CMS PAS PAS-HIG-23-007

DRAFT
CMS Physics Analysis Summary

The content of this note is intended for CMS internal use and distribution only

2023/11/09
Archive Hash:
Archive Date:

Search for anomolous Higgs boson couplings in WH→ $\ell\nu b\bar{b}$ production through Vector Boson Scattering

The CMS Collaboration

Abstract

A study of the electroweak-induced production of a W boson and a Higgs boson is presented. Events are selected that target the decay of the W boson to $\mu^+\nu_\mu$, $\mu^-\overline{\nu}_\mu$, $e^+\nu_e$, or $e^-\overline{\nu}_e$, and the decay of the Higgs boson to $b\overline{b}$. In particular, beyond standard model Higgs couplings are considered, where the Higgs decay products boosted and thus reconstructed as a single merged jet. Limits are set on the parameters that scale the Higgs couplings to the W and Z bosons in the κ framework. The hypotheses where these parameters have the same magnitude but different signs with respect to the standard model are excluded.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Jonathan Guiang, Leonardo Giannini, and Philip Chang
PDFTitle:	Search for anomolous Higgs boson couplings in WH to lnubb production through Vector Boson Scattering
PDFSubject:	CMS
PDFKeywords:	CMS, Higgs, Vector Boson Scattering

Please also verify that the abstract does not use any user defined symbols

Contents

1 Introduction

Since the discovery of the Higgs boson by the ATLAS and CMS Collaborations in 2012 [1–3], various measurements of its interactions with Standard Model (SM) particles have been performed. In particular, the interactions of the Higgs boson with the electroweak gauge bosons and charged fermions of the SM have been established with coupling strengths consistent with the SM predictions, and the main production modes of the Higgs boson have also been observed [4–18]. In this note, we focus on the production of a W boson and a Higgs boson via vector boson scattering (VBS). This is a rare SM process, with a cross section of 0.075 pb.

The scattering of massive vector bosons has been studied extensively at the LHC, including $W^{\pm}W^{\pm}$ scattering [19], $W^{+}W^{-}$ scattering [20], WZ scattering [19], and ZZ scattering [21]. These processes have a unique signature due to the presence of two final state quarks, which manifest themselves as hadronic jets with a high rapidity gap, due to the lack of color flow between the quark lines. However, the production of a W boson and Higgs boson via VBS has not yet been observed. The same process can nevertheless be used to probe scenarios beyond the Standard Model (BSM). In particular, it allows for probing the relative sign of the of the WH and ZH coupling modifiers, which are referred to as κ_W and κ_Z respectively in the so-called κ-framework [22]. This is because the electroweak production of WH + 2 jets can proceed through multiple diagrams, such as the ones shown in Figure 1, where the interference between these diagrams generates a term in the cross section that is linear in both κ_W and κ_Z. Thus, assuming the expected SM magnitudes of κ_W and κ_Z, we present a search for WH production via VBS targeting the exclusion of the BSM scenario where $\lambda_{WZ} = \frac{\kappa_W}{\kappa_Z} = -1$.

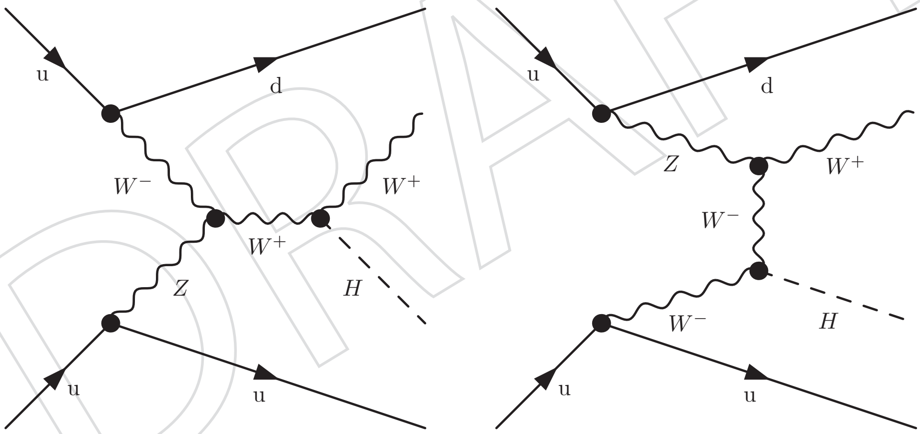

Figure 1: Two Feynman diagrams for the production of WH via vector boson scattering.

Importantly, the cross section and kinematics of the final state change for such a BSM scenario, with the Higgs and vector bosons receiving a significant boost [23]. The Higgs candidate is thus reconstructed using a single large-cone jet is used in the BSM search. The decay products of the W boson also receive a boost, so the analysis selects for a single, high-p_{T} lepton and larger missing transverse energy from the neutrino. This final states is referred to as “boosted topology” in order to distinguish it from the base where the Higgs is reconstructed as two individual jets (the “resolved topology”). Moreover, the search described in this note is fully cut-based and uses a data-driven background estimate. This simple strategy is sufficient to exclude $\lambda_{WZ} < 0$ scenarios without further optimization.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [24].

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [25]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [26].

3 Simulated samples

Signal and background processes are simulated with several Monte Carlo (MC) event generators, while the CMS detector response is modeled with GEANT4 [27].

The MC simulation of $\mathrm{pp} \to \mathrm{W}^{\pm}\mathrm{H} + \mathrm{jj}$ was generated at leading order (LO) using the MADGRAPH5_aMC@NLO generator [28] with a modified version of the Standard Model MADGRAPH model, where $\kappa_{\mathrm{W}} = -1$ and $\kappa_{\mathrm{Z}} = +1$ in order to test $\lambda_{\mathrm{WZ}} = -1$. The W and H bosons are decayed inclusively, and parton showers are handled by PYTHIA (version 8) with the CP5 tune [29]. Standard MADGRAPH phase space cuts are also applied, but two cuts are specifically tightened for generating this signal process. In particular, the pseudorapidity of any jet is required to be less than 6.5 and the invariant mass of any two jets is required to be larger than 100 GeV. The analysis has been specifically optimized for this signal sample.

In addition, a set of signal samples were generated using the reweighting feature of MADGRAPH5_aMC@NLO such that the values of κ_{W} and κ_{Z} could be varied in a two-dimensional scan. The reweighting for 625 total points was performed using $(\kappa_{\mathrm{W}} = -1, \kappa_{\mathrm{Z}} = +1)$ as a central point. These samples are generated using the same version and settings of PYTHIA as the $(\kappa_{\mathrm{W}} = -1, \kappa_{\mathrm{W}} = +1)$ sample.

The main source of background to this analysis is $t\bar{t}$ production–in particular for the boosted analysis, where one of the top decays leptonically and one of b jets can fake a boosted Higgs candidate. The $t\bar{t}$ background is generated at next-to-leading-order (NLO) accuracy in perturbative QCD with the POWHEG v2 program [30].

Other sizeable background sources are the single top production, the W+jets production and the diboson production. Diboson background events are generated with MADGRAPH5_aMC@NLO (v2.4.2) [28] at NLO with the FxFx merging scheme [31] and up to two additional partons, or with POWHEG v2.0 [32–35]. MADGRAPH5_aMC@NLO (v2.4.2) is used at LO accuracy with the MLM matching scheme [36] to generate W+jets events. Z+jets events generated with the same settings are added to background simulation, but don't contribute significantly in the analysis phase space. The single top production processes in the tW and t channels are generated to NLO accuracy with POWHEG v2.0 [32–35].

Minor backgrounds include the already mentioned Z+jets process, the production of $t\bar{t}$ + one or two bosons, including the Higgs boson, EWK production of a W or Z boson, triboson production, and the production of a Higgs bosons in the VH channel. The contributions from the $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}W}$ processes are simulated using the MADGRAPH5_aMC@NLO generator at NLO precision in QCD. Contributions from the $\mathrm{t\bar{t}WW}$, $\mathrm{t\bar{t}ZZ}$ are generated using MADGRAPH5_aMC@NLO (v2.4.2) at LO accuracy interfaced with PYTHIA. The $\mathrm{t\bar{t}}b\bar{b}$ background is generated using POWHEG (openLoops) interfaced with PYTHIA. POWHEG v2.0 [37–39] at NLO precision in QCD.

The EWK production of a Z or a W boson + jets (predominantly VBF production of a W or Z boson) are generated at LO using the MADGRAPH5_aMC@NLO v2.6.5 generator.

The VH production is generated at next-to-leading order (NLO) QCD accuracy using the POWHEG v2.0 [32–35] event generator extended with the MiNLO procedure for the quark initiated ZH and WH processes, while the gluon-induced ZH process is generated at leading order (LO) accuracy POWHEG v2.0 [32–35].

Residual processes as the EWK VV production, generated with MADGRAPH, and the VVV prodcution, generated with MADGRAPH5_aMC@NLO, are added to the simulated samples.

The simulated events at the ME level for both signal and background processes, except for the EWK V production, are interfaced with PYTHIA v8.2.2 or higher [40] to simulate the shower and hadronization of partons in the initial and final states, along with the underlying event description. The CP5 tune [29] is used everywhere. Simulated VBF signal events are interfaced with PYTHIA but, rather than the standard p_{T}-ordered parton shower, the dipole shower is chosen to model the ISR and FSR [41].

The NNPDF v3.1 NNLO parton distribution functions (PDFs) are used [42, 43].

For all samples, simulated pileup interactions are added to the hard-scattering process with multiplicity distributions matched to the data-taking year pileup profile.

4 Event reconstruction

The particle-flow algorithm [44] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. The energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

Hadronic jets are reconstructed from particle-flow objects using the anti-kT clustering algorithm [45] implemented in the FASTJET package [46, 47], with distance parameter 0.4 or 0.8. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_{T} spectrum and detector acceptance. Additional proton-proton interactions within the same or nearby bunch crossings can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate this effect, tracks identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation studies so that the average measured energy of jets becomes identical to that of particle level jets. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made [48]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures.

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is computed as the negative vector p_{T} sum of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [49]. The corrections to the energy scale of the reconstructed jets in the event are propagated to the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Events with anomalously high-$p_{\mathrm{T}}^{\mathrm{miss}}$ can arise from a variety of reconstruction failures, detector malfunctions, or non-collision backgrounds. Such events are rejected by dedicated event filters [49].

Primary vertices are reconstructed from charged-particle tracks in the event. The candidate vertex with the largest value of the sum of the p_{T}^2 of all associated physics objects is taken to be the primary pp interaction vertex. In this sum, the physics objects are the jets, clustered using the jet finding algorithm with the tracks assigned to candidate vertices as inputs, and the associated $p_{\mathrm{T}}^{\mathrm{miss}}$ [50].

The Higgs boson candidates are reconstructed as single large cone jets, which contain both the b quarks from the Higgs decay. The analysis makes use of the mass-decorrelated $X \to \mathrm{b\overline{b}}$ ParticleNet tagger [51], which combines NN components of the fatjets via a graph neural network. Additional jets originating from the hadronization of b quarks are identified using a deep neural network (DeepJet) that takes the following as input: neutral, charged particles in the jet cone, secondary vertices associated to the jet, and jet kinematic variables [52].

Muon candidates, within the geometrical acceptance of the muon detectors ($|\eta| \leq 2.4$), are reconstructed by combining the information from the silicon tracker and the muon chambers [53]. The muon candidate are also required to satisfy a set of quality criteria based on the number of hits measured in the silicon tracker and in the muon system, the properties of the muon track, and the impact parameters of the track with respect to the primary vertex of the event. Electron candidates within $|\eta| \leq 2.5$ are reconstructed by associating fitted tracks in the silicon tracker with electromagnetic energy clusters in the ECAL [54]. Electron candidates are required to satisfy identification criteria based on the shower shape of the energy deposit, the matching of the electron track to the ECAL energy cluster, the relative amount of energy deposited in the HCAL detector, and the consistency of the electron track with the primary vertex. Electron candidates in the transition region between the ECAL barrel and endcaps, $1.44 \leq |\eta| \leq 1.57$, are discarded, due to the suboptimal detector performance. Electron candidates identified as coming from photon conversions in the detector are also rejected.

Identified muons and electrons are required to be isolated from hadronic activity in the event. The isolation sum is defined by summing the p_{T} of all the PF candidates in a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ (0.3) around the muon (electron) track, where ϕ is the azimuthal angle in radians, and is corrected for the contribution of neutral particles from pileup interactions [53, 54].

5 Event selection and search strategy

The boosted VBS WH signature provides a number of signal-to-background discrimination handles. First, the scattered quarks have a characteristically large separation in pseudorapidity, $|\Delta\eta_{\mathrm{jj}}|$, as well as a large combined invariant mass, M_{jj}. In addition, when $\lambda_{\mathrm{WZ}} = \frac{\kappa_{\mathrm{W}}}{\kappa_{\mathrm{Z}}} < 0$, the W and H receive a significant boost, such that the $\mathrm{H} \to \mathrm{b\overline{b}}$ decay products are reconstructed as a merged 'fat' jet with a ΔR distance parameter of 0.8. Because background processes rarely have boosted $X \to \mathrm{b\overline{b}}$ fat jets, and more rarely $\mathrm{H} \to \mathrm{b\overline{b}}$ fat jets, the boosted Higgs alone serves as a distinguishing characteristic of the BSM VBS WH signal. Moreover, a graph neural network referred to as 'ParticleNet' has been trained to efficiently classify fat jets as having coming from particular hadronic decays of boosted particles reconstructed as single fat jets. The boosted Higgs can therefore be identified with the mass-decorrelated $X \to \mathrm{b\overline{b}}$ ParticleNet tagger (Xbb) in combination with a selection on the softdrop mass M_{SD}, the mass of the fat jet corrected with PUPPI, of the Higgs fat jet candidate. Finally, the boost in both the W and W is accessed together via the variable S_{T}, defined as follows:

$$S_{\mathrm{T}} = p_T(\ell) + E_T^{miss} + p_T(H \to b\bar{b}\ \text{fat jet}) \tag{1}$$

To start, events considered for the boosted topology are required to contain at least one fat jet that does not overlap with the lepton selected using the shared criteria described in the previous section, where fat jet and any other object are considered as overlapping if $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between them is less than 0.8. These fat jets are then required to be within the tracker acceptance $|\eta| < 2.5$, have $p_{\mathrm{T}} > 300\,\mathrm{GeV}$, and have $M > 50\,\mathrm{GeV}$ and $M_{\mathrm{SD}} > 40\,\mathrm{GeV}$, where M is the invariant mass taken from the reconstructed fat jet four-momentum. The fat jet with the highest Xbb score is taken as the boosted Higgs boson candidate. The scattered quarks are selected last, using the shared criteria described in the previous section with the additional requirement that they do not overlap with the boosted Higgs candidate.

A "Preselection" is defined to capture all of the events considered for the BSM analysis. These criteria build on the object-level selections just described, but applies more stringent selections on each object – one lepton, two jets, and one fat jet – to narrow in on a signal-like phase space. First, the combined invariant mass M_{jj} of the two scattered quarks is required to be larger than $500\,\mathrm{GeV}$. Next, the ParticleNet Xbb score of the boosted Higgs cnadidate is required to be greater than 0.3. The event is also required to have no jets passing the Medium DeepJet working point. The event must furthermore have no additional leptons passing a looser set of identification criteria than the lepton considered to be coming from the W decay. Finally, the event must have $S_T > 800\,\mathrm{GeV}$.

A set of selections are made on the handles described above towards defining a signal region in which VBS WH (where $\lambda_{WZ} < 0$) could easily be distinguished from background processes when measured in data. The signal region is therefore defined on top of the Preselection with similar, but tighter selections. First, the S_T threshold is increased to $S_T > 900\,\mathrm{GeV}$. Then, the selections on the VBS jet variables are tightened to $M_{jj} > 600\,\mathrm{GeV}$ and $|\Delta\eta_{jj}| > 4$. Finally, the selections on the boosted Higgs candidate are made much more strict, where the ParticleNet Xbb score is required to be greater than 0.9 and M_{SD} is required to be less than 150 GeV.

6 Background estimation

The background in the signal region is estimated using the "ABCD" method as follows. First, let the background yield in regions A, B, C, and D in Monte Carlo be defined as A_{MC}, B_{MC}, C_{MC}, and D_{MC}, where Region D is the signal region and Regions A, B, and C are neighboring regions where the $\Delta\eta_{jj}$ cut, the M_{SD} cut, or both are inverted. Likewise, let the the same yields in data be defined as A_{data}, B_{data}, C_{data}, and D_{data}. Under these definitions, the estimated background yield in Region D, which will be referred to as D_{data}^{pred}, can be computed with data as follows:

$$D_{data}^{pred} = C_{data} \times \frac{A_{data}}{B_{data}} \tag{2}$$

where the same can be done in MC, yielding D_{MC}^{pred}, such that the closure of the method can be measured. In doing this, it becomes clear that $\Delta\eta_{jj}$ and M_{SD} have a minor correlation in the background, leading to a systematic over-prediction taken as the systematic uncertainty on the method. However, because this correlation is well-modeled in Monte Carlo, a correction factor is also derived to correct the method when applied to data, such that the final prediction (Fig. 2) is given by

$$D_{data}^{pred} = C_{data} \times \frac{A_{data}}{B_{data}} \times \left(\frac{D_{MC}/C_{MC}}{A_{MC}/B_{MC}}\right) \tag{3}$$

The original non-closure of the method in MC is taken as the systematic uncertainty on the method.

7 Systematic uncertainties

Because the background in the signal yield is estimated from data for the merged analysis, the systematic uncertainty on the yield corresponds only to the systematic error of the ABCD method. The systematic uncertainty on the method itself has already been described. However, an additional systematic error is added to account for the uncertainty in the W+jets and diboson background compositions and how they impact the method. This is obtained by varying each background up and down by a factor of two and taking the largest difference in the estimation as the error. Thus, a total systematic uncertainty on the ABCD method of 13% is obtained.

An extensive set of systematics (ranging from 1% to 18%) is derived for the signal yield in the boosted signal region because it is taken directly from MC. Systematic errors on the renormalization and factorization scales (μ_R and μ_F), PDF, and various experimental scale factors

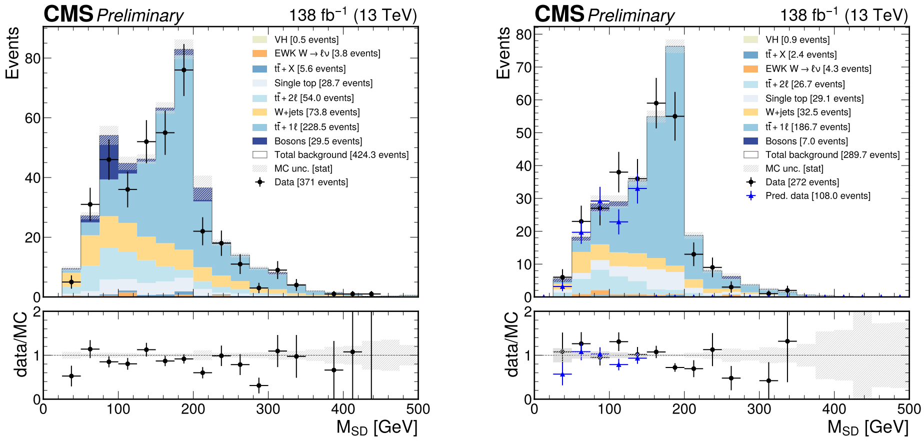

Figure 2: Data/MC comparison for the softdrop mass of the AK8 jet selected as Higgs boson candidate, in the regions used in the background estimate. In the signal regions, the blue markers show the background yield predicted from data.

are obtained by varying each individually by one standard deviation and taking the maximal difference in yield as the error. Notably, the corrections and uncertainty for the ParticleNet $X \to b\bar{b}$ discriminator distribution shape are computed using a tool originally developed for another analysis, but repurposed for general use. This tool uses a Boosted Decision Tree (BDT) to select suitable proxy jets from collision data to measure the efficiency of a given ParticleNet tagger. Ultimately, the leading systematic uncertainty (18%) on the signal yield in the boosted analysis comes from the uncertainty on the factorization scale μ_R, while the next leading is the uncertainty on the efficiency of ParticleNet (6%).

8 Results

The signal yield in the signal region is varied using the MadGraph reweighting scheme, such that many κ_W, κ_Z points can be explored. This yields a two-dimensional exclusion of each of these points with a signal strength of 1. The significance σ of each exclusion is plotted on the z-axis of a two-dimensional (κ_W, κ_Z) histogram. Because only a limited number (625 total) of reweighted points could be generated in a reasonable timeframe, the histogram is interpolated (Fig. 3) from the original histogram which had half as many bins, each centered on one of the scanned points.

9 Summary

In this note we have reported the first study of the process electroweak WH production using data collected by the CMS detector at the LHC. In total, 138 fb^{-1} of data collected between 2016 and 2018 with a center-of-mass energy of 13 TeV were analyzed. Events were selected by requiring exactly one isolated charged lepton, missing transverse momentum, two jets consistent with a vector boson scattering interaction, and an additional large-cone jet consistent with the boosted Higgs boson decay to $\mathrm{b\bar{b}}$. Scenarios where $\lambda_{WZ} < 0$ allowed by current limits are excluded at the 95% confidence level.

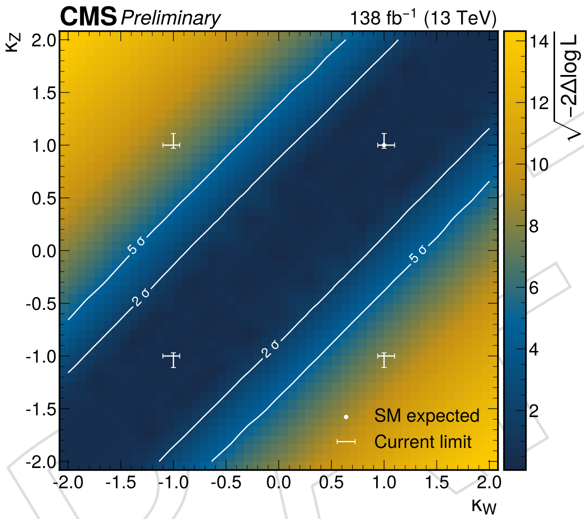

Figure 3: A 2D scan of the κ_W, κ_Z plane is plotted as a two-dimensional histogram, with the exclusion significance on the z-axis. The expected 1,2,5 σ exclusion contours are shown. The current best limits $|\kappa_W| = 1.02 \pm 0.08$ and $|\kappa_Z| = 1.04 \pm 0.07$ are plotted as capped error bars. Opposite-sign κ_W, κ_Z scenarios are excluded.

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** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[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** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] CMS Collaboration, "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector", *Phys. Lett. B* **779** (2018) 283, doi:10.1016/j.physletb.2018.02.004, arXiv:1708.00373.

[5] CMS Collaboration, "Observation of $\mathrm{t\bar{t}H}$ production", *Phys. Rev. Lett.* **120** (2018) 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.

[6] CMS Collaboration, "Observation of Higgs boson decay to bottom quarks", *Phys. Rev. Lett.* **121** (2018) 121801, doi:10.1103/PhysRevLett.121.121801, arXiv:1808.08242.

[7] CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2018) 185, doi:10.1007/JHEP11(2018)185, arXiv:1804.02716.

[8] CMS Collaboration, "Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **791** (2019) 96, doi:10.1016/j.physletb.2018.12.073, arXiv:1806.05246.

[9] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2017) 047, doi:10.1007/JHEP11(2017)047, arXiv:1706.09936.

[10] CMS Collaboration, "Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019) 421, doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.

[11] ATLAS Collaboration, "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **99** (2019) 072001, doi:10.1103/PhysRevD.99.072001, arXiv:1811.08856.

[12] ATLAS Collaboration, "Observation of $\mathrm{H} \to \mathrm{b\bar{b}}$ decays and VH production with the ATLAS detector", *Phys. Lett. B* **786** (2018) 59, doi:10.1016/j.physletb.2018.09.013, arXiv:1808.08238.

[13] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", *Phys. Lett. B* **784** (2018) 173, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.

[14] ATLAS Collaboration, "Measurements of gluon-gluon fusion and vector-boson-fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **789** (2019) 508, doi:10.1016/j.physletb.2018.11.064, arXiv:1808.09054.

[15] ATLAS Collaboration, "Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **03** (2018) 095, doi:10.1007/JHEP03(2018)095, arXiv:1712.02304.

[16] ATLAS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **98** (2018) 052005, doi:10.1103/PhysRevD.98.052005, arXiv:1802.04146.

[17] ATLAS Collaboration, "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment", *Phys. Rev. D* **101** (2020) 012002, doi:10.1103/PhysRevD.101.012002, arXiv:1909.02845.

[18] CMS Collaboration, "Evidence for Higgs boson decay to a pair of muons", *JHEP* **01** (2021) 148, doi:10.1007/JHEP01(2021)148, arXiv:2009.04363.

[19] CMS Collaboration, "Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **809** (2020) 135710, doi:10.1016/j.physletb.2020.135710, arXiv:2005.01173.

[20] CMS Collaboration, "Observation of electroweak W^+W^- pair production in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", 2023. arXiv:2205.05711. Accepted by PLB.

[21] CMS Collaboration, "Evidence for electroweak production of four charged leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **812** (2020) 135992, doi:j.physletb.2020.135992, arXiv:2008.07013.

[22] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[23] D. Stolarski and Y. Wu, "Tree-level interference in vector boson fusion production of Vh", *Phys. Rev. D* **102** (Aug, 2020) 033006, doi:10.1103/PhysRevD.102.033006.

[24] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[25] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[26] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[27] GEANT4 Collaboration, "GEANT4–a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[28] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[29] CMS Collaboration, "Extraction and validation of a new set of cms pythia8 tunes from underlying-event measurements", *The European Physical Journal C* **80** (Jan, 2020) 4, doi:10.1140/epjc/s10052-019-7499-4.

[30] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, "Top-Pair production and decay at NLO matched with parton showers", *JHEP* **04** (2015) 114, doi:10.1007/JHEP04(2015)114, arXiv:1412.1828.

[31] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[32] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[33] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[34] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[35] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM", *JHEP* **02** (2012) 088, doi:10.1007/JHEP02(2012)088, arXiv:1111.2854.

[36] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[37] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson-fusion matched with shower in POWHEG", *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.

[38] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, "$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to gosam and their merging within MiNLO", *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.

[39] H. B. Hartanto, B. Jager, L. Reina, and D. Wackeroth, "Higgs boson production in association with top quarks in the POWHEG BOX", *Phys. Rev. D* **91** (2015) 094003, doi:10.1103/PhysRevD.91.094003, arXiv:1501.04498.

[40] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[41] B. Cabouat and T. Sjöstrand, "Some dipole shower studies", *Eur. Phys. J. C* **78** (2018) 226, doi:10.1140/epjc/s10052-018-5645-z, arXiv:1710.00391.

[42] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[43] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[44] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[45] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[46] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[47] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57–61, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[48] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[49] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019), no. 07, P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[50] CMS Collaboration, "Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid", CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015.

[51] H. Qu and L. Gouskos, "ParticleNet: Jet Tagging via Particle Clouds", *Phys. Rev. D* **101** (2020), no. 5, 056019, doi:10.1103/PhysRevD.101.056019, arXiv:1902.08570.

[52] E. Bols et al., "Jet Flavour Classification Using DeepJet", *JINST* **15** (2020), no. 12, P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[53] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018), no. 06, P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[54] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.