CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2020/10/25

Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in pp collisions at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

Abstract

A search for nonresonant production of Higgs boson pairs via gluon-gluon fusion and vector boson fusion in final states with two bottom quarks and two photons is presented. This search uses data from proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ recorded by the CMS detector at the LHC from 2016 to 2018, corresponding to an integrated luminosity of $137\,\text{fb}^{-1}$. No signal is observed, and a 95% confidence level upper limit is set on the product of the inclusive Higgs boson pair production cross section and branching fraction into $\gamma\gamma\text{b}\bar{\text{b}}$. The observed (expected) upper limit is determined to be 0.67 (0.45) fb, which corresponds to 7.7 (5.2) times the standard model prediction. Assuming all other Higgs boson couplings are equal to their values in the standard model, the coupling modifiers of the trilinear self-coupling κ_λ and the coupling between a pair of Higgs bosons and a pair of vector bosons c_{2V} are constrained within the ranges $-3.3 < \kappa_\lambda < 8.5$ and $-1.3 < c_{2V} < 3.5$ at 95% confidence level. Constraints on κ_λ are also set by combining this analysis with a search for single Higgs bosons produced in association with top quark-antiquark pairs, and by performing a simultaneous fit of κ_λ and the top Yukawa coupling modifier κ_t.

1 Introduction

Following the discovery of the Higgs boson (H) by the ATLAS and CMS Collaborations [1–3], there has been significant interest in thoroughly understanding the Brout–Englert–Higgs (BEH) mechanism [4, 5]. With the only free parameter, the mass of the Higgs boson, now measured to be around 125 GeV, the Higgs boson self-coupling and the structure of the scalar Higgs field potential are completely determined in the standard model (SM). Therefore, measuring the Higgs boson's trilinear self-coupling λ_{HHH} is of particular importance because it provides valuable information for reconstructing the shape of the scalar potential and to verify that the BEH mechanism is responsible for electroweak symmetry breaking.

The trilinear self-coupling of the Higgs boson is only directly accessible via Higgs boson pair production (HH). Higgs boson pair production is a rare process that mainly occurs via gluon-gluon fusion (ggF) at the LHC. Vector boson fusion (VBF) is the second most important production mechanism. In the SM, the ggF production cross section in proton-proton (pp) collisions at $\sqrt{s} = 13$ TeV is calculated at next-to-next-to-leading order (NNLO) as $31.05^{+1.41}_{-1.99}$ fb [6]. For VBF, the production cross section is calculated to be 1.723 ± 0.036 fb [7] at next-to-next-to-next-leading order (N^{3}LO). These cross sections are calculated for Higgs bosons with a mass m_{H} = 125.09 GeV.

Contributions from physics beyond the standard model (BSM) can significantly enhance the HH production cross section as well as change the kinematical properties of the produced Higgs boson pair, and consequently that of the decay products. The modification of the properties of nonresonant HH production via ggF from BSM effects can be parametrized through an effective Lagrangian that extends the SM Lagrangian with dimension-6 operators [8]. This parametrization results in five couplings: λ_{HHH}, the coupling between the Higgs boson and the top quark (y_{t}), and three additional couplings not present in the SM. Those three couplings represent contact interactions between two Higgs bosons and two gluons ($c_{2\mathrm{g}}$), between one Higgs boson and two gluons (c_{g}), and between two Higgs bosons and two top quarks (c_2). All five of these couplings are investigated in this analysis.

The VBF HH production mode gives access to λ_{HHH} as well as to the coupling between two vector bosons and the Higgs boson (HVV) and the coupling between a pair of Higgs bosons and a pair of vector bosons ($HHVV$). While λ_{HHH} is mainly constrained from measurements of HH production via ggF, and the HVV coupling modifier (c_V) is constrained by measurements of vector boson-associated production of a single Higgs boson and the decay of the Higgs boson to a pair of bosons, the $HHVV$ coupling modifier (c_{2V}) is only directly measurable via VBF HH production. Anomalous values of c_{2V} are investigated to establish the presence of the $HHVV$ mediated process as a probe of BSM physics.

Previous searches for nonresonant production of a Higgs boson pair via ggF were performed by both the ATLAS and CMS Collaborations using the LHC data collected at $\sqrt{s} = 8$ and 13 TeV [9–19]. Statistical combinations of search results in various decay channels were also performed by the two experiments [13, 20]. The combination of searches for HH production performed by the ATLAS Collaboration using up to 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV [13] results in the most stringent upper limit at 95% confidence level (CL) on the HH production cross section to date: 231 fb (335 fb expected) which corresponds to 7.5 (11) times the SM expectation. The first search for HH production via VBF was recently carried out by the ATLAS Collaboration in the $\mathrm{b\bar{b}b\bar{b}}$ channel [21].

This document describes a search for the nonresonant production of pairs of Higgs bosons decaying to $\gamma\gamma\mathrm{b\bar{b}}$ using a data sample of 137 fb^{-1} collected by the CMS experiment in 2016, 2017, and 2018. The $\gamma\gamma b\bar{b}$ final state has a combined branching fraction of 26.33×10^{-2} % [8] for a Higgs boson mass of 125 GeV. This channel is one of the most sensitive to HH production because of the large SM branching fraction of Higgs boson decays to bottom quarks, the good mass resolution of the $H \rightarrow \gamma\gamma$ channel, and relatively low background rates.

The analysis targets the main HH production modes: ggF and VBF. Both modes are pursued following similar strategies. After reducing the nonresonant $\gamma\gamma b\bar{b}$ background and the background coming from single Higgs boson production in association with a $t\bar{t}$ pair ($t\bar{t}H$) , the events are categorized into ggF and VBF enriched signal regions using a multivariate technique. The signal is extracted from a simultaneous fit to the invariant mass of the Higgs bosons in the $b\bar{b}$ and $\gamma\gamma$ final states. The analysis described in this document improves on the previous $pp \rightarrow HH \rightarrow \gamma\gamma b\bar{b}$ search [19] by improving the b-jet energy resolution with a dedicated energy regression, introducing new multivariate methods for background rejection, optimizing the event categorization, and adding dedicated VBF categories.

Finally, the search for Higgs boson pair production is combined with an orthogonal analysis that targets $t\bar{t}H$ production, where the Higgs boson decays to a diphoton pair [22]. The $t\bar{t}H$ production cross section depends on y_t, and also includes a Higgs trilinear coupling contribution from NLO electroweak corrections [23]. The combination enables the two couplings, λ_{HHH} and y_t, to be measured simultaneously and provides constraints applicable to a wider range of theoretical models, where both couplings have anomalous values.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

Events of interest are selected using a two-tiered trigger system [24]. The first level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a time interval of less than 4 μs. The second level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage.

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [25].

3 Higgs boson pair production in effective field theory

Nonresonant ggF HH production at the LHC can be described using an effective field theory (EFT) approach [8]. Considering operators up to dimension 6, the tree-level interactions of the Higgs boson are modeled by five parameters. The Feynman diagrams contributing to ggF HH production at leading order (LO) are shown in Fig. 1. Deviations from the SM values of λ_{HHH} and y_t are parametrized as $\kappa_\lambda \equiv \lambda_{HHH}/\lambda_{HHH}^{SM}$ and $\kappa_t \equiv y_t/y_t^{SM}$, where the SM values of the couplings are defined as $\lambda_{HHH}^{SM} \equiv m_H^2/(2v^2) = 0.129$, $y_t^{SM} = \sqrt{2}\, m_t/v \approx 1.0$. Here, $v = 246$ GeV is the vacuum expectation value of the Higgs field, and $m_t \approx 173\,\text{GeV}$ is the top quark mass. The anomalous couplings c_{2g}, c_2, and c_g are not present in the SM. The corresponding part of the Lagrangian can be written as [26]:

$$\mathcal{L}_{HH} = \kappa_\lambda \lambda^{SM}_{HHH} v H^3 - \frac{m_t}{v}(\kappa_t H + \frac{c_2}{v} H^2)(\bar{t}_L t_R + \text{h.c.}) + \frac{1}{4}\frac{\alpha_S}{3\pi v}(c_g H - \frac{c_{2g}}{2v} H^2) G^{\mu\nu} G_{\mu\nu}, \quad (1)$$

where t_L and t_R are the top quark fields with left and right chiralities, respectively. The Higgs boson field is denoted as H, $G^{\mu\nu}$ is the gluon field strength tensor, α_S is the strong coupling constant, and *h.c.* denotes the Hermitian conjugate.

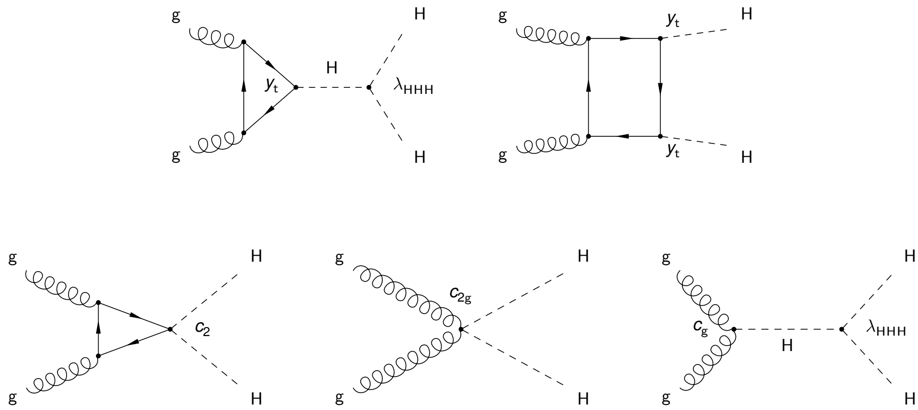


Figure 1: Feynman diagrams of the processes contributing to the production of Higgs boson pairs via ggF at LO. The top diagrams correspond to SM processes, involving the top Yukawa coupling y_t and the trilinear Higgs coupling λ_{HHH}, respectively. The bottom diagrams correspond to BSM processes: the diagram on the left involves the contact interaction of two Higgs bosons with two top quarks (c_2), the middle diagram shows the quartic coupling between the Higgs bosons and two gluons (c_{2g}), and the diagram on the right describes the contact interactions between the Higgs boson and gluons (c_g).

It has been observed in Ref. [27] that when scanning the phase space of the five parameters (κ_λ, κ_t, c_2, c_g, c_{2g}), the distributions of the main kinematic observables cluster in a small number of shapes. Twelve benchmark hypotheses have been defined to describe BSM scenarios with various combinations of EFT parameters. The parameter values for these benchmark hypotheses are summarized in Table 1. The simulated samples generated with the EFT parameters that describe the twelve benchmark hypotheses are combined to cover all possible kinematic configurations of the EFT parameter space. The specific kinematics of any point in the full 5D parameter space are obtained through a corresponding reweighting [27].

The NNLO ggF HH cross section as a function of the five BSM parameters is obtained from the LO cross section [8] by applying a global k-factor. The reweighting procedure described in Ref. [27] to obtain the distributions of the kinematic observables, however, cannot be applied for the higher order simulation because of the presence of additional partons at matrix level. Therefore, the 12 BSM signal benchmark hypotheses summarized in Table 1 are investigated using an LO Monte Carlo (MC) simulation, and only anomalous values of κ_λ and κ_t are studied with the next-to-leading order (NLO) simulation as described in Section 4.

Table 1: Coupling parameter values in the SM and in twelve BSM benchmark hypotheses identified using the method described in Ref. [27].

	1	2	3	4	5	6	7	8	9	10	11	12	SM
κ_λ	7.5	1.0	1.0	−3.5	1.0	2.4	5.0	15.0	1.0	10.0	2.4	15.0	1.0
κ_t	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0	1.0	1.5	1.0	1.0	1.0
c_2	−1.0	0.5	−1.5	−3.0	0.0	0.0	0.0	0.0	1.0	−1.0	0.0	1.0	0.0
c_g	0.0	−0.8	0.0	0.0	0.8	0.2	0.2	−1.0	−0.6	0.0	1.0	0.0	0.0
c_{2g}	0.0	0.6	−0.8	0.0	−1.0	−0.2	−0.2	1.0	0.6	0.0	−1.0	0.0	0.0

The diagrams shown in Fig. 2 contribute to the production of Higgs boson pairs via VBF at LO. In the SM, three different couplings are involved in HH production via VBF: λ_{HHH}, HVV, and $HHVV$. The Lagrangians corresponding to the left, middle, and right diagrams in Fig. 2 scale with $c_V\kappa_\lambda$, c_V^2, and c_{2V}, respectively, where c_{2V} and c_V are the $HHVV$ and HVV coupling modifiers, normalized to the SM values. A global k-factor is applied to scale the LO cross section to N^{3}LO accuracy [7].

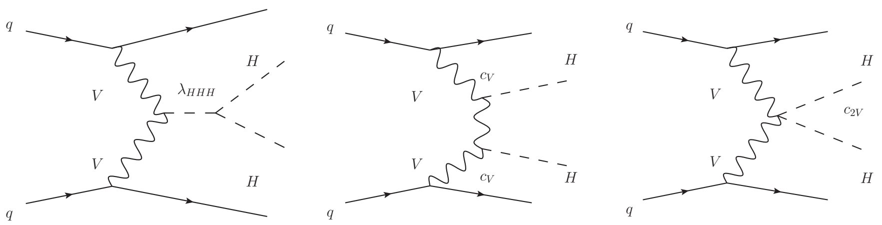


Figure 2: The Feynman diagrams that contribute to the production of Higgs boson pairs via VBF at LO. On the left the diagram involving the HHH vertex (λ_{HHH}), in the middle the diagram with two HVV vertices (c_V), and on the right the diagram with the $HHVV$ vertex (c_{2V}).

4 Data sample and simulated events

The analyzed data correspond to a total integrated luminosity of 137 fb^{-1} and were collected over a data-taking period spanning three years: 35.9 fb^{-1} in 2016, 41.5 fb^{-1} in 2017, and 59.4 fb^{-1} in 2018. Events are selected using double-photon triggers with asymmetric thresholds on the photon transverse momenta of $p_T^{\gamma 1} > 30\,\text{GeV}$ and $p_T^{\gamma 2} > 18(22)\,\text{GeV}$ for the data collected during 2016 (2017 and 2018). In addition, loose calorimetric identification requirements [28] are imposed on the photon candidates at the trigger level.

The ggF HH signal samples are simulated at NLO including the full top quark mass dependence [29] using POWHEG 2.0 [30, 31]. The samples were generated for different values of κ_λ. Samples corresponding to any point in the (κ_λ, κ_t) parameter space can be obtained from the linear combination of any three of the generated MC samples with different values of κ_λ.

In addition, LO signal samples are generated for the BSM benchmark hypotheses described in Section 3 using MADGRAPH5_aMC@NLO v2.2.2 (2016) / v2.4.2 (2017 and 2018) [32–34] interfaced with LHAPDF6 [35] and using the NLO parton distribution function (PDF) set PDF4LHC15_NLO_MC [36–40]. The simulated LO signal samples, corresponding to the 12 BSM benchmark hypotheses, are added together to increase the number of events, and then reweighted to any coupling configuration (κ_λ, κ_t, c_2, c_g, c_{2g}) using generator-level information on the HH system.

The VBF HH signal samples are generated at LO using MADGRAPH5_aMC@NLO v2.4.2 [32]. The simulated samples were generated for different combinations of the coupling modifier values (κ_λ, c_V, c_{2V}). Similarly to what is done for the ggF HH NLO samples, samples corresponding to any point in the (κ_λ, c_V, c_{2V}) parameter space can be obtained from the linear combination of any six of the generated samples.

The dominant backgrounds in this search are irreducible prompt diphoton production ($\gamma\gamma +$ jets) and the reducible background from $\gamma +$ jets events, where the jets are misidentified as isolated photons and b jets. Although these backgrounds are estimated using data-driven methods, simulated samples are used for the training of multivariate discriminants and the optimization of the analysis categories. The $\gamma\gamma +$ jets background is modeled with SHERPA v.2.2.1 [41]. It includes the Born processes with up to three additional jets as well as the box processes at LO. In addition, a b-enriched diphoton background is generated with SHERPA at LO requiring up to two b jets. The $\gamma +$ jets background is modeled with PYTHIA 8.212 [42] at LO.

Single Higgs boson production, where the Higgs boson decays to a pair of photons, is considered as a resonant background. The single H production processes are simulated using POWHEG 2.0 [30, 43–45] at NLO in quantum chromodynamics (QCD) for ggF and VBF, and MADGRAPH5_aMC@NLO v2.2.2 (2016) / v2.4.2 (2017 and 2018) for $\mathrm{t\bar{t}H}$, vector boson associated production (VH), and production associated with a single top quark. The cross sections and decay branching fractions are taken from Ref. [8].

All simulated samples are interfaced with PYTHIA 8.212 [42] for parton showering and fragmentation with the standard p_T-ordered parton shower (PS) scheme. The underlying event is modeled with PYTHIA, using the CUETP8M1 tune for 2016 and the CP5 tune for 2017 and 2018 [46, 47]. PDFs are taken from the NNPDF3.0 [40] (2016) / NNPDF3.1 [48] (2017 and 2018) set. The response of the CMS detector is modeled using the GEANT4 [49] package. The simulated events include additional pp interactions within the same or nearby bunch crossings (pileup) as observed in the data.

Additionally, the simulated VBF HH signal events were also interfaced with the PYTHIA dipole shower scheme to model initial state radiation (ISR) and final state radiation (FSR) [50]. The dipole shower correctly takes into account the structure of the color flow between incoming and outgoing quark lines, and its predictions are found to be in good agreement with NNLO QCD calculations, as reported in Ref. [51]. These simulated samples are used to derive uncertainties associated with the PYTHIA PS ISR and FSR parameters.

5 Event reconstruction and selection

The primary pp interaction vertex of the event is identified using a multivariate technique based on a boosted decision tree (BDT) [52]. The efficiency of the correct vertex assignment is greater than 99.9% thanks to the requirement of at least two jets in the $\gamma\gamma\mathrm{b\bar{b}}$ final state.

Photons are identified using a multivariate technique based on a BDT trained to separate photons from jets (photon ID) [28]. The photon ID BDT is trained using variables that describe the shape of the photon electromagnetic shower and its isolation in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ around the photon candidate direction, where ϕ is the azimuthal angle. The imperfect MC simulation modeling of these variables is corrected to match the data using a chained quantile regression (CQR) method [52] based on studies of $\mathrm{Z} \to \mathrm{ee}$ events. In the CQR method, a set of BDTs is trained to predict the cumulative distribution function (CDF) for a given input. Its prediction is conditional upon the three kinematic variables (p_T, $|\eta|$, ϕ) and the global event energy density [28]. The corrections are then applied to the simulated photons such that the predicted

CDF of the simulated variables is morphed onto the one observed in data. The photon candidates are reconstructed from energy clusters in the ECAL not linked to charged particle tracks (with the exception of converted photons). The photon energies measured by the ECAL are corrected with a multivariate regression technique based on simulation that accounts for radiation lost in material upstream of the ECAL and imperfect shower containment [28]. The ECAL energy scale in data is corrected using simulated $Z \to ee$ events, while the energy in simulated events is smeared to reproduce the resolution measured in data.

Events are required to have at least two identified photon candidates that are within the ECAL and tracker fiducial region ($|\eta| < 2.5$), excluding the ECAL barrel-endcap transition region ($1.44 < |\eta| < 1.57$). The photon candidates are required to pass the following criteria: $100 < m_{\gamma\gamma} < 180\,\text{GeV}$, $p_{T}^{\gamma 1}/m_{\gamma\gamma} > 1/3$ and $p_{T}^{\gamma 2}/m_{\gamma\gamma} > 1/4$, where p_{T}^{γ} and $m_{\gamma\gamma}$ are the transverse momenta and the invariant mass of the photon candidates. When more than two photon candidates are found, the photon pair with the highest transverse momentum $p_{T}^{\gamma\gamma}$ is chosen to construct the Higgs boson candidate.

The particle-flow (PF) algorithm reconstructs individual particles by combining information from the various subsystems of the CMS detector [53]. Jets are clustered from these candidates using the anti-k_{T} algorithm with a distance parameter $R_{j} = 0.4$ [54, 55]. Jet candidates are required to have $p_{T} > 25\,\text{GeV}$ and $|\eta| < 2.4$ (2.5) for 2016 (2017 and 2018) and to be separated from the identified photons by a distance of $\Delta R_{\gamma j} \equiv \sqrt{(\Delta\eta_{\gamma j})^2 + (\Delta\phi_{\gamma j})^2} > 0.4$. The jet pseudorapidity range is extended for the 2017 and 2018 data-taking years because of the addition of a pixel layer installed during the Phase-1 upgrade of the CMS pixel detector [56]. In addition, identification criteria are applied to remove spurious jets associated with calorimeter noise. Jets from the hadronization of b quarks are tagged by a secondary vertex algorithm, DeepJet, based on the score from a deep neural network (DNN) [57]. We will refer to the output of this DNN as the b tagging score.

In addition to standard CMS jet energy corrections [58], a b-jet energy regression [59] is used to improve the energy resolution of b jets and, therefore, the m_{jj} resolution. The energy correction is computed for each of the Higgs boson candidate jets through a regression implemented in a DNN and trained on jet properties. The regression simultaneously provides a b-jet energy correction and a resolution estimator.

An additional regression was developed specifically for the $\gamma\gamma b\overline{b}$ final states to further improve the dijet invariant mass resolution. This regression exploits the fact that there is no genuine missing energy from the hard-scattering process in the $\gamma\gamma b\overline{b}$ final state, and follows a similar approach as used in Ref. [19]. The regression targets the dijet invariant mass at generator level, and is trained using the kinematic properties of the event and the missing transverse momentum. The missing transverse momentum vector $\vec{p}_{T}^{\text{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as p_{T}^{miss} [60]. The $\vec{p}_{T}^{\text{miss}}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event. The regression is trained on a simulated sample of b-enriched $\gamma\gamma + \text{jets}$ events.

Both regression techniques were validated on data collected by the CMS detector. The two-step regression technique improves the dijet invariant mass resolution of the SM HH signal by about 20%, and the m_{jj} peak position is shifted by 5.5 GeV (5%) closer to the expected Higgs boson mass.

In events with more than two jets, the Higgs boson candidate is reconstructed from the dijet pair constructed from the two jets with the highest b tagging scores. The dijet invariant mass is required to be $70 < m_{\mathrm{jj}} < 190\,\mathrm{GeV}$.

To select events corresponding to HH production via VBF, additional requirements are imposed. The VBF process is characterized by the presence of two additional energetic jets, corresponding to two quarks from each of the colliding protons scattered away from the beam line. These "VBF-tagged" jets are expected to be in the forward and backward directions relative to the beam direction and have a large pseudorapidity separation, $|\Delta\eta_{\mathrm{jj}}^{VBF}|$, and large dijet invariant mass, m_{jj}^{VBF}. VBF-tagged jets are required to have $p_{\mathrm{T}} > 40(30)\,\mathrm{GeV}$ for the leading (subleading) jet, $|\eta| < 4.7$, and be separated from the selected photon and b jet candidates by $\Delta R\,(\text{jet}, \gamma) > 0.4$, $\Delta R\,(\text{jet}, \text{b jet}) > 0.4$. Jets must also pass an identification criterion designed to reduce the number of selected jets originating from pileup [61]. The dijet pair with the highest dijet invariant mass m_{jj}^{VBF} is selected as the two VBF-tagged jets. We will refer to these requirements as "VBF selection criteria".

6 Analysis strategy

To improve the sensitivity of the search, multivariate analysis (MVA) techniques are used to distinguish the ggF and VBF HH signal from the background. The output of the MVA classifiers is then used to define the analysis categories. The HH signal is extracted from a fit to the invariant mass of the two Higgs bosons in the final state simultaneously in all categories.

We study the properties of the HH system, built from the reconstructed diphoton and dijet candidates, to identify variables that can help us distinguish between the signal and background. The invariant mass distributions are shown in Fig. 3 for diphoton ($m_{\gamma\gamma}$) and dijet (m_{jj}) pairs for data and for signal and background simulation after requiring the selection criteria described in Section 5. The signal has a peaking distribution in $m_{\gamma\gamma}$ and m_{jj}. The data distribution, dominated by the $\gamma\gamma +$ jets and $\gamma +$ jets backgrounds, exhibits a falling spectrum due to the nonresonant nature of these processes. In this analysis, these characteristics are used to extract the signal through a simultaneous fit to $m_{\gamma\gamma}$ and m_{jj}.

The distribution of $\widetilde{M}_{\mathrm{X}}$, defined as:

$$\widetilde{M}_{\mathrm{X}} = m_{\gamma\gamma\mathrm{jj}} - (m_{\mathrm{jj}} - m_{\mathrm{H}}) - (m_{\gamma\gamma} - m_{\mathrm{H}}), \tag{2}$$

is particularly sensitive to different values of the couplings described in Section 3. The $\widetilde{M}_{\mathrm{X}}$ distribution is less dependent on the dijet and diphoton energy resolution than $m_{\gamma\gamma\mathrm{jj}}$ if the dijet and diphoton pairs originate from a Higgs boson decay [62]. In Fig. 4 the distribution of $\widetilde{M}_{\mathrm{X}}$ is shown for several BSM benchmark hypotheses affecting ggF HH production (see Table 1) and for different values of c_{2V} affecting the VBF HH production mode. The SM HH process exhibits a broad structure in $\widetilde{M}_{\mathrm{X}}$, induced by the interference between different processes contributing to HH production and shaped by the analysis selection. The signal with $c_{2V} = 0$ and $c_{2V} = 2$ has a much harder spectrum than the SM VBF HH signal.

7 $\mathrm{t\bar{t}H}$ background rejection

Single Higgs boson production is an important resonant background in the $\gamma\gamma\mathrm{b\bar{b}}$ final state, with $\mathrm{t\bar{t}H}$ production being the most dominant. To reduce $\mathrm{t\bar{t}H}$ background contamination, a dedicated classifier (ttHScore) was developed. The classifier is trained on a mixture of SM HH events and events generated for the twelve BSM benchmark hypotheses (see Table 1) as signal,

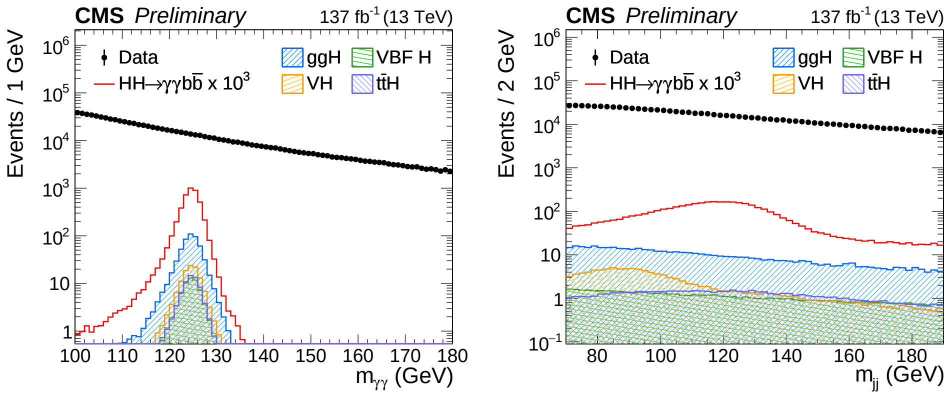


Figure 3: The distribution of $m_{\gamma\gamma}$ (left) and m_{jj} (right) in data and simulated events. Data, dominated by the $\gamma\gamma$ + jets and γ + jets backgrounds, are compared to the SM ggF HH signal samples and single H samples (tt̄H, ggH, VBF H , VH) after requiring the selection criteria described in Section 5. The error bars on the data points indicate statistical uncertainties. The HH signal has been scaled by a factor of 10^3 for display purposes.

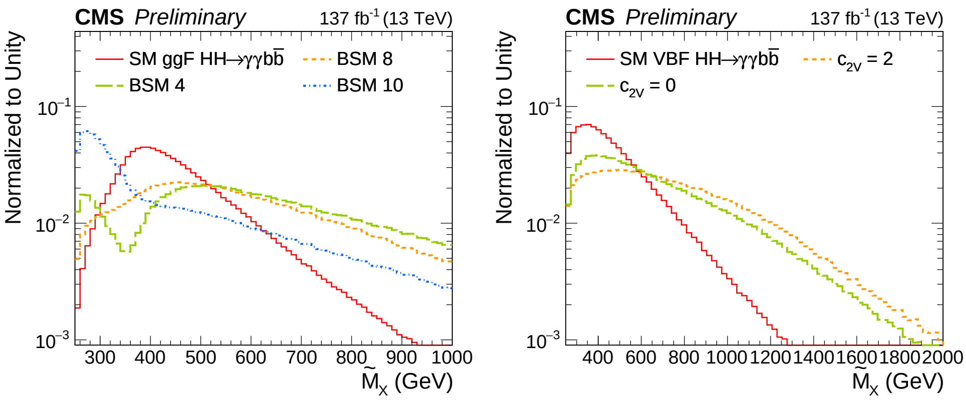


Figure 4: Distributions of $\widetilde{M}_{\mathrm{X}}$. The SM ggF HH signal is compared with several BSM hypotheses listed in Table 1 (left), and the SM VBF HH signal is compared with different anomalous values of c_{2V} (right). All distributions are normalized to unity.

and tt̄H events as background. The discriminant uses a combination of low-level information from the individual PF candidates and high-level features describing kinematical properties of the event. The kinematic variables used in the training can be classified in three groups: angular variables, variables to distinguish semileptonic decays of W bosons produced in the top quark decay, and variables to distinguish hadronic decays of W bosons. The ttHScore discriminant is implemented with a DNN combining feed-forward and recurrent layers, based on the topology-classifier architecture introduced in Ref. [63]. The network is implemented in Keras [64] using the TensorFlow [65] backend, and the hyperparameters are optimized with Bayesian optimization. The ttHScore output is shown in Fig. 5 (left) for data and simulated events. The events entering the analysis are required to pass a selection based on this classifier, which is optimized as described in Section 9.

8 Nonresonant background rejection

8.1 Background reduction in ggF HH signal region

An MVA discriminant implemented with a BDT is used to separate the ggF HH signal and the dominant nonresonant $\gamma\gamma$ + jets and γ + jets backgrounds. We select several discriminating observables for use in the training. They can be classified in three groups: kinematic variables, object identification variables, and object resolution variables. The first group exploits the kinematic properties of the HH system, the second allows us to separate the signal from the reducible backgrounds, and the third takes into account the resonant nature of the $\gamma\gamma$ and bb final states. The following discriminating variables were chosen:

- H candidate kinematic variables: $p_{\mathrm{T}}^{\gamma}/m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\mathrm{j}}/m_{\mathrm{jj}}$ for leading and subleading photons and jets, where p_{T}^{γ} and $p_{\mathrm{T}}^{\mathrm{j}}$ are the transverse momenta of the selected photon and jet candidates.
- HH transverse balance: $p_{\mathrm{T}}^{\gamma\gamma}/m_{\gamma\gamma\mathrm{jj}}$ and $p_{\mathrm{T}}^{\mathrm{jj}}/m_{\gamma\gamma\mathrm{jj}}$, where $p_{\mathrm{T}}^{\gamma\gamma}$ and $p_{\mathrm{T}}^{\mathrm{jj}}$ are the transverse momenta of the diphoton and dijet candidates.
- Helicity angles: $|\cos\theta_{\mathrm{HH}}^{\mathrm{CS}}|$, $|\cos\theta_{\mathrm{jj}}|$, $|\cos\theta_{\gamma\gamma}|$, where $|\cos\theta_{\mathrm{HH}}^{\mathrm{CS}}|$is the angle between the direction of the $\mathrm{H} \rightarrow \gamma\gamma$ candidate and the Collins–Soper reference frame [66], while $|\cos\theta_{\mathrm{jj}}|$ and $|\cos\theta_{\gamma\gamma}|$are the angles between one of the Higgs boson decay products and the direction defined by the Higgs boson candidate.
- Angular distance: minimum $\Delta R_{\gamma\mathrm{j}}$ between a photon and a jet, $\Delta R_{\gamma\mathrm{j}}^{min}$, considering all combinations between objects passing the selection criteria, $\Delta R_{\gamma\mathrm{j}}$ between the other photon-jet pair not used in the $\Delta R_{\gamma\mathrm{j}}^{min}$ calculation.
- b tagging: the b tagging score of each jet in the dijet candidate.
- photon ID: photon identification variables for leading and subleading photons.
- Object resolution: energy resolution for the leading and subleading photons and jets, the mass resolution estimators for the diphoton and dijet candidates.

The BDT is trained using the XGBOOST [67] software package using a gradient boosting algorithm. The $\gamma\gamma$ + jets and γ + jets MC samples are used as background, while an ensemble of SM HH and the 12 BSM HH benchmark hypotheses listed in Table 1 is used as signal. Training on an ensemble of BSM and SM HH signals makes the BDT sensitive to a broad spectrum of theoretical scenarios. During the training, signal events are weighted with the product of the inverse mass resolution of the diphoton and dijet systems. These resolutions are obtained using the per-object resolution estimators provided by the energy regressions developed for photons and b jets. In the training, the mass dependence of the classifier is removed by using only dimensionless kinematic variables. The inverse resolution weighting at training time improves the performance by bringing back the information about the resonant nature of the signal. Independent training and testing samples are created by splitting the signal and background samples. The classifier hyperparameters are optimized using a randomized grid search and a 5-fold cross-validation technique [68]. The BDT is trained separately for the 2016, 2017, and 2018 data-taking years, and the median energy density of the event [28] is used as an input feature to include information about different pileup conditions as observed in the three data-taking years. The distributions of the BDT output for signal and background are very similar among the different data-taking years; they are therefore merged by combining events with the same relative significance. The MVA output is transformed using a cumulative distribution of the SM ggF HH signal. This transformation is applied to all events, both in simulation and data. The distribution of the MVA output for data and simulated events is shown in Fig. 5

(right).

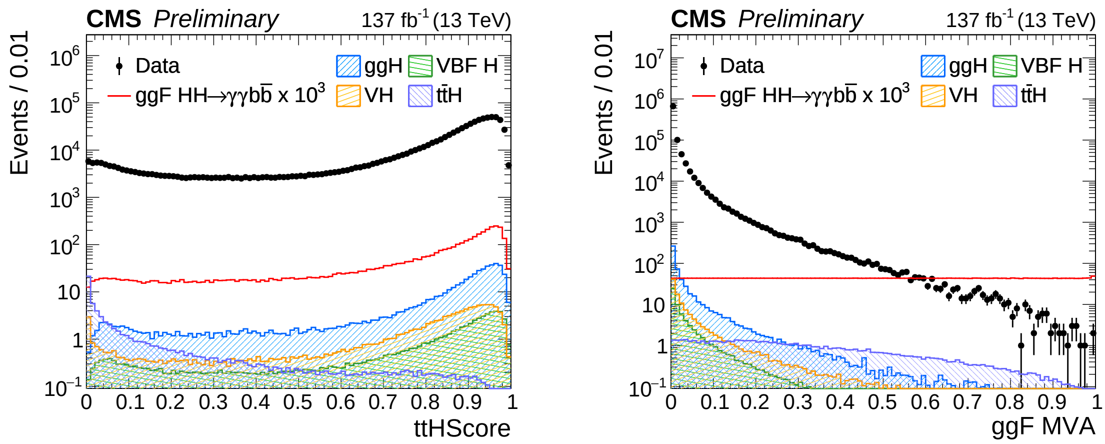


Figure 5: The distribution of the ttHScore (left) and MVA output (right) in data and simulated events. Data, dominated by $\gamma\gamma +$ jets and $\gamma +$ jets background, are compared to the SM ggF HH signal samples and single H samples ($t\bar{t}$H, ggH, VBF H , VH) after requiring the selection criteria described in Section 5. The error bars on the data points indicate statistical uncertainties. The HH signal has been scaled by a factor of 10^3 for display purposes.

8.2 Background reduction in VBF HH signal region

Similarly to the ggF HH analysis strategy, an MVA discriminant is employed to separate the VBF HH signal from the background. As for the ggF case, the $\gamma\gamma +$ jets and $\gamma +$ jets processes are the dominant sources of background. For the VBF production mode, the ggF HH events are considered as background. About a third of the ggF HH events passing the selection requirements described in Section 5 also pass the dedicated VBF selection criteria. The distinctive topology of the VBF HH process is used to separate the VBF HH signal from the various sources of background. In addition to the discriminating features of the HH signal described in Sections 6 and 8.1, the following set of VBF-discriminating features were identified:

- VBF-tagged jet kinematics: $p_{\mathrm{T}}^{VBF}/m_{\mathrm{jj}}^{VBF}$, η^{VBF} for VBF-tagged jets.
- VBF-tagged jet invariant mass: invariant mass m_{jj}^{VBF} of the VBF-tagged jets.
- Rapidity gap: product and difference of pseudorapidity between the VBF-tagged jets.
- Quark-gluon likelihood [69, 70] of the two VBF-tagged jets. A likelihood discriminator used to distinguish between jets originating from quarks and from gluons.
- HH system kinematics: $\widetilde{M}_{\mathrm{X}}$ and the transverse momentum of the pair of the reconstructed Higgs bosons.
- Angular distance: minimum ΔR between a photon and a VBF-tagged jet, and between a b jet and a VBF-tagged jet.
- Centrality variables for the reconstructed Higgs boson candidates:

$$C_{xx} = exp\left(-\frac{4}{(\eta_1^{VBF} - \eta_2^{VBF})^2}\left(\eta^{xx} - \frac{\eta_1^{VBF} + \eta_2^{VBF}}{2}\right)^2\right), \tag{3}$$

  where xx is the Higgs boson candidate reconstructed either from diphoton or dijet pairs, and η_1^{VBF} and η_2^{VBF} are the pseudorapidities of the two VBF-tagged jets.

We split events into two regions: $\widetilde{M}_{\mathrm{X}} < 500\,\mathrm{GeV}$ and $\widetilde{M}_{\mathrm{X}} > 500\,\mathrm{GeV}$. While the region of $\widetilde{M}_{\mathrm{X}} > 500\,\mathrm{GeV}$ is sensitive to anomalous values of c_{2V}, the $\widetilde{M}_{\mathrm{X}} < 500\,\mathrm{GeV}$ region retains the sensitivity to SM VBF HH production.

A multi-class BDT, using a gradient boosting algorithm and implemented in the XGBOOST [67] framework, is trained to separate the VBF HH signal from the $\gamma\gamma$ + jets, γ + jets, and SM ggF HH background. A mix of VBF HH samples with SM couplings and quartic coupling $c_{2V} = 0$ is used as signal. Training on the mix of samples makes the BDT sensitive to both SM and BSM scenarios. Although the kinematic properties of the BSM signals with anomalous values of c_{2V} are similar, the choice $c_{2V} = 0$ was motivated by the goal of setting a stringent limit on anomalous values of $c_{2V} \simeq 0$. Signal events are weighted with the inverse of the mass resolution of the diphoton and dijet systems during the training, as it is done for the ggF MVA. The BDT training is performed separately in the two $\widetilde{M}_{\mathrm{X}}$ regions. Data from the 2016, 2017, and 2018 data-taking years are merged by combining events with the same relative significance, as it is done for the ggF MVA. A cumulative transformation of the mix of VBF HH signals with SM couplings and quartic coupling $c_{2V} = 0$ is applied to all events in the two $\widetilde{M}_{\mathrm{X}}$ regions. The distribution of the MVA outputs for data and simulated events is shown in Fig. 6.

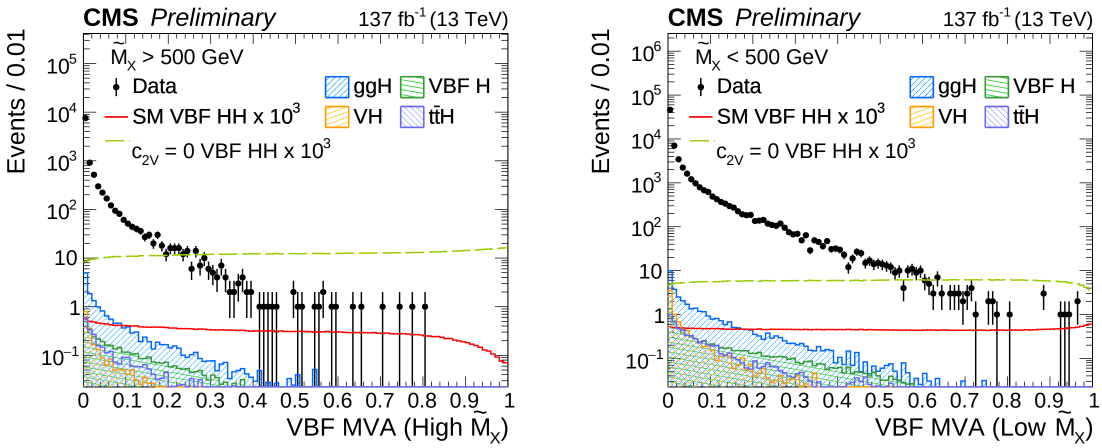


Figure 6: The distribution of the two MVA outputs is shown in data and simulated events in the two VBF $\widetilde{M}_{\mathrm{X}}$ regions: $\widetilde{M}_{\mathrm{X}} > 500\,\mathrm{GeV}$ (left) and $\widetilde{M}_{\mathrm{X}} < 500\,\mathrm{GeV}$ (right). Data, dominated by the $\gamma\gamma$ + jets and γ + jets backgrounds, are compared to the VBF HH signal samples with SM couplings and $c_{2V} = 0$, and single H samples ($\mathrm{t\bar{t}H}$, ggH, VBF H , VH) after requiring the VBF selection criteria described in Section 5. The error bars on the data points indicate statistical uncertainties. The HH signal has been scaled by a factor of 10^3 for display purposes.

9 Event categorization

In order to maximize the sensitivity of the search, events are split into different categories according to the output of the MVA classifier and the mass of the Higgs boson pair system $\widetilde{M}_{\mathrm{X}}$. The $\widetilde{M}_{\mathrm{X}}$ distribution changes significantly for different BSM hypotheses as shown in Fig. 4. Therefore a categorization of HH events in $\widetilde{M}_{\mathrm{X}}$ creates signal regions sensitive to multiple theoretical scenarios. In the search for VBF HH production, the categories in $\widetilde{M}_{\mathrm{X}}$ are defined before the MVA is trained, as described in Section 8.2. For the categories that target ggF HH production, categories in $\widetilde{M}_{\mathrm{X}}$ are defined after the MVA is trained.

The categorization is optimized by maximizing the expected significance estimated as the sum in quadrature of $S/\sqrt{B}$ over all categories in a window centred on m_{H}: $115 < m_{\gamma\gamma} < 135\,\mathrm{GeV}$.

Table 2: Summary of the analysis categories. Two VBF and 12 ggF-enriched categories are defined based on the output of the MVA classifier and the mass of the Higgs boson pair system $\widetilde{M}_X$. The VBF and ggF categories are mutually exclusive.

Category	MVA	$\widetilde{M}_X$ (GeV)
VBF CAT 0	0.52-1.00	>500
VBF CAT 1	0.86-1.00	250-500
ggF CAT 0	0.78-1.00	>600
ggF CAT 1		510-600
ggF CAT 2		385-510
ggF CAT 3		250-385
ggF CAT 4	0.62-0.78	>540
ggF CAT 5		360-540
ggF CAT 6		330-360
ggF CAT 7		250-315
ggF CAT 8	0.37-0.62	>585
ggF CAT 9		375-585
ggF CAT 10		330-375
ggF CAT 11		250-330

Here, S and B are the numbers of expected signal and background events, respectively. Simulated events are used for this optimization. The SM HH process is considered as signal, while the background consists of the $\gamma\gamma+\text{jets}$, $\gamma+\text{jets}$, and $\mathrm{t\bar{t}H}$ processes. The MVA categories are optimized simultaneously with a threshold on the value of ttHScore. Two VBF and three ggF categories are optimized based on the MVA output. For ggF HH in each MVA category a set of $\widetilde{M}_X$ categories is then optimized. The optimization procedure leads to 12 ggF analysis categories: four categories in $\widetilde{M}_X$ in each of the three categories in the MVA score. The optimized selection on ttHScore > 0.26 corresponds to 80% (85%) $\mathrm{t\bar{t}H}$ background rejection at 95% (90%) signal efficiency for the 12 ggF (2 VBF) categories. The categorization is summarized in Table 2. The VBF and ggF categories are mutually exclusive, as we only consider events that do not enter the VBF categories for the ggF categories. Events with VBF MVA scores below 0.52 (0.86) for $\widetilde{M}_X > 500$ ($\widetilde{M}_X < 500$) GeV are not considered in the VBF signal region. Because of overwhelming background contamination such events do not improve the expected sensitivity of the analysis. Similarly, events with ggF MVA scores below 0.37 are not considered in the ggF signal region.

9.1 Combination of the HH and $\mathrm{t\bar{t}H}$ signals to constrain κ_λ and κ_t

As discussed in Section 3, the HH production cross section depends on κ_λ and κ_t. The production cross section of the single H processes also depends on κ_λ as a result of NLO electroweak corrections [23]. The ggH and $\mathrm{t\bar{t}H}$ production cross sections additionally depend on κ_t. Therefore, the $\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}}$ signal can be combined with the single H production mode to provide an improved constraint on the κ_λ and κ_t parameters. In the case of anomalous values of κ_λ, the single H process with the largest modification of the cross section is $\mathrm{t\bar{t}H}$. For this reason, additional orthogonal categories targeting the $\mathrm{t\bar{t}H}$ process are included in the analysis: the "$\mathrm{t\bar{t}H}$ leptonic" and the "$\mathrm{t\bar{t}H}$ hadronic" categories, developed and optimized for the measurement of the $\mathrm{t\bar{t}H}$ production cross section in the diphoton decay channel [22]. The events that do not pass the selections for the HH categories defined in Table 2 are tested for the $\mathrm{t\bar{t}H}$ categories. This ensures the orthogonality between the events selected by the HH and $\mathrm{t\bar{t}H}$ categories.

The $\mathrm{H} \to \gamma\gamma$ candidate selection is the same as described in Section 5. The $\mathrm{t\bar{t}H}$ leptonic categories target $\mathrm{t\bar{t}H}$ events where at least one W boson, originating from the top or anti-top quark, decays leptonically. At least one isolated electron (muon) with $|\eta| < 2.4$ and $p_{\mathrm{T}} > 10(5)$ GeV, and at least one jet with $p_{\mathrm{T}} > 25$ GeV are required. The $\mathrm{t\bar{t}H}$ hadronic categories target hadronic decays of W bosons. In these categories at least three jets are required, one of which must be b-tagged, and a lepton veto is imposed. In order to maximize the sensitivity, an MVA approach is used to separate the $\mathrm{t\bar{t}H}$ events from the background, dominated by $\gamma\gamma +$ jets, $\gamma +$ jets, $\mathrm{t\bar{t}}$ + jets, $\mathrm{t\bar{t}}$ + γ, and $\mathrm{t\bar{t}}$ + $\gamma\gamma$. A BDT classifier is trained for each of the two channels using simulated events. The variables used for the training include kinematic properties of the reconstructed objects, object identification variables, and global event properties such as jet and lepton multiplicities. The BDT input variables also include the outputs of other machine learning algorithms trained specifically to target different backgrounds. The output scores of the BDTs are used to reject background-like events and to classify the remaining events in four subcategories for each of the two channels. The boundaries of the categories are optimized by maximizing the expected significance of the $\mathrm{t\bar{t}H}$ signal.

10 Signal model

In each of the HH categories, a parametric fit in the $(m_{\gamma\gamma}, m_{\mathrm{jj}})$ plane is performed. In the $\mathrm{t\bar{t}H}$ categories the $m_{\gamma\gamma}$ distribution is fitted to extract the signal. When the HH and $\mathrm{t\bar{t}H}$ categories are combined, both the HH and $\mathrm{t\bar{t}H}$ production modes are considered as signals.

The shape templates of the diphoton and dijet invariant mass distributions are constructed from simulation. In each HH and $\mathrm{t\bar{t}H}$ analysis category, the $m_{\gamma\gamma}$ distribution is fitted using a sum of at most five Gaussian functions. The number of Gaussian functions is determined by requiring a good fit to the simulated distribution, while avoiding overfitting statistical fluctuations due to the limited size of the simulated samples. The fit function for each category is normalized to the expected signal yield in that category. Figure 7 (left) shows the signal model for $m_{\gamma\gamma}$ in the category with the best resolution.

For the HH categories, the m_{jj} distributions are modeled with a double-sided Crystal Ball (CB) function, a modified version of the standard CB function [71] with two independent exponential tails. Figure 7 (right) shows the signal model for m_{jj} in the VBF and ggF categories with the best resolution.

For the HH signal, the final two-dimensional (2D) signal probability distribution function is a product of the independent $m_{\gamma\gamma}$ and m_{jj} models. The possible correlations are investigated by comparing the 2D $m_{\gamma\gamma} - m_{\mathrm{jj}}$ distribution in the simulated signal samples with the 2D probability distributions built as a product of the one-dimensional (1D) ones. With the statistical precision available in this analysis, the correlations have been found to be negligible.

11 Background model

11.1 Single Higgs background model

The SM single H background shape is constructed from the simulation following the same methodology as used for the signal model described in Section 10. For each analysis category and single H production mode, the $m_{\gamma\gamma}$ distributions are fitted using a sum of at most five Gaussian functions. The m_{jj} modeling in the HH categories depends on the production mechanism: for the ggH and VBF H processes, the m_{jj} distribution is modeled with a Bernstein polynomial; for VH production a CB function is used to model the distribution of the hadronic decays of vector bosons; for $\mathrm{t\bar{t}H}$ a Gaussian function is used. Like for the signal modeling,

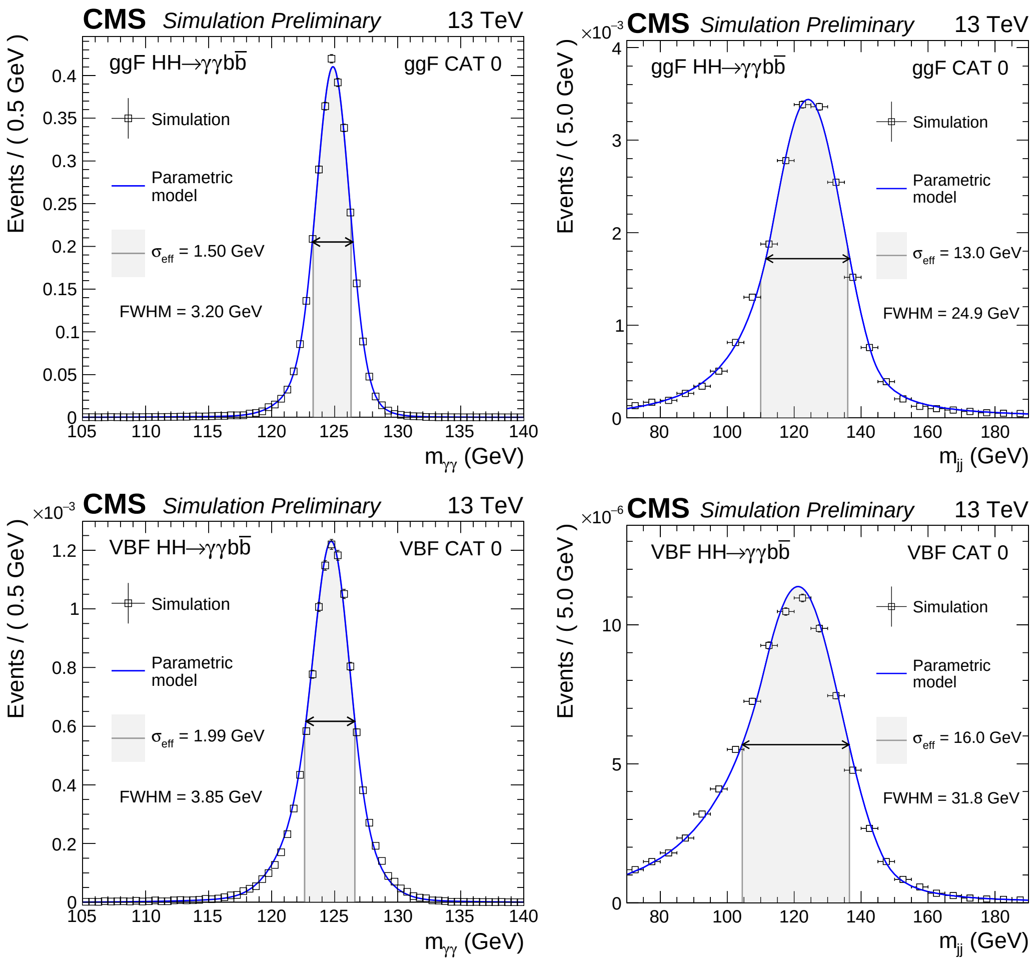


Figure 7: Parametrized signal shape for $m_{\gamma\gamma}$ (left) and $m_{\rm jj}$(right) in the best resolution ggF (top) and VBF (bottom) categories. The open squares represent weighted simulated events and the blue lines are the corresponding models. Also shown are the $\sigma_{\rm eff}$ value (half the width of the narrowest interval containing 68.3% of the invariant mass distribution) and the corresponding interval as a gray band, and the full width at half the maximum (FWHM) and the corresponding interval as a double arrow.

the final 2D SM single H model is a product of the independent models of the $m_{\gamma\gamma}$ and m_{jj} distributions.

11.2 Nonresonant background model

The model used to describe the nonresonant background is extracted from data using the discrete profiling method [72] as described in Ref. [52]. This technique was designed as a way to estimate the systematic uncertainty associated with choosing a particular analytic function to fit the background $m_{\gamma\gamma}$ and m_{jj} distributions. The method treats the choice of the background function as a discrete nuisance parameter in the likelihood fit to the data. This method is generalized to the 2D model case for the HH categories as a product of two 1D models.

MC pseudo-experiments were generated with positive and negative correlations between $m_{\gamma\gamma}$ and m_{jj} injected and then fitted with the factorized 2D model. A negligible bias has been observed, and the correlations have been found to be within the statistical precision of the analysis.

12 Systematic uncertainties

This search is statistically limited, and the total impact of systematic uncertainties on the final result is about 2%. The systematic uncertainties mainly affect the signal model and the resonant single H background, since the nonresonant background model is constructed in a data-driven way with the uncertainties associated with the choice of a background fit function taken into account by the discrete profiling method described in Section 11.2. The systematic uncertainties can affect the overall normalization, or a variation in category yields, representing event migration between the categories. Theoretical uncertainties have been applied to the HH and single H normalization. The following sources of theoretical uncertainty are considered: the QCD scale uncertainty, the uncertainty in the strong force coupling constant α_S, the impact of the PDF choice, and the uncertainty in the prediction of the branching fraction $\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$. The dominant theoretical uncertainties arise from the prediction of the SM HH and $\mathrm{t\bar{t}H}$ production cross sections. In addition, a conservative PS uncertainty is assigned to the VBF HH signal, defined as the full symmetrized difference in yields in each category obtained by varying the parton shower ISR and FSR parameters. The dominant experimental uncertainties are:

- *Photon identification BDT score*: the uncertainty arising from the photon identification BDT score is estimated by rederiving the corrections with equally sized subsets of the $\mathrm{Z} \to \mathrm{ee}$ events used to train the quantile regression corrections. Its magnitude corresponds to the standard deviation of the event-by-event differences in the corrected photon ID BDT output score obtained with the two training subsets. This uncertainty reflects the limited capacity of the network arising from the finite size of the training set. It is seen to cover the residual discrepancies between data and simulation. The uncertainty in the signal yields is estimated by propagating this uncertainty through the full category selection procedure.
- *Photon energy scale and resolution*: the uncertainties associated with the corrections applied to the photon energy scale in data and the resolution in simulation are evaluated using $\mathrm{Z} \to \mathrm{ee}$ events.
- *Per-photon energy resolution estimate*: the uncertainty in the per-photon resolution is parametrized as a rescaling of the resolution by $\pm 5\%$ around its nominal value. This is designed to cover all differences between data and simulation in the distribution,

which is an output of the energy regression.

- *Jet energy scale and smearing corrections*: The energy scale of jets is measured using the p_T balance of jets with Z bosons and photons in $Z \rightarrow ee$, $Z \rightarrow \mu\mu$ and $\gamma+$jets events, as well as using the p_T balance between jets in dijet and multijet events [70]. The uncertainty in the jet energy scale is a few percent and depends on p_T and η. The impact of jet energy scale uncertainties on the event yields is evaluated by varying the jet energy corrections within their uncertainties and propagating the effect to the final result. Correlations between years are introduced for the different jet energy scale uncertainty sources, ranging from 0 to 100%.
- *Jet b tagging*: uncertainties in the b tagging efficiency are evaluated by comparing data and simulated distributions for the b tag discriminator. These include the statistical uncertainty in the estimate of the fraction of heavy and light flavor jets in data and simulation.
- *Trigger efficiency*: the efficiency of the trigger selection is measured with $Z \rightarrow ee$ events using a tag-and-probe technique [73]. An additional uncertainty is introduced to account for a gradual shift in the timing of the inputs of the ECAL L1 trigger in the region $|\eta| > 2.0$, which caused a specific trigger inefficiency during 2016 and 2017 data taking. Both photons and, to a greater extent, jets can be affected by this inefficiency.
- *Photon preselection*: the uncertainty in the preselection efficiency is computed as the ratio between the efficiency measured in data and in simulation. The preselection efficiency in data is measured with the tag-and-probe technique from $Z \rightarrow ee$ events [73].
- *Integrated luminosity*: uncertainties are determined by the CMS luminosity monitoring for the 2016, 2017, and 2018 data-taking years [74–76]. These are partially correlated across the different data-taking years to account for common sources of uncertainty in the luminosity measurement schemes.
- *Pileup jet identification*: the uncertainty in the pileup jet classification output score is estimated by comparing the score of jets in events with a Z boson and one balanced jet in data and simulation. The assigned uncertainty depends on p_T and η, and is designed to cover all differences between data and simulation in the distribution.

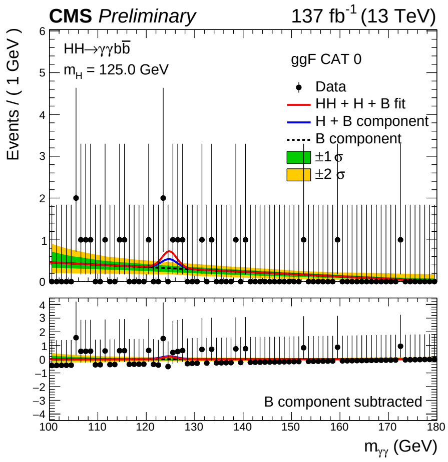
Most of the experimental uncertainties are uncorrelated among the three data-taking years. Partial correlations are introduced for the luminosity and jet energy correction uncertainties.

13 Results

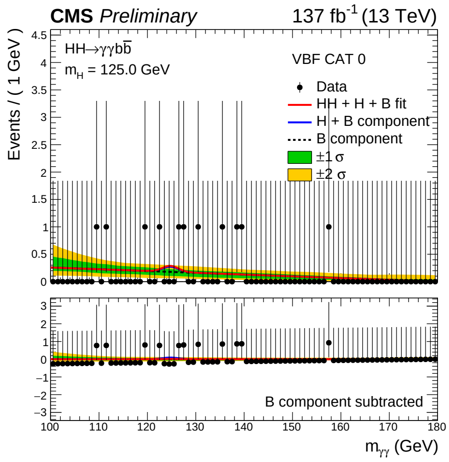
A simultaneous unbinned maximum likelihood fit to the $m_{\gamma\gamma}$ and m_{jj} distributions is performed in the 14 HH categories to extract the HH signal. The fit is performed in the mass ranges $100 < m_{\gamma\gamma} < 180\,\text{GeV}$ and $70 < m_{jj} < 190\,\text{GeV}$ for all categories apart from ggF CAT10 – CAT11. In those two categories a small but nonnegligible turn-on was observed in the m_{jj} distribution. Therefore the m_{jj}fit range is reduced to $90 < m_{jj} < 190\,\text{GeV}$; this avoids a possible bias with minimal impact on the analysis sensitivity.

In order to determine κ_λ and κ_t, the HH and $t\bar{t}H$ categories are used together in a simultaneous maximum likelihood fit. In the $t\bar{t}H$ categories, a binned maximum likelihood fit is performed to $m_{\gamma\gamma}$ in the mass range $100 < m_{\gamma\gamma} < 180\,\text{GeV}$.

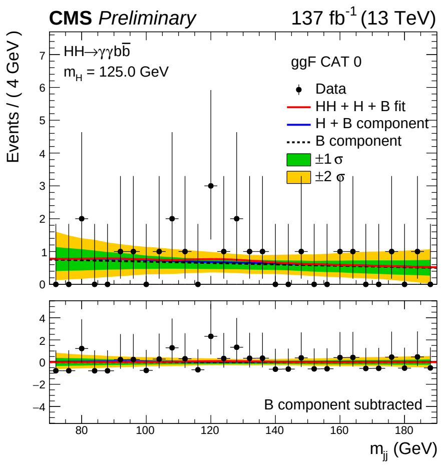
The data and the signal-plus-background model fit to $m_{\gamma\gamma}$ and m_{jj} are shown in Fig. 8 for the best resolution ggF and VBF categories. The distribution of events weighted by S/(S+B) from all HH categories is shown in Fig. 9 for $m_{\gamma\gamma}$ (left) and m_{jj} (right).



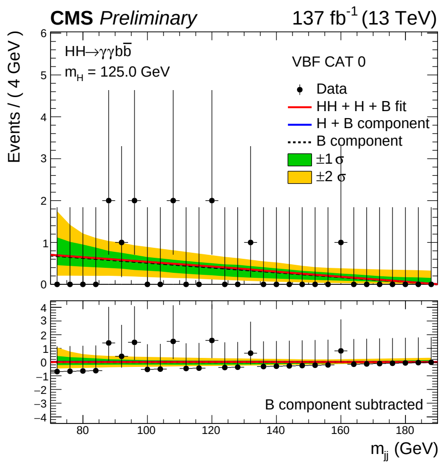
(a) ggF CAT 0



(b) VBF CAT 0



(c) ggF CAT 0



(d) VBF CAT 0

Figure 8: Invariant mass distributions $m_{\gamma\gamma}$ (top row) and m_{jj} (bottom row) for the selected events in data (black points) in the best resolution ggF (CAT0) and VBF (CAT0) categories are shown. The solid red line shows the sum of the fitted signal and background, the solid blue line shows the background component from the single Higgs boson and the nonresonant processes, and the dashed black line shows the nonresonant background component. The one (green) and two (yellow) standard deviation bands include the uncertainties in the background component of the fit. The lower panel in each plot shows the residuals after the background subtraction.

No signal has been observed. We set 95% CL upper limits on the product of the production cross section of a pair of Higgs bosons and the branching fraction into $\gamma\gamma\mathrm{b\bar{b}}$, $\sigma_{\mathrm{HH}}\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$, using the modified frequentist approach for confidence levels ($\mathrm{CL_s}$), taking the LHC profile likelihood ratio as a test statistic [77–80] in the asymptotic approximation. The observed (expected) 95% CL upper limit on $\sigma_{\mathrm{HH}}\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ amounts to 0.67 (0.45) fb. The observed (expected) limit corresponds to 7.7 (5.2) times the SM prediction. This is the most stringent limit on $\sigma_{\mathrm{HH}}\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ to date. All results were extracted assuming $m_{\mathrm{H}} = 125.00$ GeV. We observe a variation smaller than 1% in both the expected and observed upper limits when using $m_{\mathrm{H}} = 125.38$ GeV, corresponding to the most precise measurement of the Higgs boson mass to

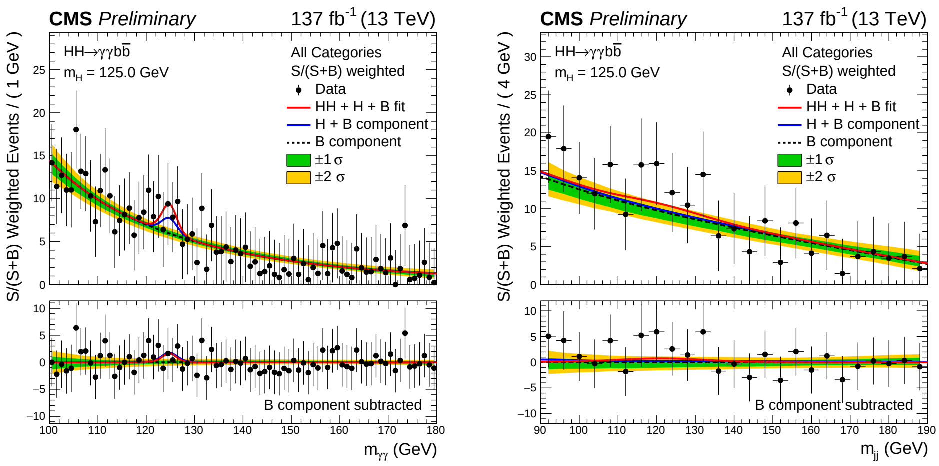


Figure 9: Invariant mass distribution $m_{\gamma\gamma}$ (left) and m_{jj} (right) for the selected events in data (black points) weighted by $S/(S+B)$, where S (B) is the number of expected signal (background) events in a $\pm\,1\sigma_{\text{eff}}$ mass window centered on m_{H}. The variable σ_{eff} is defined as the smallest interval containing 68.3% of the distribution. The solid red line shows the sum of the fitted signal and background, the solid blue line shows the background component from the single Higgs boson and the nonresonant processes, and the dashed black line shows the nonresonant background component. The one (green) and two (yellow) standard deviation bands include the uncertainties in the background component of the fit. The lower panel shows the residuals after the background subtraction.

date [81].

Limits are also derived as a function of κ_λ assuming that the top quark Yukawa coupling is SM-like ($\kappa_t = 1$). The result is shown in Fig. 10. The variation in the excluded cross section as a function of κ_λ is directly related to changes in the kinematical properties of HH production. At 95% CL the analysis constrains κ_λ to values in the interval $[-3.3, 8.5]$, while the expected constraint on κ_λ is in the interval $[-2.5, 8.2]$.

Assuming instead that an HH signal exists with the properties predicted by the SM, constraints on λ_{HHH} can be set. As discussed in Section 9.1, more stringent constraints on κ_λ can be obtained by combining the HH and $\text{t}\bar{\text{t}}\text{H}$ processes. The results are obtained both with the HH categories only, and with the HH categories combined with the $\text{t}\bar{\text{t}}\text{H}$ categories in a simultaneous maximum likelihood fit. The HH signal is considered together with the single H processes ($\text{t}\bar{\text{t}}\text{H}$, ggH, VBF H ,VH and Higgs boson production in association with a single top quark). The cross sections and branching fractions of the HH and single H processes are scaled as a function of κ_λ, while the top quark Yukawa coupling is assumed to be SM-like, $\kappa_t = 1$. One-dimensional negative log-likelihood scans for λ_{HHH} are shown in Fig. 11 for an Asimov data set generated with the SM hypothesis, $\kappa_\lambda = 1$, and for the observed data. Combining the HH analysis categories with the $\text{t}\bar{\text{t}}\text{H}$ categories improves the results, leading to $\kappa_\lambda = 0.6^{+6.3}_{-1.8}$ ($\kappa_\lambda = 1.0^{+5.7}_{-2.5}$ expected). Values of κ_λ outside the interval $[-2.7, 8.6]$ are excluded at 95% CL. The expected exclusion at 95% CL corresponds to the region outside the interval $[-3.3, 8.6]$.

The shape of the likelihood as function of κ_λ in Fig. 11 is characterized by 2 minima. This is related to an interplay between the cross section dependence on κ_λ and differences in acceptance between the analysis categories. The full degeneracy of the global minimum can be avoided thanks to the categorization in $\widetilde{M}_{\text{X}}$ and the good signal efficiency for low m_{HH}. The low mass $\widetilde{M}_{\text{X}}$ categories provide a better constraint on high values of κ_λ, while the intermediate mass

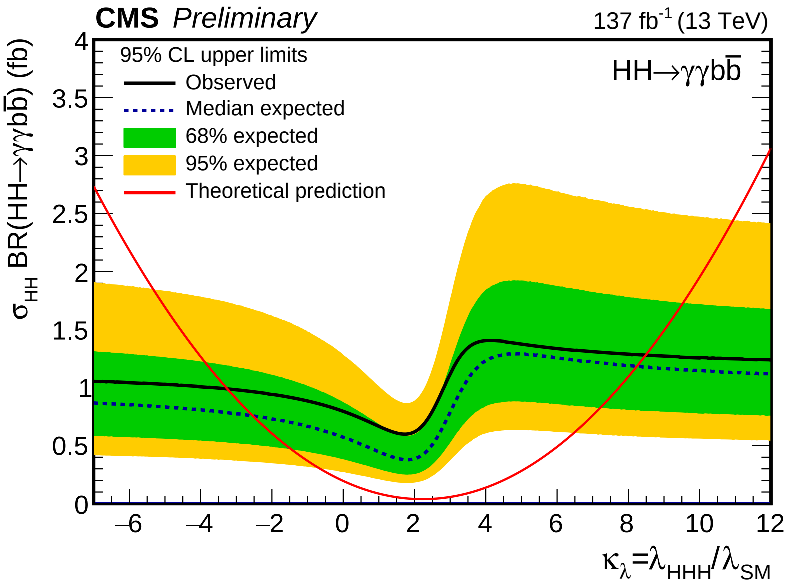


Figure 10: Expected and observed 95% CL upper limits on the product of the HH production cross section and $\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ obtained for different values of κ_λ assuming $\kappa_\mathrm{t} = 1$. The green and yellow bands represent, respectively, the one and two standard deviation extensions beyond the expected limit. The red line shows the theoretical prediction.

categories help to remove the degeneracy in the global minimum.

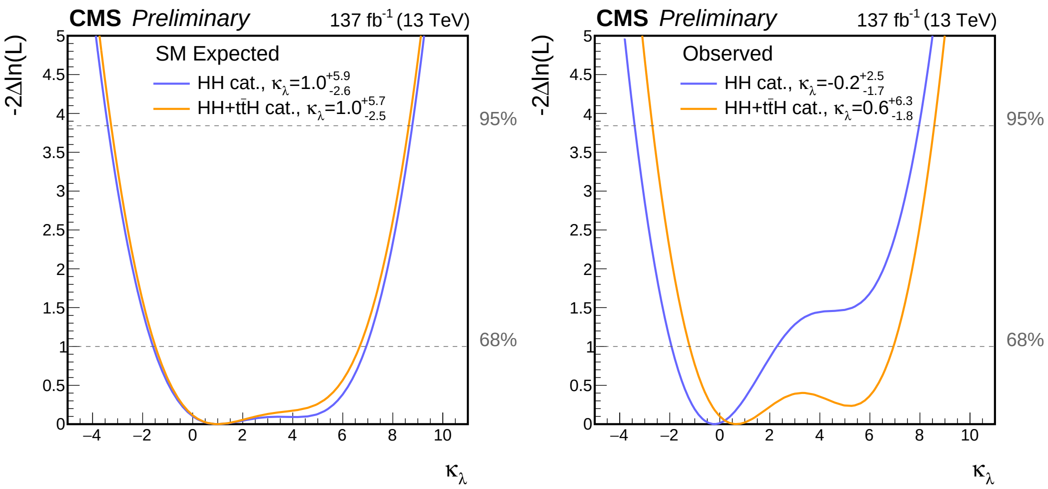


Figure 11: Negative log-likelihood as a function of κ_λ evaluated with an Asimov data set assuming the SM hypothesis (left) and the observed data (right) are shown. The 68 and 98% CL intervals are shown with the dashed gray lines. The two curves are shown for the HH (blue) and HH + $\mathrm{t\bar{t}H}$ (orange) analysis categories. All other couplings are set to their SM values.

The HH and single Higgs boson production cross sections depend not only on κ_λ, but also on κ_t. To better constrain the κ_λ and κ_t coupling modifiers, a 2D negative log-likelihood scan in the $(\kappa_\lambda, \kappa_\mathrm{t})$ plane is performed, taking into account the modification of the production cross sections and $\mathcal{B}(\mathrm{H} \to \mathrm{b\bar{b}})$, $\mathcal{B}(\mathrm{H} \to \gamma\gamma)$ for anomalous $(\kappa_\lambda, \kappa_\mathrm{t})$ values [23]. The modification of the single H production cross section for anomalous κ_λ is modeled at NLO, while the dependence on κ_t is parametrized at LO only, neglecting NLO effects. This approximation holds as long as the value of $|\kappa_\mathrm{t}|$ is close to unity, roughly in the range $0.7 < \kappa_\mathrm{t} < 1.3$. The parametric model is not reliable outside of this range. Figure 12 shows the 2D likelihood scans of κ_λ versus κ_t for an Asimov data set assuming the SM hypothesis and for the observed data. The regions of the 2D scan where the κ_t parametrization for anomalous values of κ_λ at LO is not reliable are shown with a gray band.

The inclusion of the $t\bar{t}H$ categories significantly improves the constraint on κ_t. The 1D negative log-likelihood scan as a function of κ_t with κ_λ fixed at κ_λ =1 is shown in Fig. 13 for an Asimov data set generated assuming the SM hypothesis, $\kappa_t = 1$, as well as for the observed data. The measured value of κ_t is $\kappa_t = 1.3^{+0.2}_{-0.2}$ ($\kappa_t = 1.0^{+0.2}_{-0.2}$ expected). Values of κ_t outside the interval $[0.9, 1.9]$ are excluded at 95% CL.

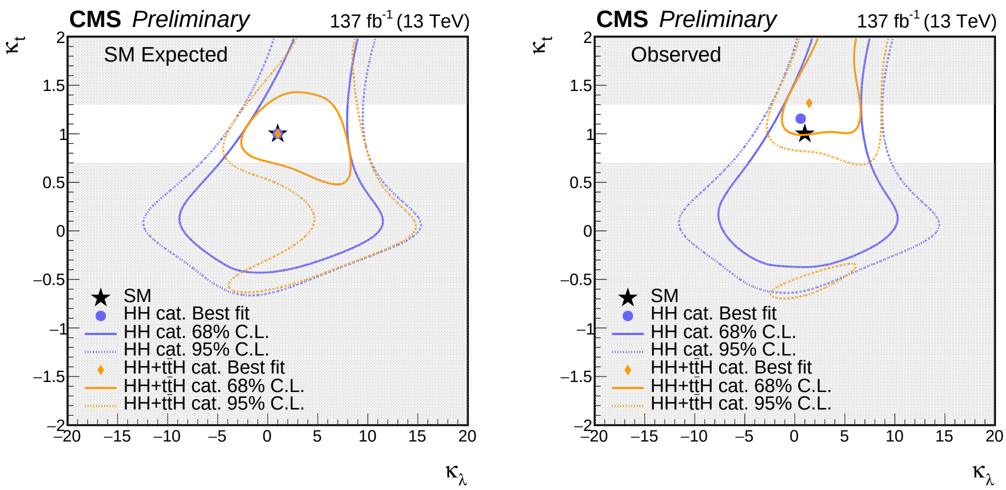


Figure 12: Negative log-likelihood contours at 68% and 95% CL in the (κ_λ, κ_t) plane evaluated with an Asimov data set assuming the SM hypothesis (left) and the observed data (right). The contours obtained using the HH analysis categories only are shown in blue, and in orange when combined with the $t\bar{t}H$ categories. The best fit value for the HH categories only (κ_λ = 0.6, κ_t = 1.2) is indicated by a blue circle, for the HH + $t\bar{t}H$ categories (κ_λ = 1.4, κ_t = 1.3) by a orange diamond, and the SM prediction (κ_λ = 1.0, κ_t = 1.0) by a black star. The regions of the 2D scan where the κ_t parametrization for anomalous values of κ_λ at LO is not reliable are shown with a gray band.

Upper limits at 95% CL are also set on the product of the HH VBF production cross section and branching fraction, $\sigma_{\text{VBF HH}}\mathcal{B}(\text{HH} \to \gamma\gamma b\bar{b})$, with the yield of the ggF HH signal constrained within uncertainties to the one predicted in the SM. The observed (expected) 95% CL upper limit on $\sigma_{\text{VBF HH}}\mathcal{B}(\text{HH} \to \gamma\gamma b\bar{b})$ amounts to 1.02 (0.94) fb. The limit corresponds to 225 (208) times the SM prediction. This is the most stringent constraint on $\sigma_{\text{VBF HH}}\mathcal{B}(\text{HH} \to \gamma\gamma b\bar{b})$ to date.

Limits are also set as a function of c_{2V}, as presented in Fig. 14. The observed excluded region corresponds to $c_{2V} < -1.3$ and $c_{2V} > 3.5$, while the expected exclusion is $c_{2V} < -0.9$ and $c_{2V} > 3.1$. It can be seen in Fig. 14 that this analysis is more sensitive to anomalous values of c_{2V} than to the region around the SM prediction. This is related to the fact that for anomalous values of c_{2V} the $\widetilde{M}_X$ spectrum is harder, which leads to an increase in the product of signal acceptance and efficiency as well as a more distinct signal topology.

In the scenario where HH production occurs via the VBF and ggF modes, we set constraints on the κ_λ and c_{2V} coupling modifiers. A 2D negative log-likelihood scan in the (κ_λ, c_{2V}) plane is performed using the 14 HH analysis categories. Figure 15 shows 2D likelihood scans for the

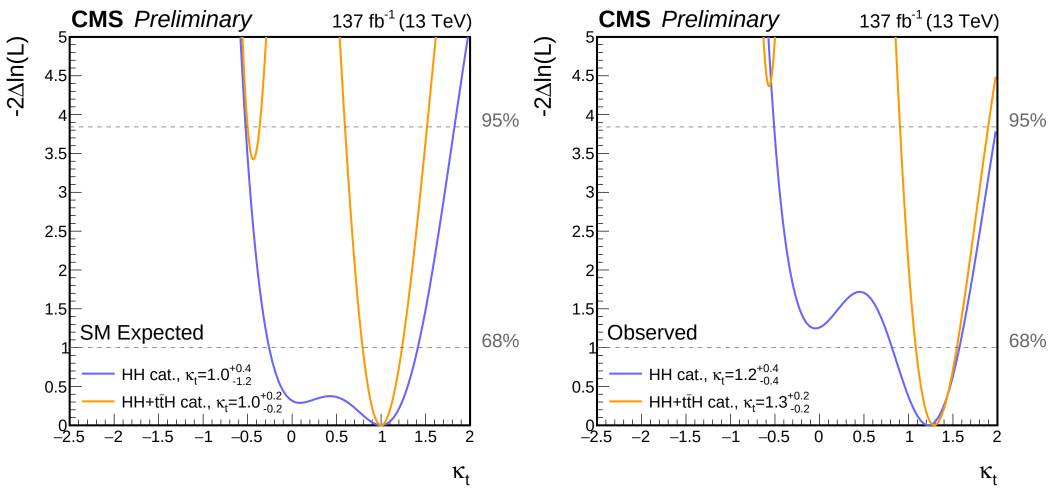


Figure 13: Negative log-likelihood scan as a function of κ_t evaluated with an Asimov data set assuming the SM hypothesis (left) and the observed data (right) are shown. The 68 and 98% CL intervals are shown with the dashed gray lines. The two curves are shown for the HH (blue) and the HH + t$\bar{t}$H (orange) analysis categories. All other couplings are fixed to their SM values.

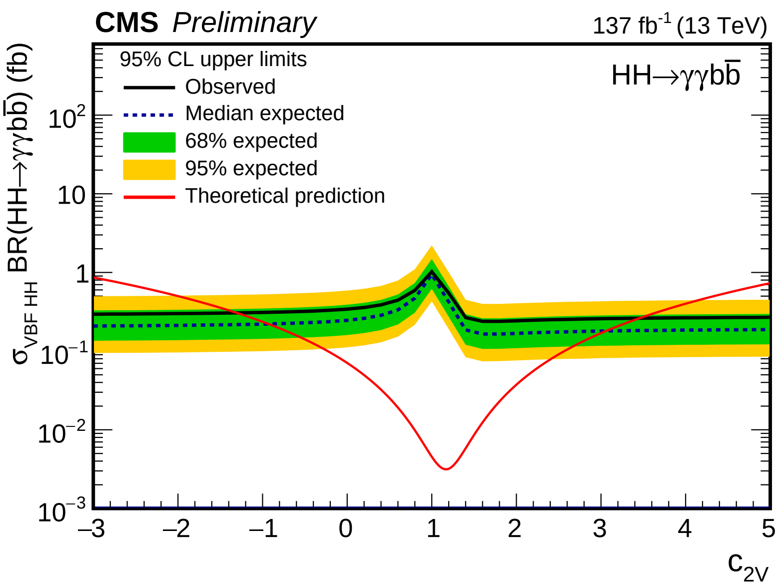


Figure 14: Expected and observed 95% CL upper limits on the product of the VBF HH production cross section and $\mathcal{B}(\mathrm{HH} \rightarrow \gamma\gamma\mathrm{b\bar{b}})$ obtained for different values of c_{2V}. The green and yellow bands represent, respectively, the one and two standard deviation extensions beyond the expected limit. The red line shows the theoretical prediction.

observed data and for an Asimov data set assuming all couplings are at their SM values.

We also set upper limits at 95% CL for the twelve BSM benchmark hypotheses defined in Table 1. The limits for different BSM hypotheses are shown in Fig. 16. In addition, limits are also calculated as a function of the BSM coupling between two Higgs bosons and two top quarks, c_2, as presented in Fig. 17. The observed excluded region corresponds to $c_2 < -0.6$ and $c_2 > 1.0$, while the expected exclusion is $c_2 < -0.4$ and $c_2 > 0.9$. The yield of the VBF HH signal is

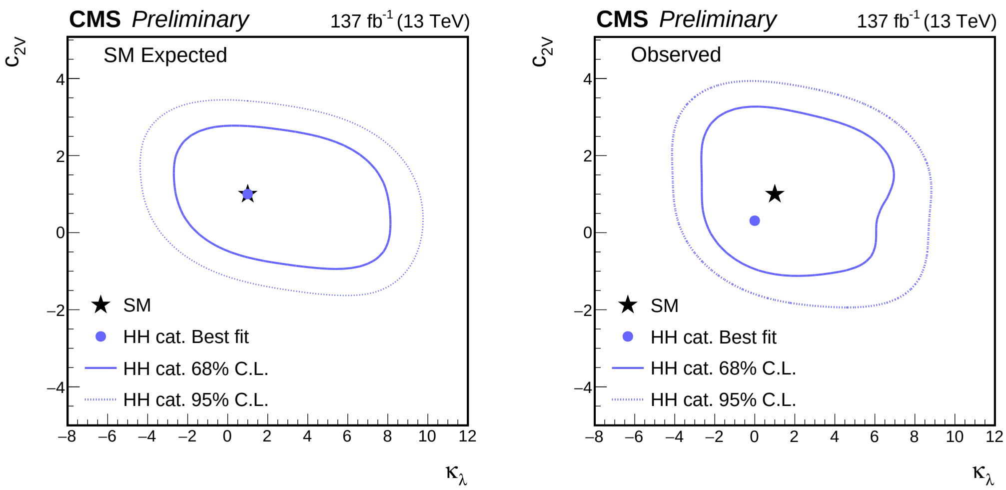


Figure 15: Negative log-likelihood contours at 68% and 95% CL in the (κ_λ, c_{2V}) plane evaluated with an Asimov data set assuming the SM hypothesis (left) and with the observed data (right). The contours are obtained using the HH analysis categories only. The best fit value (κ_λ = 0.0, c_{2V} = 0.3) is indicated by a blue circle, and the SM prediction (κ_λ = 1.0, c_{2V} = 1.0) by a black star.

constrained within uncertainties to the one predicted in the SM.

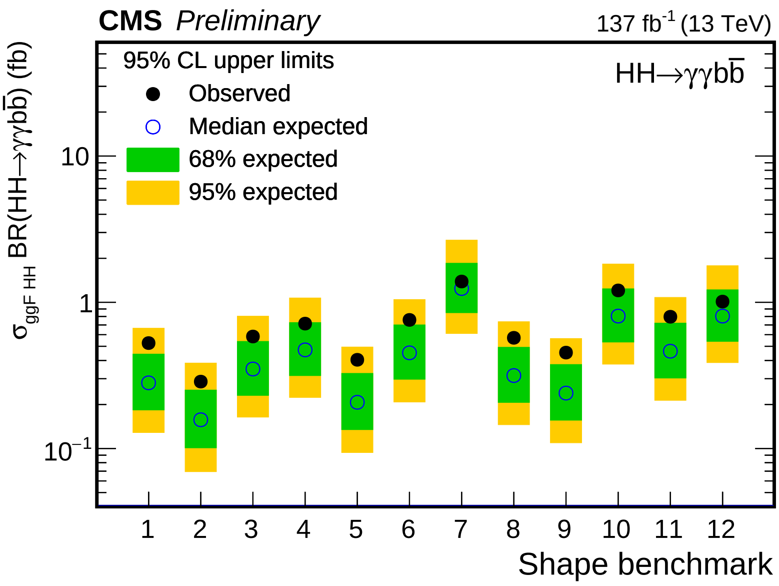


Figure 16: Expected and observed 95% CL upper limits on the product of the ggF HH production cross section and $\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ obtained for different nonresonant benchmark models (defined in Table 1). The green and yellow bands represent, respectively, the one and two standard deviation extensions beyond the expected limit.

14 Summary

A search for nonresonant Higgs boson pair production (HH) has been presented, where one of the Higgs bosons decays to a pair of bottom quarks and the other to a pair of photons. This search uses proton-proton collision data collected at $\sqrt{s} = 13\,\mathrm{TeV}$ by the CMS experiment at the LHC, corresponding to a total integrated luminosity of $137\,\mathrm{fb}^{-1}$. No signal has been

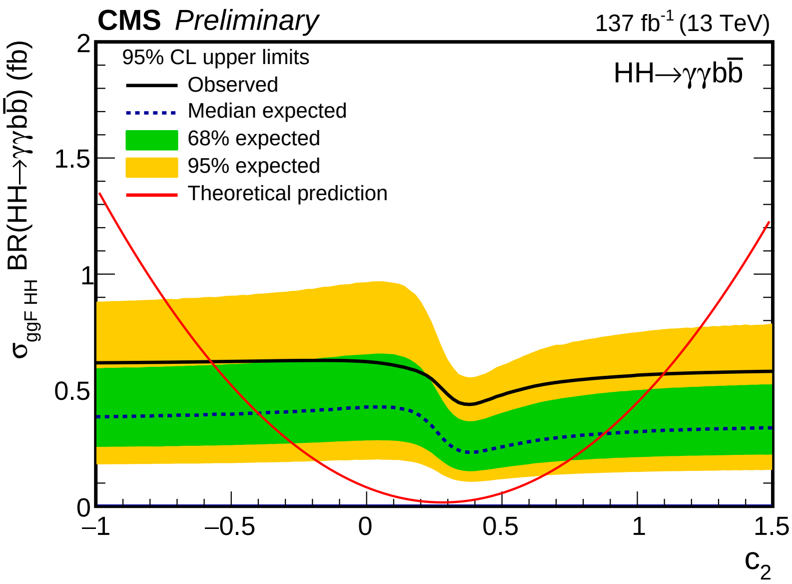


Figure 17: Expected and observed 95% CL upper limits on the product of the ggF HH production cross section and $\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ obtained for different values of the BSM coupling c_2. The green and yellow bands represent, respectively, the one and two standard deviation extensions beyond the expected limit. The red line shows the theoretical prediction.

observed. Upper limits at 95% confidence level (CL) on the product of the HH production cross section and the branching fraction into $\gamma\gamma\mathrm{b\bar{b}}$ are extracted for production in the standard model (SM) and in several scenarios beyond the standard model (BSM). The expected upper limit on $\sigma_{\mathrm{HH}}\mathcal{B}(\mathrm{HH} \to \gamma\gamma\mathrm{b\bar{b}})$ is 0.45 fb, corresponding to about 5.2 times the SM prediction, while the observed upper limit is 0.67 fb, corresponding to 7.7 times the expected value for the SM process. The presented result has the highest sensitivity to the SM HH production to date. Upper limits at 95% CL on the SM HH production cross section are also derived as a function of the Higgs boson self-coupling modifier $\kappa_\lambda \equiv \lambda_{\mathrm{HHH}}/\lambda^{\mathrm{SM}}_{\mathrm{HHH}}$ assuming that the top quark Yukawa coupling is SM-like. The coupling modifier κ_λ is constrained within a range $-3.3 < \kappa_\lambda < 8.5$, while the expected constraint is within a range $-2.5 < \kappa_\lambda < 8.2$ at 95% CL.

This search is combined with an analysis that targets top quark-antiquark associated production of a single Higgs boson decaying to a diphoton pair. In the scenario in which the HH signal has the properties predicted by the SM, the coupling modifier κ_λ has been constrained. In addition, a simultaneous measurement of κ_λ and the modifier of the coupling between the Higgs boson and the top quark κ_t is presented when both the HH and single Higgs boson processes are considered as signals.

Limits are also set on the cross section of nonresonant HH production via vector boson fusion. The most stringent limit to date is set on the product of the vector boson fusion HH production cross section and the branching fraction into $\gamma\gamma\mathrm{b\bar{b}}$. The observed (expected) upper limit at 95% CL amounts to 1.02 (0.94) fb, corresponding to 225 (208) times the SM prediction. Limits are also set as a function of the modifier of the coupling between two vector bosons and two Higgs bosons, c_{2V}. The observed excluded region corresponds to $c_{2V} < -1.3$ and $c_{2V} > 3.5$, while the expected exclusion is $c_{2V} < -0.9$ and $c_{2V} > 3.1$.

Numerous BSM hypotheses and coupling modifiers have been explored, both in the context of inclusive Higgs boson pair production and for HH production via gluon-gluon fusion and vector boson fusion. The production of Higgs boson pairs was also combined with the top quark-antiquark pair associated production of a single Higgs boson. Overall, all of the results are consistent with the SM predictions.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

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

[4] F. Englert and R. Brout, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[5] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[6] M. Grazzini et al., "Higgs boson pair production at NNLO with top quark mass effects", *JHEP* **05** (2018) 059, doi:10.1007/JHEP05(2018)059, arXiv:1803.02463.

[7] F. A. Dreyer and A. Karlberg, "Vector-Boson Fusion Higgs Pair Production at N^{3}LO", *Phys. Rev. D* **98** (2018) 114016, doi:10.1103/PhysRevD.98.114016, arXiv:1811.07906.

[8] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", Technical Report FERMILAB-FN-1025-T, CERN-2017-002-M, CERN, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[9] ATLAS Collaboration, "Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector", *Phys. Rev. Lett.* **114** (2015) 081802, doi:10.1103/PhysRevLett.114.081802, arXiv:1406.5053.

[10] ATLAS Collaboration, "Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 412, doi:10.1140/epjc/s10052-015-3628-x, arXiv:1506.00285.

[11] ATLAS Collaboration, "Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector", *Phys. Rev. D* **92** (2015) 092004, doi:10.1103/PhysRevD.92.092004, arXiv:1509.04670.

[12] ATLAS Collaboration, "Search for pair production of Higgs bosons in the $\mathrm{b\overline{b}b\overline{b}}$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **94** (2016) 052002, doi:10.1103/PhysRevD.94.052002, arXiv:1606.04782.

[13] ATLAS Collaboration, "Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector", *Phys. Lett. B* **800** (2020) 135103, doi:10.1016/j.physletb.2019.135103, arXiv:1906.02025.

[14] CMS Collaboration, "Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV", *Phys. Rev. D* **94** (2016) 052012, doi:10.1103/PhysRevD.94.052012, arXiv:1603.06896.

[15] ATLAS Collaboration, "Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment", *JHEP* **11** (2018) 040, doi:10.1007/JHEP11(2018)040, arXiv:1807.04873.

[16] CMS Collaboration, "Search for Higgs boson pair production in the $\mathrm{b\bar{b}}\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **96** (2017) 072004, doi:10.1103/PhysRevD.96.072004, arXiv:1707.00350.

[17] CMS Collaboration, "Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s}$ =13TeV", *Phys. Lett. B* **778** (2018) 101–127, doi:10.1016/j.physletb.2018.01.001, arXiv:1707.02909.

[18] CMS Collaboration, "Search for resonant and nonresonant Higgs boson pair production in the $\mathrm{b\bar{b}}\ell\nu\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **01** (2018) 054, doi:10.1007/JHEP01(2018)054, arXiv:1708.04188.

[19] CMS Collaboration, "Search for Higgs boson pair production in the $\gamma\gamma\mathrm{b\bar{b}}$ final state in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **788** (2019) 7, doi:10.1016/j.physletb.2018.10.056, arXiv:1806.00408.

[20] CMS Collaboration, "Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **122** (2019) 121803, doi:10.1103/PhysRevLett.122.121803, arXiv:1811.09689.

[21] ATLAS Collaboration, "Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **07** (2020) 108, doi:10.1007/JHEP07(2020)108, arXiv:2001.05178.

[22] CMS Collaboration, "Measurements of $\mathrm{t\bar{t}}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel", *Phys. Rev. Lett.* **125** (2020) 061801, doi:10.1103/PhysRevLett.125.061801, arXiv:2003.10866.

[23] F. Maltoni, D. Pagani, A. Shivaji, and X. Zhao, "Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC", *Eur. Phys. J. C* **77** (2017) 887, doi:10.1140/epjc/s10052-017-5410-8, arXiv:1709.08649.

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

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

[26] G. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The Strongly-Interacting Light Higgs", *JHEP* **06** (2007) 045, doi:10.1088/1126-6708/2007/06/045, arXiv:hep-ph/0703164.

[27] A. Carvalho et al., "Higgs Pair Production: Choosing Benchmarks With Cluster Analysis", *JHEP* **04** (2016) 126, doi:10.1007/JHEP04(2016)126, arXiv:1507.02245.

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

[29] G. Buchalla et al., "Higgs boson pair production in non-linear Effective Field Theory with full m_t-dependence at NLO QCD", *JHEP* **09** (2018) 057, doi:10.1007/JHEP09(2018)057, arXiv:1806.05162.

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

[31] G. Heinrich et al., "Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects", *JHEP* **06** (2019) 066, doi:10.1007/JHEP06(2019)066, arXiv:1903.08137.

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

[33] B. Hespel, D. Lopez-Val, and E. Vryonidou, "Higgs pair production via gluon fusion in the two-Higgs-doublet model", *JHEP* **09** (2014) 124, doi:10.1007/JHEP09(2014)124, arXiv:1407.0281.

[34] R. Frederix et al., "Higgs pair production at the LHC with NLO and parton-shower effects", *Phys. Lett. B* **732** (2014) 142, doi:10.1016/j.physletb.2014.03.026, arXiv:1401.7340.

[35] A. Buckley et al., "LHAPDF6: parton density access in the LHC precision era", *Eur. Phys. J. C* **75** (2015) 132, doi:10.1140/epjc/s10052-015-3318-8, arXiv:1412.7420.

[36] S. Carrazza, J. I. Latorre, J. Rojo, and G. Watt, "A compression algorithm for the combination of PDF sets", *Eur. Phys. J. C* **75** (2015) 474, doi:10.1140/epjc/s10052-015-3703-3, arXiv:1504.06469.

[37] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[38] S. Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** (2016) 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.

[39] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J. C* **75** (2015) 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.

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

[41] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

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

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

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

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

[46] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[47] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

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

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

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

[51] B. Jäger et al., "Parton-shower effects in Higgs production via Vector-Boson Fusion", *Eur. Phys. J. C* **80** (2020) 756, doi:10.1140/epjc/s10052-020-8326-7, arXiv:2003.12435.

[52] CMS Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *Eur. Phys. J. C* **74** (2014) 3076, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.

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

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

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

[56] CMS Collaboration, H. Weber, "The phase-1 upgrade of the CMS pixel detector", in *Proceedings, 2016 IEEE Nuclear Science Symposium and Medical Imaging Conference: NSS/MIC 2016: Strasbourg, France*, p. 8069719. 2016. doi:10.1109/NSSMIC.2016.8069719.

[57] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[58] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[59] CMS Collaboration, "A deep neural network for simultaneous estimation of b jet energy and resolution", arXiv:1912.06046. Submitted to Comput. Softw. Big Sci.

[60] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[61] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.

[62] N. Kumar and S. P. Martin, "LHC search for di-Higgs decays of stoponium and other scalars in events with two photons and two bottom jets", *Phys. Rev. D* **90** (2014) 055007, doi:10.1103/PhysRevD.90.055007, arXiv:1404.0996.

[63] T. Q. Nguyen et al., "Topology classification with deep learning to improve real-time event selection at the LHC", *Comput. Softw. Big Sci.* **3** (2019) 12, doi:10.1007/s41781-019-0028-1, arXiv:1807.00083.

[64] F. Chollet et al., "Keras". https://github.com/fchollet/keras, 2015.

[65] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems", 2015. Software available from tensorflow.org. http://tensorflow.org/.

[66] S. Bolognesi et al., "On the spin and parity of a single-produced resonance at the LHC", *Phys. Rev. D* **86** (2012) 095031, doi:10.1103/PhysRevD.86.095031, arXiv:1208.4018.

[67] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system", in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD, p. 785. ACM, New York, NY, USA, 2016. doi:10.1145/2939672.2939785.

[68] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning". Springer-Verlag New York, 2nd edition, 2009. doi:10.1007/978-0-387-84858-7.

[69] CMS Collaboration, "Performance of quark/gluon discrimination in 8 TeV pp data", CMS Physics Analysis Summary CMS-PAS-JME-13-002, 2013.

[70] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[71] M. J. Oreglia, "A study of the reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, Stanford University, 1980. SLAC Report SLAC-R-236.

[72] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes: the discrete profiling method", *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.

[73] CMS Collaboration, "Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **10** (2011) 132, doi:10.1007/JHEP10(2011)132, arXiv:1107.4789.

[74] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

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

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

[77] A. L. Read, "Presentation of search results: the CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[78] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[79] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[80] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

[81] CMS Collaboration, "A measurement of the Higgs boson mass in the diphoton decay channel", *Phys. Lett. B* **805** (2020) 135425, doi:10.1016/j.physletb.2020.135425, arXiv:2002.06398.