EFT interpretations of VBS final state with CMS data

CERN Summer Student Programme 2024

Student: Agustin Guillenea
Supervisors: Sitian Qian, Andrea Massironi,
Giulia Lavizzari, Jie Xiao

European Organization for Nuclear Research
EP department
Geneva, Switzerland
August 2024

Abstract

Although the Standard Model has been very successful describing the data collected by ATLAS and CMS at LHC, possible deviations and indirect measurements of high energy new physics could be explored by means of an Effective Field theory (EFT) approach. By measuring precisely the distributions in Vector Boson Scattering events, that have an excellent signal over background ratio and are less affected by fixed order predictions compared to other processes, it will be possible to test higher order operators described in an EFT framework. In this project we focus the Standard Model Effective Field Theory (SMEFT) with dimension-six operators and test the Q_W operator performing a recasting of an analysis performed by CMS Collaboration. We generated the SMEFT contribution to the signal, same-sign WW production in proton-proton collision with fully leptonic final state, and performed a profile likelihood test using the simplified likelihood framework to compare it with the experimental data. Doing this we managed to obtain a t_{C_W} distribution, being C_W the Wilson coefficient of the Q_W operator, varying its value between -1 and 1.

Contents

1. Introduction **2**

2. Theoretical framework **3**
- 2.1. Standard Model Effective Field Theory 3
- 2.2. Simplified likelihood test 5

3. Methods **6**
- 3.1. Collisions simulation 6
  - 3.1.1. Model and configuration 6
- 3.2. Analysis 7

4. Results **11**

5. Conclusion **14**

References **15**

1. Introduction

The Standard Model (SM) of particle physics is one of the most successful theories in describing the fundamental interactions of elementary particles. Over the past few decades, extensive experimental efforts, particularly at the Large Hadron Collider (LHC), have provided plenty of data that corroborates the predictions of the Standard Model. Despite its success, the Standard Model has its limitations. It does not account for certain phenomena such as dark matter, neutrino masses, or the matter-antimatter asymmetry of the universe. These gaps suggest the existence of new physics beyond the Standard Model, which might only become apparent at higher energy scales or through subtle deviations from Standard Model predictions.

One approach to explore these potential deviations is the use of Effective Field theories (EFTs). EFTs provide a powerful framework for describing new physics indirectly by parameterizing it in terms of higher-dimensional operators added to the Standard Model Lagrangian. These operators encapsulate the effects of new high-energy physics that are currently out of direct reach but might influence lower-energy observables.

In this project, the focus is on the usage of an EFT to study Vector Boson Scattering (VBS) events at the LHC. VBS takes place at the LHC when two quarks from two different colliding protons radiate vector bosons, which in turn interact. VBS processes have an excellent signal-to-background ratio and are less affected by fixed order predictions compared to other processes, making them ideal for testing high order operators described in an EFT framework. We do this by recasting the results obtained by the CMS Collaboration based on proton-proton (pp) collision data collected at $\sqrt{s} = 13$ TeV by the CMS Experiment at the CERN LHC during 2016–2018, corresponding to an integrated luminosity of 138 fb^{-1} [1]. A recasting consists in using the estimates of backgrounds and systematic uncertainties, as well as the observations in data, from an existing search involving a certain process to reinterpret the existing experimental search with a new theory. Thanks to such strategy we can focus only on the generation of signal events and on the extraction of signal yields in the desired final region. The process we are working with consists in the production of a pair of same-sign leptons via VBS in proton-proton collisions. In particular we focus on the final state with 2 muons. Overall, our final signature comprises two jets with a large invariant mass in the forward region, two same-sign charged muons and the missing transverse energy due to the two neutrinos (some representative diagrams can be seen in Figure 1). Such process is expected to be sensitive to modifications of the EW sector and of the Higgs mechanism: a delicate set of cancellations between VBS diagrams with and without the Higgs boson takes place in order for the VBS cross-section to be non-divergent. Any slight modification to such delicate equilibrium would manifest itself in anomalous contributions to the various VBS observables.

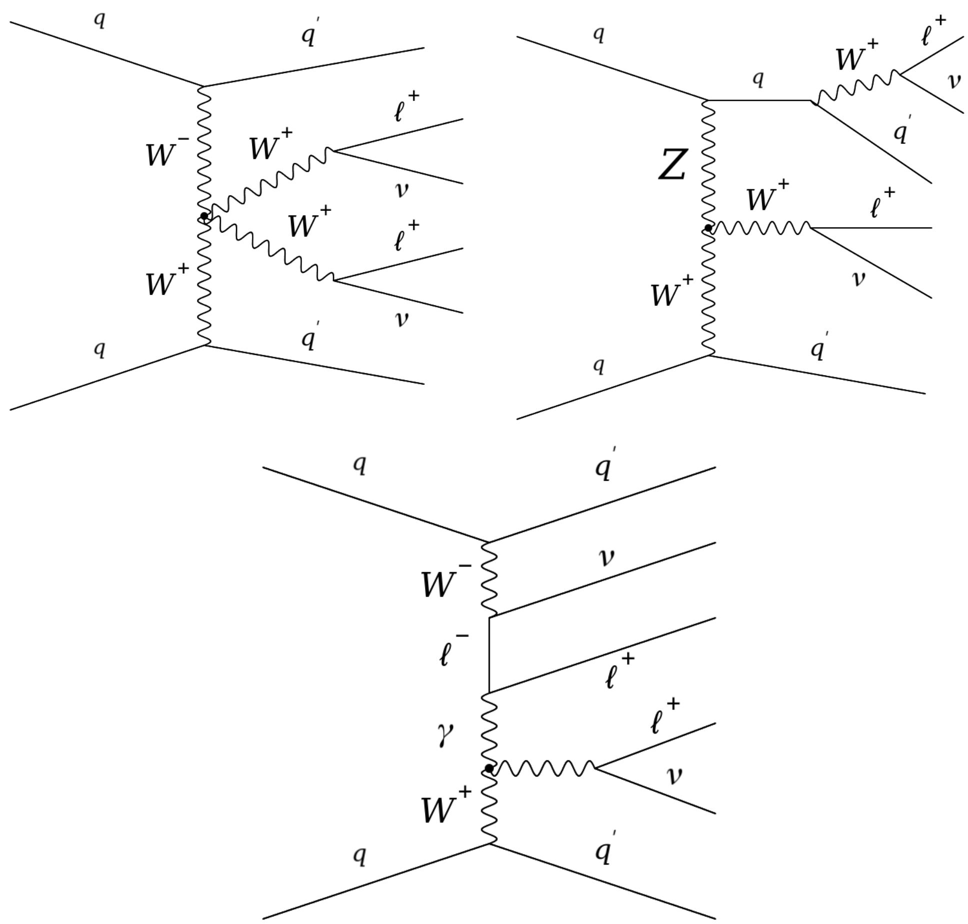

Figure 1: Representative tree level Feynman diagrams for our signal process with positive same-sign leptons in the final state.

2. Theoretical framework

In this section we will give a brief theoretical background in order to comprehend the physics model we are working with, as well as the kind of analysis we will perform using said model.

2.1. Standard Model Effective Field Theory

The analysis presented here is conducted within the framework of the Standard Model Effective Field Theory (SMEFT) [2], whose formulation employs the degrees of freedom and gauge symmetries of the SM and it is structured as an infinite series of operators sorted by canonical dimension. The Lagrangian of this theory is:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{SM} + \frac{1}{\Lambda}\sum_{\alpha} C_{\alpha}^{(5)} Q_{\alpha}^{(5)} + \frac{1}{\Lambda^2}\sum_{\alpha} C_{\alpha}^{(6)} Q_{\alpha}^{(6)} + \dots \tag{1}$$

where Λ is the energy scale that characterizes the beyond Standard Model (BSM) dynamics, $Q_{\alpha}^{(i)}$ is an effective operator of dimension i, and $C_{\alpha}^{(i)}$ is its Wilson coefficient (the sum in α implies the sum of all the effective operators with the same dimension). At the observables level, this reproduces a series expansion in (E/Λ), and the condition $(E/\Lambda) \ll 1$, indicating the near decoupling of the new physics sector, is necessarily assumed. Here we focus on the SMEFT Lagrangian with only dimension-six operators added to $\mathcal{L}_{SM}$.

In this SMEFT framework, we consider processes with Feynman diagrams containing none or only one EFT vertex. This consideration is made as it corresponds to the leading order contribution from the EFT operators including both EFT-SM-interference (linear) and pure EFT contribution (quadratic). The EFT contributions to a generic observable can be written at $\mathcal{O}(\Lambda^{-2})$ as follows:

$$\sigma = \sigma_{SM} + \sum_{\alpha} \sigma_{\alpha} \bar{C}_{\alpha} + \sum_{\alpha,\beta} \sigma_{\alpha\beta} \bar{C}_{\alpha} \bar{C}_{\beta} \tag{2}$$

where $\bar{C}_{\alpha} = (v^2/\Lambda^2)\, C_{\alpha}$, being v the vacuum expectation value, and σ_{SM}, σ_{α} and $\sigma_{\alpha\beta}$ denote the SM, linear and quadratic contributions respectively. σ_{SM} and $\sigma_{\alpha\alpha}$ are always positive quantities, while σ_{α} and $\sigma_{\alpha\beta}$ with $\alpha \neq \beta$ can take negative values.

In this project, we set the value of the Wilson coefficients of every dimension-six operator but one to zero, which allows us to test this particular operator in our process. Also, the idea is using only one operator as a proof of concept, using this reinterpretation approach that could be easily generalized by including more operators. In the Warsaw basis [3] the operator we are working with is called Q_W and it is given by:

$$Q_W = \varepsilon^{ijk} W_{\mu}^{i\nu} W_{\nu}^{j\rho} W_{\rho}^{k\mu} \tag{3}$$

As we only have the contribution of one operator, the expression for a generic observable written in equation 2 becomes:

$$\sigma = \sigma_{SM} + \sigma_W \bar{C}_W + \sigma_{WW} \bar{C}_W^2 \tag{4}$$

From now on, we will refer to the terms multiplied by σ_W and σ_{WW} in equation 6 when we talk about the linear and quadratic components respectively.

2.2. Simplified likelihood test

The objective of this project is testing the Q_W SMEFT operator, which we perform carrying out a likelihood test using the simplified likelihood framework [4–6]. Said framework is a systematic approximation scheme for experimental likelihoods such as those originating from LHC experiments. The data collected in particle physics originate from quantum processes, and have thus an intrinsic statistical uncertainty, which vanishes in the limit of large data sets. This method focus is on the systematic uncertainties, which are independent of the amount of data. More detail in how the simplified likelihood framework works in the references [4–6].

A likelihood function measures the probability of seeing the observed data under different model parameter values in order to assess how well a statistical model describes the data. If we measure the value of one kinematic observable x for each selected event of an experiment, we can represent this data in one histograms. The expectation value of the ith histogram bin can be written as [7]:

$$E[n_i] = \mu s_i + b_i \tag{5}$$

where s_i and b_i are the mean number of signal and background entries in the ith bin, and μ is a parameter that determines the strength of the signal process. To test a hypothesized value of μ we consider the profile likelihood ratio:

$$\lambda(\mu) = \frac{L\left(\mu, \hat{\hat{\vec{\theta}}}\right)}{L\left(\hat{\mu}, \hat{\vec{\theta}}\right)} \tag{6}$$

where $\vec{\theta}$ contains all our nuisance parameters [7], $\hat{\hat{\vec{\theta}}}$ denotes the value of $\vec{\theta}$ that maximizes L for a certain μ, and the denominator is the maximized unconditional likelihood function. From this definition, we can see that $0 \leq \lambda \leq 1$, with the closer λ is to one implies better agreement between the data and the hypothesized value of μ. Based on this, the quantity we are basing out statistical test on is:

$$t_\mu = -2 \ln \lambda\left(\mu\right) \tag{7}$$

Higher values of t_μ correspond to greater disagreement between the data and the value of μ. Replacing here the expression for $\lambda(\mu)$ presented in equation 6, we obtain:

$$t_\mu = -2\Delta \ln L \tag{8}$$

with $\Delta \ln L = \ln L\left(\mu, \hat{\hat{\vec{\theta}}}\right) - \ln L\left(\hat{\mu}, \hat{\vec{\theta}}\right)$. In this project we take $\mu = C_W$ (the Wilson coefficient of the Q_W operator), and aim to obtain a t_{C_W} distribution.

3. Methods

As we previously said, we only need to generate events corresponding to the SMEFT signal to perform the recasting. In order to do this we used various software packages to simulate the whole process of the collision and detection of the events, and another one to perform the subsequent analysis. In this section we present a complete overview of the conducted the procedure.

3.1. Collisions simulation

To simulate the collision we used MadGraph5_aMC@NLO [8] at LO accuracy. This tool uses a Monte Carlo method to compute the cross section of a given process and generate the hard events, that is to say, the simulation of the state of the system right after the hard scattering (in this case the VBS process). The products of this hard scattering simulation are handled to two other programs called by MadGraph to continue the complete simulation. The first one is Pythia 8.3 [9], that simulate the interactions and decays between the particles in the final state of the hard scattering, in order to make particles that would be stable enough to reach the detector. In this case, in the final state we have one neutrino (antineutrino) pair that does not decay and two same-sign muons (antimuons) that are stable enough to go through the detector and therefore Pythia does not do anything to them. However, in the final state we also have two quarks. Due to color confinement, these quarks can't be isolated, and thus Pythia makes them interact to form hadrons, and this hadrons decay until they are stable enough. Pythia also executes this for the underlying event. The last program in the collision simulation is Delphes [10], that takes the information handled by Pythia and performs a detection simulation. Because we are performing a recasting of an analysis made by the CMS Collaboration with data obtained with the CMS Detector, we used Delphes with its configuration, i.e. using the `delphes_card_CMS.dat` generated in the MadGraph folder.

3.1.1. Model and configuration

The model used to perform the SMEFT signal simulation was obtained from SMEFT-sim package [11, 12] and is called `SMEFTsim_topU3l_MwScheme_UFO_b_massless`. This package has a set of models that allow automated calculations in the SMEFT. The SMEFT Lagrangian is not invariant under flavor rotations of the fermion fields, so it has to be defined to avoid ambiguities. Also, it is needed to assign numerical values to the Lagrangian parameters. The SMEFTsim package has many independent UFO models with different flavor assumptions and input parameters implementations. The model we use has the topU3l flavour scheme and uses the m_W input scheme [11, 12]. As mentioned in the previous section, we set the value of all the Wilson coefficients of the dimension-six operators to zero, and the one associated to the Q_W operator, namely C_W, to 1. Also, we set the new physics energy scale, namely Λ, to 1000 GeV.

In order to realize the analysis we generated two different datasets, one for the linear and another one for the quadratic component, as were defined before, with 10.000 Monte Carlo events each. This samples were simulated using the NNPDF 3.0 [13] parton distribution function. The center-of-mass energy for the pp collisions generated was $\sqrt{s} = 13$ TeV, which is the center-of-mass energy of the LHC when the experimental data was measured. The exact syntax used to generate the signal for the linear SMEFT contribution is:

```
generate p p > mu+ mu+ vm vm j j SMHLOOP=0 QCD=0 NP=1 NP^2==1
add process p p > mu- mu- vm~ vm~ j j SMHLOOP=0 QCD=0 NP=1 NP^2==1
```

And the syntax used to generate the signal for the quadratic SMEFT contribution is:

```
generate p p > mu+ mu+ vm vm j j SMHLOOP=0 QCD=0 NP==1
add process p p > mu- mu- vm~ vm~ j j SMHLOOP=0 QCD=0 NP==1
```

The CMS data was taken between the years 2016 and 2018, and therefore for the recasting is necessary to have different events for each year of data taking. We did not make a different sample for each year at MadGraph level, but instead split each one of the two different samples we generated later during the analysis to take this into account.

3.2. Analysis

The analysis was carried out using the MadAnalysis 5 package. MadAnalysis 5 is a framework for phenomenological investigations at particle colliders based on a C++ kernel [14–22]. This software takes as inputs the events produced by Delphes and allows for creating the distributions of multiple observables and for filterng the sample by applying kinematic or geometric cuts according to given conditions. For our recast to make sense, the cuts have to be the same that the ones applied in the original analysis [1].

The condition established for the final state of the events used in the analysis is that they have two well identified same-sign muons with transverse momentum $p_T > 30$ GeV and pseudorapidity $|\eta| < 2.4$, and at least two jets with $p_T > 30$ GeV and $|\eta| < 4.7$. We consider the two jets with highest transverse momentum to be the VBS candidates, and we demand them to have a large dijet invariant mass $m_{jj} > 750$ GeV and a large pseudorapidity separation $|\Delta\eta_{jj}| > 2.5$. Also we require the same-sign muons invariant mass to be $m_{\ell\ell} > 20$ GeV.

Events produced via VBS are expected to present rapidity gaps between the muons and the jets, so we make use of this we require $Max(\mathcal{Z}_\ell) < 0.75$, where $\mathcal{Z}_\ell$ is the Zeppenfeld variable [28] of a lepton ℓ, defined by:

$$Z_\ell = \frac{\left|\eta_\ell - \frac{\eta_{j_1}+\eta_{j_2}}{2}\right|}{|\Delta\eta_{jj}|} \tag{9}$$

where η_ℓ is the lepton pseudorapidity and η_{j_1} and η_{j_2} are the two VBS jets pseudorapidity.

We also apply four different kind vetoes. Two to remove additional loosely identified electrons and muons with $|\eta| < 2.5$ and $|\eta| < 2.4$ respectively, and $p_T > 10$ GeV (for both). Another one is to remove events with an hadronically decaying tau lepton with $|\eta_e| < 2.3$ and $p_T > 20$ GeV. This cuts were made in the original analysis with the objective of reducing the backgrounds contributions, such as $W^{\pm}Z$ production. The last veto corresponds to the rejection of events with at least b-tagged jet to reject events from $t\bar{t}$ pair where only one W boson decays leptonically, which are the main source of non-prompt-lepton backgrounds.

The original analysis compares the experimental data with two BSM models that use different mechanisms to explain the generation of neutrino masses. One involves the extension of the SM Lagrangian adding the dimension-five Weinberg operator [23], which add a Majorana mass term to the SM left-handed neutrino. The other one adds another heavy particle, the right-handed neutrino (N), which has a Majorana mass, and only couple to the SM through mixing with SM left-handed neutrinos [24–27]. Both models have particles with Majorana masses and introduce lepton number violation (LNV). They study events with same-sign muon pairs in the final state generated through VBS processes mediated by the Weinberg operator and processes with heavy Majorana neutrino production. Both of this kind of events have only the muons and two jets in the final state, and neither have neutrinos. This absence of neutrinos in the final states would make them lepton number violated states, which is not the case for our SMEFT final states.

For each model they establish two separate regions using two different observables. For the Weinberg operator mediated processes they use the missing transverse momentum (p_T^{miss}) as a discriminant, because this signal processes do not produce final state neutrinos other than those produced in hadron decays. For the processes with heavy Majorana neutrino production they claim the azimuthal separation of the leptons ($\Delta\phi_{\ell\ell}$) is better for improving the sensitivity to the signal than p_T^{miss}, because the two muons in the final state tend to be more back-to-back than in background processes. However, p_T^{miss} is more discriminating than azimuthal separation for the Weinberg operator process because of the different event topology resulting from the discrepancy between the two t-channel propagators. Therefore, we plotted the distributions for this two variables to find out which of the previous regions we wanted to compare our signal with. The two p_T^{miss} regions defined in the original analysis correspond to events with $p_T^{miss} < 50$ GeV and events with $p_T^{miss} > 50$ GeV. On the other hand, the two $\Delta\phi_{\ell\ell}$ regions defined in the original analysis correspond to events with $\Delta\phi_{\ell\ell} < 2$ and events with $\Delta\phi_{\ell\ell} > 2$. Based on the two distributions shown in Figure 2 and Figure 3, we concluded p_T^{miss} is not as an efficient discriminant to use in our analysis as $\Delta\phi_{\ell\ell}$, so we decided using the later instead.

Although, we only used the region with $\Delta\phi_{\ell\ell} > 2$ region, adding this cut to the others mentioned before. It is worth mentioning, that the heavy Majorana nautrino model involves a process with a t-channel exchange of a high mass virtual particle, the heavy Majorana neutrino, that makes it similar to our SMEFT model considered here. Thus, it is reasonable to compare our model with this one.

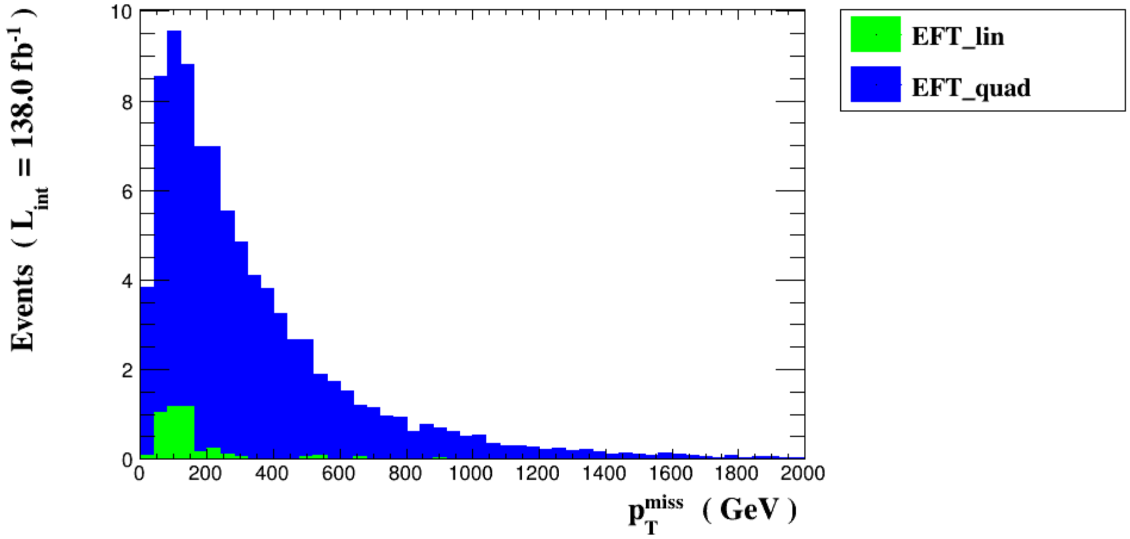

Figure 2: p_T^{miss} distribution for the signal events generated (SMEFT linear component in green and quadratic component in blue) before performing any cut. This variable was used as a discriminant in the original analysis [1].

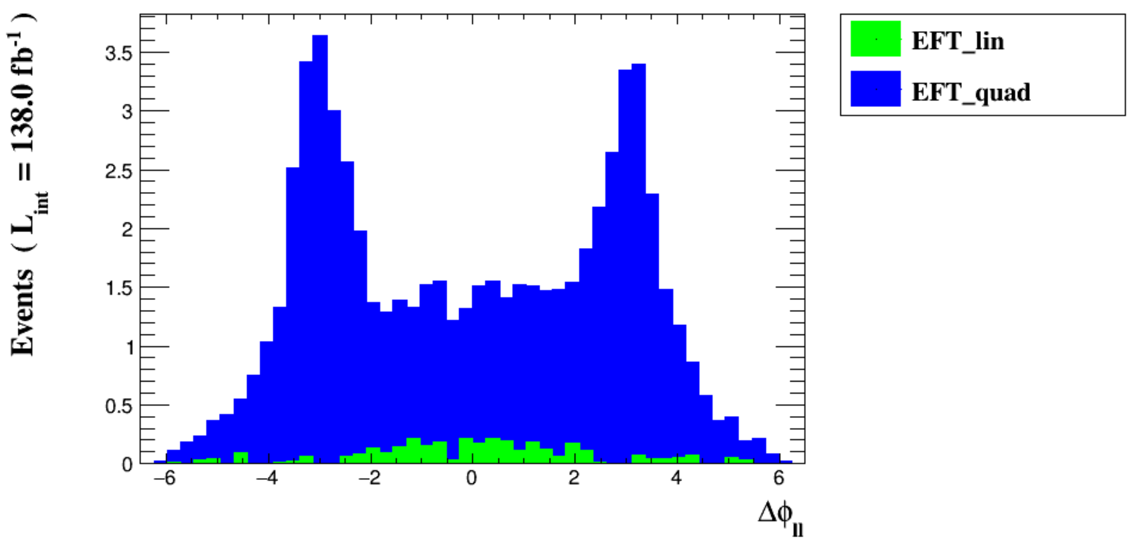

Figure 3: $\Delta\phi_{\ell\ell}$ distribution for the signal events generated (SMEFT linear component in green and quadratic component in blue) before performing any cut. This variable was used as a discriminant in the original analysis [1].

Variable	Signal region cut
Number of leptons	Exactly 2 same-sign muons
Lepton p_T	>30 GeV
Lepton $\|\eta\|$	<2.4
Jet p_T	>30 GeV
Jet $\|\eta\|$	<4.7
$m_{\ell\ell}$	>20 GeV
m_{jj}	>750 GeV
$\|\Delta\eta_{jj}\|$	>2.5
Max($\mathcal{Z}_\ell$)	<0.75
Loose e veto	Required
Loose μ veto	Required
Hadronic τ veto	Required
b-tagged jet veto	Required
$\Delta\phi_{\ell\ell}$	>2

Table 1: List of the cuts applied to define the signal region used to perform the analysis.

To summarize, all the kinematic cuts applied to define the signal region in which we perform the analysis are shown in the Table 1.

This analysis was realized using the MadAnalysis expert mode. In this mode you write an analysis file using the C++ language, and compile it with some tools generated automatically when running the program for the first time. The advantage of this is that we can create the analysis from scratch and define the selections and plot however we want. However, we had access to a version of the analysis file used in the original analysis we are recasting, and this helped to have a better understanding of how the original analysis was made, and thus performing ours in a more efficient way.

Nevertheless, some initial distributions (as those shown in Figure 2 and Figure 3) were plotted using the normal mode of MadAnalysis, where you can, through a python interface, make selection cuts and plot the distributions of some predefined variables. Such figures were obtained through a simpler pipeline, as no cuts were needed.

As mentioned in the section before, the experimental data used for the analysis we are recasting corresponds to three years of collision data collected by the CMS Experiment (2016-2018). At MadGraph level, we only generated one dataset for each of our SMEFT components, so we have to take this into account in the analysis. By default, MadAnalysis normalizes all regions to 138 fb^{-1}. To correct this we define three regions (one for each year) and apply them all the cuts in Table 1. Then we implement a per-region weight that corrects the 138 fb^{-1} normalization down to per-year luminosity. The luminosity splitting we did is 36:42:60 fb^{-1} to the 2016, 2017 and 2018 regions respectively. With the idea of using as many events as possible while avoiding statistics overlap, each event is assigned to only one of the three signal regions through the means of a random number generator.

The output of the analysis performed with MadAnalysis is given in the form of XML-like files containing the number of Monte Carlo events remaining after applying each cut, their efficiencies, along with other statistical information. This files were processed into histograms using the MadAnalysis 5 Interpreter For Expert Mode [29] python plugin, that parses all the signal regions contained in them and construct an object-base, interactable cutflow.

As a final step, we wanted to compare the results obtained with the data used by CMS Collaboration to carry out the original analysis. This data is public and available in the HEPData database [30]. To do the comparison we need to have the same histograms as those presented in [1], namely $H_T/p_T^{\mu_1}$ distributions. This variable corresponds to the quotient between the scalar sum of the p_T for all jets that satisfy the aforementioned p_T and $|\eta|$ cuts ($p_T > 30$ GeV and $|\eta| < 4.7$), and the p_T of μ_1, being μ_1 the muon with the highest p_T. In the original analysis they use this discriminating variable, because it measures the amount of hadronic activity relative to leptonic activity in a given event. They claim that, as hadronic activity is typically suppressed in VBS production processes, this variable tends to be smaller for signal than for background events [1].

With our two signals $H_T/p_T^{\mu_1}$ distributions performed a profile likelihood test, as explained in the Theoretical framework section, combining them as in equation 4, and using [4–6] as a base. To do this we had to have the same binning as the one used in Figure 2 of [1]. This corresponds to four $H_T/p_T^{\mu_1}$ regions: 0-1, 1-2, 2-5 and >5. We plotted a distribution of t_{C_W} for values of C_W between -1 and 1.

4. Results

With our two 10.000 Monte Carlo events samples for the SMEFT linear and quadratic components we achieved three different luminosity normalized $H_T/p_T^{\mu_1}$ distributions for each one, corresponding to the recreation of 2016, 2017 and 2018 luminosities for both samples. Figure 4 and Figure 5 shows the distributions obtained.

Using the information displayed on this histograms we performed the aforementioned profile likelihood test to obtain the t_{C_W} distribution presented in Figure 6.

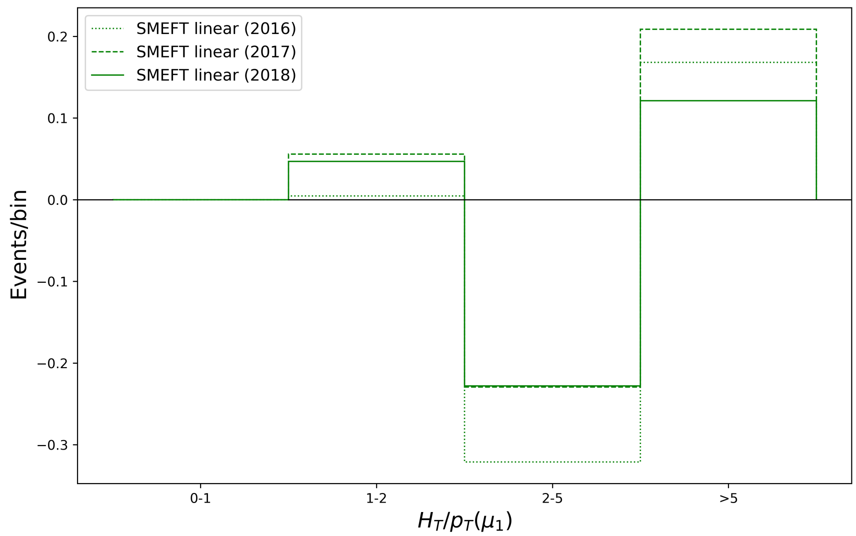

Figure 4: $H_T/p_T^{\mu_1}$ distribution obtained for the SMEFT linear component signal. The binning is the same as the used in the Figure 2 of [1]. The different normalizations on each correspond to the luminosity splitting for the years 2016 (dotted), 2017 (dashed) and 2018 (solid).

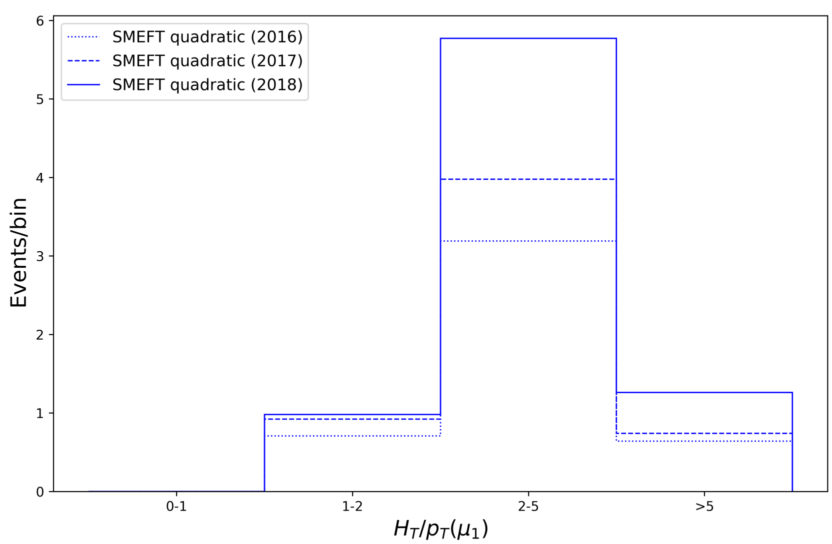

Figure 5: $H_T/p_T^{\mu_1}$ distribution obtained for the SMEFT quadratic component signal. The binning is the same as the used in the Figure 2 of [1]. The different normalizations on each correspond to the luminosity splitting for the years 2016 (dotted), 2017 (dashed) and 2018 (solid).

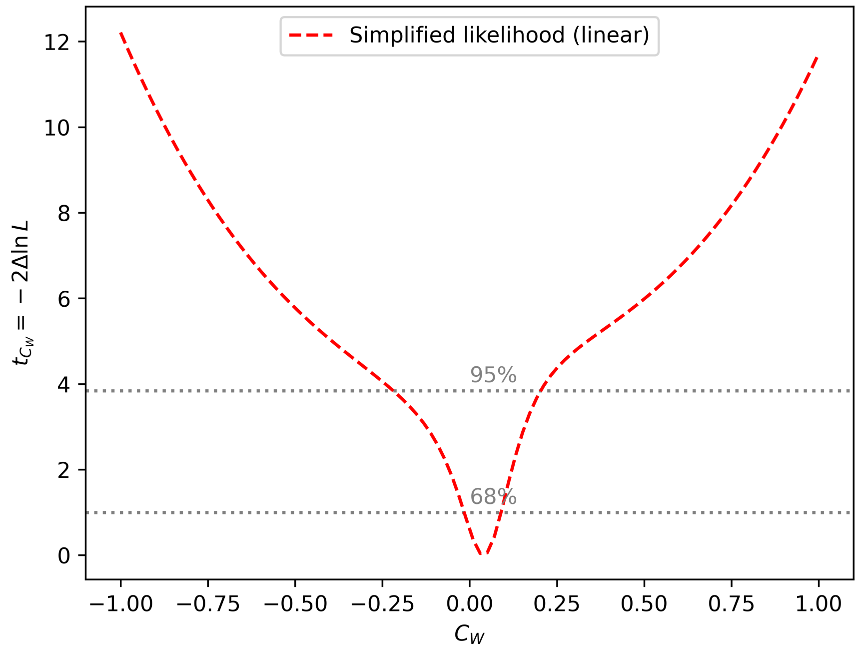

Figure 6: t_{C_W} distribution obtained performing a profile likelihood test using the simplified likelihood framework [4–6]. The two dotted horizontal lines correspond to the limit of the regions where the confidence level is 68% and 95%.

The shape of the t_{C_W} distribution obtained with our analysis is reasonable. We write the contribution of the SMEFT Q_W operator to an observable recalling the expression presented in the equation:

$$\sigma = \sigma_{SM} + \sigma_W \bar{C}_W + \sigma_{WW} \bar{C}_W^2 \tag{4}$$

The quadratic component is always positive, as seen in Figure 5. Therefore, if our signal only had a quadratic contribution, i.e. $\sigma_W = 0$, the t_{C_W} distribution should be a parabola with positive concavity and its vertex in $(C_W, t_{C_W}) = (0, 0)$. On the other hand, the linear component can be negative, as seen in Figure 4. Therefore, adding the linear contribution to the signal will move the the vertex to $C_W = -\sigma_W/2\sigma_{WW}$, which will be negative or positive depending on the sign of σ_W. This would be why the t_{C_W} distribution shown in Figure 6 resembles a positive concavity parabola near $C_W = -1$ and $C_W = 1$, but it looks like collapsing when approaching $C_W = 0$ from both sides. This would also be the reason the vertex is not located exactly on $C_W = 0$.

Additionally, employing the t_{C_W} distribution we located the values of the Wilson coefficient C_W corresponding to the intersection between the curve and the horizontal lines representing a confidence level of 68% and 95%. This limits are shown in the Table 2.

Confidence level	Lower C_W observed limit	Higher C_W observed limit
68%	-0.01575	0.08999
95%	-0.21889	0.20389

Table 2: Observed limits of the Wilson coefficient for the SMEFT operator Q_W (C_W) obtained from Figure 6 corresponding to 68% and 95% confidence levels.

5. Conclusion

In this project, we aimed to test the SMEFT dimension-six operator Q_W (equation 3) by the means of a recasting of the analysis presented in [1]. We studied the generation of two same-sign muons in VBS events. 10.000 Monte Carlo events for the collision signal, one for the linear and one for the quadratic SMEFT contributions, were simulated separately using MadGraph5_aMC@NLO [8], Pythia 8.3 [9] and Delphes [10]. With this signal we performed an analysis using MadAnalysis 5 [14–22] applying the kinematic cuts presented in Table 1 we got from the original analysis [1]. We successfully reproduced the $H_T/p_T^{\mu_1}$ distribution in Figure 2 of [1], for our two signal samples, as shown in Figure 4 and Figure 5. Using the information contained in each bin of this histograms, we performed a profile likelihood test using the simplified likelihood framework [4–6]. With this analysis we successfully constructed a $t_{C_W} = -2\Delta \ln L$ distribution varying the C_W Wilson coefficient between -1 and 1 (Figure 6). The shape obtained for this distribution was found reasonable and explained using our model. As a final step, we retrieved from the $H_T/p_T^{\mu_1}$ distributions the lower and higher values of the C_W coefficient corresponding to 68% and 95% confidence levels. This lower and higher limits are -0.01575 and 0.08999 for 68%, and -0.21889 and 0.20389 for 95% confidence levels.

Taking into consideration that this project was made within a time frame of 2 months, we can affirm that this kinds of analysis are relatively fast. Only simulating the signal sample and reusing the backgrounds saves a lot of time, and allowed us to dedicate more time to the analysis. Doing a full simulation from the collision to the detection for the background processes would have prevented this project from being realized out in this short time.

Future perspectives include a similar analysis but using other SMEFT dimension-six operators instead of Q_W. Slightly modifying the analysis, is possible to include more operators, and work with two or more at the same time. This kind of analysis also can be made using different BSM models, testing them in a similar way.

References

[1] Probing Heavy Majorana Neutrinos and the Weinberg Operator through Vector Boson Fusion Processes in Proton-Proton Collisions at $\sqrt{s} = 13\ TeV$, Phys.Rev.Lett. 131 (2023) 1, 011803, [2206.08956].

[2] I. Brivio and M. Trott, The Standard Model as an Effective Field Theory, Phys. Rept. 793 (2019) 1–98, [1706.08945].

[3] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, Dimension-Six Terms in the Standard Model Lagrangian, JHEP 1010 (2010) 085, [1008.4884].

[4] Buckley, A., Citron, M., Fichet, S. et al. The simplified likelihood framework. J. High Energ. Phys. 2019, 64 (2019).

[5] The CMS Collaboration, Simplified likelihood for the re-interpretation of public CMS results, Tech. Rep. CMS-NOTE-2017-001.

[6] S. Fichet, Taming systematic uncertainties at the LHC with the central limit theorem, Nucl.Phys. B911 (2016) 623-637, [1603.03061].

[7] G. Cowan, Use of the profile likelihood function in searches for new physics, 10.5170/CERN-2011-006.109.

[8] Alwall, J., Frederix, R., Frixione, S. et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. J. High Energ. Phys. 2014, 79 (2014).

[9] C. Bierlich et al, A comprehensive guide to the physics and usage of PYTHIA 8.3, SciPost Phys.Codeb. 2022 (2022) 8.

[10] The DELPHES 3 collaboration., de Favereau, J., Delaere, C. et al. DELPHES 3: a modular framework for fast simulation of a generic collider experiment. J. High Energ. Phys. 2014, 57 (2014).

[11] I. Brivio, Y. Jiang, M. Trott, The SMEFTsim package, theory and tools. JHEP 12 (2017) 070, [1709.06492].

[12] I. Brivio, SMEFTsim 3.0 – a practical guide. JHEP 04 (2021) 073, [2012.11343].

[13] Parton distributions for the LHC Run II, NNPDF Collaboration, JHEP 04 (2015).

[14] E. Conte, B. Fuks, and G. Serret; Comput. Phys. Commun. 184 (2013) 222.

[15] E. Conte, B. Dumont, B. Fuks and C. Wymant; Eur. Phys. J. C 74 (2014) 10, 3103.

[16] B. Dumont, B. Fuks, S. Kraml et al.; Eur. Phys. J. C 75 (2015) 2, 56.

[17] E. Conte, and B. Fuks; Int. J. Mod. Phys. A 33 (2018) 1830027.

[18] J. Y. Araz, M. Frank and B. Fuks; Eur. Phys. J. C 80 (2020) 6, 531.

[19] J. Y. Araz, B. Fuks and G. Polykratis; Eur. Phys. J. C 81 (2021) 329.

[20] J. Y. Araz, B. Fuks, M. D. Goodsell, and M. Utsch; Eur. Phys. J. C 82 (2022) 597.

[21] G. Alguero, J. Y. Araz, B. Fuks, S. Kraml; SciPost Phys. 14 (2023) 009.

[22] J. Y. Araz, A. Buckley, B. Fuks; Eur. Phys. J. C 83 (2023) no.7, 664.

[23] S. Weinberg, Baryon- and Lepton-Nonconserving Processes, Phys. Rev. Lett. 43, 1566 (1979).

[24] P. Minkowski,$\mu \to e\gamma$ at a rate of one out of 1-billion muon decays?, Phys. Lett. 67B, 421 (1977).

[25] M. Gell-Mann, P. Ramond, and R. Slansky, Complex spinors and unified theories, in Supergravity (North Holland, Amsterdam, 1979), p. 315; Conf. Proc. C 790927, 315 (1979).

[26] Proceedings of the Workshop on the Unified Theories and the Baryon Number in the Universe, edited by O. Sawada and A. Sugamoto (National Laboratories High Energy Physics, Tsukuba, Japan, 1979).

[27] R. N. Mohapatra and G. Senjanović, Neutrino Mass and Spontaneous Parity Violation, Phys. Rev. Lett. 44, 912 (1980).

[28] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, Probing color-singlet exchange in Z + 2-jet events at the CERN LHC, Phys. Rev. D 54, 6680 (1996).

[29] MadAnalysis 5 Interpreter For Expert Mode.

[30] CMS-EXO-21-003, 2022.