

A Machine Learning approach to the EFT re-interpretation of the $WZjj$ fully leptonic electroweak production

Konstantinos Bachas^{1,2}, Ioannis Karkanias¹, Eirini Kasimi¹, Christos Leonidis¹, Chara Petridou¹, Despina Sampsonidou³, Katerina Zachariadou⁴

¹Aristotle University of Thessaloniki, ²University of Thessaly, ³Tsung-Dao Lee Institute & Shanghai Jiao Tong University, ⁴University of West Attica

E-mail: dinos.bachas@cern.ch

Abstract. In this paper we study the use of Machine Learning techniques to set constraints on indirect signatures of physics beyond the Standard Model in Vector Boson Scattering (VBS), in the electroweak (EWK) production of self-interacting $W^\pm Z$ bosons in association with two jets. The WZ fully leptonic channel has been extensively studied by the ATLAS Collaboration at the LHC and we are about to provide results using the full Run 2 data corresponding to an integrated luminosity of 139fb^{-1} . The EWK production of the WZ in association with two jets has been already observed at 36fb^{-1} with an observed significance of 5.3 standard deviations. A factor of four increase in the integrated luminosity provides an opportunity to check for deviations from the Standard Model (SM) predictions, in particular for model independent, indirect searches for New Physics. Such searches can be realized in the context of an extension of the SM in terms of an Effective Field Theory (EFT) formalism, providing a way to quantify possible deviations from the Standard Model. The EFT Lagrangian besides the Standard Model terms comprises contributions from higher dimension operators, their effect being determined by the strength of their corresponding parameters (Wilson coefficients scaled to the appropriate power of Λ , indicating the scale of the appearance of New Physics). In this paper an attempt is made to search for New Physics effects in the $WZjj$ 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 $WZjj$ signal events originating from the effect of EFT operators, from $WZjj$ background events originating from strong (QCD) or EWK $WZjj$ processes. We investigate the utilization of the ensemble model response in estimating the sensitivity of $WZjj$ 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.

1. Introduction

The discovery of the Higgs boson at LHC [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 further investigate and confirm in 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 coupling to the Higgs boson, while is also responsible for maintaining unitarity in vector boson self-interactions up to the TeV scale. In this regard, measurements



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

of the self-interactions between the vector gauge bosons, involving three or four bosons and 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 coupling to Higgs from Standard Model (SM) expectations, in particular at TeV energy scales, is an indirect way to search for New Physics and is one of the challenges at the LHC that are currently being addressed by both experiments (ATLAS and CMS) with the Run 2 data and will be further explored at the Phase 1 and Phase 2 of the LHC and its experiments (Run 3 and Run 4 respectively).

At the LHC, the final state that comprises the production of 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, referred to as electroweak mediated production, involves only weak interactions and includes mostly VBS diagrams (diagrams with self-interacting gauge bosons and associated production of two jets). The first observation of the Vector Boson Scattering (VBS) has been reported by ATLAS and CMS [3, 4, 5, 6], at 13 TeV: in the WZ fully leptonic, accompanied by two jets, channel ($WZjj$) and in the same sign WW leptonic, accompanied by two jets, channel ($ssWWjj$), with 36 fb^{-1} by the ATLAS experiment, and for the same channels with the full Run 2 luminosity of 139 fb^{-1} , by CMS.

The rarity of the VBS processes -most of them with a cross section of order less than 1 fb - and the importance of the measurements in order 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, both in order 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 SM. An attractive feature of VBS is the appearance of quartic self-couplings between the gauge bosons (QGCs), which in consequence, provide 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].

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

Furthermore, there is recently a huge effort by both ATLAS and CMS experiments 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 to combine results on operator limits across experiments based on the full LHC Run-2 data, such that the constraints on the anomalous couplings are either maximized, or, if present, their values are quantified.

In view of this goal the current paper aims to suggest improvements in the search of anomalous coupling effects by introducing Machine Learning (ML) techniques and in particular, through the utilization of a ML classifier response, to distinguish between events of the SM $WZjj$ production and events due to different dimension-8 EFT operators effects. We study the $W^\pm Z$ production in association with two jets in the fully leptonic decay mode.

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.

In Section 4 the statistical model used for limit extraction on different dimension-8 operators is described, 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.

2. 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 distance. As a consequence the corresponding field operators are dimension-6 in energy and the next relevant for LHC are dimension-8 operators (fourth-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 will use the simplified parameters $f_i^{(6)} = \frac{c_i}{\Lambda^2}$ and $f_j^{(8)} = \frac{c_j}{\Lambda^4}$ 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

In order to study the effect of the dimension-8 operators in the $WZjj$ process and to extract limits for the couplings, large amounts of Monte Carlo samples in a dense grid of the parameter space are needed. 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 decomposition method is briefly explained below. 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 which 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, we consider only contributions due to the quadratic term of the EFT Lagrangian as explained below.

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

The dimension-8 operators are divided in three categories. The first category comprises operators that contain four covariant derivatives of the Higgs field which are of scalar type ($O_{S0,1}$). Operators in the second category contain two Higgs covariant derivatives and two field strength tensors and they are of the Mixed (scalar-tensor) Type ($O_{M0,1,2,3,4,5,6,7}$). Finally the third category contains those operators with four field strength tensors and they are of the Tensor Type($O_{T0,1,2,5,6,7,8,9}$). The f_S , f_M and f_T are the corresponding Wilson coefficients to the O_S , O_M and O_T operators accordingly. The operators that mostly affect the $WZjj$ process are the: $O_{S0,1}$, $O_{M0,1,2,3,4,5,6,7}$ and $O_{T0,1,2,5,6,7}$. In this paper, we have studied, indicatively, the operators O_{S1} , and $O_{T0,1}$.

2.2.2. The decomposition procedure 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 higher than four-dimension operators are added as extra terms in an expansion around the Standard Model Lagrangian. The A_{SM} is the SM amplitude while the A_i are amplitudes containing the individual higher dimension operators. For processes like the $WZjj$ 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

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

where $\sum_i c_i 2\text{Re}(A_{SM} A_i)$ is the amplitude of the interference between the SM and the EFT operator, named interference term, $\sum_i c_i^2 |A_i|^2$ is the pure EFT operator contribution, which is called quadratic term and $\sum_{ij, i \neq j} c_i c_j 2\text{Re}(A_i A_j)$ 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 estimator

$$P = \left| \frac{1}{m_{(\ell^+, \ell^-)}^2 - (m_Z^{\text{PDG}})^2 + i \Gamma_Z^{\text{PDG}} m_Z^{\text{PDG}}} \right|^2 \times \left| \frac{1}{m_{(\ell', \nu_{\ell'})}^2 - (m_W^{\text{PDG}})^2 + i \Gamma_W^{\text{PDG}} m_W^{\text{PDG}}} \right|^2, \quad (4)$$

where m_Z^{PDG} (m_W^{PDG}) and Γ_Z^{PDG} (Γ_W^{PDG}) represent the world average mass and decay width for the Z (W) boson. 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, and the configuration yielding the highest value of P is used to determine 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 $p_T > 40\text{GeV}$ 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 hemispheres, $\eta_{j1} \cdot \eta_{j2} < 0$, is required to be $m_{jj} > 500\text{GeV}$ in order to enhance the sensitivity to the $WZjj$ 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 $WZjj$ cross-section measurements by the ATLAS experiment in reference [4].

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^{\text{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
$N_{b\text{-quark}}$	$= 0$

3. The Machine Learning classifiers procedures

In this paper we investigate the effect of different EFT dimension-8 operators on the $WZjj$ process and use a Machine Learning approach in the $WZjj$ 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 $WZjj$ 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 of 36fb^{-1} and 139fb^{-1} . The latter 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

- Deep Neural Net
- XGBoost GBM (Gradient Boosting Machine)
- GLMs (Generalized Linear Models)
- Random Forest
- XRT (Extremely Randomized Trees)

Both sets use events selected in the signal region of the VBS phase space of the $WZjj$ 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 in order 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 $WZjj$. The effect from the linear term is 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 be 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 ϕ angle of the leptons (electrons and/or muons) from the Z boson and the W boson decay
- the 4-momentum η and ϕ angle of the two tagging jets

- the invariant mass, transverse momentum, η and ϕ of the Z boson
- the transverse mass, transverse momentum, η and ϕ of the W boson
- the transverse mass, the invariant mass and $\Delta\phi$ of the WZ system
- the mass, Δy and $\Delta\phi$ 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 1

3.3. Performance of the ML models

All base models and the ensemble model in each classifier 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 1. In Figures 2a and 2b 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.

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 operator T1 quadratic term sample is shown in Figure 3. 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 the T0 and S1 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.

4. 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 $WZjj$ 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 depend 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_{pred}^b in a bin b of the differential distribution is parametrized as

$$x_b^{\text{pred}}(\mathbf{c}, \boldsymbol{\theta}) = x_b^{\text{SM}} \left(1 + \sum_i \frac{c_i x_i^{\text{int}}(\theta)}{x_b^{\text{SM}}} + \sum_i \frac{c_i^2 x_i^{\text{quad}}(\theta)}{x_b^{\text{SM}}} + \sum_{i \neq j} \frac{c_i c_j x_{ij}^{\text{cross}}(\theta)}{x_b^{\text{SM}}} \right) \times \prod_i^{n_{\text{syst}}} \left(1 + \theta_j u_j^b \right) \quad (5)$$

where c are the Wilson coefficients, $\theta = (\theta_1, \dots, \theta_{n_{\text{syst}}})$ are nuisance parameters, n_{syst} is the number of nuisance parameters, x_{SM}^b is the nominal SM cross section prediction, and u_j^b is the relative size of the theory uncertainty j in bin b . Note that in the current paper we do not take

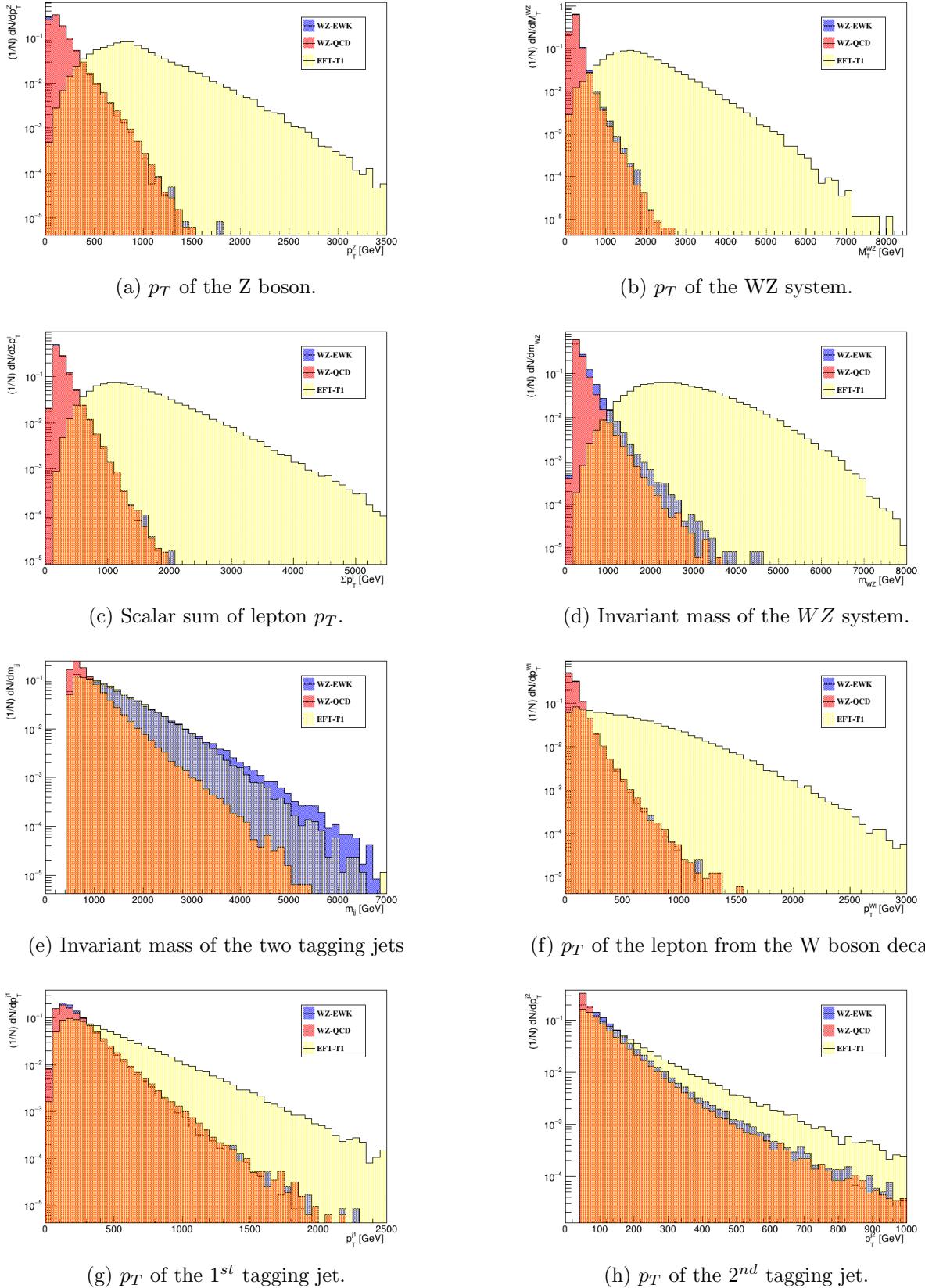
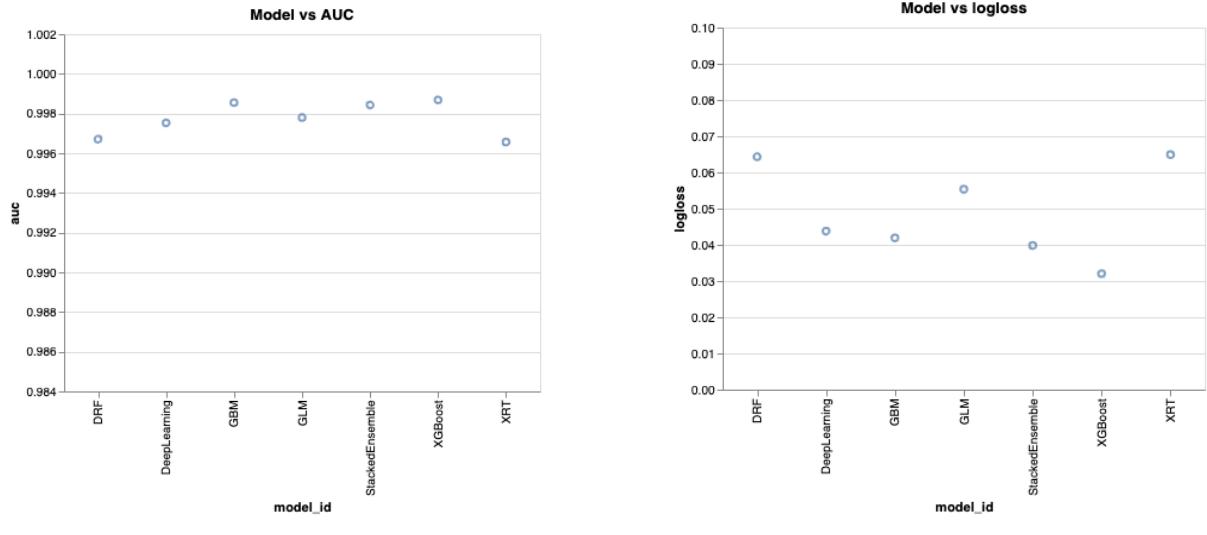


Figure 1: Kinematic input to the training of the ML models.



(a) AUC performance of the base models.

(b) Logloss performance of the base models.

Figure 2: Performance of the different ML models in terms of AUC and logistic loss metrics.

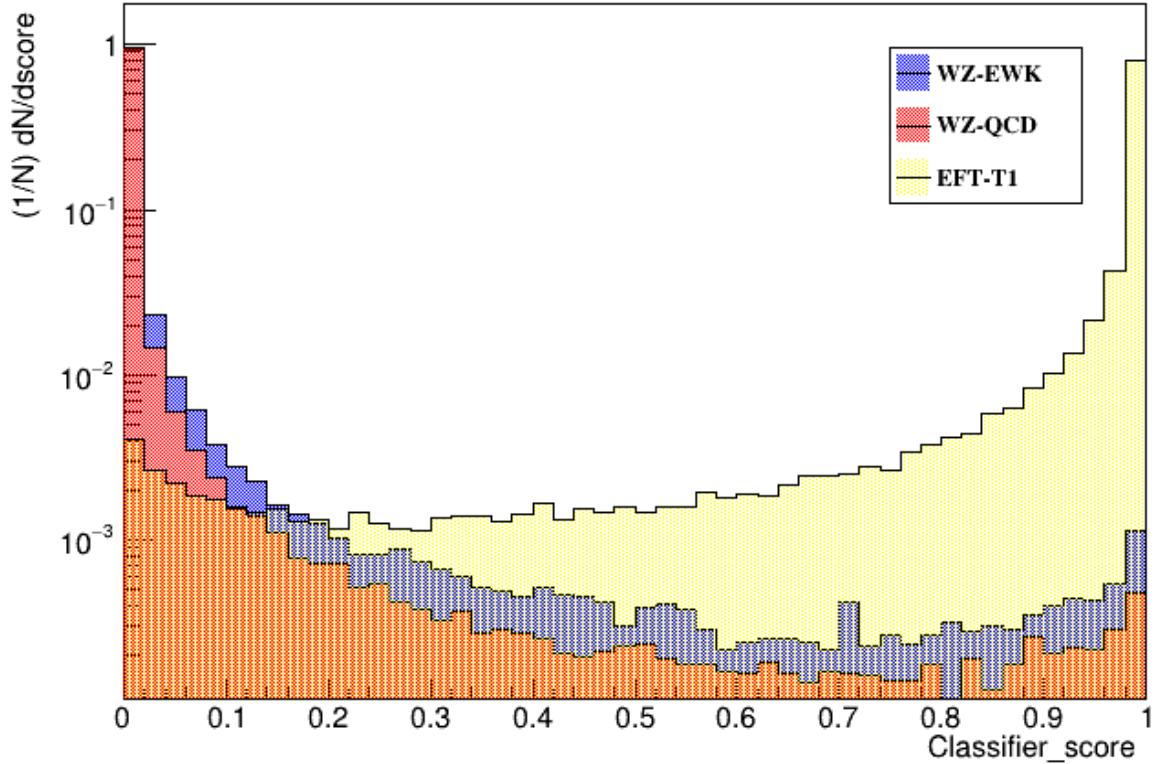


Figure 3: The classifier score distribution for the WZ-EWK, WZ-QCD and EFT-T1 operator events.

into account any systematic uncertainty and we consider only the statistical uncertainties.

The complete likelihood function is given by:

$$L(\mathbf{x}|\mathbf{c}, \boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^{n_{\text{bins}}} \det(C)}} \exp\left(-\frac{1}{2}\Delta\mathbf{x}^T(\mathbf{c}, \boldsymbol{\theta}) C^{-1} \Delta\mathbf{x}(\mathbf{c}, \boldsymbol{\theta})\right) \times \prod_i^{n_{\text{syst}}} g_i(\theta_i) \quad (6)$$

where x is the nominal (expected) Standard Model 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), 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,

$$\Delta x^b = x_{\text{meas}}^b - x_{\text{pred}}^b(\mathbf{c}, \boldsymbol{\theta}) \quad (7)$$

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

$$\lambda(c_i) = -2 \log \frac{L\left(c_i, \hat{\boldsymbol{\theta}}\right)}{L\left(\hat{c}_i, \hat{\boldsymbol{\theta}}\right)}. \quad (8)$$

where $L\left(c_i, \hat{\boldsymbol{\theta}}\right)$ is the maximum of the likelihood for fixed c_i and $L\left(\hat{c}_i, \hat{\boldsymbol{\theta}}\right)$ 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 [15], assuming that $\lambda(c_i)$ is χ^2 distributed. In this paper, we consider only the contribution from the statistical uncertainty.

5. 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 4, differential cross sections are presented for the transverse mass of the WZ system (M_T^{WZ}), the scalar sum of the transverse momentum of the three leptons ($\sum p_{Tl}$), the azimuthal angle separation between the W and Z bosons ($\Delta\phi(WZ)$) and the rapidity separation between the two tagging jets (Δy_{jj}). The SM expected differential cross section is compared with the SM plus the contribution from the f_{T1} operator in two operator value scenarios. One where $f_{T1} = 1.6$ and the other where $f_{T1} = 3.2$. It is interesting to notice that, as expected, the higher the value of the operator the higher the departure of the cross sections from the SM expectations.

Limits on the EFT couplings were evaluated for each kinematical variable and each operator separately, setting all other Wilson coefficients to zero. The results from the kinematical variables are then compared with the results obtained from the fit to the MC event using the Machine Learning score distribution shown in Figure 3. These results are presented in Table 2 for 36fb^{-1} and Table 3 for 139fb^{-1} . It is remarkable the level of improvement for the obtained limits that the ML score provides in both integrated luminosity cases which is of the order of $\times 3$ with respect to the other kinematical variables.

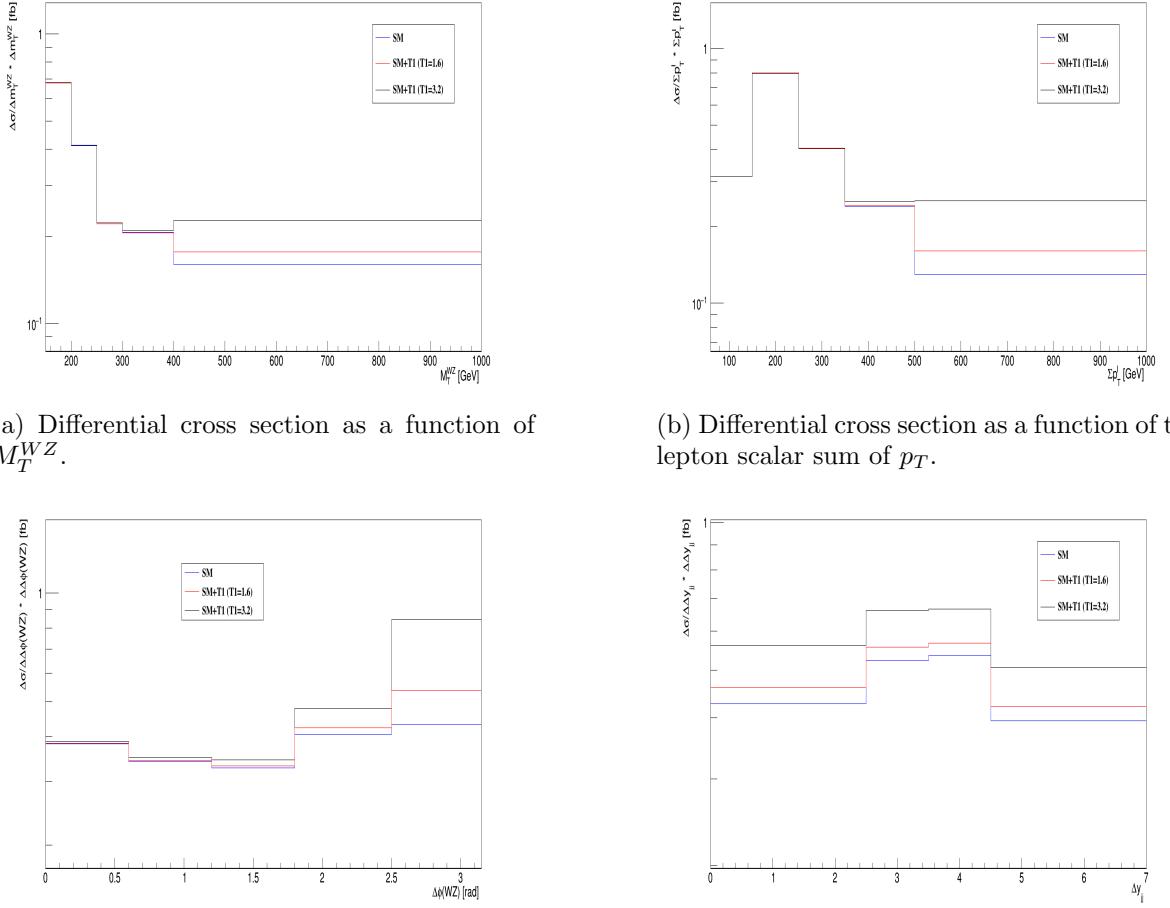


Figure 4: Differential cross sections for kinematical variables used to obtain the EFT coupling limits as described in the text. The SM expected cross section(blue) is compared with the SM plus the contribution from the f_{T1} operator in two operator value scenarios: $f_{T1} = 1.6$ (red) and $f_{T1} = 3.2$ (black)

Table 2: Expected limits at $36 fb^{-1}$

Wilson coeff.	M_T^{WZ}	$\sum p_{Tl}$	$\Delta\phi(WZ)$	Δy_{jj}	ML Model
f_{S1}	$[-167.4, 167.4]$	$[-132.3, 132.3]$	$[-182.6, 182.6]$	$[-184.3, 184.3]$	$[-54.6, 54.6]$
f_{T0}	$[-2.97, 2.97]$	$[-2.21, 2.21]$	$[-3.65, 3.65]$	$[-3.97, 3.97]$	$[-1.00, 1.00]$
f_{T1}	$[-2.04, 2.04]$	$[-1.53, 1.53]$	$[-2.39, 2.39]$	$[-2.77, 2.77]$	$[-0.47, 0.47]$

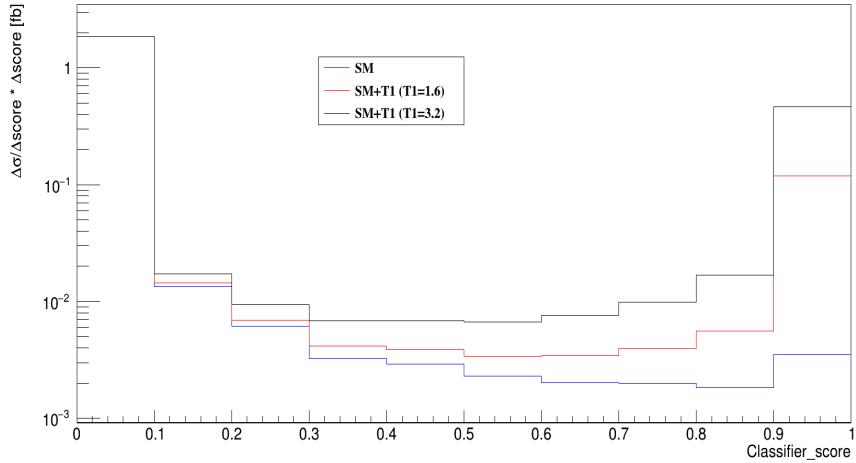


Figure 5: Differential cross section as a function of ML classifier score.

Table 3: Expected limits at 139fb^{-1}

Wilson coeff.	M_T^{WZ}	ML Model
f_{S1}	$[-124.7, 124.7]$	$[-39.0, 39.0]$
f_{T0}	$[-2.21, 2.21]$	$[-0.71, 0.71]$
f_{T1}	$[-1.52, 1.52]$	$[-0.47, 0.47]$

6. Summary

In this paper we have studied the expected EFT effects in the VBS EWK production of the $W^\pm Zjj$ process and have extracted expected limits on the EFT couplings for 3 of the most sensitive dimension-8 operators in this channel. The study is realized at particle level MC generated events. Limits are obtained through a template fit to the Standard Model MC events using traditional differential distributions of kinematical variables like M_T^{WZ} , $\sum p_T^l$, $\Delta\phi^{WZ}$ and Δy_{jj} . 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 template to fit the data, gives much better limits on EFT couplings than the most sensitive variables for dimension-8. This is the case for both integrated luminosity points that have been considered; for 36fb^{-1} and 139fb^{-1} . We plan to repeat these very promising studies using the full Run 2 reconstructed data from the ATLAS experiment.

Acknowledgments

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.

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-Sign W Boson Pair in Association with Two Jets in pp Collisions at $\sqrt{s} = 13$ 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 $\sqrt{s} = 13$ 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 pp collisions at $\sqrt{s} = 13$ 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] Céline 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. Éboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi. $pp \rightarrow jj e^\pm \mu^\pm \nu \nu$ and $jj e^\pm \mu^\mp \nu \nu$ at $\mathcal{O}(\alpha_{\text{em}}^6)$ and $\mathcal{O}(\alpha_{\text{em}}^4 \alpha_s^2)$ 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] Torbjörn Sjöstrand, 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 $\sqrt{s} = 8$ 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- k_t 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.