting to answer the several questions that the SM does not resolve. However, despite the increasingly large amount of statistics made available by the LHC, there are no indications for processes predicted by those theories. Since the number of possible BSM theories and their variations in terms of free parameters is extremely large, and only few of them can be tested, it is possible

*e-mail: g.lavizzari1@campus.unimib.it

that so far we have simply chosen the wrong models to look for. As a result, many efforts are currently directed towards the development of agnostic analyses with respect to the underlying BSM theory assumed, marking a change of paradigm in the way new physics searches are carried out at the LHC (see for example [6, 7]). A possible strategy to probe BSM physics is the so-called Standard Model Effective Field Theory (SMEFT) [8], a low-energy, generalized SM extension that can be used to parametrise small deviations of SM observables. In this perspective, the SM lagrangian is expanded in a series with respect to the inverse of the new physics energy scale, and the higher order operators represent the new physics contributions. Following the SM Lagrangian, which has dimension 4 in terms of energy, the first terms of such expansion that comply with the accidental symmetries of the SM are of dimension 6 and 8.

SMEFT is a complex, multidimensional problem: it involves thousands of parameters to constrain and each EFT operator affects each variable differently. Therefore, it is extremely hard to define a single kinematic observable that allows for detecting all operators: our goal is thus to build a strategy that maximizes the observation of anything that is not SM. Our strategy relies on unsupervised learning, following an increasingly popular choice in the field of independent searches, as an unsupervised model does not require any knowledge of a specific BSM theory. The underlying idea is to train a model on known physics, and then to use it to detect EFT events as outliers. We adopt a model based on Variational AutoEncoders (VAE) [9, 10], generative algorithms that encode the inputs over a lower-dimensional latent space, from which a point is then sampled and mapped back to the input space through a second network. As EFT events follow different patterns than the ones the VAE learnt during training, they will be reconstructed poorly. Therefore, anomalies can be singled out by building a metric that accounts for the difference between input and output event by event.

To test such strategy we employed parton-level simulations of Vector Boson Scattering events at the LHC, considering the expected integrated luminosity of Run 3, namely 350 fb^{-1}. This study has been conducted neglecting any background process and considering generator-level observables.

2 The physics usecase: Vector Boson Scattering at the LHC

A suited physics process for model independent BSM studies is the scattering of vector bosons (VBS) [11], an event that takes place at the LHC when two partons of the incoming protons radiate vector bosons, which in turn interact. Such process is expected to be sensitive to modifications of the EW sector and of the Higgs mechanism: a delicate set of cancellations between VBS diagrams with and without the Higgs boson takes place in order for the VBS cross-section to be non-divergent. Any slight modification to such delicate equilibrium would manifest itself in anomalous contributions to the various VBS observables, independently from the theory assumed to describe the new physics responsible for the discrepancies. In particular, VBS provides tree-level sensitivity to effective operators inducing modifications of triple and quartic gauge couplings, as well as of Higgs-gauge couplings away from the Higgs mass-shell, and even to contact interactions among four quarks. Mainly motivated by the sensitivity to quartic gauge couplings, previous studies of EFT effects in VBS were often restricted to dimension-8 operators. The impact of dimension-6 operators was explored systematically only recently, mainly driven by the interest in incorporating these measurements into global SMEFT analyses. In this study we will focus on the latter.

Due to their electroweak nature, VBS processes tend to suffer from a quite low signal to background ratio. However, the production of two same-sign W bosons (SSWW) decaying leptonically produces a very clean signature in the detector, shown in Fig. 1. It comprises two jets with a large invariant mass in the forward region, two same-sign charged leptons and the

missing transverse energy due to the two neutrinos. SSWW is therefore considered a golden channel for many experimental measurements, as the background induced by quantum chromodynamics (QCD) processes is relatively small compared to the EW induced production.

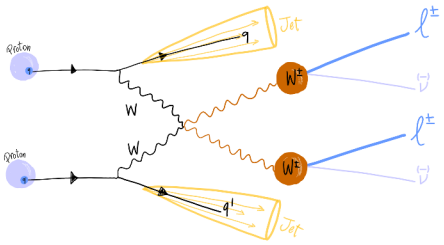

Figure 1. Topology of same sign WW scattering in the fully leptonic final state.

The observables we studied comprise several kinematic and angular distributions encoding the most important features of SSWW scattering: the invariant masses of the dilepton and dijet systems (m_{ll} and m_{jj}), the transverse momentum of the two leptons and two jets ($p_{t_{l1}}$, $p_{t_{l2}}$, $p_{t_{j1}}$, $p_{t_{j2}}$), the transverse momentum of the dilepton system ($p_{t_{ll}}$), the missing transverse momentum (MET), the pseudorapidities of leptons and jets (η_{l1}, η_{l2}, η_{j1}, η_{j2}) and the pseudorapidity and azimuthal angle differences between the two jets ($\Delta\eta_{jj}$, $\Delta\Phi_{jj}$).

3 A dimension-6 EFT interpretation of VBS

SMEFT [12–14] is a generalized SM extension that can be used to describe small deviations in the distributions of SM observables. Within the SMEFT framework, the BSM contributions are described by taylor-expanding the SM lagrangian in powers of $1/\Lambda$, where Λ is the energy scale of the new physics. The BSM effects are thus parametrized as the resulting higher dimensional operators, which respect the Lorentz and SU(3) × SU(2) × U(1) gauge invariances:

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \sum_{i,d>4} \frac{c_i}{\Lambda^{d-4}} Q^{(d_i)} \tag{1}$$

where Q_i^d is a set of dimension d operators and c_i are the so-called Wilson Coefficients, that gauge the intensity of their effect. The presence of additional terms in the Lagrangian results in a variation of the scattering matrix amplitude of known processes predicted by the SM. Processes that have a large EFT contribution can be used to measure the sensitivity to new operators, putting limits on the Wilson coefficients values and probing for the presence of new physics effects.

In this study we focus only on dimension 6 operators [15]. We employed the `SMEFTsim` package [16] to generate the EFT predictions at leading order in the U35 flavour scheme and m_W input scheme for the following operators: Q_W, $Q_{Hq}^{(1)}$, Q_{HW}, $Q_{qq}^{(1)}$, $Q_{qq}^{(1,1)}$, $Q_{qq}^{(3)}$, $Q_{qq}^{(3,1)}$. The events were generated at parton level via `MadGraph5_aMC@NLO` [17].

3.1 Dim-6 operator effects on the observables

The SMEFT Lagrangian, truncated at dimension six and considering a single non-zero operator Q_α, can be written as follows:

$$\mathcal{L}_{BSM} = \mathcal{L}_{SM} + \frac{c_\alpha^{(6)}}{\Lambda^2} Q_\alpha^{(6)} \tag{2}$$

The presence of an additional term with respect to the traditional SM Lagrangian is reflected into a variation of the amplitude of the process under study:

$$A_{BSM} = A_{SM} + \frac{c_\alpha}{\Lambda^2} \cdot \mathcal{A}_{Q_\alpha} \tag{3}$$

Squaring A_{BSM} yields a quantity proportional to the event probability:

$$|A_{BSM}|^2 = |A_{SM}|^2 + \frac{c_\alpha}{\Lambda^2} \cdot 2Re(A_{SM} A_{Q_\alpha}^\dagger) + \frac{c_\alpha^2}{\Lambda^4} \cdot |\mathcal{A}_{Q_\alpha}|^2 \tag{4}$$

Here, the first term represents the pure SM contribution, while the following ones introduce a linear and quadratic dependence on the EFT amplitudes. These operators collectively modify the kinematic distributions of the given process as shown in Fig. 2, and the extent of these modifications depends on the Wilson coefficients associated to each operator.

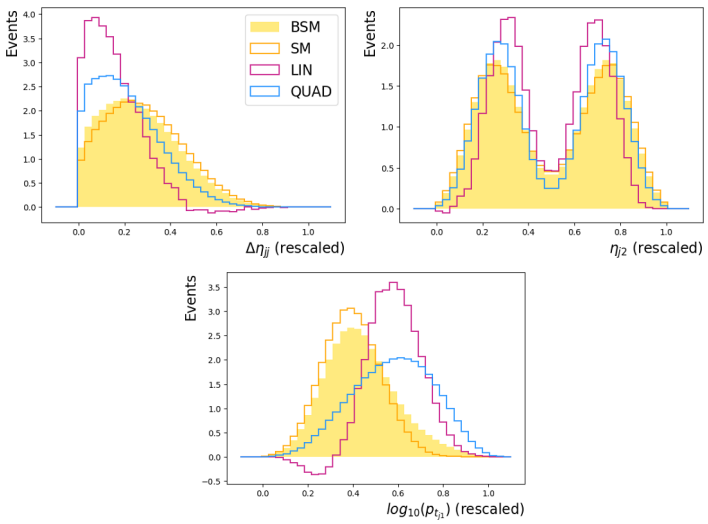

Figure 2. Distributions of pure SM, linear and quadratic contributions (solid colored lines) for different observables: $\delta\eta_{jj}$ (top left), η_{j2} (top right) and $\log_{10}(p_{T,j1})$ (bottom). The distributions are normalised to an area of 1. The weighted sum of the three contributions gives the overall BSM distribution (filled histogram).

4 The anomaly detection strategy: a VAE-based model

VAEs are a type of probabilistic generative models. They work by creating a hidden representation of input data, which is then used to recreate an output that closely matches the input.

These models were first introduced in 2013 [9] and have since found applications in various areas, such as generating content and reducing noise. However, because of their generative capabilites, VAEs can also be useful for creating reliable anomaly detection methods that can handle variations in the reference data.

4.1 VAE architecture

The VAE architecture consists of two main components, each realized as a deep neural network. The first one, known as the encoder, is responsible for mapping the input distributions into a lower-dimensional latent space. During training the latent space is forced to be regular, i.e. to be described by a Gaussian distribution: such regularity enhances the model's flexibility and expressive capacity. The second component, called decoder, operates as a complementary network performing an inverse transformation that maps points sampled from the latent space back to the original input space. This process aims to generate an output closely resembling the input.

The model is trained by minimizing a loss function comprising two key components: the Mean Squared Error (MSE) between input and output, ensuring accurate reconstruction of the inputs, and the Kullback-Leibler Divergence (KLD). The latter forces the latent distributions to be close to Gaussian distributions, therefore accounting for the regularization of the latent space.

4.2 VAEs for anomaly detection

Anomaly detection is the task of finding unusual patterns in a dataset: in our case this means to isolate events stemming from BSM processes, based on the difference in their signature with respect to the known physics. To achieve this, the VAE model is trained only on SM events, therefore it learns the underlying patterns of known physics. Once trained, the model is evaluated also on BSM events, which exhibit different structures from those learned during training. Therefore, the model does not interpret such events as well as SM ones, as shown in Fig. 3. To assess the quality of the reconstruction of an event we compute the MSE between input and output, averaged on all the observables. By selecting the events for which the MSE is greater than a certain threshold we are able to identify an region enriched in anomalies, as shown in Fig. 4.

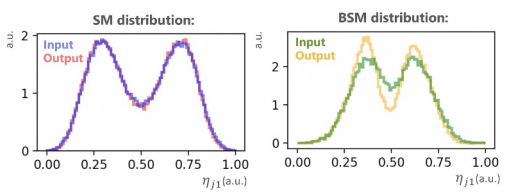

Figure 3. Comparison between the input and output of the model in the case of SM (left) and BSM (right) distributions for the pseudorapidity of the leading jet.

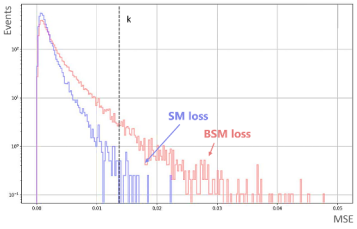

Figure 4. MSE for SM (blue) and BSM (red) events. Anomalous events lie in the tail of the loss function.

Two preprocessing strategies were employed: first we scaled the input distributions between 0 and 1 and we computed the logarithm of the kinematic variables to reduce their dynamic range. Then we split the SM sample in two subsets comprising 80% and 20% of the total 900000 events, respectively used for training and testing. The whole strategy is implemented through the scikit-learn [18] and TensorFlow [19] libraries.

4.3 Embedding a classification step in the training to optimize for discrimination

The primary limitation of this approach arises from the fact that, despite our aim being the isolation of EFT events, the training process is optimized exclusively for reconstructing a SM dataset. However, choices that enhance SM reconstruction are not always ideal for discrimination purposes: for instance, a larger latent space was found to improve SM events reconstruction, but it diminishes the model's ability to discriminate against EFT perturbations. This can be explained by the fact that a more complex latent space allows for the model to better learn the underlying data structure, and therefore the SM reconstruction improves. However, this enhanced understanding also translates to improved extrapolation capabilities, ultimately compromising discrimination power.

To address such limitation, we introduced a classification step in the training procedure by incorporating a third component into the model, namely a two-layers neural network functioning as a classifier. Encoder, decoder and classifier are separate architectures, each comprised of several densely connected layers, that are later integrated into an end-to-end model. The Encoder and Decoder consist of three layers with dimensions 20, 10, and 7, respectively, all employing a Leaky ReLU activation function. The dimension of the latent space has a profound impact on the model's performance and has been varied between 3 and 7, resulting in the effects described earlier. The Classifier comprises two layers: the first with a dimension of 50 and a linear activation function, and the second with a dimension of 1 and a hard sigmoid activation.

To train this model, we divide the training sample into two datasets. One dataset exclusively contains SM events, while the other includes both SM and EFT contributions. Initially, the VAE part of the architecture processes the SM subset, and the resulting MSE and KLD are added to the model losses for minimization. Subsequently, the sample that includes EFT contributions is processed through the model: the MSE and KLD losses obtained from this step are used as inputs for the Classifier, and its binary cross-entropy is added to the overall model loss.The total loss minimized during the training is therefore a sum of three components, namely the MSE, KLD and binary-crossentropy, therefore effectively embedding the discrimination process within the training procedure.

It's important to note that this strategy introduces an additional reliance on the modeling of new physics contributions. To maintain as much model independence as possible, we chose to use a single operator during training and later assess the model's performance with other operators as well. Figure 5 shows the architecture and anomaly detection strategy.

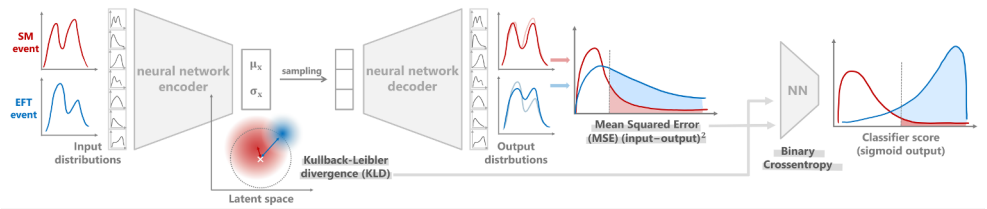

Figure 5. Structure of the VAE+DNN model and anomaly detection strategy.

5 Results

To quantify the discrimination power of the models we defined a proxy metric for the significance σ, namely by considering the number of BSM events detected in the anomaly-enriched

region, minus the number of SM events that pass that same selection:

$$\sigma(c_{op}) = \frac{|n_{BSM}(c_{op}) - n_{SM}|}{\sqrt{n_{SM}}} = \frac{|n_{LIN}(c_{op}) + n_{QUAD}(c_{op}^2)|}{\sqrt{n_{SM}}} \tag{5}$$

A model is considered sensitive to a given operator if $\sigma = 3$ for some value of c_{op} (table 1). The smaller the value of c_{op} for which we have $\sigma = 3$, the better.

Table 1. The value of c_{op} for which $\sigma(c_{op}) = 3$, considering an integrated luminosity of 350 fb^{-1}. The results for the VAE+NN refer to a model trained to reconstruct SM VBS EWK events and to discriminate them from a sample comprising contribution from the Q_W operator: during training c_W is set to 1, then the events are properly weighted by the correct value of c_{op}.

model	c_W	c_{qq}^1	$c_{qq}^{1,1}$	c_{qq}^3	$c_{qq}^{3,1}$	c_{Hq}^1	c_{HW}
VAE	0.34	0.56	0.29	0.04	0.06	-	0.41
VAE+NN	0.13	0.17	0.18	0.11	0.11	0.61	0.65

The first conclusion we can draw is that it is possible to separate EFT contributions from known physics by means of the VAE model. Furthermore, the embedding of the classifier in the model improves the discrimination without impinging excessively the generality of the approach: the performances improve significantly for Q_W, on which the model was trained, but also for $Q_{qq}^{(1)}$, $Q_{qq}^{(1,1)}$ and deteriorate only slightly for $Q_{qq}^{(3)}$, $Q_{qq}^{(3,1)}$ and Q_{HW}. Furthermore, the new model is sensitive to $Q_{Hq}^{(1)}$, which was not detected by the simple VAE since we only considered values of c_α smaller than one.

5.1 Conclusions and future perspectives

This study demonstrates the feasibility of employing a VAE-like model to isolate an EFT-enriched region. By training the model to recognize known SM physics patterns exclusively, we were able to detect BSM phenomena in a model-independent manner. Such procedure was tested using the SSWW VBS process, considering an integrated luminosity of 350 fb^{-1}. Figure 6 shows the comparison between the shape of the model score and of a simple kinematic variable with tight selections targeting a VBS-like signature. Even though the latter may be optimal for a single operator, it is not sensitive to other operators; on the other hand, the model score grants sensitivity to several EFT operators.

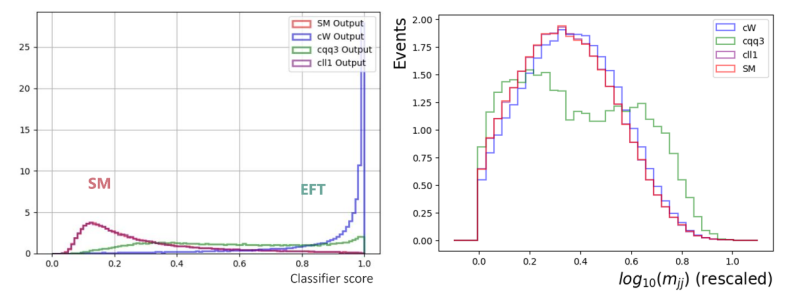

Figure 6. Comparison between the output score of the model (left) and a simple kinematic variable with tight selections to select a VBS-like signature (right).

While this study has demonstrated the potential of unsupervised models in detecting EFT deviations in the SSWW signature, a comprehensive analysis incorporating the additional SM background sources and employing fully reconstructed events is essential to evaluate the practical sensitivity of the proposed approach. Further investigations in this direction will be carried out, encompassing the primary sources of background for the $2l2\nu + 2j$ final state: these include for example events with nonprompt or fake leptons, as well as the relatively rare QCD-induced production.

References

[1] ATLAS Collaboration, Phys.Lett.B **716** (2012) 1-29.
[2] CMS Collaboration, Phys.Lett.B **716** (2012) 30-61.
[3] Perez Adan D (on behalf of the ATLAS and CMS Collaborations), Rencontres de Moriond 2022: Proceedings of the ElectroWeak Session (2022, La Thuile, Italy).
[4] LHCb Collaboration, Eur. Phys. J. C **83** (2023) 543.
[5] Koren S, *arXiv e-prints* (2020) 2009.11870v1.
[6] CMS Collaboration, Eur. Phys. J. C **81** (2021) 629.
[7] Kasieczka G et al., Rep. Prog. Phys. (2021) **84** 124201.
[8] Ellis J, Contribution to the Proceedings of the BSM-2021 Conference (2021, Zewail City, Egypt).
[9] Kingma D P, Welling M, *arXiv e-prints* (2013) 1312.6114v11.
[10] Kingma D P, Welling M, Foundations and Trends in Machine Learning: Vol. **12** (2019): No. 4.
[11] Franzosi D B, Gallinaro M et al., Rev.Phys. **8** (2022) 100071.
[12] Buchmuller W, Wyler D, Nucl.Phys.B **268** (1986) 621–653.
[13] Degrande C et al., Annals Phys. **335** (2013) 21.
[14] Brivio I, Trott M, Phys. Rept. **793** (2018) 1-98.
[15] Grzadkowski B et al., JHEP **10** (2010) 85.
[16] Brivio I, JHEP **04** (2021) 73.
[17] Alwall J et al., JHEP 07 (2014) **079**.
[18] Pedregosa F et al., JMLR **12** (2011) 2825-2830.
[19] Abadi M et al., *arXiv e-prints* (2016) 1603.04467v2.