International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319-6734, ISSN (Print): 2319-6726
www.ijesi.org ||Volume 10 Issue 10 Series II || October 2021 || PP 30-39

Application of Machine Learning techniques in indirect searches for New Physics in the Effective Field Theory approach

Konstantinos Bachas[1,2], Ioannis Karkanias[1], Eirini Kasimi[1], Christos Leonidis[1], Chara Petridou[1], Katerina Zachariadou[3]

[1](*Physics Department, Aristotle University of Thessaloniki, Greece)*
[2](*Physics Department, University of Thessaly, Greece)*
[3](*Department of Electrical and Electronics Engineering, University of West Attica, Greece)*

ABSTRACT : *The electroweak (EWK) production of self-interacting $W^{\pm}Z$ bosons in association with two jets process has already been observed by the ATLAS and CMS Collaborations, at the LHC, while the existing full Run 2 data corresponding to an integrated luminosity of 139 fb^{-1} and the prospects of the forthcoming Run 3, provide the foreground for improved sensitivity searches for New Physics in a model independent way. Such searches can be realized in the context of an extension of the Standard Model (SM) in terms of a Effective Field Theory (EFT) formalism, providing a way to quantify possible deviations from SM. Such deviations have been estimated so far with traditional methods based on total cross section and kinematical distributions. In this publication an attempt is made to search for New Physics effects in the $W^{\pm}Zjj$ production, using state-of-the-art machine learning models where diverse network architectures are effectively combined into ensembles trained on the outcomes of base learners maximizing performance. The base learners are trained to identify pure $W^{\pm}Zjj$ signal events originating from the effect of EFT operators, from $W^{\pm}Zjj$ background events originating from strong (QCD) or EWK $W^{\pm}Zjj$ processes. We investigate the utilization of the ensemble model response in estimating the sensitivity of $W^{\pm}Zjj$ events in some of the dimension-8 EFT operators and compare the results to sensitive kinematic variables traditionally used to constrain the EFT operator effects.*

KEYWORDS – *Dimension-8 operators, Effective Field Theories, Machine Learning, Vector Boson Scattering, $W^{\pm}Zjj$ channel*

Date of Submission: 08-10-2021 Date of Acceptance: 22-10-2021

I. INTRODUCTION

The discovery of the Higgs boson at Large Hadron Collider (LHC) at CERN [1, 2] has highlighted the importance of understanding the details of the Electroweak Symmetry Breaking mechanism (EWSB) as one of the most pressing issues to investigate further and confirm the Standard Model of the Electroweak interactions. The EWSB mechanism is responsible for the mass acquired by the W and Z vector bosons through their couplings to the Higgs boson. It is also responsible for maintaining unitarity in the production cross sections of vector boson production through self-interactions up to the TeV scale. In this regard, measurements of self-interaction between the vector gauge bosons, involving three or four gauge bosons (manifesting themselves through Triple or Quartic gauge-boson couplings, (TGCs and QGCs), or as s- or t-channel interactions with the Higgs boson) are of great importance. In particular, the rare interactions involving the scattering of two massive vector bosons (Vector Boson Scattering, VBS) provide a relatively clean and rich ground for such measurements. Deviations of TGCs and QGCs or s- and t-channel couplings to Higgs from Standard Model (SM) expectations, in particular at TeV energy scales, as an indirect way to search for New Physics is one of the challenges currently addressed by both ATLAS and CMS experiments at LHC and will be further explored in the future with the upgraded LHC accelerator (after its Phase 1 and Phase 2 upgrades).

At the LHC, two gauge bosons and two jets (*VVjj*) can be produced via two classes of mechanisms. The first class, referred here as QCD mediated production, involves both strong and electroweak interactions. The second-class, named electroweak mediated production, involves only weak interactions and includes mostly VBS Feynman diagrams (diagrams with self-interacting gauge bosons and associated production of two jets).

The first observation of the Vector Boson Scattering has been reported at 13 TeV by ATLAS and CMS experiments, with the luminosities of 36 fb^{-1} and 139 fb^{-1} (the full Run 2 luminosity), respectively, for two

production channels: the $W^{\pm}Z$ fully leptonic, accompanied by two jets ($W^{\pm}Zjj$) and the same sign WW leptonic, accompanied by two jets ($ssWWjj$) [3-6].

The rarity of the VBS processes (most of them with a cross section of order less than one fb) and their importance to investigate the validity of the Standard Model down to the sub-femto barn region make these challenging measurements very interesting from a theoretical point of view; to motivate, deliver and check the higher-order corrections to the Standard Model predictions and to provide an EFT description of the physics beyond the S.M. In Figure 1 representative Feynman diagrams show the electroweak production of a pair of W and Z bosons associated with two jets.

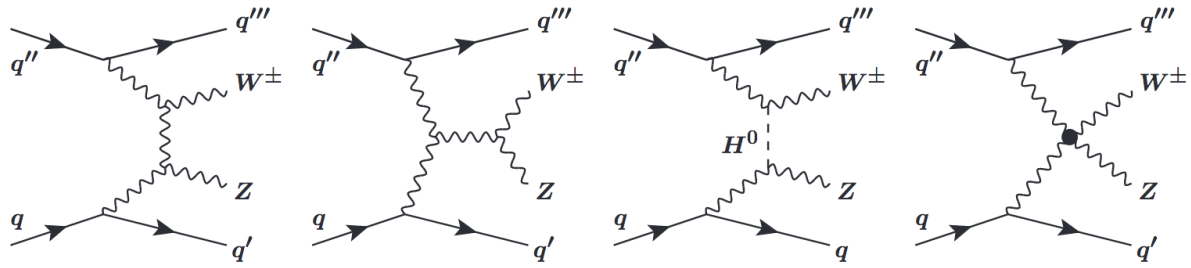

Figure 1: Feynman diagrams of the vector boson scattering process at the LHC in the $W^{\pm}Zjj$ final state, including triple and quartic gauge boson vertices as well as the Higgs boson exchange diagrams.

An attractive feature of VBS is the appearance of quartic self-couplings between the gauge bosons (QGCs), which provides the possibility for a theoretical interpretation of the VBS data in terms of anomalous QGCs (aQGCs), as well.

The Standard Model Effective Field Theory (SMEFT) [7] is a theoretical framework that describes beyond the Standard Model effects, which introduce new-physics states at a mass scale Λ, large compared to the electroweak scale. In the EFT description of the VBS processes, it is possible to construct an effective Lagrangian with dimension-8 operators that provide modifications to the VBS production cross sections via the presence of aQGCs [8].

It is expected that the presence of aQGCs affect both the fiducial cross section of the VBS processes and the shape of the distributions of kinematical variables. The more significant the effect on the shape of the distributions is, the stronger the constraints on anomalous couplings, or the greater the probability of unraveling the presence of New Physics in the data. Therefore, it is desirable to devise a kinematical variable or a combination of them that will be most affected by the presence of anomalous couplings.

Recently, ATLAS and CMS experiments have made an enormous effort to combine limits on EFT couplings from all relevant processes, utilizing the most sensitive variables into global fits to data. The goal is to profit from the complementarity between processes for different operators and combine results on operator limits across experiments based on the complete LHC Run-2 data to either maximize the constraints on the anomalous couplings or, if they exist, quantify their values.

Given this goal, the basis of this research is to suggest improvements in searching for anomalous coupling effects by introducing Machine Learning (ML) techniques and in particular, through the utilization of an ML classifier response, to distinguish between events of the SM $W^{\pm}Zjj$ production and events due to different dimension-8 EFT operator effects. The focus is to study the $W^{\pm}Z$ production in the fully leptonic decay mode, associated with two jets.

The paper is organized as follows: in Section 2 the EFT framework and the relevant dimension-8 operators used in this publication are briefly described. A short description of the decomposition method utilized for the production of the various signal samples is given. In addition, the methodology for the EFT couplings' limit extraction is presented, and the procedure for producing the Monte Carlo (MC) SM and EFT samples is

explained. The fiducial phase space where the events are selected, which resembles the one used by the ATLAS experiment for the same process, is also described. Section 3 describes the Machine Learning procedures, followed for the training models on events from signal and background processes, and gives a brief overview of their performance. Section 4 describes the statistical model used for limit extraction on different dimension-8 operators, while in Section 5, the results on EFT couplings limits for different dimension-8 operators using both the ML model score distribution and various sensitive kinematical variables are presented.

II. THE EFT MODEL AND THE MONTE CARLO SAMPLES

2.1 THE EFT MODEL

The EFT Lagrangian can be written as an expansion in inverse distance (or equivalently in energy), where the first terms that conserve baryon and lepton-number have coefficients quadratic in the distance. Consequently, the corresponding field operators are dimension-6 in energy and the subsequent relevant for LHC are dimension-8 operators (fouth-power in distance/energy). Therefore, the effective Lagrangian can be written in terms of higher dimension operators and their respective Wilson coefficients as:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} O_i + \sum_j \frac{c_j^{(8)}}{\Lambda^4} O_j + \dots \quad (1)$$

where, $O^{i,j}$ are the i, j dimension-6, 8 operators respectively and involve SM fields with respective couplings $c_i^{(6)}$ and $c_j^{(8)}$, while Λ is the energy scale where the new processes turn-on. For simplicity, as coefficients we use the simplified parameters $f_i^{(6)}=c_i/\Lambda^2$ *and* $f_j^{(8)}=c_j/\Lambda^4$ for the dimension-6 and 8 operators respectively (Wilson coefficients). It is important to note that the energy scale E of the considered process must be $E < \Lambda$.

2.2 MONTE CARLO SAMPLES

The study of the effect of the dimension-8 operators in the $W^{\pm}Zjj$ process and the extraction of limits for the couplings, require large amounts of Monte Carlo samples in a dense grid of the parameter space. However, instead of following this resource-intensive procedure, one can profit from the decomposition method implemented in the MadGraph event generator [9] to circumvent the technical requirement of the dense grid in the parameter space. The following paragraphs briefly explain the decomposition method.

The EFT events used in the current study have been produced using the Eboli-Garcia model [8], which is implemented in the MadGraph event generator. Each EFT sample represents events that are the outcome of a single dimension-8 EFT operator at a given parameter value, while samples were produced for each of the relevant for the given process, EFT operators. In this paper, only contributions due to the quadratic term of the EFT Lagrangian are considered, as explained below.

2.2.1 THE EBOLI-GARCIA MODEL

The Eboli-Garcia model describes the anomalous quartic interactions using dimension-8 effective operators at leading order, assuming that the recently observed Higgs boson belongs to a $SU(2)_L$ doublet and is an extension of the Standard Model.

The dimension-8 operators are divided into three categories. They contain the following operators:

1) Operators that consist of four covariant derivatives of the Higgs field, of the Scalar type (O_S) named as $O_{S0,1}$ as two are the ones that may affect the $W^{\pm}Zjj$ process

2) Operators that contain two Higgs covariant derivatives and two field strength tensors, being of the Mixed Type (O_M) named as $O_{M0,1,2,3,4,5,6,7}$, as eight are the ones that may affect the $W^{\pm}Zjj$ process

3) Operators with four field strength tensors, being of the Tensor type (O_T) named as $O_{T0,1,2,5,6,7,8,9}$, as ten are the ones that may affect the $W^{\pm}Zjj$ process

Accordingly, the f_S, f_M, and f_T are the corresponding Wilson coefficients to the O_S, O_M, and O_T operators. This paper presents the study for the most sensitive operators O_{S1}, $O_{M0,1}$ and $O_{T0,1,2}$ for the $W^{\pm}Zjj$ process.

2.2.2 THE DECOMPOSITION METHOD

The amplitude of a process described with an EFT Lagrangian can be written as:

$$|A_{SM} + \sum_i c_i A_i| \quad (2)$$

given that in the EFT approach, the operators having dimension greater than four are added as extra terms in an expansion around the Standard Model Lagrangian. The A_{SM} is the SM amplitude while the A_i*'s* are amplitudes containing the individual higher dimension operators. For processes like the $W^{\pm}Zjj$ production, where we can assume that dimension-6 operators contribute very little, the amplitude expansion for the process can be approximated with dimension-8 operators only. The total squared amplitude at the EFT point i, is then given by:

$$|A_{SM} + \sum_i c_i A_i|^2 = |A_{SM}|^2 + \sum_i c_i 2Re(A_{SM} A_i) + \sum_i c_i^2 |A_i|^2 + \sum_{ij, i \neq j} c_i c_j 2Re(A_i A_j) \quad (3)$$

where $\Sigma_i\ c_i 2Re(A_{SM}A_i)$ is the amplitude of the interference between the SM and the EFT operator, named interference term, $\Sigma_i\ c_i^2|A_i|^2$ is the pure EFT operator contribution, which is called quadratic term and $\Sigma_{ij,\, i \neq j}\ c_i c_j 2Re(A_iA_j)$ is the amplitude of the interference between two EFT operators, which is called cross term. In the case of dimension-8 operators the contribution of the SM-EFT interference term to the total and differential cross sections was found to be less than 1‰ and therefore the contribution from this term was omitted in the representation of the "signal" events.

2.3 FIDUCIAL PHASE SPACE

In the current study, in order to be as realistic as possible, an attempt is made to generate events as close to those reconstructed by the ATLAS detector. To this end events are generated at particle level using the PYTHIA [10] showering model with the ATLAS tune and the fiducial phase space as defined by the ATLAS selection criteria. Furthermore, the so-called "dressed" kinematics of the final state charged leptons is used, accounting for the effect of final state QED radiation. This is done by adding to the generated lepton the energy from radiated photons within a $\Delta R < 0.1$ cone around the lepton. Leptons originating from a τ-lepton decay are not considered. The dressed leptons are matched to the boson they originate from through the so-called "resonant-shape" algorithm [11], where leptons are either associated to the *W* or *Z* boson depending on the value of the appropriate estimator. All possible combinations of two same-flavour, opposite-charge leptons are considered as potential *Z* boson decay products, with the remaining lepton associated to the *W* boson, while the configuration yielding the highest estimator value, as in [11], is kept as the chosen assignment.

Jets are reconstructed using the *anti-k_T* [12] algorithm from all stable particles within a radius parameter $R = 0.4$ from the seed parton, excluding particles associated with the *W* and *Z* decays. At least two particle level jets with $p_T > 40\ GeV$ and $|\eta_j| < 4.5$ are required. The angular distance between all selected leptons and jets is required to be $\Delta R(j,l) > 0.3$. If the $\Delta R(j,l)$ requirement is not satisfied, the jet is discarded. The invariant mass, m_{jj}, of the two highest-p_T jets in opposite hemi-spheres, $\eta_{j1} \cdot \eta_{j2} < 0$, is required to be: $m_{jj} > 500\ GeV$, in order to enhance the sensitivity to the $W^{\pm}Zjj$ process. These two jets are referred to as tagging jets. The fiducial phase space definition is summarized in Table 1.

Table 1: Phase-space definitions as used for the fiducial $W^{\pm}Zjj$ cross-section measurements by the ATLAS experiment in reference [4].

Variable	Fiducial $WZjj$
Lepton $\lvert\eta\rvert$	< 2.5
p_T of ℓ_Z, p_T of ℓ_W [GeV]	$> 15, > 20$
m_Z range [GeV]	$\lvert m_Z - m_Z^{\mathrm{PDG}}\rvert < 10$
m_{T}^W [GeV]	> 30
$\Delta R(\ell_Z^-, \ell_Z^+)$, $\Delta R(\ell_Z, \ell_W)$	$> 0.2, > 0.3$
p_T two leading jets [GeV]	> 40
$\lvert\eta_j\rvert$ two leading jets	< 4.5
Jet multiplicity	≥ 2
$\eta_{j1} \cdot \eta_{j1}$	< 0
m_{jj} [GeV]	> 500
$\Delta R(j, \ell)$	> 0.3
$N_{b-\mathrm{quark}}$	$= 0$

III. THE MACHINE LEARNING CLASSIFIERS PROCEDURES

In this paper we investigate the effect of different EFT dimension-8 operators on the $W^{\pm}Zjj$ process and use a Machine Learning approach in the $W^{\pm}Zjj$ VBS region to tackle a binary classification problem, that is to distinguish events because of EFT effects from SM events. The goal is to build an ML classifier response distribution and use it as a template to eventually fit the data and set limits on EFT couplings, improving if possible the current sensitivity which is obtained from templates of traditional variables like M_T^{WZ}, M_{WZ} or p_T^Z. The steps followed towards achieving this goal are:

- Select events at MC generator level in the $W^{\pm}Zjj$ VBS phase space following the existing analysis procedures published by the ATLAS collaboration for this physics process as in reference [4]. This is realized by the use of the so-called Rivet routine [13] to obtain the fiducial phase space used by ATLAS as described above.
- Train a set of diverse ML classifiers and obtain their response (score distribution)
- Create Asimov data that correspond to an integrated luminosity 139 fb^{-1} which is equivalent to the full integrated luminosity collected by the ATLAS experiment during Run 2
- Use the score distribution of each of the EFT coupling values as the discriminant variable and perform a template fit to the Asimov data in order to obtain limits on each of the EFT couplings, respectively.
- Get limits on different EFT operators and at the same time compare with traditional kinematical variables sensitive to QGCs.

3.1 MACHINE LEARNING MODEL ARCHITECTURES AND THE TRAINING PROCEDURE

Two sets of ML classifiers are trained. Each set of classifiers comprises 5 families of diverse model architectures. Specifically, the ML model families utilized are

1. Deep Neural Net
2. XGBoost GBM (Gradient Boosting Machine)
3. GLMs (Generalized Linear Models)
4. Random Forest
5. XRT (Extremely Randomized Trees)

Both sets use events selected in the signal region of the VBS phase space of the $W^{\pm}Zjj$ process. The first classifier set is trained with events, which are odd numbered in the list of events while the second classifier set is trained on the complementary even numbered events.

For each of these base models we perform a hyper-parameter search to obtain a setup that exhibits the highest performance in terms of Area Under Curve (AUC) metric. The best model of each family is retained and an ensemble model is built out of these best 5 models.

The Ensemble Model uses the so-called stacking technique [14] to find the optimal combination of the base learners. It uses a meta-learning algorithm to learn how to best combine the predictions from the individual models.

Finally, for each set of classifiers, out of the best performing base models and the ensemble model we retain the one, which has the best AUC. This model from each set is then applied on the complementary set of events to obtain the score, that is the even numbered events for the 1st set and the odd numbered events for the 2nd set. In this way, we ensure that no events from the training set of events are used to evaluate the model performance on the final sample.

From the machine learning perspective this is a binary classification problem. One class comprises EFT events generated with MadGraph taking into account only the quadratic term of the EFT amplitude for 3 different dimension-8 EFT operators, while the other class comprises events from the SM processes of the EWK and QCD production of WZ_{jj}. The effect from the linear term in not taken into account because its contribution has been checked and is found to be negligible compared to the quadratic term.

Each sample is split in even and odd numbered events as described above, therefore the training sample for each of the two sets of models is 50% of the total and the rest 50% is held out for testing the model performance and obtaining the score distribution which will then used as the template for the fit to the Asimov data. No events from the test sample are used in the training process. In the training sets, 80% of the events is used for the training, and 20% is held out for the internal validation and tuning of the model hyper-parameters.

3.2 INPUT DATA TO THE ML MODEL TRAINING PROCEDURE

The inputs to the ML model training process comprise lepton, jet and boson kinematical variables at particle level. In particular, the following variables are used:

- The 4-momentum, pseudorapidity (η) and azimuthal *(φ)* angle of the leptons (electrons and/or muons) from the Z and the W bosons decay
- the 4-momentum and angle of the two tagging jets
- the invariant mass, transverse momentum, and of the Z boson
- the transverse mass, transverse momentum, and of the W boson
- the transverse mass, the invariant mass and of the *WZ* system
- the mass, and rapidity (*y)* of the di-jet system
- the scalar sum of the transverse momentum of the leptons
- the number of jets in the event

Example distributions of the variables above normalized to the same number of events are shown for comparison of their shapes in Figure 2.

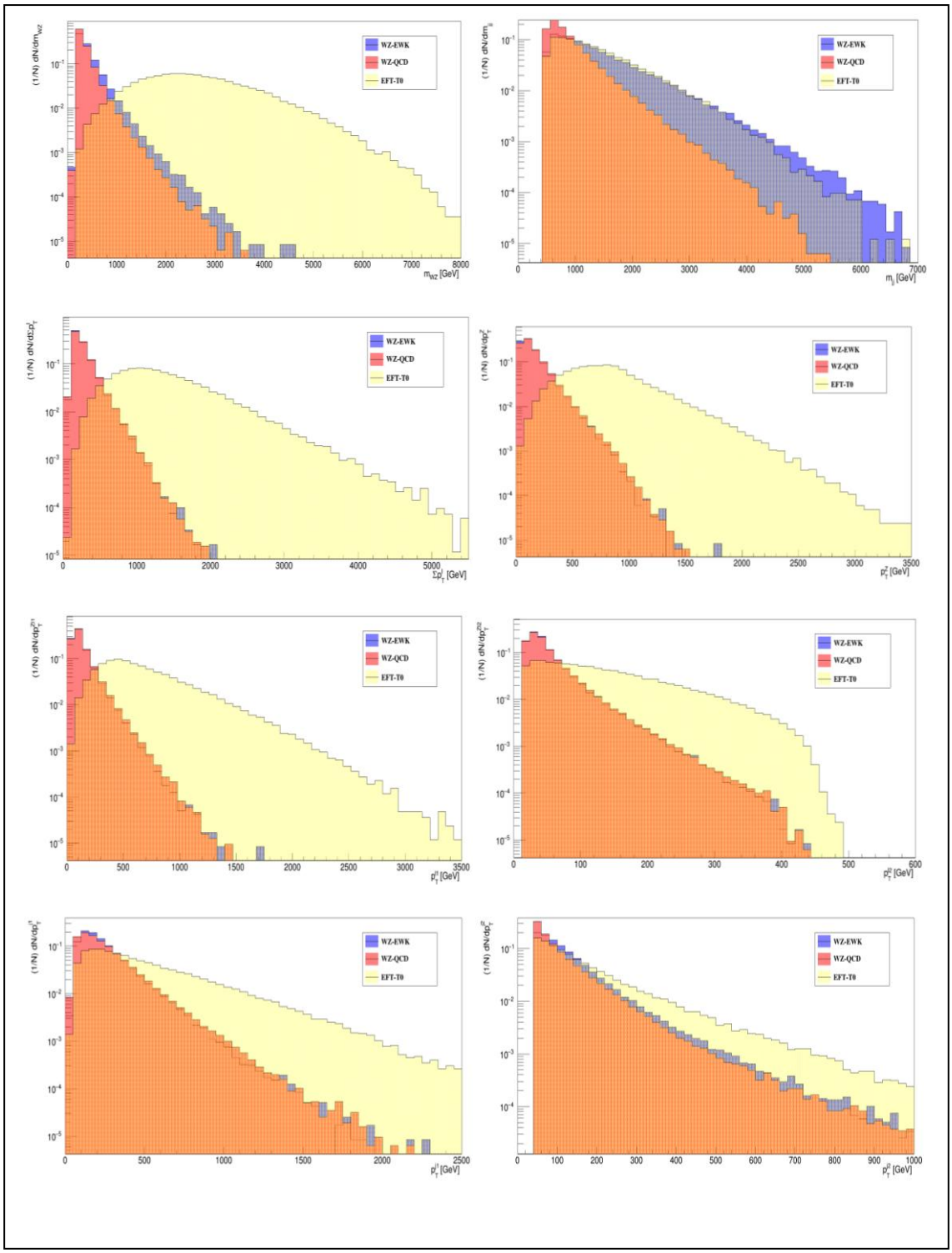

Figure 2: Kinematical variables used as input to the ML models. From top left to right m_{WZ}, m_{jj}, $\Sigma p_T^{\,l}$, $p_T^{\,Z}$, p_T of the highest p_T (leading) lepton, p_T of the second highest p_T (sub-leading) lepton, p_T of the leading jet ($p_T^{\,j1}$) and p_T of the sub-leading jet ($p_T^{\,j2}$). Shape comparison between the SM QCD and EWK produced $W^{\pm}Zjj$ processes and the one with non-zero Wilson coefficient for the O_{T0} operator.

3.3 PERFORMANCE OF THE ML MODELS

All base models and the ensemble model in each classifier's set achieve a very high AUC and logistic loss metric performance. This is expected because of the already well separation between the EWK+QCD versus the EFT events in several variables as shown in Figure 2. In Figures 3a and 3b, the performance for the best performing model of each family is shown from the best performing set of classifiers, however, it should be noted that the same performance is achieved in both sets.

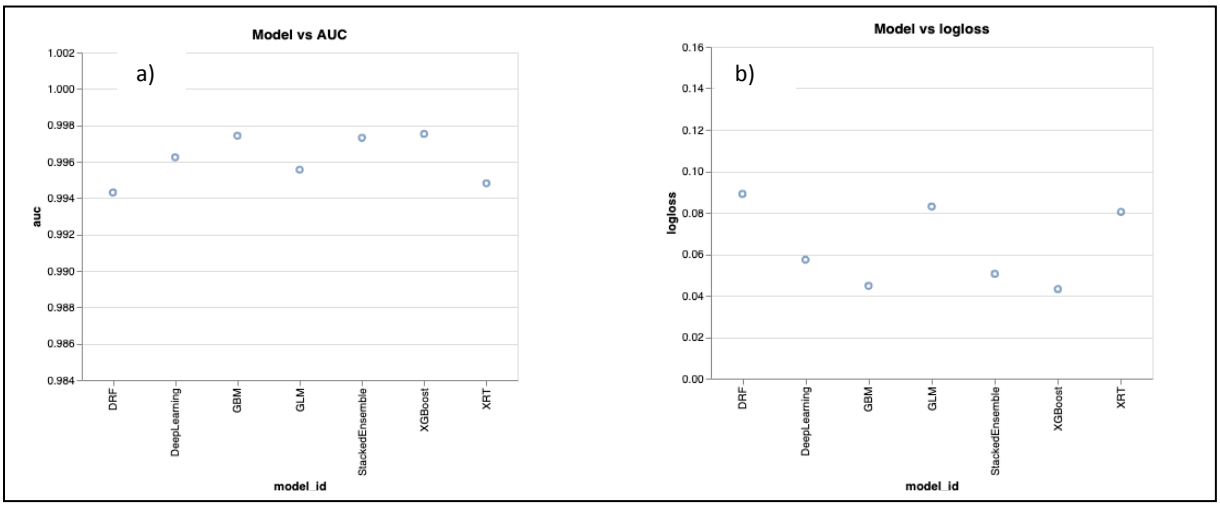

Figure 3: Performance of the different ML models in terms of AUC (a) and logistic loss (b) metrics.

The score distribution obtained from the best performing model in each of the two sets of classifiers applied on the corresponding two test sample events in the *WZ-EWK*, *WZ-QCD* and EFT for the operator T_0 quadratic term sample, are shown in Figure 4. The score distribution shows a very clean separation between the two classes of events. The same kind of separation in the score distribution is obtained also for other operators operators for which the same training procedure has been followed. This score distribution is the template which is used for the statistical fit to the Asimov data in order to obtain the limits on the EFT couplings as described in the next section.

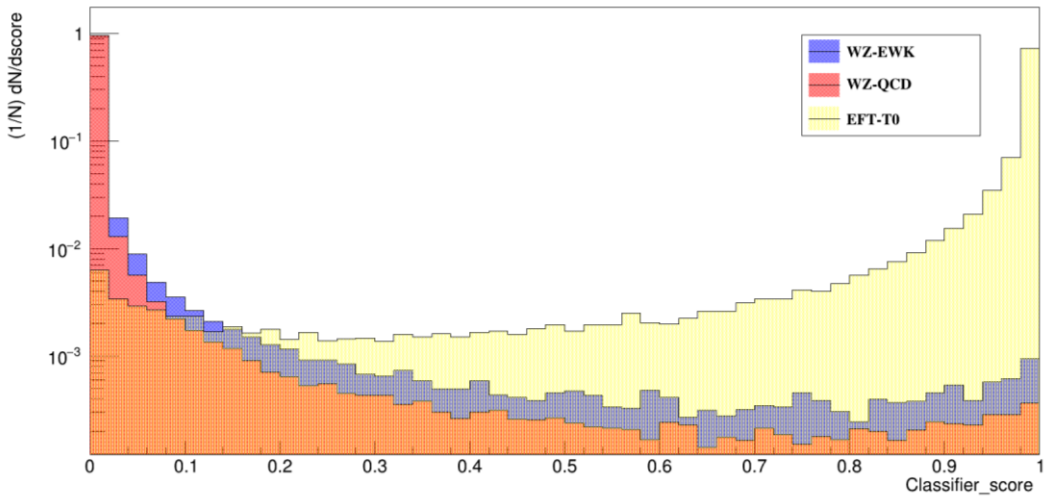

Figure 4: The classifier score distribution for events from the *WZ-EWK*, *WZ-QCD* and *EFT* for the *T0* operator samples

IV. STATISTICAL MODEL

The binned likelihood function used to extract expected limits from differential cross-section distributions is based on a multivariate Gaussian distribution. Standard Model generated Monte Carlo events, at truth level, are compared to respective differential cross sections of various dimension-8 EFT operators (one at a time) in order to study the sensitivity of various kinematical variables to EFTs as well as the sensitivity of the

score variable, built based on Machine Learning techniques, as explained in the previous section. This reinterpretation of the various differential distributions of the electroweak $W^{\pm}Zjj$ process and the improvements found with the use of the ML score variable are shown in section 5. In the likelihood function the experimental uncertainties -although not used at the current stage- are encoded in a covariance matrix, while theory uncertainties can be introduced as nuisance parameters constrained by Gaussian distributions.

The prediction of the EFT differential cross sections depends on the set of Wilson coefficients c, according to the decomposition method, and is also subject to theory systematic uncertainties, which are parametrized by nuisance parameters. The predicted fiducial cross section x_b^{pred} in a bin b of the differential distribution is parametrized as:

$$x_b^{\mathrm{pred}}(c,\theta) = x_b^{\mathrm{SM}}\left(1 + \sum_i \frac{c_i x_i^{int}(\theta)}{x_b^{\mathrm{SM}}} + \sum_i \frac{c_i^2 x_i^{quad}(\theta)}{x_b^{\mathrm{SM}}} + \sum_{i\neq j} \frac{c_i c_j x_{ij}^{cross}(\theta)}{x_b^{\mathrm{SM}}}\right) \times \prod_i^{n_{syst}} \left(1 + \theta_j u_j^b\right) \quad (4)$$

where c are the Wilson coefficients, $\theta=(\theta_{1,2.....}, \theta_{nsyst})$ are nuisance parameters, n_{syst} is the number of nuisance parameters, x_{SM}^{bis} the nominal SM cross prediction, and u_j^b the relative size of the theory uncertainty j in bin b. The complete likelihood function is given by:

$$L(x|c,\theta) = \frac{1}{\sqrt{(2\pi)^{n_{bins}}\, det(C)}} \exp\left(-\frac{1}{2}\Delta x^T(c,\theta) C^{-1} \Delta x(c,\theta)\right) \times \prod_i^{n_{syst}} g_i(\theta_i) \quad (5)$$

where x is the nominal (expected) Standard Model differential cross sections of the $W^{\pm}Zjj$ process, C is the covariance matrix which represents the correlation between the statistical and systematic uncertainties of the differential cross-section distributions (to be obtained in the future by unfolded data distributions and their statistical and systematic uncertainties), g_i correspond to the Gaussian constraints on nuisance parameters and Δx represents the difference between measurement and prediction and its components. Δx^b is the difference between predicted and measured cross section:

$$\Delta x^b = x_{\mathrm{meas}}^b - x_{\mathrm{pred}}^b(c,\theta) \quad (6)$$

In order to estimate the confidence interval for a Wilson coefficient ci, a profile likelihood ratio test statistics is constructed from the likelihood:

$$\lambda(c_i) = -2log\frac{L\left(c_i,\hat{\hat{\theta}}\right)}{L(\hat{c_i},\hat{\theta})} \quad (7)$$

where $L\left(c_i,\hat{\hat{\theta}}\right)$ is the maximum of the likelihood for fixed c_i and $L(\hat{c_i},\hat{\theta})$ is the value at the absolute maximum of the likelihood. Maximum likelihood fits are performed for individual Wilson coefficients by setting other coefficients to zero and maximizing the likelihood with respect to the nuisance parameters. Confidence intervals are derived using Wilks' theorem [15], assuming that $\lambda(c_i)$ is χ^2 distributed. In this paper, only statistical uncertainties corresponding to the full Run 2 statistics are considered.

V. RESULTS

Several kinematical variables were studied for all dimension-8 operators that are expected to modify the $W^{\pm}Zjj$ differential cross sections, and the most affected by the presence of EFTs are selected for limit extraction. In Figure 5, differential cross sections are presented for the transverse mass M_T^{WZ} of the $W^{\pm}Z$ system, the scalar sum of the transverse momentum of the three leptons, Σp_T^l, the azimuthal angle separation between the W and Z bosons, $\Delta\varphi^{WZ}$, and the rapidity separation between the two tagging jets Δy_{jj}. The expected differential cross section based on the Standard Model is compared to the respective ones comprising contributions from the f_{T0} operator for two different coupling coefficient values. One where $f_{T0} = 2.4$ and the other where, $f_{T0} = 4.8$. It is interesting to note that, as expected, the higher the value of the coefficient the higher the departure of the differential cross sections from the SM expectations.

Expected limits on the EFT couplings were evaluated for each kinematical variable and each operator coefficient separately, setting all other Wilson coefficients to zero. The results from the different kinematical variables are then compared with the results obtained from the fit to the MC events using the Machine Learning score distribution shown in Figure 5. These results are presented in Table 2 for the integrated luminosity of 139 fb^{-1}. It is remarkable the level of improvement for the obtained limits that the ML score variable provides for the Run 2 integrated luminosity, which is about a factor of two better with respect to the best limit obtained using traditional kinematical variables.

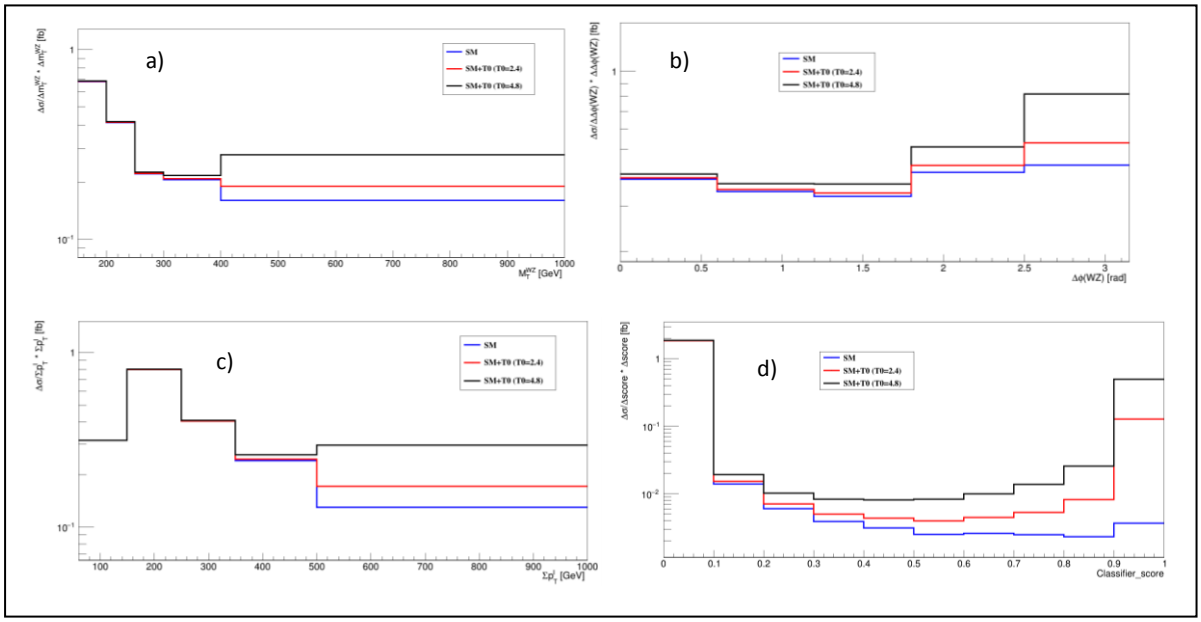

Figure 5: Differential cross section as a function of a) M_T^{WZ}, b) $\Delta\varphi^{WZ}$ c) the lepton scalar sum of the lepton p_T and d) the ML classifier score.

Table 2: Expected limits for the most sensitive operators of the $W^{\pm}Zjj$ process, for the full Run 2 luminosity of 139 fb^{-1}. Limits are presented for all sensitive kinematical variable and for the ML model (last column)

Wilson Coeff.	M_T^{WZ}	Σp_{Tl}	$\Delta\varphi^{WZ}$	Δy_{jj}	*ML Model*
f_{M0}	[-14.09, 14.09]	[-16.82,16.82]	[-28.36, 28.36]	[-28.14, 28.14]	[-7.67, 7.67]
f_{M1}	[-21.80, 21.80]	[-26.36, 26.36]	[-43.50, 43.50]	[-43.56, 43.56]	[-12.01, 12.01]
f_{S1}	[-77.36, 77.36]	[-96.16, 96.16]	[-142.64,142.64]	[-131.76, 131.76]	[-39.32, 39.32]
f_{T0}	[-1.36, 1.36]	[-1.61, 1.61]	[-2.82, 2.82]	[-3.00, 3.00]	[-0.73, 0.73]
f_{T1}	[-0.93, 0.93]	[-1.12, 1.12]	[-1.87, 1.87]	[-2.11, 2.11]	[-0.47, 0.47]
f_{T2}	[-2.70, 2.70]	[-3.26, 3.26]	[-5.49, 5.49]	[-6.11, 6.11]	[-1.39, 1.39]

VI. CONCLUSION

In this paper, we have studied the expected EFT effects in the electroweak production of the $W^{\pm}Zjj$ process and have extracted expected limits on the EFT couplings for the six most sensitive dimension-8 operators for this channel.

The study is realized at particle level Monte Carlo generated events, after PYTHIA showering. Limits are obtained through a template fit to the Standard Model MC events using differential cross section distributions of traditional kinematical variables like the transverse mass M_T^{WZ} of the $W^{\pm}Z$ system, the scalar sum of the transverse momentum of the three leptons Σp_T^l, the azimuthal angle separation between the W and Z bosons, $\Delta\varphi^{WZ}$, and the rapidity separation between the two tagging jets Δy_{jj}. All chosen variables are the most sensitive in one or more dimension-8 operators. In comparison to the results obtained from these variables, we studied the sensitivity of score templates constructed by state-of-the-art Machine Learning classifiers, which have been trained to distinguish between EFT and SM events. The types of ML models used are based on five model families and one ensemble model build from the best of each family. In the fiducial phase space as defined by ATLAS it has been found that using the ML score as a template to fit the data, gives much better limits on EFT couplings than the most sensitive variables for dimension-8. This study provides an interesting ground for applying similar techniques to the current Run 2 and the forthcoming Run 3 LHC data, in order to dramatically increase the sensitivity to unravel possible New Physics effects beyond the Standard Model, in the near future.

Acknowledgements

This research is carried out/funded in the context of the project "Search for new physics with the Atlas experiment at the LHC with indirect methods utilizing novel statistical analysis techniques and in the context of Effective Field Theories" under the call for proposals "Supporting researchers with emphasis on new researchers" (EDULLL 34). The project is co-financed by Greece and the European Union (European Social Fund-ESF) by the Operational Programme Human Resources Development, Education and Lifelong Learning

2014-2020. The authors would also like to thank kindly the ATLAS Experiment for the long term participation and collaboration.

REFERENCES

[1]. ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B, 716:1*, 2012.

[2]. CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B, 716:30*, 2012.

[3]. ATLAS Collaboration. Observation of Electroweak Production of a Same-SignW Boson Pair in Association with Two Jets in pp Collisions at √s= 13 TeV with the ATLAS Detector.Phys. *Rev.Lett., 123:161801*, 2019.

[4]. ATLAS Collaboration. Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at √s= 13 TeV with the ATLAS detector. *Phys. Lett. B, 793:469*, 2019.

[5]. CMS Collaboration. Measurement of electroweak WZ boson production and search for new physics in WZ+ two jets events in p collisions at √s= 13 TeV *Phys. Lett. B, 795:281*, 2019.

[6]. CMS Collaboration. Observation of Electroweak Production of Same-Sign W Boson Pairs in the Two Jet and Two Same-Sign Lepton Final State in Proton-Proton Collisions at 13 TeV. *Phys. Rev. Lett.,120:081801*, 2018.

[7]. Ćeline Degrande, Nicolas Greiner, Wolfgang Kilian, Olivier Mattelaer, Harrison Mebane, Tim Stelzer, Scott Willenbrock, and Cen Zhang. Effective field theory: A modern approach to anomalous couplings. *Annals of Physics, 335:21–32*, 2013.

[8]. O. J. P.´Eboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi. pp → j j $e^{\pm}$ $\mu^{\pm}$ ν ν and j j $e^{\pm}$ $\mu^{\mp}$ ν ν at $O(\alpha^6_{em})$ and $O(\alpha^4_{em}\alpha^2_s)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC. *Phys. Rev.D, 74:073005*, Oct 2006.

[9]. Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics, 2011(6)*, Jun 2011.

[10]. Torbjorn Sjostrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten,Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2.Comput. Phys. Commun., 191:159, 2015.

[11]. ATLAS Collaboration. Measurements of $W^{\pm}Z$ production cross sections in pp collisions at √s= 8 TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings. *Phys. Rev. D,93:092004*, 2016.

[12]. Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *JHEP,04:063*, 2008.

[13]. Christian Bierlich, Andy Buckley, Jonathan Butterworth, Christian Holm Christensen, Louie Corpe, David Grellscheid, Jan Fiete Grosse-Oetringhaus, Christian Gutschow, Przemyslaw Karczmarczyk, Jochen Klein, and et al. Robust independent validation of experiment and theory: Rivet version 3. *SciPost Physics, 8(2)*, Feb 2020.

[14]. Mark J. van der Laan, Eric C Polley, and Alan E. Hubbard. *Super learner.Statistical Applications in Genetics and Molecular Biology, 6(1)*, 2007.

[15]. S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math.Stat., 9:60–62*, 1938.

Konstantinos Bachas, et. al. "Application of Machine Learning techniques in indirect searches for New Physics in the Effective Field Theory approach ." *International Journal of Engineering Science Invention (IJESI)*, Vol. 10(10), 2021, PP 30-39. Journal DOI- 10.35629/6734