CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2023/11/30

Search for anomalous Higgs boson couplings in WH$\rightarrow \ell\nu b\bar{b}$ 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. The data collected by the CMS experiment in 2016-2018 from proton-proton collisions at a centre-of-mass energy of 13 TeV are used, corresponding to an integrated luminosity of 138 fb^{-1}. Selected events target the leptonic decay of the W boson and the decay of the Higgs boson to $b\bar{b}$. In particular, beyond standard model (BSM) Higgs couplings are considered, such that the Higgs boson decay products are boosted and thus reconstructed as a single large-cone jet. Limits are set on the parameters that scale the Higgs couplings to the W and Z bosons in the κ framework. The BSM scenario where $\lambda_{WZ} = \kappa_W/\kappa_Z = -1$ is excluded at a confidence level (CL) larger than 99.99%, with an expected and observed significance beyond $5\,\sigma$. The hypotheses where the κ_W and κ_Z parameters are compatible with the current measurements and have different signs are similarly excluded at a CL larger than 99.99%.

1 Introduction

Since the discovery of the Higgs (H) 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 document, 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 predicted by MADGRAPH (v2.6.1) at leading order (LO) of perturbative quantum chromodynamics (QCD). Nevertheless, it can be used to probe physics scenarios beyond the standard model (BSM). In particular, it allows for probing the relative sign of the of the WH and ZH couplings, whose multiplicative modifiers with respect to the SM are referred to as κ_W and κ_Z respectively in the so-called κ-framework [19]. This is because the electroweak (EW) production of WH + 2 jets can proceed through multiple diagrams, such as the ones shown in Fig. 1, where the interference between these diagrams generates a term in the cross section that is linear in both the WH and ZH couplings. Thus, assuming the expected SM magnitudes of κ_W and κ_Z, we present a search for WH production via VBS targeting the BSM scenario where $\lambda_{WZ} = \kappa_W/\kappa_Z = -1$. Additionally, we perform a scan in the κ_W and κ_Z two-dimensional plane in order to exclude the opposite-sign ($\lambda_{WZ} < 0$) scenarios compatible with the current measurements.

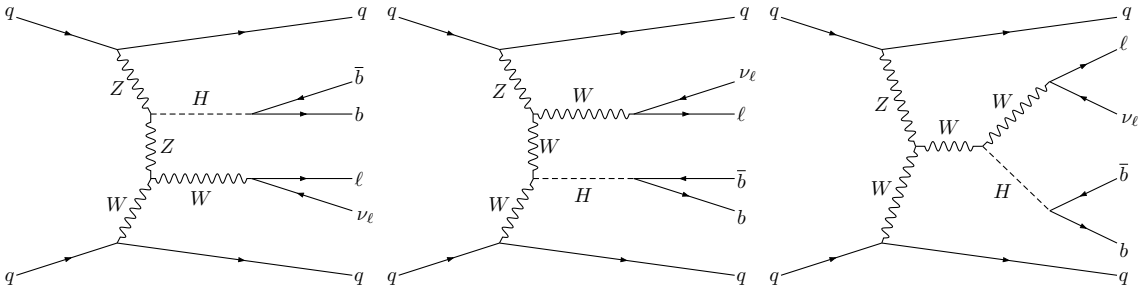

Figure 1: Feynman diagrams for the production of $WH \to \ell\nu b\bar{b}$ via VBS.

The cross section and kinematics of the final state change for $\lambda_{WZ} < 0$, where the Higgs boson and W boson both receive a significant Lorentz boost [20]. We specifically target the leptonic decay of the W boson and the decay of the Higgs boson to $b\bar{b}$. Due to the boost, the Higgs boson candidate can be reconstructed as a single large-cone jet. Since the decay products of the W boson also receive a boost, the analysis selects for one isolated high-p_T lepton and missing transverse momentum from the neutrino. The search described in this document is based on data from proton-proton (pp) collisions, at a center-of-mass energy of 13 TeV. The data were collected with the CMS experiment at the LHC from 2016 to 2018, corresponding to an integrated luminosity of 138 fb^{-1}. As mentioned in Ref. [20], the exclusion of $\lambda_{WZ} < 0$ scenarios is relevant to models with higher isospin representations [21], such as the Georgi-Machacek model [22].

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. [23].

Events of interest are selected using a two-tiered trigger system. The first level, 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 [24]. 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 [25].

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 [26]. The background MC samples listed below are only used to optimize the event selection, as the final background estimation is solely derived from data.

The MC simulation of $pp \rightarrow W^{\pm}H + jj$ is generated at leading order (LO) using the MADGRAPH5_aMC@NLO (v2.4.2) generator [27] with a modified version of the SM MADGRAPH model, where $\kappa_W = -1$ and $\kappa_Z = +1$ in order to test $\lambda_{WZ} = -1$. The decay of the bosons are modeled using PYTHIA (v8.2.2) [28], considering all decay modes of the W boson and only the $b\bar{b}$ final state of the H boson. Two MADGRAPH phase space cuts are applied: 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 matrix element reweighting feature of MADGRAPH5_aMC@NLO [29] such that the values of κ_W and κ_Z could be varied in a two-dimensional scan. The reweighting is performed for 625 points using $(\kappa_W = -1, \kappa_Z = +1)$ as a central point and varying the two coupling modifiers in the range $[-2, 2]$. These samples are generated using the same version and settings of PYTHIA as the $(\kappa_W = -1, \kappa_W = +1)$ sample. Despite the analysis being specifically optimized for $\lambda_{WZ} = -1$, it remains valid also for the majority of the BSM points, due to the analogous kinematics.

The main source of background to this analysis is the $t\bar{t}$ production, where one of the top quarks decays leptonically and one of b jets is mis-reconstructed as a boosted Higgs candidate. The $t\bar{t}$ background [30] is generated at next-to-leading-order (NLO) accuracy in perturbative QCD with the POWHEG (v2.0.0) program [31–33].

Other sizeable background sources are the single top production, the W+jets production and the diboson production. Diboson events are generated with MADGRAPH5_aMC@NLO at NLO + next-to-next-to-leading logarithmic (NNLL) accuracy with the FxFx merging scheme [34] and up to two additional partons, or with POWHEG to NLO accuracy. MADGRAPH5_aMC@NLO is used at LO accuracy with the MLM matching scheme [35] to generate W+jets events. Z+jets events are generated with the same settings, although their contribution is rather small after the analysis selection. The single top production processes in the tW and t channels are generated to NLO accuracy with POWHEG.

Minor backgrounds include the production of $t\bar{t}$ + one or two bosons, including the Higgs bo-

son, electroweak (EW) production of a W or Z boson, triboson production, and the production of a Higgs bosons in the VH (V=W/Z boson) channel. The contributions from the $t\bar{t}Z$ and $t\bar{t}W$ processes are simulated using the MADGRAPH5_aMC@NLO generator at NLO precision in QCD. Contributions from the $t\bar{t}WW$ and $t\bar{t}ZZ$ are generated using MADGRAPH5_aMC@NLO at LO accuracy interfaced with PYTHIA. The $t\bar{t}b\bar{b}$ background is generated using POWHEG (openLoops) at NLO precision in QCD interfaced with PYTHIA. The EW production of a Z or a W boson + jets, predominantly vector boson fusion production of a W or Z boson, are generated at LO using the MADGRAPH5_aMC@NLO (v2.6.5) generator. Additionally, dedicated EW diboson + additional jets production simulated with Madgraph at LO and triple boson production simulated with MADGRAPH5_aMC@NLO at NLO are considered. The VH production is generated at NLO QCD accuracy using the POWHEG event generator extended with the MiNLO procedure for the quark initiated ZH and WH processes [36], while the gluon-induced ZH process is generated at LO accuracy using the POWHEG generator.

The simulated events at the matrix-element level for background processes, except for the EW V production, are interfaced with PYTHIA v8.2.2 or higher to simulate the shower and hadronization of partons in the initial and final states, along with the underlying event description. Simulated VBS signal and EW V production events are interfaced with PYTHIA but, rather than the standard p_T-ordered parton shower, the dipole shower is chosen to model the initial-state radiation (ISR) and final-state radiation (FSR) [37]. The CP5 tune [38] for PYTHIA and the NNPDF v3.1 NNLO parton distribution functions (PDFs) [39, 40] are used for all the simulated samples. Additional pp interactions in the same and nearby bunch crossings, referred to as pileup, are also simulated. The distribution of the number of pileup interactions in the simulation is adjusted to match the one observed in the data, with an average pileup of about 34 interactions per bunch crossing.

4 Event reconstruction

The event reconstruction is handled by the particle-flow (PF) algorithm [41], which 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 reconstructed particles, referred to as PF candidates, are classified into electrons, muons, photons, charged hadrons, and neutral hadrons. The proton-proton (pp) interaction vertices are reconstructed using tracks. The primary vertex is taken to be the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [42]. All other vertices are identintified as pileup vertices. Charged PF candidates are associated with pp interaction vertices using their associated tracks.

Muon candidates are reconstructed within the geometrical acceptance of the muon detectors ($|\eta| \leq 2.4$) by combining the information from the silicon tracker and the muon chambers [43]. The muon candidates 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 [44]. 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 [43, 44].

Hadronic jets are reconstructed from PF candidates using the anti-kT clustering algorithm [45] implemented in the FASTJET package [46, 47], with distance parameter 0.4 or 0.8, defined as AK4 or AK8 jets. The 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 5-10% of the true momentum over the whole p_T spectrum and detector acceptance. Pileup interactions may 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 [46].

The Higgs boson candidates are reconstructed as single large-cone jets, which contain both of the b quarks from the Higgs boson decay. The analysis makes use of the mass-decorrelated $X \to b\bar{b}$ ParticleNet tagger [49], a graph neural network that uses the components of a large-cone jet to classify it as having originated from the decay of a boosted particle to $b\bar{b}$. The analysis relies also on the ability reconstruct the Higgs boson candidate mass. The soft-drop (SD) algorithm [50] with angular exponent $\beta = 0$ and soft radiation fraction z = 0.1, also known as the modified mass-drop algorithm [51], is applied to the Higgs boson jet candidate to remove soft and wide-angle radiation. The pileup per particle identification algorithm [52] is used to mitigate the pileup effects on the softdrop mass.

Additional AK4 jets originating from the hadronization of b quarks are identified using a deep neural network (DeepJet) based on properties such as the presence of secondary vertices and the kinematics of the jets and their constituents [53].

The missing transverse momentum vector $\vec{p}_T^{\,miss}$ is computed as the negative vector p_T sum of all the PF candidates in an event, and its magnitude is denoted as p_T^{miss} [54]. The corrections to the energy scale of the reconstructed jets in the event are propagated to the $\vec{p}_T^{\,miss}$. Events with anomalously high-p_T^{miss} can arise from a variety of reconstruction failures, detector malfunctions, or non-collision backgrounds. Such events are rejected by dedicated event filters [54].

5 Event selection and search strategy

The boosted VBS WH production provides a clean signature to reduce the contributions from background processes. Together with those features, the relatively large signal cross section for BSM scenarios allows for a simple analysis strategy. Since only leptonic W boson decays are considered, the selection criteria of online HLT triggers require the presence of one isolated lepton, electron or muon, with high p_T. The H boson Lorentz boost allows for the reconstruction of the $H \to b\bar{b}$ candidate as a single large-cone or AK8 jet, with a clustering distance ΔR of 0.8.

The boosted Higgs boson can therefore be identified with the mass-decorrelated $X \to b\bar{b}$ ParticleNet tagger (Xbb) in combination with a selection on the its softdrop mass M_{SD}. The boost in the W and H boson systems is exploited together via the variable S_T, defined as follows:

$$S_T = p_T(\ell) + p_T^{miss} + p_T(H \to b\bar{b}\ jet). \quad (1)$$

Finally, signal events are expected to exhibit a typical VBS signature, with a large rapidity separation between the scattered quarks.

Single-lepton triggers are used to select events, where the leptons are required to be isolated from other PF objects. The single-muon trigger requires $p_T > 24(27)$ GeV in the 2016, 2018 (2017) data taking years, while the single-electron trigger requires $p_T > 27(32)$ GeV in the 2016 (2017, 2018) data taking years. Events selected under each trigger are then required to contain respectively a muon with $p_T > 26$ or 30 GeV or an electron p_T with $p_T > 30$ or 35 GeV depending on the data-taking year. No additional leptons passing a looser set of identification criteria than the lepton considered to be coming from the W boson decay are allowed.

Events considered in this analysis are required to contain at least one AK8 jet that does not overlap with the selected lepton, where an AK8 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 AK8 jets are then required to be within the tracker acceptance $|\eta| < 2.5$, have $p_T > 300$ GeV, and have a mass $M > 50$ GeV and $M_{SD} > 40$ GeV, where M is the mass component of the AK8 jet four-vector, or ungroomed mass. Among the AK8 jets that pass these criteria, the AK8 jet with the highest Xbb score is taken as the boosted Higgs boson candidate. The VBS jet candidates are then selected. The event must have at least two AK4 jets that do not overlap with the lepton and Higgs boson candidate. These AK4 jets must also have $p_T > 30$ GeV and $|\eta| < 4.7$ to be considered in this analysis. If there are exactly two such jets, they are taken as the VBS quark candidates. However, if there are more than two, then the following prescription is applied. First, the jets are split by whether they are located in the positive or negative η-hemisphere, then sorted by the magnitude of their respective three-momenta. If two of the jets are in different η-hemispheres, the leading jet from each η-hemisphere is selected. Otherwise, the two jets with the largest three-momenta are taken as the VBS candidates.

After the basic event features are identified, stricter criteria are applied to define a phase space suitable for the signal region optimization and the background estimation. Specifically, the VBS jets invariant mass M_{jj} is required to be larger than 600 GeV and the ParticleNet Xbb score of the Higgs boson candidate is required to be larger than 0.9. The event is also required to have no extra jets passing the Medium DeepJet working point. Finally, S_T is required to be larger than 900 GeV.

The signal region selection uses $M_{SD} < 150$ GeV to better isolate the Higgs boson candidate, and $|\Delta\eta_{jj}| > 4$ to select more VBS-like events, where $\Delta\eta_{jj}$ is the difference in pseudorapidity between the VBS candidates. These two variables are independent to a large extent and allow for an estimation of the overall background from data. The main selection criteria are summarized in Table 1.

6 Background estimation

The background in the signal region is estimated using the "ABCD" method as follows. First, let the background yield taken from the Monte Carlo simulation in regions A, B, C, and D simulation 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 $\Delta\eta_{jj}$ requirement, the M_{SD} requirement,

Selection	Description
Trigger	Single lepton or sigle muon trigger
Lepton selection	1 isolated lepton with p_{T} larger than the trigger threshold, and no additional leptons
Higgs boson candidate	1 AK8 jet with $\lvert\eta\rvert < 2.5$, $p_{\mathrm{T}} > 300\,\mathrm{GeV}$, $M > 50\,\mathrm{GeV}$, $M_{\mathrm{SD}} > 40\,\mathrm{GeV}$, highest ParticleNet Xbb score
VBS jets candidates	2 jets with $\lvert\eta\rvert < 4.7$, $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, chosen as the largest momentum jets in each η-hemisphere
b-jet veto	no jets with $\lvert\eta\rvert < 2.5$, $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, passing the Medium DeepJet working point
Signal region	$M_{\mathrm{jj}} > 600\,\mathrm{GeV}$, $S_{\mathrm{T}} > 900\,\mathrm{GeV}$, ParticleNet Xbb > 0.9 for the Higgs candidate, $\lvert\Delta\eta_{\mathrm{jj}}\rvert > 4$, $M_{\mathrm{SD}} < 150\,\mathrm{GeV}$

Table 1: Summary of selections for the signal region.

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^{pred}_{data}, can be computed with data as follows:

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

where the same can be performed in MC, yielding D^{pred}_{MC}. The goodness of the prediction can then be quantified by comparing D^{pred}_{MC} with the MC estimate in the signal region D_{MC}. This method works if the variables defining the ABCD plane are uncorrelated, but $\Delta\eta_{\mathrm{jj}}$ and M_{SD} have a minor correlation in the background, leading to a systematic over-prediction. However, because this correlation is well-modeled in simulation, a correction factor is also derived to correct the method when applied to data, such that the final prediction (Fig. 2) is given by

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

The original non-closure of the method in MC can be taken as the systematic uncertainty in the method. Fig. 2 shows that the regions have different background compositions, so it is appropriate to extend the systematic uncertainty in the ABCD method to cover the uncertainty in the background composition. This is done by varying the sub-leading backgrounds within uncertainty, representing scenarios in which the contributions from each background are larger or smaller, and recalculating the closure of the method. Based on how the closure varies for each scenario, a final systematic of 13% is selected to cover the uncertainty in the background composition.

7 Systematic uncertainties

Since the background estimation is estimated from data, the systematic uncertainty in the yield corresponds only to the systematic uncertainty of the ABCD method (13%) described in the previous section. However, an extensive set of systematic uncertainties (ranging from 1% to 18%) is derived for the signal yield in the signal region, since the yield is evaluted from simulation. Among the various systematic uncertainty sources we consider theoretical uncertainties,

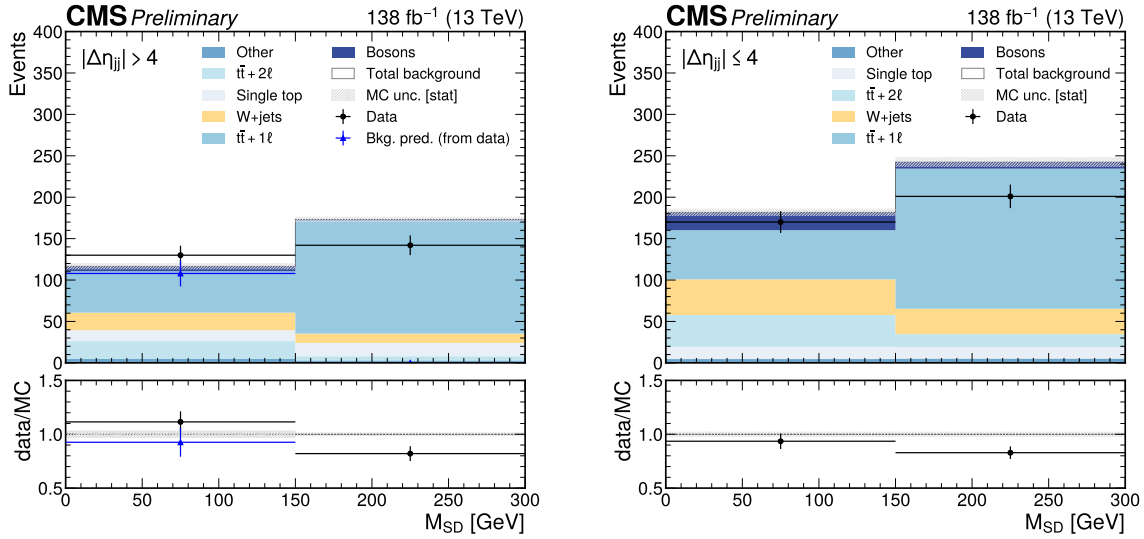

Figure 2: Data/MC comparison for the softdrop mass of the AK8 jet selected as Higgs boson candidate, in the regions used in the background estimate. Regions D and A, which both have $|\Delta\eta_{jj}| > 4$, are plotted on the left. Regions C and B, which both have $|\Delta\eta_{jj}| \leq 4$, are plotted on the right. In the signal region, region D, the blue markers show the background yield predicted from data. The small contributions from VH production, $t\bar{t} + X$ ($X = H, W, Z, b\bar{b}$, jets), and EW production of a leptonically decaying W are merged into a single category labeled "Other."

namely the renormalization and factorization scales (μ_R and μ_F) uncertainty, the parton distribution functions (PDFs) uncertainty, the ISR and FSR uncertainty, and experimental uncertainties including the lepton calibration, the jet energy scale and resolution uncertainties, the ParticleNet $X \to b\bar{b}$ tagger efficiency uncertainty and the integrated luminosity measurement [55–57]. The leading systematic uncertainty (18%) on the signal yield comes from the uncertainty in the factorization scale μ_R, while the next leading is the uncertainty in the efficiency of ParticleNet (6%).

8 Results

Results are obtained from a maximum likelihood fit of the signal and predicted background yield in the signal region to the data yield, as tabulated in Table 2. The statistical and systematic uncertainties in the signal and predicted background are treated as nuisance parameters in the fit. The exclusion significance (σ) and confidence level (CL) are extracted following the procedure described in Section 3.2 of Ref. [58]. We find that the BSM scenario where $\lambda_{WZ} = \kappa_W/\kappa_Z = -1$ can be excluded at a CL higher than 99.99%, with an expected and observed significance beyond $5\,\sigma$.

An analogous fit is performed after varying the expected signal yield in the signal region for each point the 2D κ_W-κ_Z plane. The significance σ of the exclusion for each point is plotted on the z-axis of a two-dimensional (κ_W, κ_Z) histogram. As shown in Fig. 3, all of the opposite sign scenarios with κ_W and κ_Z values compatible with the current measurements are excluded with a CL higher than 99.99%.

Type	Yield	±	stat.	±	syst.
Signal	366	±	2.9	±	68.1
Background	108	±	14.4	±	13.8
Observed	130				

Table 2: The background yield estimated from data and signal yield predicted by MC simulation in the signal region are tabulated with their associated statistical and pre-fit systematic uncertainties. The systematic uncertainty for signal quoted here is the sum in quadrature of all of the independent systematics multiplied by the total yield. The observed data yield is also tabulated.

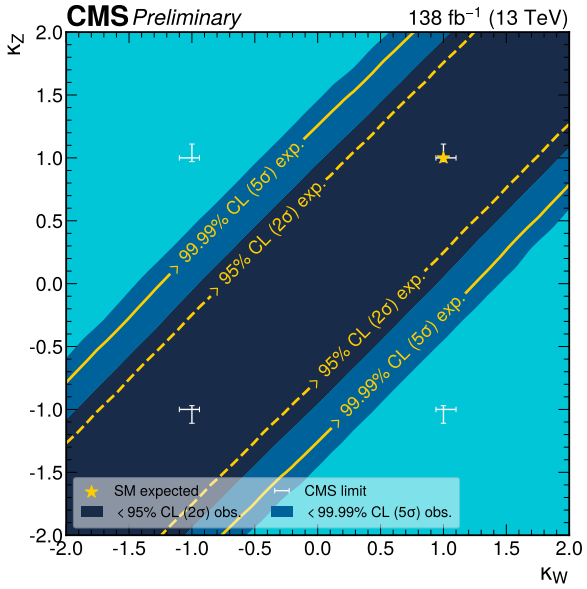

Figure 3: Exclusion significance in the 2D κ_W-κ_Z plane. The observed and expected $2\,\sigma$ and $5\,\sigma$ exclusion contours are shown. The $2\,\sigma$ and $5\,\sigma$ contours are given by $-2\Delta \log L$ less than 6.18 and 28.7 respectively. The current best limits $|\kappa_W| = 1.02 \pm 0.08$ and $|\kappa_Z| = 1.04 \pm 0.07$ are also plotted [59]. Opposite-sign κ_W-κ_Z (or negative λ_{WZ}) scenarios are excluded at an expected and observed significance beyond $5\,\sigma$.

9 Summary

In this document, we have reported the first study of electroweak WH production using data recorded with the CMS detector at the LHC. We focused on the BSM scenario where $\lambda_{WZ} = \kappa_W/\kappa_Z = -1$ and extended the search to a grid in κ_W-κ_Z phase space. 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 $b\bar{b}$. The BSM scenario where $\lambda_{WZ} = -1$ is excluded at a confidence level (CL) higher than 99.99%, with an expected and observed significance beyond $5\,\sigma$. Similarly, all the opposite sign scenarios with κ_W and κ_Z values compatible with the current measurements are excluded with a CL higher than 99.99%.

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] LHC Higgs Cross Section Working Group, "Higgs Properties: Report of the LHC Higgs Cross Section Working Group", in *Handbook of LHC Higgs Cross Sections*, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka, eds., number 3 in CERN Yellow Reports: Monographs. CERN, 7, 2013. arXiv:1307.1347.

[20] D. Stolarski and Y. Wu, "Tree-level interference in vector boson fusion production of Vh", *Phys. Rev. D* **102** (2020), no. 3, 033006, doi:10.1103/PhysRevD.102.033006, arXiv:2006.09374.

[21] I. Low and J. Lykken, "Revealing the electroweak properties of a new scalar resonance", *Journal of High Energy Physics* **2010** (Oct, 2010) 53, doi:10.1007/JHEP10(2010)053, arXiv:1005.0872.

[22] H. Georgi and M. Machacek, "Doubly charged higgs bosons", *Nuclear Physics B* **262** (1985), no. 3, 463, doi:10.1016/0550-3213(85)90325-6.

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

[24] 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.

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

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

[27] 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.

[28] 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.

[29] P. Artoisenet, V. Lemaître, F. Maltoni, and O. Mattelaer, "Automation of the matrix element reweighting method", *Journal of High Energy Physics* **2010** (Dec, 2010) 68, doi:10.1007/JHEP12(2010)068, arXiv:1007.3300.

[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] 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.

[32] 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.

[33] 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.

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

[35] 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.

[36] 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.

[37] 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.

[38] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020), no. 1, 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

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

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[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, 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] 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.

[50] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[51] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[52] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[53] 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.

[54] 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.

[55] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

[56] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[57] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[58] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.

[59] CMS Collaboration, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery.", *Nature* **607** (2022), no. 7917, 60, doi:10.1038/s41586-022-04892-x, arXiv:2207.00043.