Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration*

CERN, Switzerland

**A R T I C L E   I N F O**

Article history:
Received 13 June 2018
Received in revised form 1 December 2018
Accepted 20 December 2018
Available online 20 February 2019

Editor: M. Doser

Keywords:
Higgs
WW

**A B S T R A C T**

Measurements of the production of the standard model Higgs boson decaying to a W boson pair are reported. The $W^+W^-$ candidates are selected in events with an oppositely charged lepton pair, large missing transverse momentum, and various numbers of jets. To select Higgs bosons produced via vector boson fusion and associated production with a W or Z boson, events with two jets or three or four leptons are also selected. The event sample corresponds to an integrated luminosity of 35.9 fb$^{-1}$, collected in pp collisions at $\sqrt{s} = 13$ TeV by the CMS detector at the LHC during 2016. Combining all channels, the observed cross section times branching fraction is $1.28^{+0.18}_{-0.17}$ times the standard model prediction for the Higgs boson with a mass of 125.09 GeV. This is the first observation of the Higgs boson decay to W boson pairs by the CMS experiment.

© 2019 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/). Funded by SCOAP3.

1. Introduction

In the standard model (SM) of particle physics, the origin of the masses of the W and Z bosons is based on the spontaneous breaking of the electroweak symmetry. This symmetry breaking is achieved through the introduction of a complex doublet scalar field [1–6], leading to the prediction of the existence of one physical neutral scalar particle, commonly known as the Higgs boson (H). The observation of a new particle at a mass of approximately 125 GeV with Higgs boson-like properties was reported by the ATLAS [7] and CMS [8,9] Collaborations during the first running period of the CERN LHC in proton-proton (pp) collisions at center-of-mass energies of 7 and 8 TeV. Subsequent publications from both collaborations, based on the 7 and 8 TeV data sets [10–13], established that all measured properties of the new particle, including its spin, parity, and coupling strengths to SM particles, are consistent within the uncertainties with those expected for the SM Higgs boson. A combination of the ATLAS and CMS results [14,15] further confirmed these observations and resulted in determining the boson mass to be $m_H = 125.09 \pm 0.21$ (stat)$ \pm 0.11$ (syst) GeV.

The Higgs boson decay to a pair of W bosons was studied by the ATLAS and CMS Collaborations using the 7 and 8 TeV data sets in leptonic final states, exploring several production mechanisms [16–18]. The probability of observing a signal at least as large as the one seen, under the background-only hypothesis, corresponded to a significance of 6.5 and 4.3 standard deviations (s.d.) for ATLAS and CMS respectively, while the expected significance for a SM Higgs boson was 5.8 (5.9) s.d. for the CMS (ATLAS) collaboration. A later CMS combination [12], that includes Higgs boson production in association with a top quark pair, reported an observed significance of 4.7 s.d. for this decay. The same decay channel was used by the ATLAS and CMS Collaborations to search for the Higgs boson off-shell production [19,20] and to perform fiducial and differential cross section measurements [21,22].

In 2015, the LHC restarted at $\sqrt{s} = 13$ TeV, delivering high luminosity pp collisions. The new data are used to further constrain the properties of the Higgs boson: any significant deviation from the SM predictions would be a clear sign of new physics. This paper presents the analysis of the $H \rightarrow WW$ decay at 13 TeV, using a data sample corresponding to a total integrated luminosity of 35.9 fb$^{-1}$, collected during 2016. The same final state was recently studied by ATLAS [23] using 2015 and 2016 data.

Glue fusion (ggH) is the dominant production mode for a Higgs boson with a mass of 125 GeV in pp collisions at $\sqrt{s} = 13$ TeV. The large Higgs boson branching fraction to a W boson pair makes this channel suitable for a precision measurement of the Higgs boson production cross section, and also allows studies of subleading production channels, such as Higgs boson production via vector boson fusion (VBF) and associated production with a vector boson (VH). These channels are also studied in this paper, contributing to the precision in the measurement of the Higgs boson couplings.

The leptonic decays of the two W bosons provide the cleanest decay channel, despite the presence of neutrinos in the final state that prevents the full reconstruction of the Higgs boson mass. The different-flavor (DF) leptonic decay mode $e\mu$ has the largest branching fraction, is the least affected by background processes, and therefore is the most sensitive channel of the analysis. The same-flavor (SF) $e^+e^-$ and $\mu^+\mu^-$ final states are also considered, although their sensitivity is limited by the contamination from the Drell–Yan (DY) background with missing transverse momentum due to instrumental effects.

Events with a pair of oppositely charged leptons (electrons and/or muons) and missing transverse momentum, due to the presence of neutrinos in the final state, are selected. This signature is common to other SM processes that contribute to the background in this analysis. The main contribution comes from nonresonant production of W boson pairs (WW), an irreducible background that shares the same final state and can only be separated from the signal using kinematic distributions. Backgrounds coming from top quark events (tt and tW) are also important, followed by other processes, such as W+jets and WZ production processes. The DY process is the dominant source of background in the dielectron and dimuon final states, while it is subdominant in the electron-muon final state, since its contribution arises from the leptonic decays of the $t\bar{t}$ lepton pair.

The events are categorized by jet multiplicity to better handle the $t\bar{t}$ background. In addition, dedicated categories are designed to enhance the sensitivity to the VBF and VH production mechanisms.

2. The CMS detector

The CMS detector is a multipurpose apparatus designed to study high transverse momentum ($p_T$) physics processes in proton-proton and heavy ion collisions, and is described in detail in Ref. [24] together with a definition of the coordinate system used. A superconducting solenoid occupies its central region, providing a magnetic field of 3.8 T parallel to the beam direction. Charged particle trajectories are measured by the silicon pixel and strip trackers, which cover a pseudorapidity region of $|\eta| < 2.5$. A lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter surround the tracking volume and cover $|\eta| < 3$. The steel and quartz fiber Chenerev hadron forward calorimeter extends the coverage to $|\eta| < 5$. The muon system consists of gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid, and covers $|\eta| < 2.4$. The first level of the CMS trigger system, composed of custom hardware processors, is designed to select the most interesting events in less than 4 $\mu$s, using information from the calorimeters and muon detectors. The high-level trigger processor farm further reduces the event rate to about 1000 Hz, before data storage.

3. Data and simulated samples

The events used in this analysis are selected by high-level trigger algorithms that require the presence of one or two high-$p_T$ electrons or muons passing loose identification and isolation requirements. In single-lepton triggers, relatively tight lepton identification criteria are applied. The $p_T$ threshold is 25 GeV in the central region ($|\eta| < 2.1$) and 27 GeV for $2.1 < |\eta| < 2.5$ for electrons, while it is 24 GeV for muons ($|\eta| < 2.4$). In the dielectron trigger, the minimum required $p_T$ is 23 GeV for the leading and 12 GeV for the subleading electron. In the dimuon trigger, the minimum $p_T$ is 17 GeV for the leading and 8 GeV for the subleading muon. In the two dilepton $e\mu$ triggers used in the analysis, the minimum $p_T$ requirements are either 8 GeV for the muon and 23 GeV for the electron, or 23 GeV for the muon and 12 GeV for the electron. The combination of single-lepton and dilepton triggers provides an overall trigger efficiency in excess of 98% for selected signal events.

Several event generators are used to optimize the analysis and estimate the expected yields of signal and backgrounds, as well as their associated systematic uncertainties. Different Higgs boson production mechanisms are simulated. Both ggH and VBF are generated with POWHEG v2 [25–28], which describes the full next-to-leading order (NLO) perturbative quantum chromodynamics (QCD) properties of these processes. In addition, the ggH process is reweighted to match the Higgs boson $p_T$ and the number of associated jets to the prediction of POWHEG NNLOPS [29], which provides a next-to-next-to-leading order (NNLO) description for the inclusive Higgs boson production, NLO for the exclusive $H + 1$ jet production, and leading order (LO) for the exclusive $H + 2$ jets production. The reweighting is performed by computing the ratio of the Higgs boson $p_T$ distribution from the NNLOPS generator to that from the POWHEG generator in each jet multiplicity bin, and applying this ratio to the ggH POWHEG simulation. The MINLO HV [30] extension of POWHEG is used to simulate the associated production of the Higgs boson with vector bosons ($W^+H$, $W^-H$, $ZH$), which simulates the $VH + 0$ and $1$ jet processes with NLO accuracy. Higgs boson production in association with top or bottom quarks, such as $t\bar{t}H$ and $b\bar{b}H$ production mechanisms, are considered as well, although they only contribute to a minor extent in the phase space selected by this analysis. For the simulation of $t\bar{t}H$ production the POWHEG generator is used, while the MAGIC @NLO v2.2.2 generator [31] is used to simulate the $b\bar{b}H$ production. The Higgs boson is generated with a mass of 125.09 GeV and is made to decay into a pair of W bosons, considering only lepton W boson decays ($e$, $\mu$, or $\tau$). For Higgs bosons produced via $ggH$ [32] and VBF [33] processes, their decay into two W bosons and subsequently into leptons is simulated using JHUGEN v5.2.5 [34,35]. For the associated production mechanisms, including gluon fusion produced $ZH$, the Higgs boson decay and the associated vector boson inclusive decays are simulated by PYTHIA 8.212 [36]. The simulated signal samples are normalized using cross sections [37] and decay rates [38] computed by the LHC Higgs Cross Section Working Group. In particular the most recent next-to-next-to-next-to-leading order calculations for the inclusive gluon fusion production are used [37]. Additional simulated samples, where the Higgs boson decays into a pair of tau leptons, are also produced for each of the aforementioned production mechanisms. Unless stated otherwise, the $H \to \tau\tau$ events passing the selection are considered signal events in the signal yield determination. However, their expected contribution in the signal phase space is small compared to $H \to W^+W^-$. AWW production in order to suppress the top quark background processes, the analysis is performed defining event categories with different number of high-$p_T$ jets ($p_T > 30$ GeV). The classification of the events in bins of jet multiplicity spoils the convergence of fixed-order calculations of the $q\bar{q} \to WW$ process and requires the use of dedicated resummation techniques for an accurate prediction of the differential distributions [41,42]. The simulated $q\bar{q} \to WW$ events are therefore reweighted to reproduce the $p_T^{WW}$ distribution from the $p_T^{WW}$-resummed calculation. The LO cross section for the $gg \to WW$ process is obtained directly from MCFM. For this process, the difference between LO and NLO cross sections is significant; a $K$ factor of 1.4 is calculated [43]
and applied to the gg → WW simulation. Given the theoretical uncertainties in the K factor, and that it is mildly sensitive to the invariant mass of the WW system (m_{WW}) in the phase space of interest, an m_{WW}-independent calculation is used.

Single top quark and tt processes are generated using POWHEG v2. The cross sections of the different single top quark processes are estimated at NLO accuracy [44], while the tt cross section is computed at NNLO accuracy, with next-to-next-to-leading-logarithmic soft-gluon resummation [45].

The DY production of Z/γ* is generated using MADGRAPH5_AMC@NLO at NLO accuracy using the FxFx jet matching and merging scheme with a merging scale μ_F = 30 GeV [46], and the Z/γ* p_T distribution reweighted to match the distribution observed in data in dimuon events.

The Wγ* background was simulated with POWHEG at NLO accuracy, down to a minimum invariant mass of the virtual photon of 100 MeV. The effect of the γ* mass cutoff was estimated with a MADGRAPH5_AMC@NLO Wγ LO sample, in which the photon pair production was simulated by PYTHIA in the parton shower approximation. The impact from events in which the γ* mass is below 100 MeV was found to be one order of magnitude smaller than the uncertainties quoted in this analysis, thus their contribution was neglected.

Other multiboson processes, such as WZ, ZZ, and VVV (V = W, Z), are also simulated with MADGRAPH5_AMC@NLO at NLO accuracy.

All processes are generated using the NNPDF 3.0 [47,48] parton distribution functions (PDFs), with the accuracy matching that of the matrix element calculations. All the event generators are interfaced to PYTHIA for the showering of partons and hadronization, as well as the simulation of the underlying event (UE) and multiple-parton interactions based on the CUET8PM1 tune [49].

To estimate the systematic uncertainties related to the choice of the UE and multiple-parton interactions tune, the signal processes and the WW background are also generated with alternative tunes, which are representative of the uncertainties in the CUET8PM1 tuning parameters. The systematic uncertainty associated with showering and hadronization is estimated by interfacing the same samples with the HERWIG++ 2.7 generator [50,51], using the UE-EE-5C tune for the simulation of UE and multiple-parton interactions [49].

For all processes, the detector response is simulated using a detailed description of the CMS detector, based on the GEANT4 package [52]. Additional simulated minimum bias pp interactions from PYTHIA are overlapped with the event of interest in each collision to reproduce the number of interactions per bunch crossing (pileup) measured in data. The average number of pileup interactions is about 27 per event for the 2016 data set used in this analysis.

4. Analysis strategy

A particle-flow (PF) algorithm [53] is used to reconstruct the observable particles in the event. Energy deposits (clusters) measured by the calorimeters and charged particle tracks identified in the central tracking system and the muon detectors are combined to reconstruct individual particles.

Among the vertices reconstructed in the event, the one with the largest value of summed physics-object p_T is taken to be the primary pp interaction vertex. The physics objects include those returned by a jet-finding algorithm [54,55] applied to all charged tracks assigned to the vertex, and the associated missing transverse momentum, defined as the negative vector sum of the p_T of those objects.

Electrons are reconstructed by matching clusters in the ECAL to tracks in the silicon tracker [56]. In this analysis, electron candidates are required to have |η| < 2.5. Additional requirements are applied to reject electrons originating from photon conversions in the tracker material or jets misreconstructed as electrons. Electron identification criteria rely on observables sensitive to the bremsstrahlung along the electron trajectory, the geometrical and momentum-energy matching between the electron track and the associated energy cluster in the ECAL, as well as ECAL shower shape observables and association with the primary vertex.

Muon candidates are reconstructed in the geometrical acceptance |η| < 2.4 by combining information from the silicon tracker and the muon system. Identification criteria based on the number of measurements in the tracker and in the muon system, the fit quality of the muon track, and its consistency with its origin from the primary vertex are imposed on the muon candidates to reduce the misidentification rate.

Prompt leptons coming from EW interactions are usually isolated, whereas misidentified leptons and leptons coming from jets are often accompanied by charged or neutral particles, and can arise from a secondary vertex. Hence charged leptons are required to satisfy the isolation criterion that the p_T sum over charged PF candidates associated with the primary vertex, exclusive of the lepton itself, and neutral PF particles in a cone of a radius ΔR = √((Δφ)^2 + (Δη)^2) = 0.4 (0.3), where φ is the azimuthal angle in radians, centered on the muon (electron) direction is below a threshold of 15 (6)% relative to the muon (electron) p_T. To mitigate the effect of the pileup on this isolation variable, a correction based on the average energy density in the event [57] is applied. Additional requirements on the transverse (|d_{xy}|) and longitudinal (|d_z|) impact parameters with respect to the primary vertex are included. Electrons detected by the ECAL barrel are required to have |d_z| < 0.10 cm and |d_{xy}| < 0.05 cm, while electrons in the ECAL endcap must satisfy |d_z| < 0.20 cm and |d_{xy}| < 0.10 cm. For muons, the |d_z| parameter is required to be less than 0.10 cm, while |d_{xy}| is required to be less than 0.01 cm for muons with p_T < 20 GeV and less than 0.02 cm for p_T > 20 GeV.

The jet reconstruction starts with all PF candidates, and removes the charged ones that are not associated with the primary vertex to mitigate the pileup impact. The remaining charged PF candidates and all neutral candidates are clustered by the anti-k_T algorithm [54] with a distance parameter of 0.4. To reduce further the residual pileup contamination from neutral PF candidates, a correction based on the jet area [57] is applied. The jet energy is calibrated using both simulation and data following the technique described in Ref. [58]. To identify jets coming from b quarks (b jets), a multivariate (MVA) b tagging algorithm is used [59]. In this analysis, the chosen working point corresponds to about 80% efficiency for genuine b jets, and to a mistagging rate of about 10% for light-quark or gluon jets and of 35 to 50% for c jets. A per-jet scale factor is computed and applied to account for b tagging efficiency and mistagging rate differences between data and simulation.

The missing transverse momentum vector (p_T^{miss}), whose magnitude is denoted as p_T^{miss}, is reconstructed as the negative vectorial sum in the transverse plane of all PF particle candidate momenta. Since the presence of pileup induces a degradation of the p_T^{miss} measurement, affecting mostly backgrounds with no genuine p_T^{miss}, such as DY production, another p_T^{miss} that is constructed from only the charged particles (track p_T^{miss}) is used in events with an SF lepton pair (ee or μμ). To suppress the remaining off-peak DY contribution in categories containing events with an SF lepton pair, a dedicated MVA selection based on a boosted decision tree algorithm (BDT) is used, combining variables related to lepton kinematics and p_T^{miss}. The BDT is trained on simulated samples.
separately for different jet multiplicity categories, and the output discriminator is used to define a phase space enriched in signal events and reduced DY background contamination.

Events are required to pass the single-lepton or dilepton triggers. For each event, this analysis requires at least two high-\( p_T \) lepton candidates with opposite sign, originating from the primary vertex, categorized as dielectron, dimuon, or \( e\mu \) pairs. Only jets with \( p_T > 30 \text{ GeV} \) (20 GeV for b jets) and \(|\eta| < 4.7 \) (|\( \eta \)| < 2.4 for b jets) are considered in the analysis. Jets are ignored if they overlap with an isolated lepton within a distance of \( \Delta R = 0.3 \). In addition, the following kinematic selection is applied in the \( e\mu \) final state: one electron and one muon are required to be reconstructed in the event with a minimum \( p_T \) of 13 GeV for the electron and 10 GeV for the muon, the higher \( p_T \) threshold for the electron resulting from the trigger definition. One of the two leptons should also have a \( p_T \) greater than 25 GeV. In the case of \( SF \ e^+e^- \) and \( \mu^+\mu^- \), the leading lepton is required to have \( p_T \) greater than 25 GeV when it is an electron, or 20 GeV when it is a muon. The subleading electron is required to have \( p_T \) greater than 13 GeV, while for the muon a minimum \( p_T \) of 10 GeV is required. Both leptons are required to be well identified, isolated, and prompt.

Given the large background contribution from \( t\bar{t} \) production in both DF and SF final states, events are further categorized based on the number of jets in the event, with the 0-jet category driving the sensitivity of the analysis. A categorization of the selected events is performed, targeting different production mechanisms and different flavor compositions of the WW decay products.

5. Analysis categories

5.1. Different-flavor \( ggH \) categories

The categories described in this section target the \( ggH \) production mechanism and select the DF \( e\mu \) final state. The main background processes are the nonresonant WW, top quark (both single and pair production), DY to \( \tau \) lepton pairs, and W+jets when a jet is misidentified as a lepton. Smaller background contributions come from WZ, ZZ, \( V\gamma \), \( V\gamma^* \), and triboson production. The WW background process can be distinguished from the signal by the different kinematic properties of the lepton system, since it is dominated by the on-shell W boson pairs that do not arise from a scalar resonance decay. The top quark background process is diluted by defining different categories that depend on the number of jets in the event, and reduced by vetoing any b-tagged jet with \( p_T \) > 20 GeV.

The W+jets contribution (also referred to as nonprompt lepton background), where one jet mimics the signature of an isolated prompt lepton, is an important background process especially in the 0- and 1-jet \( ggH \)-tagged DF categories. This background is reduced by taking advantage of the charge symmetry of the signal, and the charge asymmetry of the W+jets process, in which the production of \( W^+ \) is favored over \( W^- \). Also, the fact that the probabilities for a jet to mimic an electron or a muon are different, and the fact that the misidentification rate is larger for lower-\( p_T \) leptons, are exploited. Following these physics motivations the 0- and 1-jet \( ggH \)-tagged DF categories are further split into four categories according to the lepton flavor, charge and \( p_T \) ordering: \( e^+\mu^- \), \( e^-\mu^+ \), \( \mu^+e^- \), and \( \mu^-e^+ \), where the first lepton is the one with the higher \( p_T \). In addition, the four categories are divided according to whether the subleading lepton \( p_T \) is above or below 20 GeV. This eight-fold partitioning of the 0- and 1-jet \( ggH \)-tagged categories provides an improvement in terms of the expected significance of about 15% with respect to the inclusive 0- and 1-jet categories.

To suppress background processes with three or more leptons in the final state, no additional identified and isolated leptons with \( p_T > 10 \text{ GeV} \) are allowed in the events for the dilepton categories. The dilepton invariant mass \( (m_{\ell\ell}) \) is required to be higher than 12 GeV, to reject low-mass resonances and background that comes from events with multiple jets that all arise through the strong interaction (referred to the multijet background). To suppress the background arising from DY events decaying to a \( \tau \) lepton pair, which subsequently decays to the \( e\mu \) final state, and to suppress processes without genuine missing transverse momentum, a minimum \( p_T^{\text{miss}} \) of 20 GeV is required. In the two-lepton categories, the DY background is further reduced by requiring the dilepton \( p_T (p_T^{\ell\ell}) \) to be higher than 30 GeV, as on average \( e\mu \) lepton pairs from \( Z \to \tau^+\tau^- \) decays have lower \( p_T \) than the ones from \( H \to WW \) decays. These selection criteria also reduce contributions from \( H \to WW \to \tau^+\tau^-\nu\bar{\nu} \) and \( H \to \tau^+\tau^- \). Finally, to further suppress contributions from \( Z \to \tau^+\tau^- \) and W+jets events, where the subleading lepton does not arise from a W boson decay, the transverse mass built with \( p_T^{\text{miss}} \) and the subleading lepton, defined as:

\[
m_{\ell T}^{2\text{,}p_T^{\text{miss}}} = \sqrt{2p_T^{\ell}p_T^{\text{miss}}[1 - \cos\Delta\phi(t^{\ell}, p_T^{\text{miss}})]},
\]

is required to be greater than 30 GeV. Here \( \Delta\phi(t^{\ell}, p_T^{\text{miss}}) \) is the azimuthal angle between the subleading lepton momentum and \( p_T^{\text{miss}} \).

Although the invariant mass of the Higgs boson cannot be reconstructed because of the undetected neutrinos, the expected kinematic properties of the Higgs boson production and decay can be exploited. The spin-0 nature of the SM Higgs boson results in the preferential emission of the two charged leptons in the same hemisphere. Moreover, the invariant mass of the two leptons in the signal is relatively small with respect to the one expected for a lepton pair arising from other processes, such as nonresonant WW and top quark production. On the other hand, several of the smaller remaining background processes, such as nonprompt leptons, DY to \( \tau^+\tau^- \), and \( \gamma\gamma \) resemble the same \( m_{\ell\ell} \) phase space as the Higgs boson signal. These can be partially disentangled from the signal by reconstructing the Higgs boson transverse mass as:

\[
m_T = \sqrt{2p_T^{\ell\ell}p_T^{\text{miss}}[1 - \cos\Delta\phi(\ell\ell, p_T^{\text{miss}})]},
\]

where \( \Delta\phi(\ell\ell, p_T^{\text{miss}}) \) is the azimuthal angle between the dilepton momentum and \( p_T^{\text{miss}} \). These additional background processes populate different regions of the two-dimensional plane in \( m_{\ell\ell} \) and \( m_T \). A shape analysis based on a two-dimensional binned template fit of \( m_T \) versus \( m_{\ell\ell} \) is performed to extract the Higgs boson signal in the DF ggH categories.

The observed events as a function of \( m_{\ell\ell} \) and \( m_T \) are shown in Figs. 1, 2, and 3, after the template fit to the \( (m_{\ell\ell}, m_T) \) distribution. The 0- and 1-jet categories are split into \( p_T^{Z} < 20 \text{ GeV} \) and \( p_T^{Z} > 20 \text{ GeV} \) subcategories, to show the different purity of the two regions. In these figures the postfit number of events is shown, i.e., each signal and background process is normalized to the result of a simultaneous fit to all categories, assuming that the relative proportions for the different Higgs boson production mechanisms are those predicted by the SM. The events in each bin of one of the two variables are obtained by integrating over the other, and weighted using the ratio of fitted signal (S) to the sum of signal and background (S + B). S/(S + B) ratio in each \( m_T \) bin. This ratio is then used to perform a weighted sum of the \( m_{\ell\ell} \) distributions in each \( m_T \) bin. A similar weighting procedure is applied when merging the distributions of a given variable in different categories. The
weighting procedure is used only for visualization purposes, and is not used for signal extraction.

The full list of DF ggH categories and their selection requirements is shown in Table 1.

5.2. Different-flavor VBF category

The VBF process is the second largest Higgs boson production mechanism at the LHC. This mode involves the production of a Higgs boson in association with two jets with large rapidity separations. After the common preselection, the VBF analysis requires events with exactly two jets with \( p_T > 30 \text{ GeV} \), a pseudorapidity separation \( (|\Delta \eta|) \) between the two jets greater than 3.5, and an invariant mass \( (m_{jj}) \) greater than 400 GeV. The rejection of events with more than two jets reduces the tt \( b \bar{b} \) background contribution without affecting the signal efficiency, thus improving the signal sensitivity. The VBF analysis is based on the shape of the \( m_{jj} \) distribution, and is split into two signal regions, one with \( 400 < m_{jj} < 700 \text{ GeV} \) and the other with \( m_{jj} > 700 \text{ GeV} \), to profit from the higher purity of the \( m_{jj} > 700 \text{ GeV} \) region.

5.3. Different-flavor VH with two jets category

The VH process involves the production of a Higgs boson in association with a W or Z boson. The 2-jet VH-tagged category targets final states where one vector boson (W or Z) decays into two resolved jets. This category with hadronically decaying vector bosons is affected by large backgrounds compared to the leptonic decays, but profits from a higher branching fraction. The 2-jet VH-tagged analysis reverses the pseudorapidity separation requirement of the VBF selection \( (|\Delta \eta| < 3.5) \) and requires \( m_{jj} \) to be between 65 and 105 GeV. In addition, the two leading jets are required to be central \( (|\eta| < 2.5) \) to profit from more stringent b jet veto requirements, given that b tagging can only be performed for central jets. A cut on \( \Delta R_{\ell \ell} < 2 \) is applied to suppress tt \( b \) background, taking advantage of the spin-0 nature of the Higgs boson.

Fig. 1. Postfit number of weighted events \( (N_w) \) as a function of \( m_{jj} \) and \( m_T \) for DF events with 0 jets and \( p_T < 20 \text{ GeV} \) (upper row) or \( p_T > 20 \text{ GeV} \) (lower row). The number of events is weighted according to the \( S/(S+B) \) ratio in each bin of one of the two variables, integrating over the other one. The various lepton flavors and charge subcategories are also merged and weighted according to their \( S/(S+B) \) value. The contributions of the main background processes (stacked histograms) and the Higgs boson signal (superimposed and stacked red histograms) remaining after all selection criteria are shown. The dashed gray band accounts for all systematic uncertainties on the signal and background yields after the fit.
Fig. 2. Same as previous figure, for DF events with one jet.

Fig. 3. Postfit number of weighted events ($N_w$) as a function of $m_{\ell\ell}$ and $m_T$ for DF events with at least 2 jets. The number of events is weighted according to the $S/(S+B)$ ratio in each bin of one of the two variables, integrating over the other one.
that results in leptons being preferentially emitted in nearby directions. This kinematic property is further enhanced in this category due to the boost of the Higgs boson recoiling against the associated vector boson.

The analysis is based on the shape of the $m_{T\ell}$ discriminant distribution, presented in Fig. 5. The list of event requirements applied is presented in Table 3.

5.4. Same-flavor ggH categories

Similarly to the DF ggH-tagged analysis described in Section 5.1, an analysis targeting ggH in the SF $e^+e^-$ and $\mu^+\mu^-$ channels is performed. The main challenge in this final state is the large DY background contribution. In order to control it, a BDT is trained to build a discriminator, called DYMVA, to identify DY events.

A categorization based on the $p_T$ of the subleading lepton is introduced to better control the nonprompt lepton background, and a categorization in the number of jets is used to control the top quark backgrounds. The full list of event requirements is shown in Table 4.

This is an event-counting analysis, and the event requirements are chosen to maximize the expected signal significance in each category. The DY background estimations in these channels are based exclusively on control samples in data, as described in Section 6.
Table 2
Analysis categorization and event requirements for the 2-jet VH-tagged category, in the DF dilepton final state. The phase space defined by the 2-jet VH-tagged requirements corresponds to the events shown in Fig. 4.

<table>
<thead>
<tr>
<th>Category Subcategory</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>-</td>
</tr>
<tr>
<td>2-jet VH-tagged</td>
<td></td>
</tr>
</tbody>
</table>
| eµ low mjj           | m_{ll} > 12 GeV, p_T1 > 25 GeV, p_T2 > 13 (10) GeV for e(µ)
|                      | p_T^miss > 20 GeV, p_T^1 > 30 GeV
|                      | no additional leptons with p_T > 10 GeV
|                      | lepton e and muon with opposite charges
|                      | exactly two jets with p_T > 30 GeV
|                      | 60 < m_T < 125 GeV
|                      | leptons not between the two leading jets
|                      | 400 < m_{jj} < 700 GeV and |Δη|_jj > 3.5
|                      | no b-tagged jets with p_T > 20 GeV
| eµ high mjj          | exactly two jets with p_T > 30 GeV
|                      | 60 < m_T < 125 GeV
|                      | leptons not between the two leading jets
|                      | m_{jj} > 700 GeV and |Δη|_jj > 3.5
|                      | no b-tagged jets with p_T > 20 GeV

Table 3
Analysis categorization and event requirements for the 2-jet VH-tagged category, in the DF dilepton final state. The phase space defined by the 2-jet VH-tagged requirements corresponds to the events shown in Fig. 5.

<table>
<thead>
<tr>
<th>Category Subcategory</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>-</td>
</tr>
<tr>
<td>2-jet VH-tagged</td>
<td></td>
</tr>
</tbody>
</table>
| eµ                  | m_{ll} > 12 GeV, p_T1 > 25 GeV, p_T2 > 13 (10) GeV for e(µ)
|                      | p_T^miss > 20 GeV, p_T^1 > 30 GeV
|                      | no additional leptons with p_T > 10 GeV
|                      | electron and muon with opposite charges
|                      | at least two jets with p_T > 30 GeV
|                      | two leading jets with |η| < 2.5
|                      | 60 < m_T < 125 GeV and |ΔR|_ll < 2
|                      | no b-tagged jets with p_T > 20 GeV
|                      | 65 < m_{jj} < 105 GeV and |Δη|_jj < 3.5

Fig. 5. Postfit number of events as a function of m_{ll} for DF events in the 2-jets VH-tagged category.

5.5. Associated WH production with three leptons in the final state

The three-lepton WH-tagged analysis selects events that have the leading lepton with p_T1 > 25 GeV, the subleading lepton with p_T2 > 20 GeV, and the trailing lepton with p_T3 > 15 GeV. Events with a fourth lepton with p_T > 10 GeV are discarded. A veto is applied to events with SF lepton pairs of opposite charge that are compatible with coming from the decay of a Z boson. Events containing jets with p_T > 30 GeV or b-tagged jets with p_T > 20 GeV are also vetoed, to suppress the t¯t background. The azimuthal angle between p_T^miss and the three-lepton system p_T, |Δφ(ℓℓ, p_T^miss)|, is used to reduce the contamination of nonprompt lepton backgrounds. The rest of the three-lepton WH-tagged selection is in common with the other categories. These requirements are summarized in Table 5.

The events are further divided into two categories: same-sign SF (SSSF) lepton pairs, μ±μ±e±/e±e±μ±, and opposite-sign SF (OSSF) lepton pairs, μ±μ±e±/e±e±μ∓. The two selections have different signal-over-background ratios, with the SSSF being the purest of the two. The main background contribution in both cases is the contamination from nonprompt leptons. In the OSSF category, events are required to have p_T^miss > 50 GeV to reduce the DY background.

The analysis is based on the minimum |ΔR| between oppositely charged leptons. The distribution of this variable is presented in Fig. 6, separately for the SSSF and OSSF categories.

5.6. Associated ZH production with four leptons in the final state

The ZH final state is targeted by requiring exactly four isolated leptons with tight identification criteria and zero total charge, and large p_T^miss from the undetected neutrinos. The major background processes are ZZ and t¯t production.

Among the four leptons, the pair of SF leptons with opposite charge, and with the invariant mass closest to the Z boson mass, is chosen as the Z boson candidate. The remaining dilepton system, denoted as X, can be either SF or DF. Events are therefore divided into two categories, distinguishing between the cases in which the Z boson candidate contains two DF leptons (XDF) or two SF leptons (XSF), as shown in Table 6.

The signal fraction is equally distributed in the two regions. In the XSF region, ZZ, DY, and t¯t production are the major background sources, while in the XDF region, t¯tZ and ZZ backgrounds...
are dominant. Backgrounds with two Z bosons fall predominantly into the XSF region, and enter the XDF selection only through the leptonic decays of the $\tau$ leptons. This makes the XDF region much cleaner than the XSF one.

Given the low expected signal yields in the XDF and XSF categories, the result in this case is extracted from event-counting in each category.

### Table 4

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>-</td>
<td>$m_{\ell\ell} &gt; 12$ GeV, $p_{T1} &gt; 25$ (20) GeV for e (\mu), $p_{T2} &gt; 13$ (10) GeV for e (\mu), track $p_T^{\text{miss}} &gt; 20$ GeV, $p_T^{\ell} &gt; 30$ GeV no additional leptons with $p_T &gt; 10$ GeV two electrons or two muons with opposite charges</td>
</tr>
<tr>
<td>0-jet ggH-tagged</td>
<td>$e^+e^-$ $p_{T2} &lt; 20$ GeV</td>
<td>$p_{T1} &lt; 20$ GeV, $\Delta\phi(\ell\ell) &lt; 1.7$ no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV</td>
</tr>
<tr>
<td></td>
<td>$\mu^+\mu^-$ $p_{T2} &lt; 20$ GeV</td>
<td>DYMVA &gt; 0.991, $m_{\ell\ell} &lt; 55$ GeV, $m_1 &gt; 50$ GeV, $p_{T1} &lt; 20$ GeV, $\Delta\phi(\ell\ell) &lt; 1.7$ no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV</td>
</tr>
<tr>
<td>1-jet ggH-tagged</td>
<td>$e^+e^-$</td>
<td>DYMVA &gt; 0.991, $m_{\ell\ell} &lt; 55$ GeV, $m_1 &gt; 50$ GeV, $p_{T1} &lt; 20$ GeV, $\Delta\phi(\ell\ell) &lt; 1.7$ no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV</td>
</tr>
<tr>
<td></td>
<td>$\mu^+\mu^-$</td>
<td>DYMVA &gt; 0.991, $m_{\ell\ell} &lt; 55$ GeV, $m_1 &gt; 50$ GeV, $p_{T1} &lt; 20$ GeV, $\Delta\phi(\ell\ell) &lt; 1.7$ no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>-</td>
<td>$p_{T1} &gt; 25$ GeV, $p_{T2} &gt; 20$ GeV, $p_{T3} &gt; 15$ GeV no additional leptons with $p_T &gt; 10$ GeV min-(m_{\ell\ell}-m_1) &lt; 12 GeV, total lepton charge sum ±1</td>
</tr>
<tr>
<td>3-lepton WH-tagged</td>
<td>OSSF</td>
<td>no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV $p_T^{\text{miss}} &gt; 50$ GeV, min-(m_{\ell\ell}-m_1) &lt; 100 GeV Z boson veto: $</td>
</tr>
<tr>
<td></td>
<td>SSSF</td>
<td>no jets with $p_T &gt; 30$ GeV no b-tagged jets with $p_T &gt; 20$ GeV $p_T^{\text{miss}} &gt; 50$ GeV, $\Delta\phi(\ell\ell, p_T^{\text{miss}}) &gt; 2.5$</td>
</tr>
</tbody>
</table>

Fig. 6. Postfit $\Delta R_{\ell\ell}$ distribution for events in the three-lepton WH-tagged category, split into the OSSF (left) and SSSF (right) subcategories.

6. Background estimation

6.1. Nonprompt lepton background

Events in which a single W boson is produced in association with jets may populate the signal region when a jet is misidentified as a lepton. These events contain a genuine lepton and $p_T^{\text{miss}}$
from the W boson decay as well as a second nonprompt lepton from a misidentified jet, likely arising from a B hadron decay. A similar background arises from semileptonic decays of top quark pairs, especially in the 1- and 2-jets categories. At a lower rate, multijet production and fully hadronic top quark pair decays also contribute. These backgrounds are particularly important for events with low-$p_T$ leptons and low $m_{l\ell}$, and hence in the signal region of the analysis.

The nonprompt lepton background is suppressed by the identification and isolation requirements imposed on the electrons and muons, while the remaining contribution is estimated directly from data. A control sample is defined using events in which one lepton passes the standard lepton identification and isolation criteria and another lepton candidate fails these criteria but passes a looser selection, resulting in a sample of “pass-fail” lepton pairs. The pass-fail sample is dominated by nonprompt leptons. The efficiency ($\epsilon_{\text{misID}}$) for a jet that satisfies this looser selection to pass the standard selection is estimated directly from data in an independent sample dominated by events with nonprompt leptons from multijet processes. The contamination of prompt leptons from electroweak processes in such a sample is removed using the simulation. The uncertainty from this subtraction is propagated to $\epsilon_{\text{misID}}$. The efficiency $\epsilon_{\text{misID}}$ is parameterized as a function of the $p_T$ and $\eta$ of the leptons, and is used to weight the events in the pass-fail sample by $\epsilon_{\text{misID}}/(1 - \epsilon_{\text{misID}})$, to obtain the estimated contribution from this background in the signal region. The contamination of prompt leptons in the “pass-fail” sample is corrected for using their probability to pass the standard selection given that they pass the looser selection, as measured in a Drell–Yan data control sample. The systematic uncertainty associated with the determination of $\epsilon_{\text{misID}}$ on the composition of the jet that is misidentified as a lepton. Its impact is estimated in two independent ways, which are combined to yield a conservative result. First, a closure test performed on simulated $W$+jets events with $\epsilon_{\text{misID}}$ estimated from simulated QCD multijet events provides an overall normalization uncertainty. Second, a shape uncertainty is derived by varying the jet $p_T$ threshold in the differential measurement of $\epsilon_{\text{misID}}$ in bins of the $\eta$ and $p_T$ of the lepton. The threshold is varied by a quantity that reflects the difference in the fake lepton $p_T$ spectrum between $W$+jets and $t\bar{t}$ events. The total uncertainty in $\epsilon_{\text{misID}}$, including the statistical precision of the control sample, is about 40%. This uncertainty fully covers any data/simulation differences in control regions in which two same-sign leptons are required.

### Table 6

Analysis categorization and event requirements for the ZH-tagged category, in the four-lepton final state. Here, X is defined as the remaining lepton pair after the Z boson candidate is chosen. The component leptons of X can be either same-flavor (XSF) or different-flavor (XDF).

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preselection</td>
<td>–</td>
<td>four tight and isolated leptons, with zero total charge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_T &gt; 25$ GeV for the leading lepton</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_T &gt; 15$ GeV for the second leading lepton</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_T &gt; 10$ GeV for the remaining two leptons</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no additional leptons with $p_T &gt; 10$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Z dilepton mass $&gt; 4$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X dilepton mass $&gt; 4$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no b-tagged jets with $p_T &gt; 20$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$10 &lt; m_X &lt; 50$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$35 &lt; p_T^{\text{miss}} &lt; 100$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>four-lepton invariant mass $&gt; 140$ GeV</td>
</tr>
<tr>
<td>4-lepton ZH-tagged</td>
<td>XSF</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$10 &lt; m_X &lt; 70$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_T^{\text{miss}} &gt; 20$ GeV</td>
</tr>
<tr>
<td></td>
<td>XDF</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$10 &lt; m_X &lt; 70$ GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_T^{\text{miss}} &gt; 20$ GeV</td>
</tr>
</tbody>
</table>

### Table 7

Data-to-simulation scale factors for the top quark background normalization in seven different control regions.

<table>
<thead>
<tr>
<th>Final state</th>
<th>Category</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>0-jet ggH-tagged</td>
<td>0.94 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>1-jet ggH-tagged</td>
<td>0.94 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>2-jet ggH-tagged</td>
<td>0.98 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>2-jet VH-tagged</td>
<td>0.98 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>2-jet VBF-tagged</td>
<td>1.01 ± 0.04</td>
</tr>
<tr>
<td>SF</td>
<td>0-jet ggH-tagged</td>
<td>1.03 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>1-jet ggH-tagged</td>
<td>0.98 ± 0.02</td>
</tr>
</tbody>
</table>

### 6.2. Top quark background

Background contamination from single top quark processes, in particular $tW$ associated production, and from $t\bar{t}$ production, arises because of the inefficiency of $b$ jet identification and the relatively large top quark cross sections at 13 TeV. The shapes of the top quark background distributions in the various categories are obtained from simulation, taking into account the measured $b$ jet identification inefficiencies. The normalizations are obtained from control regions enriched in top quark events. The background estimation is obtained separately for the 0-, 1-, and 2-jet ggH-tagged categories, the 2-jet VBF- and VH-tagged categories, and for DF and SF final states.

The control region for the 0-jet ggH-tagged category is defined the same way as the signal region, except for the requirement that at least one jet with $20 < p_T < 30$ GeV is identified as a $b$ jet by means of the $b$ tagging algorithm. For the 1-jet ggH-tagged top quark enriched region, exactly one jet with $p_T > 30$ GeV identified as a $b$ jet is required. In the 2-jet top quark enriched regions (either ggH-, VH-, or VBF-tagged), two jets with $p_T > 30$ GeV must be present in the event and at least one has to be identified as a $b$ jet. To reduce other backgrounds in the top quark control regions, the dilepton mass is required to be higher than 50 GeV. The derived scale factors are shown in Table 7. The normalization of the top quark background in the three- and four-lepton categories is taken from simulation with its NNLO cross section uncertainty.

The top quark $p_T$ in $t\bar{t}$ events is reweighted in simulated samples in order to have a better description of the $p_T$ distribution observed in data, as described in previous CMS analyses [60]. The difference between applying this reweighting, or not, is taken as a systematic shape uncertainty. The theoretical uncertainty related to the single top quark and $t\bar{t}$ cross sections is also taken into account. It is evaluated by varying the ratio between the single
top quark and t\bar{t} cross section by its uncertainty, which is 8% at 13 TeV [18]. A 1% theoretical uncertainty arising from PDF uncertainties and QCD scale variations affects the uncertainty on the signal region to control region ratio. All the experimental uncertainties described in Section 7 are also included as uncertainties on the top quark background shape.

6.3. Drell–Yang background

The $\tau^+\tau^-$ background is relevant for DF categories and, like the signal, populates the low-\textit{m}_T and low-m_\text{ll} phase space. The kinematic variables of this background are predicted by the simulation after reweighting the Z boson $p_T$ spectrum to match the distribution measured in the data. The normalization is estimated in data control regions by selecting events with $m_\ell < 60$ GeV and $30 < m_\ell < 80$ GeV. Normalization scale factors are extracted, separately for the 0-, 1-, 2-jet $g\gamma H$-tagged, the 2-jet VBF- and VH-tagged categories, and are shown in Table 8.

The effect of missing higher-order corrections in the DY simulation is estimated by varying the renormalization and factorization scales by a factor of two up and down. This effect is treated as a shape uncertainty and amounts to 1–2% in the DY yield. A 2% theoretical uncertainty arising from PDF uncertainties and scale variations affects the uncertainty on the signal region to control region ratio. All experimental uncertainties described in Section 7 are considered as shape uncertainties for this background process.

In the SF categories, a dominant source of background is $\text{DY} \rightarrow e^+e^-$ and $\text{DY} \rightarrow \mu^+\mu^-$. The contribution of the DY background outside the Z boson mass region (dubbed the out region, which corresponds to the signal region of the analysis) is estimated by counting the number of events in the Z boson mass region in data (in region), subtracting the non-Z-boson contribution from it, and scaling the yield by a ratio $R_{\text{out/in}}$. This ratio is defined as the fraction of events outside and inside the Z boson mass region in Monte Carlo (MC) simulation, $R_{\text{out/in}} = N_{\text{out}}/N_{\text{in}}$.

The Z boson mass region is defined as $|m_\ell\ell - m_Z| < 7.5$ GeV. Such a tight mass window is chosen to reduce the non-Z-boson background contributions, which can be split into two categories. The first one is composed of the background processes, such as top quark pair and $W^+W^-$ production, with equal decay rates into the four lepton-flavor final states ($ee\mu\mu$, $e\mu\mu\mu$, and $\mu\mu\mu\mu$). Their contributions to the Z boson mass region in data, $N_{\ell\ell}^{\text{background in}}$, can be estimated from the number of events in the $e^+e^\mp$ final state, $N_{\ell\ell}^{\text{out}}$, applying a correction factor that accounts for the differences in the detection efficiency between electrons and muons ($k_e$ and $k_\mu$):

$$N_{\ell\ell}^{\text{background in}} = \frac{1}{2} k_{\ell\ell} (N_{\ell\ell}^{\text{out}} - N_{\ell\ell}^{\text{out}}(VV)).$$  \hspace{1cm} (3)

where $\ell\ell$ stands for $ee$ or $\mu\mu$. $N_{\ell\ell}^{\text{out}}(VV)$ is the number of events, estimated from simulation, arising from WZ and ZZ decays and contributing to the $e\mu$ final state. The factor of $1/2$ comes from the relative branching fraction between the $e\ell\ell$ and $e\mu\mu$ final states. The second category is composed of background processes, such as WZ and ZZ (denoted as VV), with subsequent decay mostly into SF final states via the on-shell Z boson, which are determined from simulation. The number of events arising from these background processes contributing to the same flavor final state is denoted as $N_{\ell\ell}^{\text{out}}(VV)$.

Finally, the number of DY events in the signal region is estimated from the number of events in the SF final state, $N_{\ell\ell}^{\text{in}}$, separately for electrons and muons according to the following formula:

$$N_{\ell\ell}^{\text{out/in}} = R_{\text{out/in}} (N_{\ell\ell}^{\text{in}} - N_{\ell\ell}^{\text{background in}} - N_{\ell\ell}^{\text{out}}(VV)).$$  \hspace{1cm} (4)

The difference of the $R_{\text{out/in}}$ values from the data and simulation is taken as a systematic uncertainty, and amounts to 10–25%.

6.4. The WZ and $W\gamma^*$ background

The $W\gamma^*$ EW production is included in the simulation as part of the WZ production, and the two processes are separated using a 4 GeV threshold on the $Z/\gamma^*$ mass at the generator level. For the final states with two leptons, the WZ and $W\gamma^*$ processes may contribute to the signal region whenever one of the three leptons is not identified. Therefore, it is important to observe the process in data to validate the simulation.

The yield of the WZ background is measured in data by selecting events with three isolated leptons, two electrons and one muon ($ee\mu$), or two muons and one electron ($\mu\mu\ell$). The SF lepton pair is identified as the Z boson candidate, and its invariant mass is required to be within the Z boson mass window defined in Section 6.3. This phase space is used to derive a scale factor for the WZ simulation, which is found to be $1.14 \pm 0.18$, from the weighted average of the scale factors in the $ee\mu$ and $\mu\mu\ell$ regions with their statistical uncertainties.

A $W\gamma^*$-enriched control region is defined by selecting events with two muons with invariant mass below 4 GeV, likely arising from a $\gamma^*$ decay, and a third isolated electron or muon passing a tight identification requirement. The dimuon invariant mass region close to the $J/\psi$ resonance mass is discarded. This control region is used to derive a scale factor for the $W\gamma^*$ simulation, which is found to be $0.9 \pm 0.2$, with the uncertainty coming from the event counts in the $\mu\mu\ell$ and $e\mu\mu$ samples.

All experimental uncertainties described in Section 7 are considered as shape and yield uncertainties for the WZ and $W\gamma^*$ background determination. Moreover the effects of scale and PDF uncertainties on the normalization (3% from scale variations and 4% for PDF variations) and acceptance (3%) are included.

6.5. Nonresonant WW and other backgrounds

The nonresonant WW background populates the entire two-dimensional phase space in $m_\ell\ell$ and $m_T$, while the Higgs boson signal is concentrated at low $m_\ell\ell$ values, and $m_T$ values around the Higgs boson mass. The yield of this background is hence estimated directly from the fit procedure, separately for each category. The derived scale factors are shown in Table 9.

<p>| Table 8 | Data-to-simulation scale factors for the DY → $\tau^+\tau^-$ background normalization in the DF control regions. |</p>
<table>
<thead>
<tr>
<th>Final state</th>
<th>Category</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>0-jet $g\gamma H$-tagged</td>
<td>0.94 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>1-jet $g\gamma H$-tagged</td>
<td>1.02 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>2-jet $g\gamma H$-tagged</td>
<td>0.99 ± 0.09</td>
</tr>
<tr>
<td></td>
<td>2-jet VBF-tagged</td>
<td>0.99 ± 0.13</td>
</tr>
<tr>
<td></td>
<td>2-jet VBF-tagged</td>
<td>1.04 ± 0.16</td>
</tr>
</tbody>
</table>

<p>| Table 9 | Scale factors for the nonresonant WW background normalization. |</p>
<table>
<thead>
<tr>
<th>Final state</th>
<th>Category</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>0-jet $g\gamma H$-tagged</td>
<td>1.16 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>1-jet $g\gamma H$-tagged</td>
<td>1.05 ± 0.13</td>
</tr>
<tr>
<td></td>
<td>2-jet $g\gamma H$-tagged</td>
<td>0.8 ± 0.04</td>
</tr>
<tr>
<td></td>
<td>2-jet VH-tagged</td>
<td>0.6 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>2-jet VBF-tagged</td>
<td>0.5 ± 0.05</td>
</tr>
<tr>
<td>SF</td>
<td>0-jet $g\gamma H$-tagged</td>
<td>1.13 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>1-jet $g\gamma H$-tagged</td>
<td>1.03 ± 0.18</td>
</tr>
</tbody>
</table>
In the $q_T \to WW$ process, the $p_{T}^{\text{miss}}$ spectrum in simulation is reweighted to match the resummed calculation [41,42]. The modeling of the shape uncertainties related to missing higher orders is done in two pieces: the first varies the factorization and renormalization scales by a factor of two up and down and takes the envelope; the second independently varies the resummation scale by a factor of two up and down. The cross section of the gluon-induced WW process is scaled to NLO accuracy and the uncertainty on this $K$ factor is 15% [61]. In categories with at least two jets, the EW WW production is also taken into account. The theoretical uncertainty in the LO cross section of this process amounts to 11%, and is estimated by varying the renormalization and factorization scales by a factor of two up and down, including also the effect of PDF variations.

The WZ and $Z\gamma^*$ backgrounds in the three-lepton WH-tagged analysis are estimated using dedicated control regions from which the scale factors of $1.09 \pm 0.06$ and $1.61 \pm 0.18$, respectively, are derived. The ZZ background in the four-lepton ZH-tagged analysis is also estimated using a control region from which a scale factor of $0.96 \pm 0.07$ is derived.

All remaining backgrounds from diboson and triboson production are estimated according to their expected theoretical cross sections and the shape is taken from simulation.

7. Statistical procedure and systematic uncertainties

The statistical methodology used to interpret subsets of data selected for the $H \to WW$ analysis and to combine the results from the independent categories has been developed by the ATLAS and CMS Collaborations in the context of the LHC Higgs Combination Group. A general description of the methodology can be found in Ref. [62].

The number of events in each category and in each bin of the discriminant distributions used to extract the signal is modeled as a Poisson random variable, with a mean value that is the sum of the contributions from the processes under consideration. Systematic uncertainties are represented by individual nuisance parameters with log-normal distributions. The uncertainties affect the overall normalizations of the signal and backgrounds, as well as the shapes of the predictions across the distributions of the observables. Correlations between systematic uncertainties in different categories are taken into account.

The various control regions described in Section 6 are used to constrain individual backgrounds and are included in the fit in the form of single bins, representing the number of events in each of the control regions.

The remaining sources of systematic uncertainties of experimental and theoretical nature are described below. Effects due to the experimental uncertainties are estimated by scaling or smearing the targeted variable in the simulation and recalculating the analysis results. All experimental sources of systematic uncertainty, except for the integrated luminosity, have both a normalization and a shape component. The following experimental uncertainties are taken into account:

- The uncertainty in the measured luminosity, which is 2.5% [63].
- The trigger efficiency uncertainty associated with the combinations of single-lepton and dilepton triggers, which is 2% [64].
- The uncertainties in the lepton reconstruction and identification efficiencies, which vary within 2–5% for electrons [56] and 1–2% for muons [65], depending on $p_T$ and $\eta$.
- The muon momentum and electron energy scale and resolution uncertainties, which amount to 0.6–1.0% for electrons and 0.2% for muons.
- The jet energy scale uncertainties, which vary in the range 1–13%, depending on the $p_T$ and $\eta$ of the jet [66].
- The $p_{T}^{\text{miss}}$ resolution uncertainty includes the propagation of lepton and jet energy scale and resolution uncertainties to $p_{T}^{\text{miss}}$, as well as the uncertainties on the energy scales of particles that are not clustered into jets, and the uncertainty on the amount of energy coming from pileup interactions.
- The scale factors correcting the $b$ tagging efficiency and mistagging rates, which are varied within their uncertainties.
- The associated systematic uncertainty, which varies between 0.5–1.0% [59].

The uncertainties in the signal and background production rates due to the limited knowledge of the processes under study include several components, which are assumed to be independent: the choices of PDFs and the strong coupling constant $\alpha_s$, the UE and parton shower model, and the effects of missing higher-order corrections via variations of the renormalization and factorization scales. As most of the backgrounds are estimated from control regions in data, these theoretical uncertainties mostly affect the Higgs boson signal and they are implemented as normalization-only uncertainties unless stated otherwise.

The PDFs and $\sigma_S$ uncertainties are further split between the cross section normalization uncertainties computed by the LHC Higgs Cross Section Working Group [38] for the Higgs boson signal and their effect on the acceptance [67]. The signal cross section normalization uncertainties amount to 3% for the ggH and 2% for the VBF Higgs boson production mechanism, between 1.6% and 1.9% for VH processes, and 3.6% for tH production. The acceptance uncertainties are less than 1% for all production mechanisms.

The effect of missing higher order QCD corrections on the ggH production mechanism is split into nine individual components as identified in Ref. [37], chapter 1.4. Each component is propagated such that both the integrated effect and the correlations across different categories are properly taken into account. The overall effect on the ggH cross section is about 10%. The effect of missing higher-order corrections in the VBF and VH simulations is less than 1%, while it amounts to about 8% for the $t\bar{t}H$ simulation.

The UE uncertainty is estimated by varying the CUE8T8PM1 tune in a range corresponding to the envelope of the single tuned parameters post-fit uncertainty, as described in Section 3. The dependence on the parton shower (PS) model is estimated by comparing samples processed with different programs, as described in Section 3. The effect on the expected ggH signal yields after preselection is about 5% for the UE tuning and about 7% for the PS description, and is partially accounted for by the lepton identification scale factors and uncertainties. The remaining contribution is migration between jet categories and is anticorrelated between the 0-jet category and the categories with jets. Such effects are of the order of 15–25% for the parton shower (VBF categories being the most affected) and 5–17% for UE (2-jet VH-tagged category being the most affected). The anticorrelation between jet categories reduces the impact of these uncertainties on the final results.

Finally, the uncertainties arising from the limited number of events in the simulated samples are included independently for each bin of the discriminant distributions in each category.

8. Results

The signal strength modifier ($\mu$), defined as the ratio between the measured signal cross section and the SM expectation in the $H \to WW \to 2\ell 2\nu$ decay channel, is measured by performing a binned maximum likelihood fit using simulated binned templates for signal and background processes.
The combined results obtained using all the individual analysis categories are described in this section. A summary of the expected fraction of different signal production modes in each category is shown in Fig. 7, together with the total number of expected $H \to WW$ events. The chosen categorization proves effective in tackling the different production mechanisms, especially ggH, VBF, and VH. The measurements assume a Higgs boson mass of $m_H = 125.09$ GeV, as reported in the ATLAS and CMS combined Higgs boson mass measurement [14]. The results reported below show a very weak dependence on the Higgs boson mass hypothesis, with the expected signal yield varying within 1% when the signal mass hypothesis is varied within its measured uncertainty.

The number of expected signal and background events, and the number of observed events in data, in each category after the full event selection are shown in Tables 10 and 11.

Postfit event yields are also shown in parentheses, and correspond to the result of a simultaneous fit to all categories, assuming that the relative proportions of the different production mechanisms are those predicted by the SM.

### 8.1. Signal strength modifiers

The signal strength modifier is extracted by performing a simultaneous fit to all categories assuming that the relative proportions of the different production mechanisms are the same as the SM ones. As such, the value of $\mu$ provides an insight into the compatibility between this measurement and the SM. The combined observed signal strength modifier is:

$$\mu = 1.28^{+0.18}_{-0.17} = 1.28 \pm 0.10 \, \text{(stat)} \pm 0.11 \, \text{(syst)}^{+0.10}_{-0.07} \, \text{(theo)}, \quad (5)$$

where the statistical, systematic, and theoretical uncertainties are reported separately. The statistical component is estimated by fixing all the nuisance parameters to their best fit values and recomputing the likelihood profile. The breakdown of a given group of uncertainties (systematic or theoretical) is obtained by fixing all the nuisance parameters in the group to their best fit values, and recomputing the likelihood profile. The corresponding uncertainty is then taken as the difference in quadrature between the total uncertainty and the one obtained fixing the group of nuisance parameters. The expected and observed likelihood profiles as functions of the signal strength modifier are shown in Fig. 8, with the 68% and 95% confidence level (CL) indicated. The observed significance in the asymptotic approximation [68] of the Higgs boson production for the combination of all categories is 9.1 s.d., to be compared with the expected value of 7.1 s.d. As such, this is the first observation of the Higgs boson decay to W boson pairs with the CMS experiment.

A breakdown of the impact on $\mu$ of the different systematic uncertainties is shown in Table 12. The contributions of the normalizations that are left floating in the fit enter the statistical error on $\mu$.

In order to assess the compatibility of the observed signal with the SM predictions in each category of the analysis and to ascertain the compatibility between the different categories, a simultaneous fit in which the signal strength modifier is allowed to float independently in each category is performed. The observed signal strength modifier for each category used in the combination is reported in Fig. 9 (left). Results are generally consistent with unity, with the largest deviation showing up in the 2-jet VH-tagged category.

### Table 10

<table>
<thead>
<tr>
<th>Category</th>
<th>0-jet DF ggH-tagged</th>
<th>1-jet DF ggH-tagged</th>
<th>0-jet SF ggH-tagged</th>
<th>1-jet SF ggH-tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td>ggH</td>
<td>483.1 (642.1)</td>
<td>269.1 (339.3)</td>
<td>231.2 (324.6)</td>
<td>82.0 (92.8)</td>
</tr>
<tr>
<td>VBF</td>
<td>5.6 (7.4)</td>
<td>22.1 (29.4)</td>
<td>1.5 (2.5)</td>
<td>5.9 (5.3)</td>
</tr>
<tr>
<td>WH</td>
<td>12.4 (16.4)</td>
<td>15.8 (20.5)</td>
<td>3.3 (4.3)</td>
<td>2.9 (3.8)</td>
</tr>
<tr>
<td>ZH</td>
<td>5.2 (6.9)</td>
<td>5.0 (6.7)</td>
<td>2.6 (3.4)</td>
<td>1.4 (1.8)</td>
</tr>
<tr>
<td>tH</td>
<td>&lt;0.1 (−0.1)</td>
<td>0.2 (0.2)</td>
<td>−0.1 (&lt;−0.1)</td>
<td>−0.1 (&lt;−0.1)</td>
</tr>
<tr>
<td>bH</td>
<td>3.4 (4.4)</td>
<td>1.5 (2.0)</td>
<td>1.7 (2.3)</td>
<td>0.5 (0.7)</td>
</tr>
<tr>
<td>Signal</td>
<td>509 (677)</td>
<td>313 (398)</td>
<td>240 (337)</td>
<td>93 (108)</td>
</tr>
<tr>
<td>± total unc.</td>
<td>± (1.31)</td>
<td>± (1.24)</td>
<td>± (1.24)</td>
<td>± (1.13)</td>
</tr>
<tr>
<td>WW</td>
<td>7851 (9088)</td>
<td>3553 (3727)</td>
<td>1596 (1805)</td>
<td>373 (365)</td>
</tr>
<tr>
<td>Top quark</td>
<td>2505 (2422)</td>
<td>5395 (5224)</td>
<td>334 (339)</td>
<td>452 (443)</td>
</tr>
<tr>
<td>Nonprompt</td>
<td>1555 (1006)</td>
<td>781 (482)</td>
<td>301 (260)</td>
<td>111 (97)</td>
</tr>
<tr>
<td>DY</td>
<td>154 (154)</td>
<td>283 (302)</td>
<td>437 (459)</td>
<td>178 (216)</td>
</tr>
<tr>
<td>VV, VV*</td>
<td>368 (385)</td>
<td>327 (338)</td>
<td>101 (104)</td>
<td>43 (43)</td>
</tr>
<tr>
<td>Vγ</td>
<td>213 (210)</td>
<td>137 (128)</td>
<td>23 (26)</td>
<td>17 (19)</td>
</tr>
<tr>
<td>Other diboson</td>
<td>5.1 (5.3)</td>
<td>3.5 (3.7)</td>
<td>9.3 (9.4)</td>
<td>2.0 (2.1)</td>
</tr>
<tr>
<td>Triboson</td>
<td>9.3 (9.6)</td>
<td>16 (17)</td>
<td>1.2 (1.2)</td>
<td>1.3 (1.3)</td>
</tr>
<tr>
<td>Background</td>
<td>12660 (13280)</td>
<td>10496 (10222)</td>
<td>2803 (3004)</td>
<td>1177 (1186)</td>
</tr>
<tr>
<td>± total unc.</td>
<td>± (1.141)</td>
<td>± (1.178)</td>
<td>± (1.97)</td>
<td>± (1.83)</td>
</tr>
</tbody>
</table>

Data: 13964, 10591, 3364, 1308
Table 11
Number of expected signal and background events and number of observed events in the 2-jet, 3-lepton, and 4-lepton categories after the full event selection. Postfit event yields are also shown in parentheses, corresponding to the result of a simultaneous fit to all categories assuming that the relative proportions for the different production mechanisms are those predicted by the SM. The individual signal yields are given for different production mechanisms. For the 3-lepton WH-tagged category, the “Other diboson” background includes mainly WZ production, with a 10% contribution from ZZ events. For the 4-lepton WH-tagged category, tW and tZ are included in the top quark process, while the “Other diboson” background mainly comes from ZZ production. The total uncertainty accounts for all sources of uncertainty in signal and background yields after the fit.

<table>
<thead>
<tr>
<th>Category</th>
<th>2-jet DF</th>
<th>3-lepton</th>
<th>4-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>ggH</td>
<td>(100.6)</td>
<td>(17.4)</td>
<td>(14.6)</td>
</tr>
<tr>
<td>VBF</td>
<td>(13.3)</td>
<td>(0.6)</td>
<td>(24.5)</td>
</tr>
<tr>
<td>WH</td>
<td>(9.3)</td>
<td>(0.2)</td>
<td>(4.6)</td>
</tr>
<tr>
<td>ZH</td>
<td>(4.3)</td>
<td>(0.1)</td>
<td>(21)</td>
</tr>
<tr>
<td>tH</td>
<td>(2.1)</td>
<td>(0.2)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>bH</td>
<td>(0.7)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Signal</td>
<td>(130)</td>
<td>(25)</td>
<td>(7.4)</td>
</tr>
<tr>
<td></td>
<td>total unc.</td>
<td>(±16)</td>
<td>(±3)</td>
</tr>
<tr>
<td>WW</td>
<td>(860)</td>
<td>(34)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Top quark</td>
<td>(5187)</td>
<td>(229)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Nonprompt</td>
<td>(305)</td>
<td>(37)</td>
<td>(21)</td>
</tr>
<tr>
<td>DY</td>
<td>(110)</td>
<td>(10)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>VZ/Vγ</td>
<td>(137)</td>
<td>(6.9)</td>
<td>(10)</td>
</tr>
<tr>
<td>Vγ</td>
<td>(53)</td>
<td>(2.8)</td>
<td>(4.6)</td>
</tr>
<tr>
<td>Other diboson</td>
<td>(23)</td>
<td>(0.3)</td>
<td>(13)</td>
</tr>
<tr>
<td>Triboson</td>
<td>(15)</td>
<td>(0.3)</td>
<td>(2.0)</td>
</tr>
<tr>
<td>Background</td>
<td>(6671)</td>
<td>(253)</td>
<td>(348)</td>
</tr>
<tr>
<td></td>
<td>total unc.</td>
<td>(±502)</td>
<td>(±17)</td>
</tr>
<tr>
<td>Data</td>
<td>6802</td>
<td>285</td>
<td>386</td>
</tr>
</tbody>
</table>

Fig. 8. Observed and expected likelihood profiles for the global signal strength modifier. Dashed curves correspond to the likelihood profiles obtained including only the statistical uncertainty. The crossings with the horizontal line at $-2\Delta \ln L = 1$ (3.84) define the 68 (95)% CL interval.

category (i.e., the category targeting the associated production of a Higgs boson with a vector boson decaying hadronically). The level of compatibility of the signal strength modifiers in each category with the combined signal strength modifier corresponds to an asymptotic p-value of 0.34.

Given the sensitivity of the analysis to various production mechanisms, a fit is performed in which a different signal strength modifier is assigned to each production mechanism, i.e., $\mu_{ggH}$, $\mu_{VBF}$, $\mu_{VH}$, and $\mu_{ZH}$. A simultaneous fit to all categories is performed, and results are shown in Fig. 9 (right). The biggest deviation from unity is observed for the ZZ production mechanism, which is probed mainly by the 2-jet VH-tagged and 3-lepton WH-tagged categories. The level of compatibility of the signal strength modifiers associated with different production mechanisms with the combined signal strength modifier corresponds to an asymptotic p-value of 0.70.

A similar simultaneous fit has been performed to measure the cross section corresponding to five Higgs boson production mechanisms, using a simplified fiducial phase space, as specified in the “stage-0” simplified template cross section framework [37]. The cross sections corresponding to five Higgs boson production processes ($\sigma_{ggH}$, $\sigma_{VBF}$, $\sigma_{WH}$, $\sigma_{ZH}$, $\sigma_{VH}$) are measured requiring the generator-level Higgs boson rapidity to be $|y_H| < 2.5$. This analysis has a negligible acceptance for Higgs boson production above $|y_H| = 2.5$. The $H \rightarrow \tau \tau$ events are considered as background in this fit. The measured cross sections and their ratio with the SM predictions, for the production channels in which the analysis has sensitivity, are shown in Fig. 10. The observed deviation of the $\sigma_{VH}$ process with respect to the SM prediction corresponds to an asymptotic p-value of 0.02, and is driven by the excess of events already observed for $\mu_{ZH}$. Compared to the $\mu_{VH}$ fit, in this case the signal strength modifier for the hadronic decay of the associated W boson is fitted separately from the leptonic one, and is driven away from the SM prediction by the excess observed in the 2-jet VH-tagged category.

Table 12
Impact of the main systematic uncertainties on the signal strength $\mu$.

<table>
<thead>
<tr>
<th>Type</th>
<th>Source</th>
<th>Impact (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical</td>
<td>Signal</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>WW</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Top quark</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>PS and UE</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Sample size of simulation data</td>
<td>2</td>
</tr>
<tr>
<td>Experimental</td>
<td>Electrons</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Luminosity</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Muons</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>b-tagging</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Nonprompt</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Jets</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>$\mu^{\text{min}}$</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>VZ/Vγ scale factor</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>DY SF scale factor</td>
<td>2</td>
</tr>
</tbody>
</table>

8.2. Higgs boson couplings

Given its large cross section times branching fraction, the H \rightarrow WW channel has the potential for constraining the Higgs boson couplings to vector bosons and fermions. A fit is performed to probe these couplings. One signal strength modifier (\mu) is used to scale fermion-induced production mechanisms, i.e., ggH, t\bar{t}H, and b\bar{b}H, and another one (\muV) scales the production mechanisms associated with vector bosons, i.e., VBF and VH. The two-dimensional likelihood profile is shown in Fig. 11 (left), where the 68% and 95% CL contours in the (\mu, \muV) plane are displayed. The best fit values for the signal strength modifiers are \mu = 1.37^{+0.21}_{-0.20} and \muV = 0.78^{+0.60}_{-0.57}.

The determination of the Higgs boson coupling constants is a way to verify the theoretical predictions and to search for deviations with respect to the SM expectations. These couplings can be parametrized using two coupling modifiers associated either with fermion or vector boson vertices, using the so-called \kappa-framework [37]. The two coupling modifiers are used to scale the expected product of cross section and branching fraction to match the observed signal yields in the data, according to the following formula:

\[ \sigma \mathcal{B}(X \rightarrow H \rightarrow WW) = k_H^2 \sigma_{SM} \mathcal{B}_{SM}(X \rightarrow H \rightarrow WW), \]

where \kappa_H = \kappaF(\kappaF, \kappaV) is the Higgs boson total width modifier, defined as a function of the two fit parameters \kappaF and \kappaV. The \kappaF coupling modifier is equal to \kappaF for the ggH, t\bar{t}H, and b\bar{b}H production modes, and to \kappaV for the VBF and VH production modes. No processes other than SM ones are considered to contribute to the total width modifier. The two-dimensional likelihood profile obtained using this approach, and the corresponding 68% and 95% CL contours, are shown in Fig. 11 (right). The best fit values for the coupling modifiers, obtained with one-dimensional fits in which the other coupling is profiled, are \kappaF = 1.52^{+0.48}_{-0.41} and \kappaV = 1.10^{+0.08}_{-0.08}. The fact that \kappaV is larger than 1 while the signal strength modifier \muV is below 1 is due to the former being constrained not only by the production, but also by the decay of the Higgs boson, and thus being affected by the fact that the global observed signal strength is larger than 1.

9. Summary

Measurements of the properties of the SM Higgs boson decaying to a W boson pair at the LHC have been reported. The data samples used in the analysis correspond to an integrated luminosity of 35.9 fb\(^{-1}\) collected by the CMS detector in proton-proton collisions at \sqrt{s} = 13 TeV.

The W^+W^- candidates are selected in events with large missing transverse momentum and exactly two, three, or four leptons. In the case of events with two leptons, different categories are defined according to the lepton pair flavor, e\mu, e, or \mu\mu. The analysis has specific categories for gluon fusion production, vector
boson fusion, and vector boson associated production, with up to two jets in the final state.

The probability of observing a signal at least as large as the one seen by combining all channels, under the background-only hypothesis, corresponds to an observed significance of 9.1 standard deviations for m_T = 125.09 GeV, to be compared with the expected value of 7.1 standard deviations. The observed global signal strength modifier is \( \sigma/\sigma_{\text{SM}} = \mu = 1.28^{+0.18}_{-0.17} = 1.28 \pm 0.10 \) (stat) \( \pm 0.11 \) (syst). Measurements of the signal strength modifiers associated with the main Higgs boson production mechanisms are also performed, as well as measurements of the Higgs boson couplings to fermions and vector bosons. The measured Higgs boson production and decay properties are found to be consistent, within their uncertainties, with the SM expectation.

Acknowledgements

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staff at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COC ICIA (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); NKFIA (Hungary); DAAD and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSI P and NRF (Republic of Korea); LAS (Lithuania); MIE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, ROSATOM, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); SEID, CPAN, PCTI, and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatief door Wetenschap en Technologie (IWT-Belgium); the F.R.S.-FNRS and FWO (Belgium) under the “Excellence of Science – EOS” – be.h project n. 30820817; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Lendület (“Momentum”) Programme and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Programme ÚNKP, the NKFIA research grants 123842, 123959, 124845, 124850 and 125105 (Hungary); the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant MDM-2015-0509 and the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References


CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at √s = 8 TeV, J. Instrum. 10
The CMS Collaboration

A.M. Sirunyan, A. Tumasyan

Yerevan Physics Institute, Yerevan, Armenia


Institut für Hochenergiephysik, Wien, Austria

V. Chekhovsky, V. Mossolov, J. Suarez Gonzalez

Institute for Nuclear Problems, Minsk, Belarus


Universiteit Antwerpen, Antwerpen, Belgium


Vrije Universiteit Brussel, Brussels, Belgium


Université Libre de Bruxelles, Bruxelles, Belgium


Ghent University, Gent, Belgium


Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil


Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil


a Universidade Estadual Paulista, São Paulo, Brazil
b Universidade Federal do ABC, São Paulo, Brazil

A. Aleksandrov, R. Hadjiiska, P. Iaydjiev, A. Marinov, M. Misheva, M. Rodozov, M. Shopova, G. Sultanov

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria

A. Dimitrov, L. Litov, B. Pavlov, P. Petkov

University of Sofia, Sofia, Bulgaria

W. Fang, X. Gao, L. Yuan

Beihang University, Beijing, China


Institute of High Energy Physics, Beijing, China

Y. Ban, G. Chen, A. Levin, J. Li, L. Li, Q. Li, Y. Mao, S.J. Qian, D. Wang, Z. Xu

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China

Y. Wang

Tsinghua University, Beijing, China

C. Avila, A. Cabrera, C.A. Carrillo Montoya, L.F. Chaparro Sierra, C. Florez, C.F. González Hernández, M.A. Segura Delgado

Universidad de Los Andes, Bogota, Colombia

B. Courbon, N. Godinovic, D. Lelas, I. Puljak, T. Sculac

University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia

Z. Antunovic, M. Kovac

University of Split, Faculty of Science, Split, Croatia

V. Brigljevic, D. Ferencek, K. Kadija, B. Mesic, A. Starodumov, T. Susa

Institute Rudjer Boskovic, Zagreb, Croatia


University of Cyprus, Nicosia, Cyprus
M. Finger, M. Finger Jr.
Charles University, Prague, Czech Republic

E. Ayala
Escuela Politecnica Nacional, Quito, Ecuador

E. Carrera Jarrin
Universidad San Francisco de Quito, Quito, Ecuador

H. Abdalla, A.A. Abdelalim, A. Mohamed
Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt

S. Bhowmik, A. Carvalho Antunes De Oliveira, R.K. Dewanjee, K. Ehaataht, M. Kadastik, M. Raidal, C. Veelken
National Institute of Chemical Physics and Biophysics, Tallinn, Estonia

P. Eerola, H. Kirschenmann, J. Pekkanen, M. Voutilainen
Department of Physics, University of Helsinki, Helsinki, Finland

Helsinki Institute of Physics, Helsinki, Finland

T. Tuuva
Lappeenranta University of Technology, Lappeenranta, Finland

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

Laboratoire Leprince-Ringuet, Ecole polytechnique, CNRS/IN2P3, Université Paris-Saclay, Palaiseau, France

Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France

S. Gadrat
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France

A. Khvedelidze
Georgian Technical University, Tbilisi, Georgia
P. Azzi\textsuperscript{a}, N. Bacchetta\textsuperscript{a}, D. Bisello\textsuperscript{a,b}, A. Boletti\textsuperscript{a,b}, A. Bragagnolo, R. Carlin\textsuperscript{a,b}, P. Checchia\textsuperscript{a}, M. Dall’Oss\textsuperscript{a,b}, P. De Castro Manzano\textsuperscript{a}, T. Dorigo\textsuperscript{a}, U. Dosselli\textsuperscript{a}, F. Gasparini\textsuperscript{a,b}, U. Gasparini\textsuperscript{a,b}, S.Y. Hoh, S. Lacaprara\textsuperscript{a}, P. Lujan, M. Margoni\textsuperscript{a,b}, A.T. Meneguzzo\textsuperscript{a,b}, J. Pazzi\textsuperscript{a,b}, N. Pozzobon\textsuperscript{a,b}, P. Ronchese\textsuperscript{a,b}, R. Rossin\textsuperscript{a,b}, F. Simonetto\textsuperscript{a,b}, A. Tiko, E. Torassa\textsuperscript{a}, S. Ventura\textsuperscript{a}, M. Zanetti\textsuperscript{a,b}, P. Zotto\textsuperscript{a,b}

\textsuperscript{a} INFN Sezione di Padova, Padova, Italy
\textsuperscript{b} Università di Padova, Padova, Italy
\textsuperscript{c} Università di Trieste, Trieste, Italy

A. Braghieri\textsuperscript{a}, A. Magnani\textsuperscript{a}, P. Montagna\textsuperscript{a,b}, S.P. Ratti\textsuperscript{a,b}, V. Re\textsuperscript{a}, M. Ressegotti\textsuperscript{a,b}, C. Riccardi\textsuperscript{a,b}, P. Salvini\textsuperscript{a}, I. Vai\textsuperscript{a,b}, P. Vitulo\textsuperscript{a,b}

\textsuperscript{a} INFN Sezione di Pavia, Pavia, Italy
\textsuperscript{b} Università di Pavia, Pavia, Italy

L. Alunni Solestiz\textsuperscript{a,b}, M. Biasini\textsuperscript{a,b}, G.M. Bilei\textsuperscript{a}, C. Cecchi\textsuperscript{a,b}, D. Ciangottini\textsuperscript{a,b}, L. Fanò\textsuperscript{a,b}, P. Lariccia\textsuperscript{a,b}, R. Leonard\textsuperscript{a,b}, E. Manoni\textsuperscript{a}, G. Mantovani\textsuperscript{a,b}, V. Mariani\textsuperscript{a,b}, M. Menichelli\textsuperscript{a}, A. Rossi\textsuperscript{a,b}, A. Santocchia\textsuperscript{a,b}, D. Spiga\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Perugia, Perugia, Italy
\textsuperscript{b} Università di Perugia, Perugia, Italy

K. Androsov\textsuperscript{a}, P. Azzurri\textsuperscript{a}, G. Bagliesi\textsuperscript{a}, L. Bianchini\textsuperscript{a}, T. Boccali\textsuperscript{a}, L. Borrello, R. Castaldi\textsuperscript{a}, M.A. Ciocci\textsuperscript{a,b}, R. Dell’Orso\textsuperscript{a}, G. Fed\textsuperscript{a}, F. Fior\textsuperscript{a,c}, L. Giannini\textsuperscript{a,c}, A. Giassi\textsuperscript{a}, M.T. Grippo\textsuperscript{a}, F. Ligabue\textsuperscript{a,c}, E. Manca\textsuperscript{a,c}, G. Mandorli\textsuperscript{a,c}, A. Messineo\textsuperscript{a,b}, F. Palla\textsuperscript{a}, A. Rizzi\textsuperscript{a,b}, P. Spagnolo\textsuperscript{a}, R. Tenchini\textsuperscript{a}, G. Tonelli\textsuperscript{a,b}, A. Venturi\textsuperscript{a}, P.G. Verdini\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Pisa, Pisa, Italy
\textsuperscript{b} Università di Pisa, Pisa, Italy
\textsuperscript{c} Scuola Normale Superiore di Pisa, Pisa, Italy

L. Barone\textsuperscript{a,b}, F. Cavallari\textsuperscript{a}, M. Cipriani\textsuperscript{a,b}, N. Daci\textsuperscript{a}, D. Del Re\textsuperscript{a,b}, E. Di Marco\textsuperscript{a,b}, M. Diemoz\textsuperscript{a}, S. Gelli\textsuperscript{a,b}, E. Longo\textsuperscript{a,b}, B. Marzocchi\textsuperscript{a,b}, P. Meridiani\textsuperscript{a}, G. Organtini\textsuperscript{a,b}, F. Pandolfi\textsuperscript{a}, R. Paramatti\textsuperscript{a,b}, F. Preiato\textsuperscript{a,b}, S. Rahatlou\textsuperscript{a,b}, C. Rovelli\textsuperscript{a}, F. Santanastasio\textsuperscript{a,b}

\textsuperscript{a} INFN Sezione di Roma, Rome, Italy
\textsuperscript{b} Sapienza Università di Roma, Rome, Italy

N. Amapane\textsuperscript{a,b}, R. Arcidiacono\textsuperscript{a,c}, S. Argiro\textsuperscript{a,b}, M. Arneodo\textsuperscript{a,c}, N. Bartosik\textsuperscript{a}, R. Bellan\textsuperscript{a,b}, C. Biino\textsuperscript{a}, N. Cartiglia\textsuperscript{a}, F. Cenna\textsuperscript{a,b}, S. Cometti, M. Costa\textsuperscript{a,b}, R. Covarelli\textsuperscript{a,b}, N. Demaria\textsuperscript{a}, B. Kiani\textsuperscript{a,b}, C. Mariotti\textsuperscript{a}, S. Maselli\textsuperscript{a}, E. Migliore\textsuperscript{a,b}, V. Monaco\textsuperscript{a,b}, E. Montel\textsuperscript{a,b}, M. Monteno\textsuperscript{a}, M.M. Obertino\textsuperscript{a,b}, L. Pacher\textsuperscript{a,b}, N. Pastrone\textsuperscript{a}, M. Pelliccioni\textsuperscript{a}, G.L. Pinna Angioni\textsuperscript{a,b}, A. Romero\textsuperscript{a,b}, M. Ruspa\textsuperscript{a,c}, R. Sacchi\textsuperscript{a,b}, K. Shchelina\textsuperscript{a,b}, V. Sola\textsuperscript{a}, A. Solano\textsuperscript{a,b}, D. Soldi, A. Staiano\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Torino, Torino, Italy
\textsuperscript{b} Università di Torino, Torino, Italy
\textsuperscript{c} Università del Piemonte Orientale, Novara, Italy

S. Belforte\textsuperscript{a}, V. Candelise\textsuperscript{a,b}, M. Casarsa\textsuperscript{a}, F. Cossutti\textsuperscript{a}, G. Della Ricca\textsuperscript{a,b}, F. Vazzoler\textsuperscript{a,b}, A. Zanetti\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Trieste, Trieste, Italy
\textsuperscript{b} Università di Trieste, Trieste, Italy


Kyungpook National University, Daegu, Republic of Korea

H. Kim, D.H. Moon, G. Oh

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Republic of Korea

J. Goh\textsuperscript{29}, T.J. Kim

Hanyang University, Seoul, Republic of Korea

Korea University, Seoul, Republic of Korea

H.S. Kim

Sejong University, Seoul, Republic of Korea


Seoul National University, Seoul, Republic of Korea


University of Seoul, Seoul, Republic of Korea

Y. Choi, C. Hwang, J. Lee, I. Yu

Sungkyunkwan University, Suwon, Republic of Korea

V. Dudenas, A. Juodagalvis, J. Vaitkus

Vilnius University, Vilnius, Lithuania


National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia

A. Castaneda Hernandez, J.A. Murillo Quijada

Universidad de Sonora (UNISON), Hermosillo, Mexico


Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico

S. Carrillo Moreno, C. Oropeza Barrera, F. Vazquez Valencia

Universidad Iberoamericana, Mexico City, Mexico

J. Eysermans, I. Pedraza, H.A. Salazar Ibarguen, C. Uribe Estrada

Benemérita Universidad Autonoma de Puebla, Puebla, Mexico

A. Morelos Pineda

Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico

D. Krofcheck

University of Auckland, Auckland, New Zealand

S. Bheesette, P.H. Butler

University of Canterbury, Christchurch, New Zealand

A. Ahmad, M. Ahmad, M.I. Asghar, Q. Hassan, H.R. Hoorani, A. Saddique, M.A. Shah, M. Shoaib, M. Waqas

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan


National Centre for Nuclear Research, Swierk, Poland

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland


Laboratório de Instrumentação e Física Experimental de Partículas, Lisboa, Portugal


Joint Institute for Nuclear Research, Dubna, Russia

V. Golovtsov, Y. Ivanov, V. Kim, E. Kuznetsova, P. Levchenko, V. Murzin, V. Oreshkin, I. Smirnov, D. Sosnov, V. Sulimov, L. Uvarov, S. Vavilov, A. Vorobyev

Petersburg Nuclear Physics Institute, Gatchina (St. Petersburg), Russia


Institute for Nuclear Research, Moscow, Russia

V. Epshteyn, V. Gavrilov, N. Lychkovskaya, V. Popov, I. Pozdnyakov, G. Safronov, A. Spiridonov, A. Stepennov, V. Stolin, M. Toms, E. Vlasov, A. Zhokin

Institute for Theoretical and Experimental Physics, Moscow, Russia

T. Aushev

Moscow Institute of Physics and Technology, Moscow, Russia

M. Chadeeva, P. Parygin, D. Philippov, S. Polikarpov, E. Popova, V. Rusinov

National Research Nuclear University ‘Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia

V. Andreev, M. Azarkin, I. Dremin, M. Kirakosyan, S.V. Rusakov, A. Terkulov

PN. Lebedev Physical Institute, Moscow, Russia

A. Baskakov, A. Belyaev, E. Boos, M. Dubinin, L. Dudko, A. Ershov, A. Gribushin, V. Klyukhin, O. Kodolova, I. Lokhutin, I. Miagkov, S. Obraztsov, S. Petrushanko, V. Savrin, A. Snigirev

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia

V. Blinov, T. Dimova, L. Kardapoltsev, D. Shtol, Y. Skovpen

Novosibirsk State University (NSU), Novosibirsk, Russia


State Research Center of Russian Federation, Institute for High Energy Physics of NRC “Kurchatov Institute”, Protvino, Russia

A. Babaev, S. Baidali, V. Okhotnikov

National Research Tomsk Polytechnic University, Tomsk, Russia

P. Adzic, P. Cirkovic, D. Devetak, M. Dordevic, J. Milosevic

University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

C. Albajar, J.F. de Trocóniz

Universidad Autónoma de Madrid, Madrid, Spain


Universidad de Oviedo, Oviedo, Spain


Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain


CERN, European Organization for Nuclear Research, Geneva, Switzerland


Paul Scherrer Institut, Villigen, Switzerland


ETH Zurich - Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland


Universität Zürich, Zurich, Switzerland


National Central University, Chung-Li, Taiwan
R. Bartek, A. Dominguez  
Catholic University of America, Washington DC, USA

A. Buccilli, S.I. Cooper, C. Henderson, P. Rumerio, C. West  
The University of Alabama, Tuscaloosa, USA

D. Arcaro, T. Bose, D. Gastler, D. Rankin, C. Richardson, J. Rohlf, L. Sulak, D. Zou  
Boston University, Boston, USA

Brown University, Providence, USA

University of California, Davis, Davis, USA

University of California, Los Angeles, USA

University of California, Riverside, Riverside, USA

University of California, San Diego, La Jolla, USA

University of California, Santa Barbara - Department of Physics, Santa Barbara, USA

California Institute of Technology, Pasadena, USA

Carnegie Mellon University, Pittsburgh, USA

J.P. Cumalat, W.T. Ford, F. Jensen, A. Johnson, M. Krohn, S. Leontsinis, E. MacDonald, T. Mulholland, K. Stenson, K.A. Ulmer, S.R. Wagner  
University of Colorado Boulder, Boulder, USA

Cornell University, Ithaca, USA

Purdue University, West Lafayette, USA

T. Cheng, J. Dolen, N. Parashar

Purdue University Northwest, Hammond, USA


Rice University, Houston, USA


University of Rochester, Rochester, USA


Texas A&M University, College Station, USA


University of Tennessee, Knoxville, USA


Texas Tech University, Lubbock, USA


Vanderbilt University, Nashville, USA

M.W. Arenton, P. Barria, B. Cox, R. Hirosky, M. Joyce, A. Ledovskoy, H. Li, C. Neu, T. Sinthuprasith, Y. Wang, E. Wolfe, F. Xia

University of Virginia, Charlottesville, USA


Wayne State University, Detroit, USA


University of Wisconsin - Madison, Madison, WI, USA
69 Also at Sinop University, Sinop, Turkey.
70 Also at Mimar Sinan University, Istanbul, Istanbul, Turkey.
71 Also at Texas A&M University at Qatar, Doha, Qatar.
72 Also at Kyungpook National University, Daegu, Korea.