International Journal of Engineering and Science Invention (IJESI)
ISSN (Online): 2319-6734, ISSN (Print): 2319-6726
www.ijesi.org ||Volume 13 Issue 11 November 2024 || PP 07-16

Search for New Physics in the Effective Field Theory approach using Machine Learning techniques

Amperiadou Dimitra[1], Bachas Konstantinos[2], Petridou Chariclia[1], Kasimi Eirini[1]
[1]*(Physics Department, Aristotle University of Thessaloniki, Greece)*
[2]*(Physics Department, University of Thessaly, Greece)*

ABSTRACT: *The simultaneous production of $W^{\pm}$ and Z bosons through the Vector Boson Scattering (VBS) process has been studied by the ATLAS and CMS experiments with the full Run 2 data and an integrated luminosity of 140 fb^{-1} at a center-of-mass 13 TeV. This process can provide the foreground for searches beyond the Standard Model (BSM) through deviations from the predicted vector boson self-couplings. The way to search for such deviations is through an Effective Field Theory (EFT) where higher order operators, dimension-6 and dimension-8 are considered. Apart from the traditional methods based on observing deviations from the SM in the total cross section and kinematical distributions, there have already been efforts to study dimension-8 operators using state-of-the-art Machine Learning models where base learners are combined in an ensemble model. As an extension of this previous study, additional models are being trained and used in this paper, in order to identify pure $W^{\pm}Zjj$ signal events with the effect of EFT operators, from $W^{\pm}Zjj$ background events originating from strong (QCD) or electroweak (EWK) $W^{\pm}Zjj$ processes. Specifically, the Naive Bayes, the Decision Tree and the Random Forest network architectures are being investigated individually in order to study dimension-8 operators by extracting the 95 % Confidence Limits (CL) for their Wilson coefficients.*

Date of Submission: 09-11-2024 Date of Acceptance: 22-11-2024

I. INTRODUCTION

The discovery of the Higgs boson by the ATLAS and CMS experiments [1,2] at LHC complements the Standard Model (SM) and it also contributes to the understanding of the Electroweak Symmetry Breaking (EWSB), the mechanism that is responsible for the $W^{\pm}$ and Z bosons mass acquisition through their interactions with the Higgs boson. The study of this mechanism relies on the observation of diboson production of $W^{\pm}Z$ which involves the creation of a W and Z boson pair, and the Vector Boson Scattering (VBS) that refers to the scattering of these bosons in high-energy collisions. These processes appear as Triple or Quartic gauge-boson couplings, (TGCs and QGCs) and they are really rare due to their low cross sections. However, their investigation is important because of their sensitivity to the dynamics of electroweak symmetry breaking, their association to the nature of the Higgs mechanism and the fact that they can set the ground for indirect searches for New Physics.

In proton-proton collisions at the LHC, the production of two gauge bosons along with two jets (VVjj) occurs through two main types of mechanisms. The first, known as QCD-mediated production, involves both strong and electroweak interactions. The second type, called electroweak-mediated production, relies solely on weak interactions and is primarily characterized by VBS Feynman diagrams, where gauge bosons interact with each other, accompanied by the production of two jets (Figure 1).

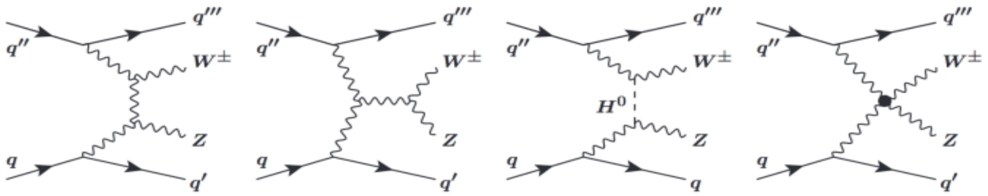


Figure 1: Feynman diagrams of the vector boson scattering process at the LHC in the $W^{\pm}Zjj$ final state, including triple and quartic gauge boson vertices as well as the Higgs boson exchange diagrams.

The VBS process in $W^{\pm}Zjj$ channel with fully leptonic final states has been already observed from ATLAS and CMS experiments at $\sqrt{\square} = 13\ TeV$ [3,4].

The evidence of New Physics through the investigation of these processes can be provided by the observation of anomalous Triple or Quadratic (aTGCs and aQGCs) gauge -boson couplings. Specifically, an Effective Field Theory (EFT) is used in order to parametrize these anomalous couplings and to search for deviations from the SM predictions [5]. In the EFT interpretation, the new physics effects are introduced at a mass scale Λ, larger than the electroweak scale, by constructing an effective Lagrangian with terms of dimension-6 and dimension-8 operators that can provide modifications to the VBS production cross sections via the presence of aTGCs or aQGCs. Also, the presence of anomalous couplings is expected to affect the shape of the distributions of the kinematical variables. In this paper, we focus on the investigation of aQGCs, so they are the dominant ones. It is important to investigate kinematical variables that are strongly affected by the appearance of aQGCs, because these will lead to stronger constraints on the anomalous couplings or to evidence for New Physics.

The goal of this paper is the study of Machine Learning (ML) models as a mean for the extraction of the 95% CL of the Wilson coefficients of dimension-8 operators in the EFT Lagrangian. Specifically, some ML models have been trained with truth level Monte Carlo (MC) samples, based on a classification problem, in order 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 results from the ML classifiers are compared to the limits extracted when a typical kinematical variable, the transverse mass of the system $W^{\pm}Z$ (M_T^{WZ}), is used.

In this paper, three different ML models are being trained and tested with MC samples. First, a binning optimization study has been performed in order to extract the binning that gives the better sensitivity for the most sensitive operator. The output of the predictions of the ML models, hereafter called the classifier score, was then used in order to perform a likelihood scan and extract limits for the Wilson coefficients of the dimension-8 operators, using the best binning from the optimization procedure.

II. THE EFFECTIVE FIELD THEORY INTERPRETATION

2.1 THE EFT MODEL

The Standard Model (SM) is the most accurate and well tested theory that we have so far which describes the particles and their interactions we observe in nature. However, it cannot describe phenomena like dark matter, neutrino masses, matter/anti-matter asymmetry, and cosmic inflation. Therefore, it should be extended in order to account for them.

The Effective Field Theory plays a crucial role in the search for New Physics providing a model-independent method assuming that the New Physics appears in higher energy scales. One of the most popular frameworks for Beyond Standard Model (BSM) searches is the Standard Model Effective Field Theory (SMEFT), which extends the theory beyond the SM by adding higher-dimension operators in the SM Lagrangian. These operators have coefficients of inverse powers of mass according to dimensional analysis and are suppressed if the mass is much larger than the experimentally accessible energies. The lowest dimensional operators will be the most dominant extensions to the theory. The effective Lagrangian can be written in terms of higher dimension operators and their respective Wilson coefficients,

$$L_{eff} = L_{SM} + \Sigma_i \frac{c_i^6}{\Lambda^2} O_i + \Sigma_j \frac{c_j^8}{\Lambda^4} O_j + \cdots \qquad (1)$$

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 appear. For simplicity, as coefficients we use the parameters $fi^{(6)}=c_i/\Lambda^2$ 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 $E < \Lambda$.

This paper mainly focuses on the dimension-8 operators and how they affect the $W^{\pm}Zjj$ VBS process in the search for possible existence of anomalous quartic gauge couplings, aQGCs. The goal is to extract the 95% CL for the Wilson coefficients of the dimension-8 operators.

2.2 DIMENSION-8 OPERATORS

In the EFT extension, anomalous quartic interactions are described 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 SM. These operators can be classified into different categories based on their field content. There are operators that come from the vector boson field strength terms, the covariant derivative acting on the Higgs field or mixed terms as well. As a result, there are three categories of interaction terms: the longitudinal (S_0, S_1, S_2), the transverse (T0, T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9) and the mixed interaction terms (M_0, M_1, M_2, M_3, M_4, M_5, M_6, M_7). These are also described in detail in [6]. In this paper the operators S_1, $M_{0,1}$ and $T_{0,1,2}$ for the $W^{\pm}Zjj$ process are going to be studied.

2.3 THE MONTE CARLO SAMPLES AND THE DECOMPOSITION METHOD

In order to study the effect of dimension-8 operators on the $W^{\pm}$ Zjj VBS process, MC samples have been produced with events from WZ-EWK process, WZ-QCD and for the EFT events used in the current study, the Eboli-Garcia model [7] has been used, which is implemented in the MadGraph 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 order to avoid generating events with different values of the Wilson coefficients for the tested operators every time, the decomposition method is used. It provides the ability to use a simple scaling based on a formula to provide the expected contribution and various values of the Wilson coefficients. It is based on the fact that we can represent the amplitude of a process described with an EFT Lagrangian as,

$$|A_{SM} + \Sigma_i c_i A_i| \qquad (2)$$

, where A_{SM} is the SM amplitude while the A_i are amplitudes containing the individual higher dimension operators. In the expansion of the amplitude, we only take into account the dimension-8 operators. The above formula can then be expanded to,

$$|A_{SM} + \Sigma_i c_i A_i|^2 = |A_{SM}|^2 + \Sigma_i c_i 2\Re(A_{SM}A_i) + \Sigma_i c_i^2 |A_i|^2 + \Sigma_{i,j,i\neq j} c_i c_j 2\Re\left(A_i A_j\right) \qquad (3)$$

, where $\Sigma_i c_i 2\Re(A_{SM}A_i)$ is the interference term between SM and the EFT operator, $\Sigma_i c_i^2 |A_i|^2$is the quadratic term of the EFT operator representing the contribution of the pure EFT operator and $\Sigma_{i,j,i\neq j} c_i c_j 2\Re\left(A_i A_j\right)$is the amplitude of the interference between two EFT operators, which is called cross term.

In this analysis, only the quadratic term has been considered since 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%.

2.4 FIDUCIAL PHASE SPACE

For the event generation, PYTHIA [8] showering model has been used with the ATLAS tune and the fiducial phase space as defined by the ATLAS selection criteria in order to make the study more realistic. The definition of the phase space is also used and well described in [9] and the table below (Table 1) provides a summary of it.

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

Table 1: Fiducial Phase Space for WZjj cross-section measurements as defined by the ATLAS experiment [3].

III. THE MACHINE LEARNING MODELS

3.1 CLASSIFICATION PROBLEMS

Classification is the process of predicting the class of given data points. Classes are sometimes called targets, labels or categories. Classification predictive modeling is the task of approximating a mapping function from input variables to discrete output variables. In such problems, the output or "response" variable is a categorical value e.g yes/no or it can also be multiclass.

In this paper, a binary classification problem is being studied where the goal is to distinguish events generated with EFT effects from SM events. For this purpose, as an extension to the work of reference [9], three ML models are being trained and their performance is being compared. The best model is then used 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}.

3.2 CLASSIFICATION ALGORITHMS IN MACHINE LEARNING

In ML, classification problems are usually solved by supervised learners, where the output variable is known. For our case, three models have been tested: the Naive Bayes Classifier, the Decision Tree and the Random Forest.

The performance of ML classifiers is evaluated using metrics of the Area Under Curve (AUC) and the LogisticLoss (logloss). In ML, AUC refers to the Area Under the Receiver Operating Characteristic (ROC) curve. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings of a classifier. The goal of the ROC curve is to assess how well a model distinguishes between two classes (e.g., positive vs. negative outcomes) given some threshold of the classifier score. The closest to 1 the AUC is, the better the model performs [10].

On the other hand, Logistic Loss, also known as Log Loss or Binary Cross-Entropy Loss, is the loss function used in binary classification tasks, especially for logistic regression models and neural networks. It measures how well a model's predicted probabilities match the actual labels of the data. For binary classification, the closest to zero the logloss is, the better the predictions from the model are [10].

The steps that have been followed for this analysis are the following:

- Select events at the MC generator level within the $W^{\pm}Zjj$ vector boson scattering (VBS) phase space, following the established analysis procedures published by the ATLAS collaboration for this process (as outlined in reference [4]). This is achieved using the Rivet routine [13] to match the fiducial phase space used by ATLAS.
- Split the data sets in two folds (event numbered and odd numbered events). One-fold is used for the training and the other for the testing of the ML models.
- Train three different ML models and evaluate their output (score distributions) and their performance metrics.
- Use an Ensemble model that uses multiple learning algorithms to obtain better predictive performance in order to perform a binning optimization study in the score distribution using all the SM processes (EWK and QCD) and the quadratic term of the T1 operator, which is the most sensitive operator to the $W^{\pm}Zjj$ process.
- Generate Asimov data corresponding to an integrated luminosity of 140 fb^{-1} which represents the total luminosity collected by the ATLAS experiment for Run 2.
- Use the score distribution for each EFT coupling as the discriminating variable between EFT events and SM events and perform a fit to the Asimov data to derive limits on each of the EFT couplings.
- Obtain constraints on different EFT operators and simultaneously compare them with traditional kinematic variables that are sensitive to quartic gauge couplings (QGCs).

3.2.1 THE NAIVE BAYES MODEL

Naive Bayes is a classification algorithm that relies on strong assumptions of the independence of covariates in applying the Bayes Theorem. In other words, each feature contributes to the predictions with no relation between each other. Although this is not the most realistic approach, the Naive Bayes Classifier is easy to implement and computationally efficient [11].

Two important parameters regarding the output of the Naive Bayes classifier are the a-priori and conditional probabilities of each class of the response. The a-priori probability is the estimated probability of a particular class before observing any of the predictors. It is based on the overall distribution of classes in the training data and calculated by a Gaussian fit P(C). The conditional probability is the probability of observing a particular feature (or evidence) given that the class is already know P(Xi|C). For our case, the a-priori probability would be the probability of an event coming from EFT effect while the conditional probability would be the probability of an event coming from EFT effect given a set of parameters, which are the kinematical variables.

After training the model, it can be used to predict the conditional probabilities for each event for the two classes (EFT and SM) using the data set for the testing. The distribution of the conditional probabilities for each event being classified in the EFT category can be used as the score, as an input for the extraction of the 95% C.L of the Wilson Coefficients.

3.2.2 THE DECISION TREE MODEL

The Decision Tree algorithm creates a tree structure where each internal node represents a test on one or more attributes. Each branch emerging from a node represents the outcome of a test, and each leaf node represents a class label or a predicted value [12].

The algorithm takes a dataset consisting of numerical features and a binary target variable. For this study, the numerical features are the kinematical variables while the binary target variable is the EFT or the SM. For optimization reasons, especially on large data sets, the algorithm applies a binning to discretize continuous features. After that, the selection thresholds are set based on some metrics for each node (e.g entropy). While splitting the data sets into smaller ones regarding the attributes and the thresholds, the goal is to end up with pure sub data sets, minimizing the entropy. The process is repeated recursively for each subset, creating a new internal node or leaf node until a stopping criterion is met (e.g., all instances in a node belong to the same class or a predefined depth is reached).

In decision trees, particularly for classification tasks, not only the predicted class of each input (event) can be predicted, but also the probabilities associated with each class. These probabilities represent the likelihood that an instance belongs to each class, based on the training data. The class with the biggest probability assigned is then the predicted class for the given input.

For this study, the extraction of the limits is based on the probabilities of the events belonging to the class labeled as “EFT”.

3.2.3 THE RANDOM FOREST MODEL

Random Forest is a powerful classification and regression tool. When given a set of data, the model generates a forest of classification or regression trees, rather than a single classification or regression tree. Each of these trees is a weak learner built on a subset of rows and columns. More trees will reduce the variance. Both classification and regression take the average prediction over all of their trees to make a final prediction, whether predicting for a class or numeric value [13].

Also in this model, the score is considered to be the probability of each event to belong in the “EFT” class after the predictions.

3.3 INPUT VARIABLES TO THE ML MODELS

The input data (x values) in the ML models are some of the kinematical variables of the process WZjj decaying to leptons. These are used in order to train the ML model during the training process and also do the predictions for the classes (y), that would be either SM or EFT. Specifically, the following variables have been used.

The variables related to the kinematics of the vector bosons W and Z:

- 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 invariant mass M, transverse momentum p_T, of the Z boson
- the transverse mass M_T^{WZ}, transverse momentum p_T of the W boson
- the transverse mass M_T^{wz}, the invariant mass M^{WZ} of the WZ system
- the scalar sum of the transverse momentum of the leptons

The variables related to the tagging jets:

- the 4-momentum and angle of the two tagging jets
- the mass m_{jj}, and rapidity y_{jj} of the di-jet system
- the number of jets N_{jets} in the event

3.4 THE TRAINING OF THE MACHINE LEARNING MODELS

The ML model with the best performance has been trained for each operator, where the data sets were split in two subsets based on the event number. The first time, the subset with the odd numbered events have been used for the training and the even numbered events for the testing and vice versa for the second time.

In Figure 2, the ROC curve and the AUC for each model from the training for T1 operator are shown. Also, Table 2 summarizes the logloss and the AUC metrics for all the models. From the results of the metrics, the Random Forest model can be considered as the classifier with the best performance. Therefore, this model will be used for the extraction of the score distributions that is going to be used for the estimation of limits of the Wilson coefficients.

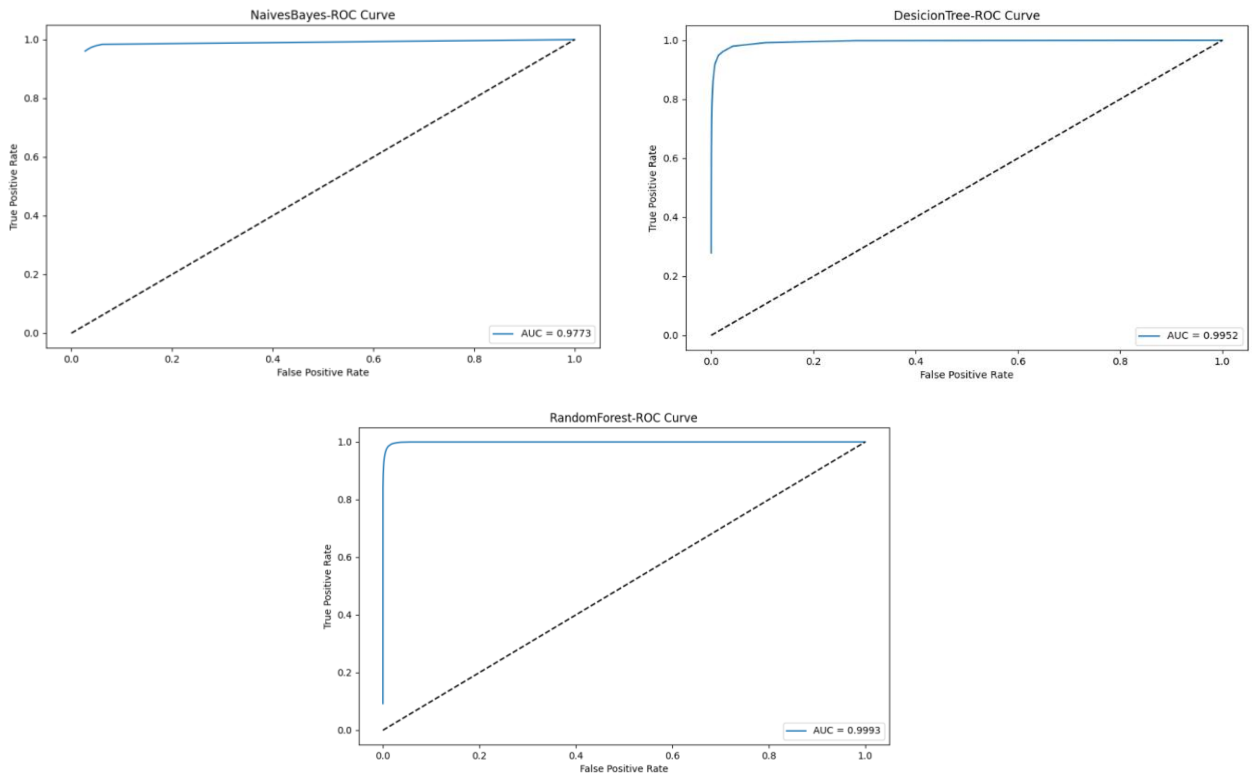


Figure 2: The ROC plots with the AUC values for the Naive Bayes (upper left), the Decision Tree (upper right) and the Random Forest (down) models.

Model	AUC	LogLoss
Naive Bayes	0.9973	0.704
Decision Tree	0.9952	0.0556
Random Forest	0.9993	0.029

Table 2: The AUC and the LogLoss of the three models from the training for the T1 operator.

IV. STATISTICAL MODEL

The expected limits for the Wilson coefficients have been extracted using the differential cross section distributions both for the M_T^{WZ} and the probabilities from the ML model with the binned Likelihood function based on a multivariate Gaussian.

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 likelihood function L(x|c, θ) representing the probability of the nominal SM differential cross-section with the given Wilson coefficient c and nuisance parameter and is given by the following formula,

$$L(x|c,\theta) = \left(\frac{1}{\sqrt{(2\pi)^{n_{bins}}\ \det(C)}}\right) e^{\left(-0.5\ \Delta x(c,\theta)^T\ C^{-1} \Delta x(c,\theta)\right)} \times\ \Pi_i{}^{n_{syst}} g_i(\theta)$$

where x is the nominal (expected) SM differential cross sections of the WZjj 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), gi 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 in every bin.

$$\Delta x^{\,b} = x_{meas}{}^{b} - x_{pred}{}^{b}$$

For the estimation of the 95% CL of the Wilson coefficients, the profile likelihood method has been used. The profile likelihood ratio test statistics is constructed from the likelihood:

$$\lambda(ci) = -2log\frac{L\left(ci, \hat{\hat{\theta}}\right)}{L\left(\hat{c_l}, \hat{\theta}\right)}$$

where $L\left(c_i\hat{\hat{\vartheta}}\right)$ is the maximum of the likelihood for a given fi and $L\left(\hat{c}_L\hat{\vartheta}\right)$ 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, assuming that λ (ci) is χ^2 distributed [14].

V. RESULTS

5.1 OPTIMAL BINNING STUDY WITH THE ENSEMBLE MODEL

The binned Profile-likelihood that is performed using the decomposition method has been shown to be heavily dependent on the binning of the histogram that is used in the fit. Hence, it is interesting to investigate which binning of the histogram of the score yields the most optimal results for the limits of the Wilson coefficients.

The ensemble model described in [9], has been used in order to perform the optimal binning study with the operator O_{T1}, which has been proven to be the most sensitive one. Ensemble ML methods use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms. It has been shown that these algorithms give the best predictions so far.

In Figure 3, the distribution of the score for the three samples is shown from the ensemble model in logarithmic scale. This distribution has been used in order to perform a templated fit in the Asimov data, in order to extract the limits of the EFT coupling. This model is the same as the one used in [9].

The various binning options that have been used to perform binned profile-likelihood are shown in the following table (Table 3) with the results of the limits for the O_{T1} operator. The binning choice resulted in the narrowest positive-to-negative 95% C.L. Wilson coefficient width is chosen as the optimized binning.

From the results shown in Table 2 it can be understood that by splitting the bins in a way that more bins are taken close to the values that the signal appears, the binned Profile Likelihood performs better. Also, by slightly changing the binning in the values close to the signal the result does not change. Therefore, the chosen binning is: {0., 0.9, 0.94, 0.96, 0.98, 1.}.

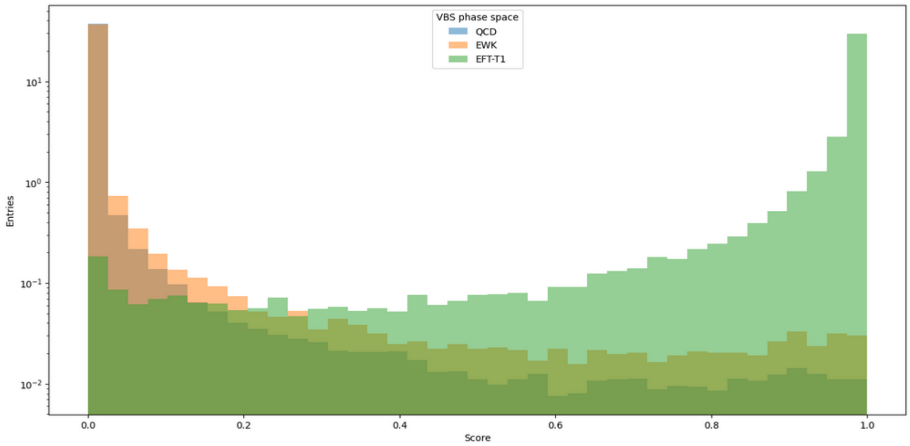


Figure 3: The score distribution from the Ensemble Model for the O_{T1} operator.

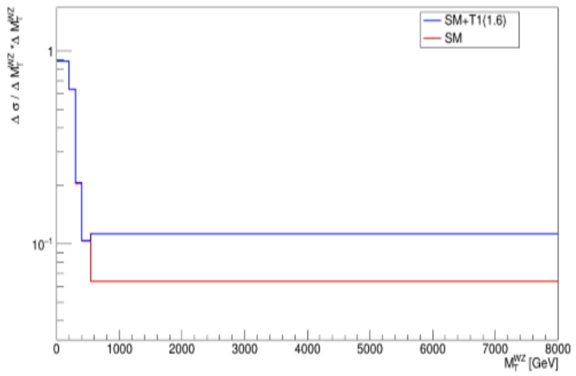
Binning	95% C.L of Wilson coefficient
{0., 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 0.9 1.}	-0.452, 0.452
{0., 0.2, 0.4, 0.6, 0.8, 1.}	-0.507, 0.507
{0., 0.9, 0.94, 0.96, 0.98, 1.}	-0.34, 0.34
{0., 0.9, 0.98, 1.}	-0.34, 0.34

Table 3: The results from the optimal binning study for the O_{T1} operator with the Ensemble model.

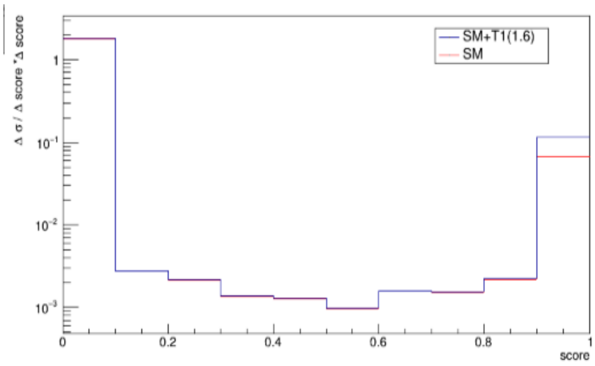
5.2 THE RESULTS FOR THE LIMITS ON THE WILSON COEFFICIENTS

In this section, the results for the limits using the score from the Random Forest model as a discriminative variable are presented for the most sensitive dimension-8 operators. Also, the kinematic variable of the transverse mass of the $W^{\pm}Z$ system has been tested for comparison.

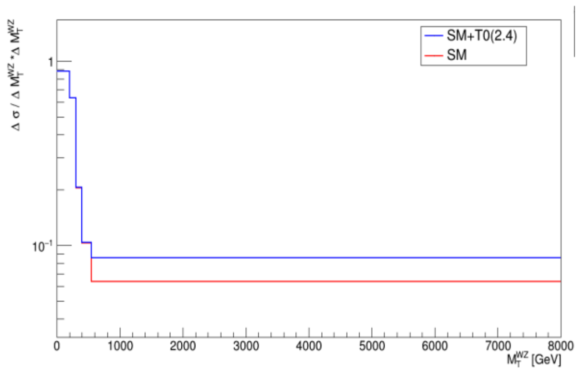
The differential cross section based on the SM is compared to the respective ones considering contribution from the EFT operators O_{T0} and O_{T1}. The distributions of the differential cross section as a function of the score and the transverse mass of the $W^{\pm}Z$ system are shown in Figure 4.



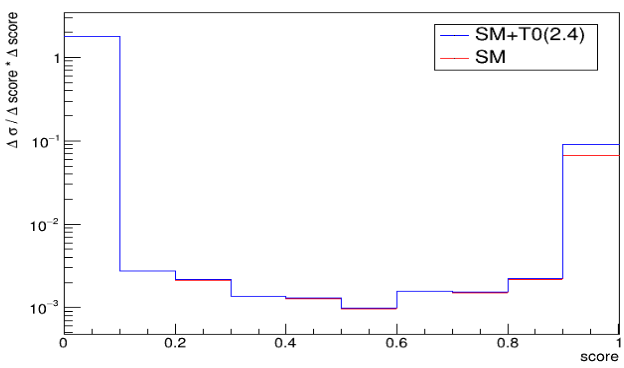
(a)



(b)



(c)



(d)

Figure 4: The differential cross section as a function of M_T^{WZ} and score for the operator T1 (a) and (b) and for the operator T0 (c) and (d).

The expected limits on the EFT couplings were extracted for each of the understudied dimension-8 operators from the M_T^{WZ} and the score distributions resulting from the Random Forest model. All the other Wilson coefficients were set to zero. The results are presented in the table below (Table 4).

Wilson Coefficient	M_T^{WZ}	ML Score
fT1	{-0.925 , 0.925}	{-0.950 , 0.950}
fT0	{-1.376 , 1.376}	{-1.402, 1.402}
fT2	{-4.503, 4.503}	{-4.458,4.458}
fM0	{-14.066,14.066}	{-14.993,14.993}
fS1	{-77.280, 77.280}	{-74.634, 74.634}

Table 4: Expected limits for the most sensitive operators of $W^{\pm}Zjj$ process for the full Run 2 luminosity of 139 fb^{-1}. Limits are presented for the kinematical variable M_T^{WZ} and the score from the Random Forest ML model.

V.CONCLUSIONS

In this study, we examined the expected effects of Effective Field Theory on the electroweak production of the $W^{\pm}Zjj$ process and derived expected limits on EFT couplings for the six most sensitive dimension-8 operators.

In this analysis, MC generated events at truth level with PYTHIA showering have been used in order to distinguish SM events from events affected by the dimension-8 EFT operators. The 95% CL limits of the Wilson coefficients of these operators were extracted through a template fit to the SM MC events using differential cross section distributions of the transverse mass of the $W^{\pm}Z$ system and the score from the Random Forest ML model which has been trained in order to classify the events between SM and EFT categories. Also, an optimal binning study has been conducted based on an Ensemble ML model described in [9] that has been already tested for the extraction of the limits of these operators. The optimal binning study indicates the effect of the chosen binning for the score distribution on results for the limits of the operators. The best binning that leads to the narrowest interval for the most sensitive operator O_{T1} is the one that allows for more detailed binning close to the region where we expect more signal events (close to one). This binning was used for the score distribution from the Random Forest model, too. The usage of the score distribution from the Random Forest ML model instead of the M_T^{WZ} does not indicate any improvements in the limits, considering the Ensemble model as the most optimal one.

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 $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector. Phys. Lett. B, 793:469, 2019.
[4]. CMS Collaboration. Measurement of electroweak WZ boson production and search for new physics in WZ+ two jets events in pp collisions at $\sqrt{s}$= 13 TeV Phys. Lett. B, 795:281, 2019.
[5]. Ćeline Degrande, Nicolas Greiner, Wolfgang Kilian, Olivier Mattelaer, Harrison Mebane, Tim Stelzer, Scott Willenbrock, and CenZhang. Effective field theory: A modern approach to anomalous couplings. Annals of Physics, 335:21–32, 2013.
[6]. Kasimi, E. Simultaneous production of two bosons in the ATLAS experiment at the LHC, doctoral diss, Aristotle University of Thessaloniki, 2023.
[7]. O. P. Eboli, M. C. Gonzalez-Garcia, and . K. Mizukoshi. $pp \to j\, j\, e^{\pm} \mu^{\pm} \nu\, \nu$ and $j\, j\, e^{\pm} \mu^{\mp} \nu\, \nu$ 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.
[8]. 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.
[9]. Bachas, K., Karkanias, I., Kasimi, E., Leonidis, C., Petridou, C., & Zachariadou, K. Application of Machine Learning techniques in indirect searches for New Physics in the Effective Field Theory approach. International Journal of Engineering Science Invention, 2021 .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
[10]. Rainio, O., Teuho, J. & Klén, R. Evaluation metrics and statistical tests for machine learning. Sci Rep **14**, 6086 (2024). https://doi.org/10.1038/s41598-024-56706-x
[11]. Scikit learn - classification with naïve Bayes. Tutorialspoint. (n.d.). https://www.tutorialspoint.com/scikit_learn/scikit_learn_classification_with_naive_bayes.htm
[12]. L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classification and Regression Trees", Wadsworth, Belmont, CA, 1984.
[13]. Testas, A. (2023). Random Forest classification with Scikit-learn and pyspark. Distributed Machine Learning with PySpark, 243–258. https://doi.org/10.1007/978-1-4842-9751-3_9
[14]. Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. The European Physical Journal C, 71(2), Feb 2011.