CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2024/03/25

Search for the nonresonant and resonant production of a Higgs boson in association with an additional scalar boson in the $\gamma\gamma\tau\tau$ final state

The CMS Collaboration

Abstract

Proton-proton interactions resulting in final states with two photons and two tau leptons are studied in a search for the production of two scalar bosons. The search considers both nonresonant production of a Higgs boson pair, HH, and resonant production via a new boson X which decays either to HH or to H and a new scalar Y. The analysis uses 138 fb^{-1} of data collected at a center-of-mass energy of 13 TeV with the CMS detector at the LHC from 2016 to 2018. An observed (expected) upper limit at the 95% confidence level (CL) on the HH production cross section is found to be 930 (740) fb, corresponding to 33 (26) times the standard model prediction. The observed (expected) constraint on the Higgs boson self-coupling is $-13\ (-11) < \kappa_\lambda < 18\ (16)$ at the 95% CL. Observed (expected) upper limits at the 95% CL for the $X \to HH$ cross section are found to be within 140 to 2200 (200 to 1700) fb depending on m_X. In the $X \to YH$ scenario, the most significant excess is found for $m_X = 525$ GeV and $m_Y = 115$ GeV in the $Y \to \gamma\gamma$ decay channel and has a local (global) significance of 3.4 (0.1) standard deviations. In the $Y \to \tau\tau$ decay channel, an excess at $m_X = 320$ GeV and $m_Y = 60$ GeV is found with a local (global) significance of 2.6 (2.2) standard deviations.

1 Introduction

The Higgs boson (H) holds a unique position in the standard model of particle physics (SM) as the only fundamental scalar (spin-0) boson. Since its discovery in 2012 at the CERN LHC by the ATLAS and CMS Collaborations [1–3], there has been an extensive research program to characterize its properties. So far, the properties have been measured to be consistent with SM predictions. Nevertheless, the measurements of several Higgs boson properties are extremely challenging at the LHC, leading to relatively weak constraints. One of the least constrained of these properties is the strength of the Higgs trilinear self-interaction, λ_{HHH}. Within the SM, the value of λ_{HHH} and subsequently the structure of the Higgs field potential are completely determined, given the Higgs boson mass, M_{H}, and the vacuum expectation value of the Higgs potential. Measurements of λ_{HHH} thus provide a unique consistency test of the SM, whilst offering valuable information concerning the shape of the scalar potential.

At the LHC, λ_{HHH} is directly constrained through measurements of Higgs boson pair production (HH). The primary production mechanism is via gluon-gluon fusion (ggF), with representative leading order (LO) diagrams shown in Fig. 1 (top row). The destructive interference between diagrams leads to a small SM cross section at $\sqrt{s} = 13\,\mathrm{TeV}$ of $31.1^{+2.1}_{-7.2}$ fb, calculated for $M_{\mathrm{H}} = 125\,\mathrm{GeV}$ at next-to-next-to-leading order (NNLO) with resummation at next-to-next-to-leading-logarithm accuracy and scale and top quark mass effects at next-to-leading order (NLO) [4–11]. This value is three orders of magnitude smaller than the single H production cross section, explaining why HH production is yet to be observed experimentally. The ultimate sensitivity is achieved via the statistical combination of final states, where CMS and ATLAS have recently reported upper limits on the HH cross section of 3.4 and 2.4 times the SM prediction respectively at the 95% confidence level (CL) [12, 13].

The cross section and kinematical distributions of HH production can be significantly affected in many theories of physics beyond-the-SM (BSM). The lack of discovery of BSM particles at the LHC may suggest that these particles if they exist, have masses large enough to avoid direct production at current LHC energies. This motivates the use of an effective field theory (EFT) framework to parametrize the contributions of new physics in terms of higher-dimensional operators, suppressed by a large mass scale Λ. In this analysis, BSM contributions to nonresonant ggF HH production are parametrized through an effective Lagrangian that extends that of the SM with dimension-6 operators [14, 15]. This effective Lagrangian parametrizes BSM effects in terms of five couplings involving the Higgs boson. Two are present in the SM and are expressed as ratios to their SM values: the trilinear self-coupling ($\kappa_\lambda = \lambda_{\mathrm{HHH}}/\lambda^{\mathrm{SM}}_{\mathrm{HHH}}$) and the top quark Yukawa coupling ($\kappa_{\mathrm{t}} = y_{\mathrm{t}}/y^{\mathrm{SM}}_{\mathrm{t}}$), while three are contact interactions not present in the SM: 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). Diagrams involving each of these five couplings are shown in Fig. 1 (bottom row).

Moreover, many BSM theories predict the existence of additional scalar particles whose production can lead to the same final states studied in HH production. For example, the "Randall-Sundrum" (RS) bulk model introduces compactified extra dimensions in which particles may propagate, in turn providing an explanation for the weakness of the gravitational force and a solution to the SM hierarchy problem [16–18]. The model predicts the existence of a spin-0 resonance referred to as the Radion, and a spin-2 first Kaluza-Klein excitation of the Graviton. Both resonances can decay to HH with a sizeable branching fraction, thus motivating the search for the resonant production of a Higgs boson pair.

Other examples of BSM theories are supersymmetric theories (SUSY), which extend upon the SM by introducing a fermionic (bosonic) partner for each boson (fermion) in the SM. One such

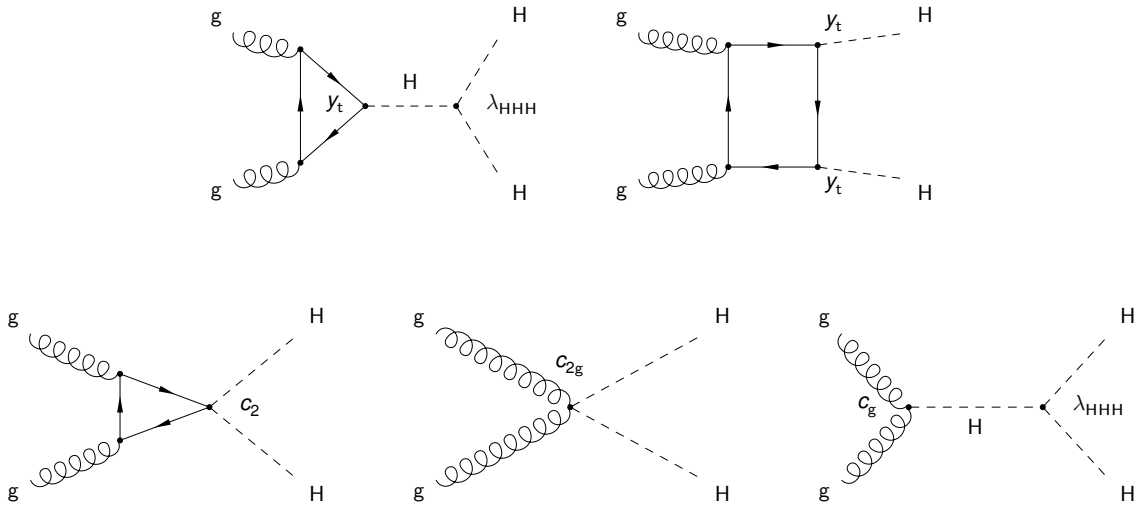

Figure 1: Leading order Feynman diagrams of the nonresonant HH production via ggF. The two diagrams in the upper row correspond to SM processes, involving the top Yukawa coupling y_t and the trilinear Higgs boson self-coupling λ_{HHH}. The diagrams in the lower row correspond to BSM processes involving contact interactions introduced in the effective field theory, namely c_2, c_{2g} and c_g.

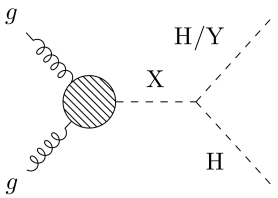

Figure 2: Feynman diagram of the resonant production of a pair of SM Higgs bosons ($X \to HH$) or a SM Higgs boson and a new scalar particle ($X \to YH$).

theory, known as the next-to-minimal supersymmetric SM (NMSSM) [19], introduces an additional complex Higgs doublet and a complex singlet field. This scenario amounts to five neutral Higgs bosons, one of which can be associated with the 125 GeV Higgs boson, H. Within the NMSSM, the heavier Higgs bosons X may decay to lighter Higgs bosons (H or Y), inspiring the search for the $X \to YH$ process at the LHC. A Feynman diagram for the resonant production of two scalar bosons is shown in Fig. 2.

Recent results from the CMS experiment using data collected between 2016 and 2018 have yielded excesses in specific final states. An excess with a local (global) significance of 3.8 (2.8) standard deviations was observed in the $X \to Y(bb)H(\gamma\gamma)$ channel at $m_X = 650\,\text{GeV}$ and $m_Y = 90\,\text{GeV}$ [20]. A search for a heavy Higgs boson decaying to a W boson pair reported an excess around 650 GeV, with a local (global) significance of 3.8 (2.6) standard deviations [21]. The search for a scalar particle decaying to a tau lepton pair yielded excesses at around 100 GeV and 1.2 TeV with local (global) significances of 3.1 (2.7) and 2.8 (2.2) standard deviations respectively [22]. In addition, a search for a light Higgs boson decaying to a photon pair observed an excess with a local (global) significance of 2.9 (1.3) standard deviations at 95 GeV [23]. All in all, the compatibility in mass across channels further motivates searching for new scalar particles in different final states, particularly those involving photons and tau leptons where excesses

have been observed.

This note presents a number of searches for the production of two scalar bosons in the $\gamma\gamma\tau\tau$ final state using proton-proton (pp) collision data collected by the CMS experiment between 2016 and 2018 at $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $138\,\text{fb}^{-1}$. Despite the small SM HH $\to \gamma\gamma\tau\tau$ branching fraction, the diphoton pair offers a clean experimental signature to trigger on with a good mass resolution, whilst the additional tau leptons in the event help further isolate signal from background. In total, there are five search channels.

1. Nonresonant HH production via ggF.
2. Resonant HH production via a spin-0 particle, $X^{(0)} \to HH$.
3. Resonant HH production via a spin-2 particle, $X^{(2)} \to HH$.
4. Resonant $X \to YH$ production, in which Y decays to a tau lepton pair, $X \to Y(\tau\tau)H(\gamma\gamma)$.
5. Resonant $X \to YH$ production, in which Y decays to a photon pair, $X \to Y(\gamma\gamma)H(\tau\tau)$.

In addition to a search for HH production, the nonresonant channel is used to place direct constraints on κ_λ, under the assumption that all other couplings of the Higgs boson have values as predicted by the SM. This channel is also used to constrain EFT parameters by placing limits on the cross section of ggF HH production for thirteen non-SM benchmark scenarios involving the five couplings: κ_λ, κ_t, c_{2g}, c_g and c_2. The benchmark scenarios are those proposed in Refs. [24, 25].

In the resonant search channels, no assumptions are made regarding the specific BSM theory to which new particles X and Y belong other than that they have a narrow width, are spin-0 except for $X^{(2)}$ which is spin-2, and that Y can decay to two photons or two tau leptons. The $X^{(0)} \to HH$ and $X^{(2)} \to HH$ channels consider the mass of X to be in the range, $260 \leq m_X \leq 1000\,\text{GeV}$. The $X \to Y(\tau\tau)H(\gamma\gamma)$ channel considers $300 \leq m_X \leq 1000\,\text{GeV}$ and $50 \leq m_Y \leq 800\,\text{GeV}$, whilst the $X \to Y(\gamma\gamma)H(\tau\tau)$ channel considers the same m_X range, with a slightly reduced range in m_Y of $70 \leq m_Y \leq 800\,\text{GeV}$ which is imposed by a triggering requirement of $m_{\gamma\gamma} > 65\,\text{GeV}$. In the $X \to YH$ channels, Y is the lighter scalar particle so m_X is required to be greater than $m_Y + M_H$.

In this analysis, results are extracted by performing fits to the diphoton invariant mass distribution, $m_{\gamma\gamma}$, in event categories designed to be enriched in signal events from one of the five search channels. The first four search channels amount to fitting a peak at M_H since the diphoton system is produced via the decay of a SM Higgs boson, $H \to \gamma\gamma$. In the $X \to Y(\gamma\gamma)H(\tau\tau)$ search, the strategy is different as the peak formed by $Y \to \gamma\gamma$ events must be fit over the wide range of considered m_Y values. To account for variations in the trigger efficiencies and event kinematics for different mass regimes, the $X \to Y(\gamma\gamma)H(\tau\tau)$ channel is effectively split into two channels.

1. Low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search considers m_Y in the range $70 \leq m_Y \leq 125\,\text{GeV}$.
2. High-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search considers m_Y in the range $125 \leq m_Y \leq 800\,\text{GeV}$.

This note is organized as follows: Section 2 gives a description of the CMS detector. Following this, an overview of the analysis strategy is provided in Section 3, and the data and simulations used are detailed in Section 4. The event reconstruction and selection are described in

Section 5, followed by Sections 6 and 7 which explain the event categorization procedure for the nonresonant search channel and the resonant search channels, respectively. The signal and background modeling is described in Section 8, and the treatment of systematic uncertainties is detailed in Section 9. Finally, the results for each search channel are presented in Section 10, and a summary of the whole analysis is provided in Section 11.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematical variables, can be found in Ref. [26].

Events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of 4 μs [27]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [28].

A particle-flow (PF) algorithm [29] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. Photons and electrons are reconstructed, at least in part, from energy deposits in the ECAL. Reconstruction begins with a clustering algorithm that groups ECAL crystals with energies exceeding a given threshold. Then, clusters within a certain geometric area are merged into "superclusters" to include photon conversions and bremsstrahlung losses. Photon candidates are identified by clusters in the ECAL not linked to any charged particle trajectories seeded in the pixel detector. The energy of the photon is then estimated by summing the energy of each crystal in the supercluster, calibrated and corrected for response variations in time. An electron is identified by an ECAL cluster that is linked to a charged particle trajectory, and its energy is estimated from a combination of the energy of the ECAL supercluster, and the momentum of the electron at the primary interaction vertex as determined by the tracker. Further detail about these reconstruction algorithms can be found in Ref. [30].

The energy scale and resolution of photons and electrons are corrected in two stages. First, a multivariate regression technique based on simulation alone is employed to account for energy not being entirely deposited inside operational ECAL crystals. Then, a further set of corrections is derived using $\mathrm{Z} \to \mathrm{ee}$ events such that simulation matches data. Both of these procedures are described in Ref. [30].

The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

In the barrel section of the ECAL, an energy resolution of about 1% is achieved for uncon-

verted or late-converting photons in the tens of GeV energy range. The energy resolution of the remaining barrel photons is about 1.3% up to $|\eta| = 1$, changing to about 2.5% at $|\eta| = 1.4$. In the endcaps, the energy resolution is about 2.5% for unconverted or late-converting photons, and between 3 and 4% for the other ones [31]. The diphoton mass resolution, as measured in $H \to \gamma\gamma$ decays, is typically in the 1–2% range, depending on the measurement of the photon energies in the ECAL and the topology of the photons in the event [32].

The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The momentum resolution for electrons with $p_T \approx 45\,\text{GeV}$ from $Z \to ee$ decays ranges from 1.6 to 5%. It is generally better in the barrel region than in the endcaps, and also depends on the bremsstrahlung energy emitted by the electron as it traverses the material in front of the ECAL [30, 33].

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. The single muon trigger efficiency exceeds 90% over the full η range, and the efficiency to reconstruct and identify muons is greater than 96%. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution, for muons with p_T up to 100 GeV, of 1% in the barrel and 3% in the endcaps. The p_T resolution in the barrel is better than 7% for muons with p_T up to 1 TeV [34].

For each event, hadronic jets are clustered from these reconstructed particles using the infrared and collinear safe anti-k_T algorithm [35, 36] with a distance parameter of 0.4. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional proton-proton interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate this effect, tracks identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions [37]. Jet energy corrections are derived from simulation studies so that the average measured energy of jets becomes identical to that of particle level jets. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made [38]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures [37].

Hadronic τ decays (τ_h) are reconstructed from jets, using the hadrons-plus-strips algorithm [39], which combines 1 or 3 tracks with energy deposits in the calorimeters, to identify the tau decay modes. Neutral pions are reconstructed as strips with dynamic size in η-ϕ from reconstructed electrons and photons, where the strip size varies as a function of the p_T of the electron or photon candidate. A correction to the τ_h energy scale in simulation is derived by comparing simulation and data in Z/γ^* events [39].

To distinguish genuine τ_h decays from jets originating from the hadronization of quarks or gluons, and from electrons, or muons, the DEEPTAU algorithm is used [40]. Information from all individual reconstructed particles near the τ_h axis is combined with properties of the τ_h candidate and the event. The algorithm provides three discriminants, D_e, D_μ, and D_{jet} which are designed to distinguish τ_h candidates from electrons, muons, and jets respectively. The efficiency for identifying a genuine τ_h using these discriminants is measured in data with $Z \to ee$ events for taus with $p_T < 140\,\text{GeV}$ and $W^* \to \tau\nu_\tau$ events for $p_T > 140\,\text{GeV}$. Corrections are applied to simulated events such that the efficiencies measured with simulation match those measured with data.

The rate of a jet to be misidentified as τ_h by the DEEPTAU algorithm depends on the p_T and quark flavor of the jet. In simulated events from W boson production in association with jets, it has been estimated to be 0.43% for a genuine τ_h identification efficiency of 70%. The misidentification rate for electrons (muons) is 2.60 (0.03)% for a genuine τ_h identification efficiency of 80 (>99)%.

The missing transverse momentum vector $\vec{p}_T^{\,miss}$ 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} [41]. The $\vec{p}_T^{\,miss}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event.

3 Analysis strategy

Proton-proton collision events deemed compatible with the $\gamma\gamma\tau\tau$ final state are selected starting from a diphoton high-level trigger with asymmetric photon p_T thresholds of 30 and 18 GeV. For all except the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search channel, the subleading photon threshold is increased to 22 GeV for 2017 and 2018. The triggers also apply loose isolation requirements. Offline, the photon candidates are required to pass a loose selection on kinematical variables, with different thresholds for the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$, and a photon identification criterion. The diphoton mass, $m_{\gamma\gamma}$, is required to be greater than 65 GeV for the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search, and greater than 100 GeV for all other channels. In addition, events must contain at least one reconstructed tau lepton candidate to enter the analysis.

In each search channel, a separate machine learning (ML) classifier is used to further isolate events with signal-like characteristics from background. The classifiers are trained independently in each channel using Monte Carlo (MC) simulation of signal and background events. For the nonresonant search channel, a Boosted Decision Tree (BDT) discriminant is applied which uses input features related to the kinematical properties of the event constituents. Sequential boundaries are placed on the BDT output to define event categories of different signal purity, where the boundary positions are chosen to maximize signal sensitivity. Further details are provided in Section 6.

In the resonant channels, we search for signal events produced over a wide range of resonance mass hypotheses. As the kinematical properties of the signal events vary dramatically over the considered mass range, it would be sub-optimal to train a single classification algorithm using all mass points to identify signal from background. However, it would not be feasible to train a separate classification algorithm at each mass point due to the large number of mass points considered. A Parametric Neural Network (pNN) [42] is thus developed which has additional training features, m_X (and m_Y for the $X \to YH$ searches), that are treated as conditional parameters. The pNN is trained with signal simulation from multiple mass hypotheses simultaneously, allowing it to implicitly learn how the kinematical distributions evolve with m_X (and m_Y) and thereby change the way it discriminates between signal and background based upon the resonance mass. In this approach, only one pNN is needed for each of the resonant search channels. Furthermore, provided that enough mass hypotheses are used in training, the pNN can be used to place a sensitive upper limit for any mass hypothesis within the considered range, regardless of whether it was seen in training.

To place an upper limit, a set of analysis categories is defined by placing boundaries on the pNN output, where the pNN is evaluated at the values of m_X (and m_Y) that correspond to the mass hypothesis being considered. The boundary positions are chosen to provide good signal sensitivity whilst maintaining a sufficient background yield to be able to fit the background

model successfully. For every mass hypothesis considered, a new set of categories with new boundaries is defined using the pNN output for that particular hypothesis. Since only a finite number of upper limits can be calculated, a set of mass hypotheses must be chosen.

The granularity in m_X (and m_Y) that defines the set of chosen mass hypotheses is motivated by the experimental resolution of m_X (and m_Y). We consider setting upper limits on hypotheses within the mass range using categories defined for the closest matching hypotheses in the set. The granularity is set such that the expected upper limit for any hypothesis within the mass range is no different than 10% of the limit one would attain by creating a new set of categories where the pNN was evaluated at that specific mass hypothesis.

In each search channel, maximum likelihood fits are performed to the $m_{\gamma\gamma}$ distributions of the corresponding event categories. A likelihood function is defined for each analysis category using an extended probability density function (pdf), where the signal and background components of the model are defined as analytic functions of $m_{\gamma\gamma}$. The smoothly falling background (continuum) is modelled directly from data. The signal and the single H production, which acts as a resonant background, are modelled from simulation. In the $X \to Y(\gamma\gamma)H(\tau\tau)$ search, there is an additional background from Drell-Yan (DY) where two electrons from a Z decay are misidentified as photons. This background is modelled from data using an ABCD method [43]. The fits are used to extract 95% CL upper limits on the relevant production cross sections.

4 Data and simulated samples

This analysis uses data collected by the CMS detector during LHC pp collisions corresponding to a total integrated luminosity of $138\,\text{fb}^{-1}$, in which 36.3, 41.5, and $59.8\,\text{fb}^{-1}$ were collected during 2016, 2017, and 2018, respectively [44–46]. In 2018, the trigger used in the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search was not introduced until after the start of that year's data taking, resulting in a lower integrated luminosity of $54.7\,\text{fb}^{-1}$.

Monte-Carlo (MC) simulation of ggF HH signal is generated with POWHEG 2.0 using the full top quark mass dependence [25] and with NLO accuracy [47–54]. Both SM ($\kappa_\lambda = 1$) and BSM ($\kappa_\lambda = 0, 2.45, 5$) scenarios are studied. The dependence of the ggF HH cross section on κ_λ and κ_t can be expressed, to any order in perturbative QCD, by three terms, corresponding to diagrams involving κ_λ, κ_t, and their interference [49]. Therefore, any linear superposition of three samples involving unique κ_λ values can be used to obtain the signal sample corresponding to an arbitrary point in the $(\kappa_\lambda, \kappa_t)$ parameter space.

Following the recommendations from Refs. [24, 25], a re-weighting of the NLO κ_λ samples is applied to produce 13 EFT benchmarks hypotheses, each corresponding to a different point in the 5D parameter space defined by the couplings: κ_λ, κ_t, c_2, c_g and c_{2g}. The list of benchmark hypotheses generated is provided in Table 1.

In the resonant searches, we consider the production of X via gluon-gluon fusion. The signal simulation samples for these searches are generated using MADGRAPH5_aMC@NLO (version 2.6.5) at leading order accuracy [55] in perturbative quantum chromodynamics (QCD).

The dominant backgrounds for this analysis are irreducible prompt diphoton production ($\gamma\gamma +$ jets) and the reducible background from $\gamma +$ jets events, where jets can be misidentified as photons or tau leptons. The $\gamma\gamma +$ jets process is modelled at LO with SHERPA v.2.2.1 [56] and up to three additional jets as well as the box processes. To model $\gamma +$ jets, we use instead PYTHIA 8 [57] at LO.

Table 1: Parameter values of nonresonant BSM benchmark hypotheses. The first column corresponds to the SM sample, while the next 13 correspond to the benchmark hypotheses identified using the method from Refs. [24, 25].

	SM	1	2	3	4	5	6	7	8	8*a*	9	10	11	12
κ_λ	1.0	7.5	1.0	1.0	-3.5	1.0	2.4	5.0	1.0	15.0	1.0	10.0	2.4	15.0
κ_t	1.0	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0	1.0	1.0	1.5	1.0	1.0
c_2	0.0	-1.0	0.5	-1.5	-3.0	0.0	0.0	0.0	0.5	0.0	1.0	-1.0	0.0	1.0
c_g	0.0	0.0	-0.8	0.0	0.0	0.8	0.2	0.2	4/15	-1.0	-0.6	0.0	1.0	0.0
c_{2g}	0.0	0.0	0.6	-0.8	0.0	-1.0	-0.2	-0.2	0.0	1.0	0.6	0.0	-1.0	0.0

Other nonresonant backgrounds include $V + \gamma$, $t\bar{t}$, $t\bar{t} + \gamma$ and $t\bar{t} + \gamma\gamma$ where the vector boson(s) (from direct production or decay of a top quark) can decay into a tau lepton candidate(s) and associated jets can be misidentified as photons. These processes are simulated at NLO in QCD using MADGRAPH5_aMC@NLO v2.6.5 [55, 58, 59] for all except $t\bar{t}$ which uses MADGRAPH5_aMC@NLO v2.6.1.

Single Higgs boson production in the SM, where the Higgs boson decays to a pair of photons, is considered a resonant background. We include the following production modes: gluon-gluon fusion (ggH), vector-boson fusion (VBF), vector boson associated production (VH) and production associated with a top quark-antiquark pair $t\bar{t}H$. These processes are simulated using MADGRAPH5_aMC@NLO v2.4.2, with cross sections and decay branching ratios taken from Ref. [14]. No other processes from single H are considered, as their contributions are negligible.

All simulated samples are interfaced using PYTHIA 8 with the CP5 tune [60, 61] for parton showering, fragmentation with the standard p_T-ordered parton shower (PS) scheme and the underlying event description. Parton distribution functions are taken from the NNPDF 3.1 set [62]. The detector response is modelled using the GEANT4 package [63]. The simulation includes a description of pileup interactions and a re-weighting is applied to the simulated events such that the distribution of the number of interaction vertices matches that seen in data.

5 Event selection

Events are required to have at least two photon candidates. The variables used to select photon candidates include:

- Charged-hadron isolation ($\mathcal{I}_{ch}$) – the p_T sum of charged hadrons within a cone of $R = 0.3$ centered around the photon.
- Photon isolation ($\mathcal{I}_{ph}$) – the p_T sum of all other particles within a cone of $R = 0.3$ centered around the photon.
- Track isolation ($\mathcal{I}_{tk}$) – the p_T sum of all charged-particle tracks within a cone of $R = 0.3$ centered around the photon.
- Hadronic over EM energy ratio (H/E) – the ratio of the energy in the HCAL tower behind the photon supercluster to the energy of the photon's supercluster.
- Conversion safe electron veto (CSEV) – indicates whether the supercluster of a photon candidate is near a track that is compatible with an electron.
- R_9 – the energy sum of the 3×3 crystals centered around the most energetic crystal

in the cluster divided by the energy of the photon.

- $\sigma_{\eta\eta}$ – the energy-weighted spread of the 5×5 crystals centered around the most energetic crystal in the cluster.

Selected photons are required to have $p_T > 25\,\text{GeV}$, $\mathcal{I}_{\text{ch}} < 20\,\text{GeV}$, $\mathcal{I}_{\text{ch}}/p_T^{\gamma} < 0.3$, H/E < 0.08. Further requirements are made separately for barrel and endcap photons and for different ranges of R_9. These requirements are shown in Table 2. Photon candidates are also required to satisfy a loose identification criterion based on a BDT classifier trained to separate photons from jets [30]. The BDT is trained using variables that describe the shape of the photon's electromagnetic shower and isolation.

Table 2: Additional photon requirements for barrel and endcap photons at different ranges of R_9, intended to mimic the HLT requirements.

	R_9	$\sigma_{i\eta i\eta}$	$\mathcal{I}_{\text{ph}}$ (GeV)	$\mathcal{I}_{\text{tk}}$ (GeV)
Barrel	[0.50, 0.85]	<0.015	<4.0	<6.0
	>0.85	—	—	—
Endcaps	[0.80, 0.90]	<0.035	<4.0	<6.0
	>0.90	—	—	—

Diphoton candidates are required to have an invariant mass ($m_{\gamma\gamma}$) in the range 100–180 GeV for the searches targeting processes with a $H \to \gamma\gamma$ decay, while this range is extended to 65-1000 GeV for the search targeting $X \to YH$ production with $Y \to \gamma\gamma$ decay. At least one photon in the pair must have $p_T > 30\,\text{GeV}$. We additionally impose, in all but the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search, mass-dependent photon p_T requirements of $p_T/m_{\gamma\gamma} > 1/3$ and $>1/4$ for the highest p_T (leading) and second-highest p_T (subleading) photons, respectively. All events are required to have at least one diphoton candidate. In the case where more than one candidate exists, the candidate with the highest scalar sum p_T of the photons is chosen.

When constructing a $H/Y \to \tau\tau$ candidate, both hadronic and leptonic decay modes of the tau lepton are considered. Therefore, three primary physics objects are considered: e, μ and τ_h.

Electron candidates are required to pass the requirements: $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$, excluding the barrel-endcap transition region. Furthermore, the electron's impact parameter with respect to the primary interaction vertex in the transverse plane (d_{xy}) is required to be less than 0.045 cm, and the distance between the vertex and electron along the beam axis (d_z) is required to be less than 0.2 cm.

Muon candidates are selected after applying the cuts: $p_T > 15\,\text{GeV}$, $|\eta| < 2.4$, $d_{xy} < 0.045\,\text{cm}$, $d_z < 0.2\,\text{cm}$. In addition, isolation and identification requirements are imposed on both [34, 64]. Finally, an isolation requirement of $\Delta R > 0.2$ with respect to selected photons is placed.

The hadronically-decaying taus are selected according to: $p_T > 20\,\text{GeV}$, $|\eta| < 2.3$, $d_z < 0.2\,\text{cm}$. Additionally, an isolation requirement of $\Delta R > 0.2$ with respect to the selected photons and leptons is placed. Finally, a selection is applied using the DEEPTAU discriminants according to predefined working points. For genuine τ_h candidates with $p_T \in [30, 70]\,\text{GeV}$ from $H \to \tau\tau$ events, these working points correspond to identification efficiencies of 99%, 99.95%, and 80% for D_e, D_μ, and D_{jet} respectively [40].

If only a τ_h is reconstructed in the event, with no additional electrons or muons (that pass selection), further tau lepton candidates are searched for as isolated tracks. These tracks must be associated with PF candidates, originate from the PV, have $p_T > 5\,\text{GeV}$, $d_{xy} < 0.2\,\text{cm}$ and $d_z < 0.1$ cm. An isolation cut of $\Delta R > 0.2$ is applied with respect to the selected photons,

electrons, muons and τ_h.

A $\tau\tau$ candidate can be identified from any of the following pairs of reconstructed objects: $\tau_h\tau_h$, $\tau_h\mu$, τ_he, μe, $\mu\mu$, ee, τ_h + isolated track. Any event with an opposite-sign same flavor (OSSF) ee or $\mu\mu$ pair is rejected if either: $m_{ll} \in m_Z \pm 10\,\text{GeV}$ or $m_{ll\gamma} \in m_Z \pm 25\,\text{GeV}$, where $m_{ll\gamma}$ is calculated with respect to both the leading and subleading photons. This selection rejects events consistent with a $Z \to ll$ or $Z \to ll\gamma$ decay. Events are selected if they have a valid $\tau\tau$ candidate or if they have a single τ_h.

In this analysis, information about jets in the event is used by the multivariate techniques described in Sections 6 and 7 to reject backgrounds like $t\bar{t}$. Considered jets are required to have $p_T > 25\,\text{GeV}$, $|\eta| < 2.4$, and to be isolated from photons, leptons, or taus, which is attained by requiring $\Delta R > 0.4$ between the jet and those objects.

6 Nonresonant search event categorization

After the selection described in Section 5 is applied, events are categorized according to a discriminant designed to separate signal from background. In the nonresonant search, this discriminant is the output of a BDT trained with the XGBOOST [65] framework on simulation samples of signal and background processes. The nonresonant background simulation samples include γ + jets, $\gamma\gamma$ + jets, $t\bar{t}$ + jets, $t\bar{t} + \gamma$, $t\bar{t} + \gamma\gamma$, and $V + \gamma$ processes, while the resonant background simulation samples include ggH, VBF, VH, and $t\bar{t}H$. All backgrounds are normalized to their SM cross sections during training.

Input features of the BDT include kinematical properties of the physics objects in each event: jets, leptons, photons, jet and lepton multiplicities, p_T^{miss}, and the b-tagging scores of jets from the DEEPJET algorithm [66]. In channels that include more than one tau lepton candidate, a likelihood-based estimate of the invariant mass $m_{\tau\tau}$ [67] is also included. This estimation combines the measurement of p_T^{miss} and measurements of the visible $\tau\tau$ decay products. A further consideration in the input feature selection is the effect the BDT selection could have on the $m_{\gamma\gamma}$ spectrum. The background modelling later described in Section 8.3 requires that the $m_{\gamma\gamma}$ distribution of the continuum background be smoothly falling such that the analytical functions chosen can model it well. It is important that no peaking structure is induced in the distribution which could lead to a mistakenly-found excess. Therefore, $m_{\gamma\gamma}$ is not included as an input feature, and the p_T of the photon candidates are divided by $m_{\gamma\gamma}$ to reduce correlations between the photon's p_T and $m_{\gamma\gamma}$. The $m_{\gamma\gamma}$ distributions of the continuum background in simulation after the BDT selection are studied and no peaking structure is found.

The output score distribution from the BDT after training is shown in Fig. 3. Events are then categorized according to boundaries placed on the output score. The location of the boundaries is chosen to optimize an expected upper limit calculated with signal and background yields from simulation and without systematic uncertainties included. Events are removed from consideration if they have a BDT score <0.974. Two categories are defined with the rest of the events. The category purest in nonresonant ggF HH production, defined with a BDT score ≥ 0.989, has a signal efficiency with respect to the inclusive phase space of 10.5%. The second category, defined with a BDT score <0.989, has a signal efficiency of 3.5%. The possibility of having a greater number of categories was investigated but the improvement to the expected upper limit was found to be negligible.

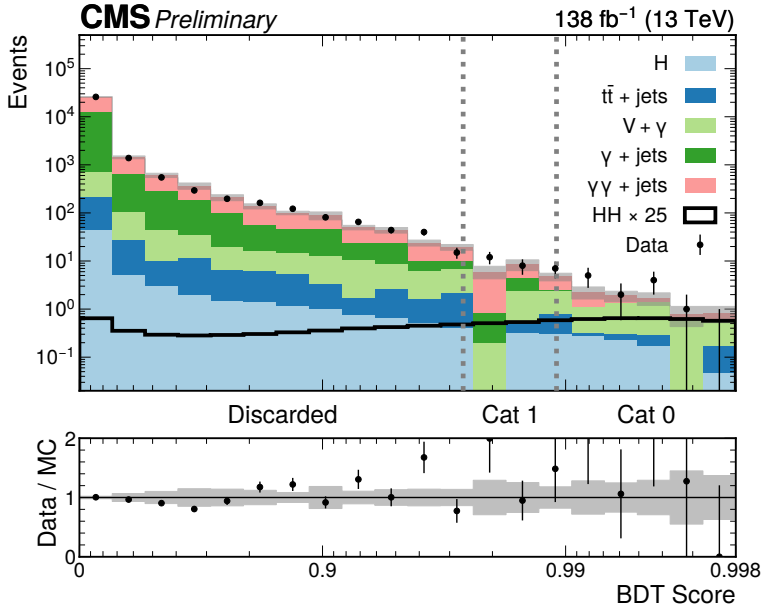

Figure 3: Distribution of the BDT scores used for the nonresonant analysis event categorization from data (black points) and predictions from MC simulation (coloured histograms). The "H" process includes ggH, VBF, VH, and $t\bar{t}H$. The background histograms are stacked, while the signal distribution is shown separately. The normalization of the signal distribution is set to 25 times the SM prediction and the background MC simulation is normalized to data. The ratio of the data to the sum of the background predictions is shown in the lower panel. Statistical MC uncertainties for the background are represented by the grey-shaded bands. The grey dotted lines represent the boundaries that define the analysis categories.

7 Resonant searches event categorization

In the resonant search, the discriminants used for event categorization are the outputs of pNNs which were first introduced in Section 3. Events passing the initial selection are used to train a pNN in each resonant search channel. The background simulation includes $\gamma\gamma$ + jets, $t\bar{t} + \gamma$, $t\bar{t} + \gamma\gamma$, and $V + \gamma$. Two processes, γ + jets and $t\bar{t}$ + jets, are excluded from training due to their limited sample size which leads to premature overtraining.

All of the events are assigned weights such that the sum of weights over a process corresponds to the expected number of events, where for the signal, the cross section is arbitrarily set to be 1 fb. Each signal sample (one for every mass hypothesis) is then reweighted such that they all have an equal sum of weights. In each training batch, there are an equal number of signal and background events. These events are randomly sampled from the entire data set with probabilities given by their weights. Once a batch is created, the weight of every event is set to be unity so that the events are unweighted in the loss function.

An un-parameterized neural network (NN) that discriminates background from signal for a single mass hypothesis ($m_X = m_i$) is trained on simulated events of all considered backgrounds and the signal hypothesis of interest. The output score of this NN can be written as $f^i(\vec{x})$ where $\vec{x}$ is a vector of input features. When training a pNN, the target function is instead $f(\vec{x}; m_X)$ where $f(\vec{x}; m_X = m_i) = f(\vec{x})^i$, i.e. a single network which adapts the way it discriminates based upon the target mass hypothesis which is passed in as an input feature.

To achieve the intended parametric behaviour, a NN is trained on simulated events of all considered backgrounds and signal hypotheses simultaneously and the parameters are included as additional input features which receive special treatment. The signal simulated events are assigned values of m_X (and m_Y) which correspond to the simulation sample they originated from. For the sake of assigning m_X (and m_Y) to the background events, each background event in a batch is paired with a signal event, and the m_X (and m_Y) values from the signal event are given to the background event. This ensures that on a batch-by-batch basis, there is no direct discrimination power provided by the parameters. Therefore, the network's dependence on the parameters is motivated purely by the correlation between the parameters and the rest of the variables.

The input features for the resonant pNNs are the same as those used for the nonresonant channel BDT, with the addition of variables related to the reconstructed resonance mass. For example, the invariant mass of the $\gamma\gamma\tau\tau$ system is used as a proxy for m_X. As in the nonresonant channel, the continuum background distributions of $m_{\gamma\gamma}$ are studied in simulation to confirm no peaking structure is induced for the pNN selection.

For each mass hypothesis, the output scores are transformed such that the distribution of the background simulated events is flat. This choice helps with the interpolation of the signal efficiency because a set of boundaries on a pNN score now corresponds to the same efficiency on the background simulated events, regardless of m_X (and m_Y). Given also that the network input distributions of the signal evolve in a continuous way with m_X (and m_Y), it is expected that the signal efficiencies for a given background efficiency, or equivalently, a set of pNN boundaries, will evolve in a predictable manner. In Section 8, one finds that this is indeed the case.

The transformed pNN output score distributions evaluated at $m_X = 260\,\text{GeV}$ and $m_X = 800\,\text{GeV}$ are shown in Fig. 4 for the $X^{(2)} \to HH$ channel. The targeted signal events are also shown. The plots demonstrate how the classification problem evolves as a function of m_X. At the high

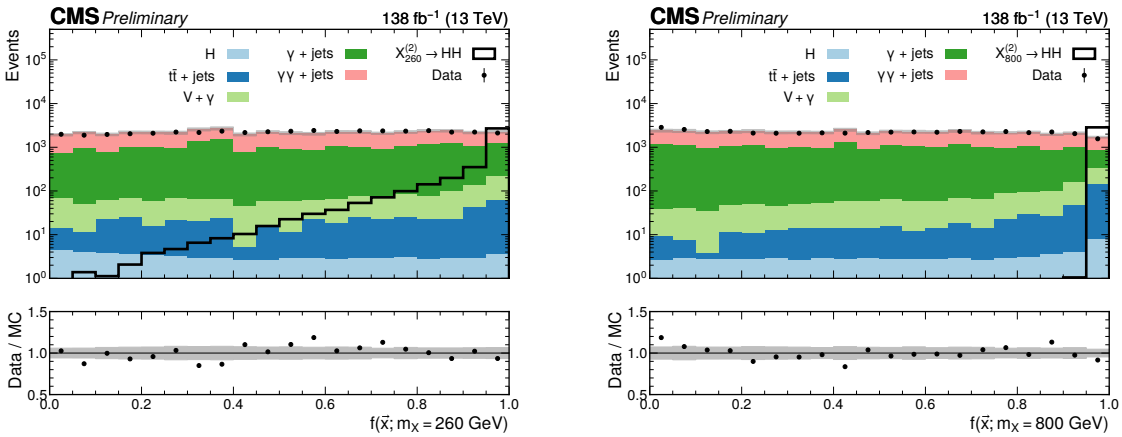

Figure 4: Transformed output of the pNN used in the $X^{(2)} \to HH$ search, evaluated at $m_X = 260\,\text{GeV}$ (left) and $m_X = 800\,\text{GeV}$ (right). The filled histograms represent the background simulation, and the data are shown by the black points. The "H" process includes ggH, VBF, VH, and $t\bar{t}H$. The targeted signal distributions for which the pNN is evaluated are shown by the black unfilled histograms. The background MC simulation is normalized to data and the signal is normalized to an arbitrary cross section for representation purposes. The ratio of the data to the sum of the background predictions is shown in the lower panel. Statistical MC uncertainties for the background are represented by the grey-shaded bands.

m_X value of 800 GeV, the signal distribution has migrated towards higher scores, showing it is easier to discriminate a high m_X signal from background. In addition, the background composition in the signal-like bins changes. For example, the relative contribution to the background from single H production is larger for $m_X = 800\,\text{GeV}$, compared to $m_X = 260\,\text{GeV}$.

Examples of transformed pNN output score distributions for all resonant search channels are shown in Fig. 5. For each channel, the pNN is evaluated at the mass hypothesis where the largest excess with respect to the background-only hypothesis is observed. In the plots, the simulation with a mass point closest to the excess is used to illustrate the signal distribution.

Event categories are defined by placing boundaries on the transformed pNN output scores. This is done separately in each channel for each mass hypothesis considered in the grid. Due to the good signal to background discrimination which is achieved in the resonant search channels, a grid-search approach to finding the optimal category boundaries leads to boundaries at very high values of the pNN score. In this case, there is an insufficient number of data events passing the selection in order to model the background $m_{\gamma\gamma}$ distribution correctly.

This motivates an alternative approach where category definitions are based on the expected number of background events. These definitions are later translated into boundaries on the pNN distribution using the background simulation (now including γ + jets and $t\bar{t}$ + jets). The translation is done in such a way that the events are ordered by their pNN scores, i.e. the events in the first category have the highest pNN scores.

A minimum number of 10 expected events was chosen since it was the lowest number (and therefore most sensitive) which still led to acceptable bias in the significance of a possible excess. Event categories are sequentially defined, starting with 10 expected background events, and continuing to add categories with 10 events, evaluating the expected limits with every category added. This continues until none of the limits improve by more than 1%, at which point a category with twice the number of events is considered instead. This continues until the sum

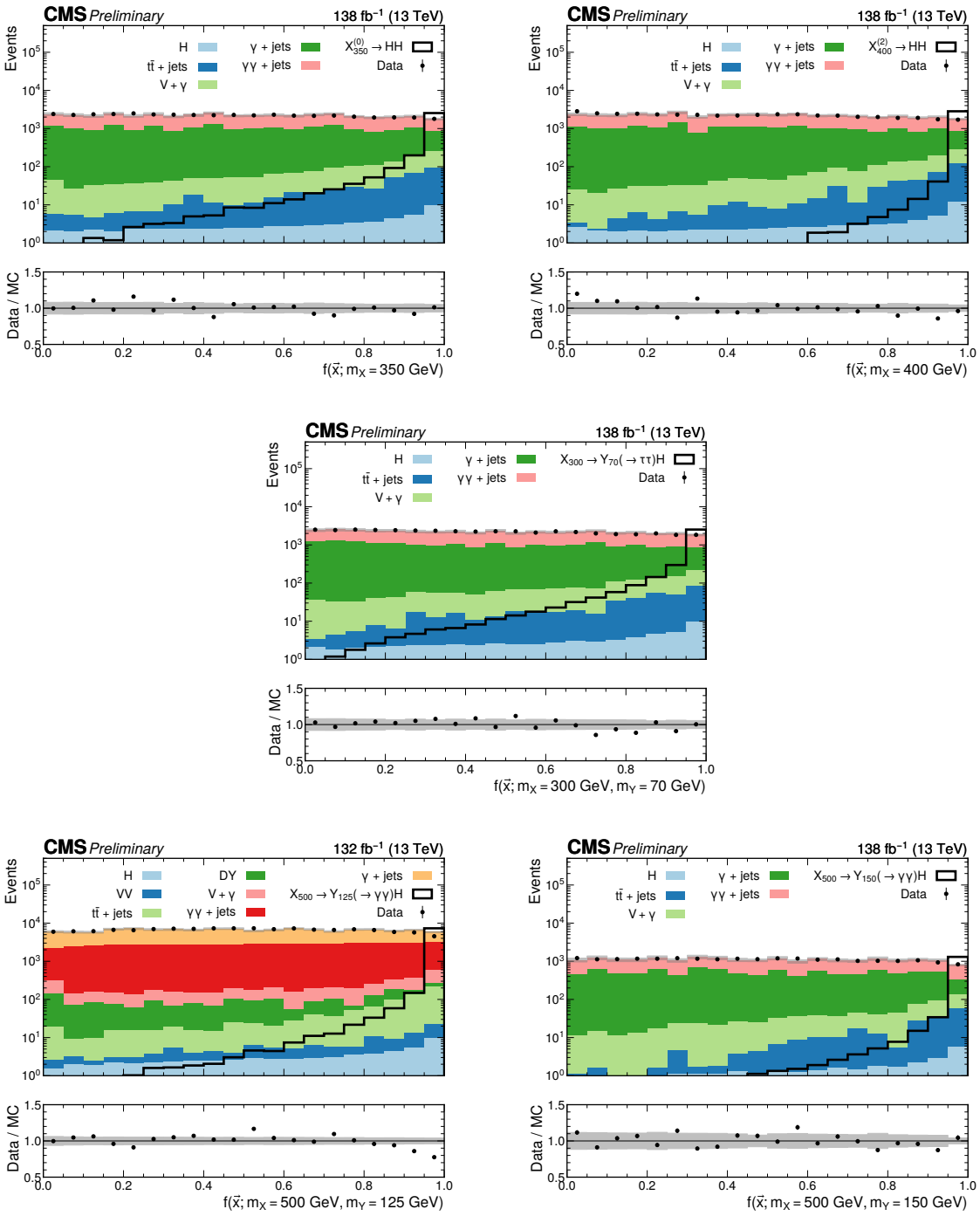

Figure 5: Transformed output of the pNNs used in the $X^{(0)} \rightarrow HH$ (top left), $X^{(2)} \rightarrow HH$ (top right), $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (middle), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom left) and high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom right) searches. The pNNs are evaluated at the mass points where the largest excess with respect to the background-only hypothesis is observed. If the MC simulation at this mass point is not available, then the sample produced at a mass point closest to the excess is shown. The filled histograms represent the background simulation, and the data are shown by the black points. The "H" process includes ggH, VBF, VH, and $t\bar{t}H$. The targeted signal distributions for which the pNN is evaluated are shown by the black unfilled histograms. The background MC simulation is normalized to data and the signal is normalized to an arbitrary cross section for representation purposes. The ratio of the data to the sum of the background predictions is shown in the lower panel. Statistical MC uncertainties for the background are represented by the grey-shaded bands.

Table 3: Number of expected events in each category for each analysis. There is one last category in addition to the ones shown here that contains the remainder of the events.

Category	0	1	2	3	4	5	6
$X^{(0)} \to HH$	10	10	10	10	20	80	-
$X^{(2)} \to HH$	10	10	10	10	20	80	-
$X \to Y(\tau\tau)H(\gamma\gamma)$	10	10	10	10	20	80	320
Low-Mass $X \to Y(\gamma\gamma)H(\tau\tau)$ $Y \to \gamma\gamma$	10	10	10	10	20	80	320
High-Mass $X \to Y(\gamma\gamma)H(\tau\tau)$	10	10	10	10	20	80	320

of events across all categories is greater than the number in data after preselection. The results of this optimization procedure are shown in Table 3.

An advantage of this categorization approach is that this definition applies to a search for any mass hypothesis. This is particularly helpful when placing a limit for a mass hypothesis where a simulation sample was not generated since an optimization specific to that hypothesis is difficult without the corresponding signal simulation.

8 Signal and background modeling

This section describes the construction of the signal and background pdfs used in the maximum likelihood fits. For the nonresonant HH, $X^{(0)} \to HH$, $X^{(2)} \to HH$ and $X \to Y(\tau\tau)H(\gamma\gamma)$ searches, the pdfs are defined over the range of $100 < m_{\gamma\gamma} < 180\,\text{GeV}$. In the high-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search, the pdfs are defined in the range of $100 < m_{\gamma\gamma} < 1000\,\text{GeV}$, whilst in the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search, the pdfs are defined in the range of $65 < m_{\gamma\gamma} < 150\,\text{GeV}$. These ranges ensure there is a sufficient number of background events that populate the $m_{\gamma\gamma}$ distribution in order to accurately model the continuum background component. The lower bound at 65 GeV in the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search is chosen to avoid the trigger efficiency turn-on curve at low $m_{\gamma\gamma}$.

The shape and normalization of the signal models are derived assuming $M_H = 125.38\,\text{GeV}$, which is the most precise measurement of the Higgs boson mass reported by CMS [32].

8.1 Signal modeling in the nonresonant search channel

The shape of the $m_{\gamma\gamma}$ distribution for signal events is derived using simulation events. This is done independently in each nonresonant search analysis category, for each year of data-taking. The signal pdf is constructed by fitting the $m_{\gamma\gamma}$ distribution with a Double Crystal Ball (DCB) function. A DCB consists of a Gaussian core which switches to a power law function on the left and another power law function on the right of the distribution:

$$f(m_{\gamma\gamma}; \overline{m}_{\gamma\gamma}, \sigma, \beta_l, m_l, \beta_r, m_r) = N \cdot \begin{cases} \exp(-x^2/2) & \text{for } -\beta_l < x < \beta_r \\ A(\beta_l, m_l) \cdot (B(\beta_l, m_l) - x)^{-m_l} & \text{for } x < -\beta_l \\ A(\beta_r, m_r) \cdot (B(\beta_r, m_r) - x)^{-m_r} & \text{for } x > \beta_r \end{cases} \tag{1}$$

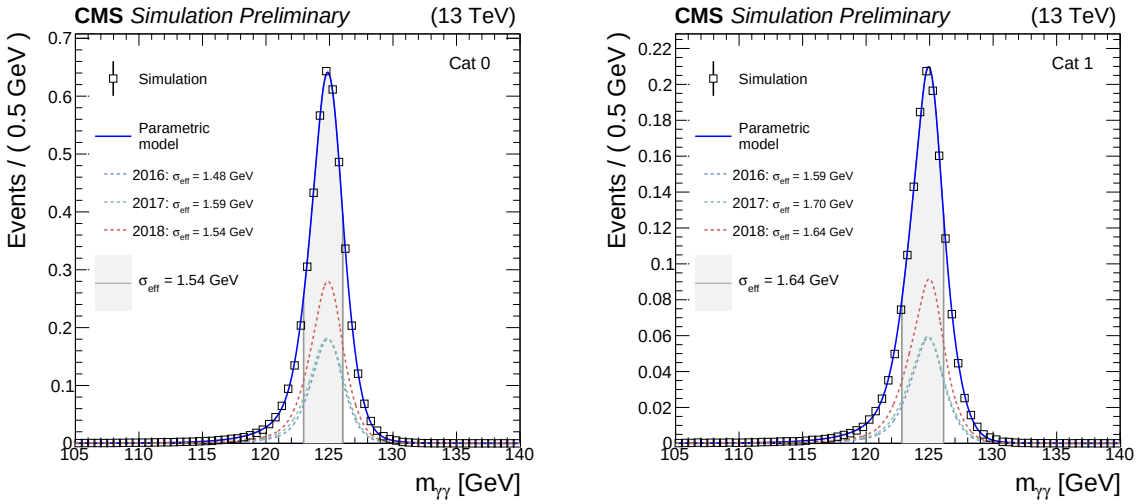

Figure 6: Signal pdfs for the nonresonant search analysis categories, shown for each year of simulated data, and for the sum of all years together. The pdfs are normalized to 25 times the expected event yields in the SM. The open squares represent the weighted simulation events and the blue line is the corresponding pdf. The grey shaded areas correspond to the σ_{eff}, defined as half the width of the narrowest interval containing 68.3% of the $m_{\gamma\gamma}$ distribution. The contribution from each year of data-taking is illustrated with the dotted lines.

where

$$x = \frac{m_{\gamma\gamma} - \overline{m}_{\gamma\gamma}}{\sigma} \tag{2}$$

$$A(\beta, m) = \left(\frac{m}{|\beta|}\right)^{m} \cdot \exp\left(-\frac{|\beta|^2}{2}\right) \tag{3}$$

$$B = \frac{m}{\beta} - |\beta| \tag{4}$$

The normalization of the signal pdf in analysis category, i, is defined by the formula,

$$N_i = \sigma(\text{pp} \to \text{HH}) \cdot \mathcal{B}(\text{HH} \to \gamma\gamma\tau\tau) \cdot \epsilon_i \cdot \mathcal{L} \cdot \mathcal{C}(\vec{\theta}), \tag{5}$$

where $\sigma(\text{pp} \to \text{HH})$ is the nonresonant ggF HH cross section on which an upper limit is placed, $\mathcal{B}(\text{HH} \to \gamma\gamma\tau\tau)$ is the SM branching fraction of HH to the $\gamma\gamma\tau\tau$ final state, ϵ_i is the efficiency for signal events to be reconstructed in analysis category i which is derived directly from simulation, and $\mathcal{L}$ is the integrated luminosity. The final term, $\mathcal{C}(\vec{\theta})$, corresponds to the corrections to the signal yield estimates from nuisance parameters, $\vec{\theta}$, representing systematic uncertainties in the analysis (see Section 9). Figure 6 shows the signal pdfs for the two nonresonant analysis categories, where the pdfs are normalized to 25 times the SM prediction for $\sigma(\text{pp} \to \text{HH})$.

8.2 Signal modeling in the resonant search channels

Extracting limits for mass hypotheses that simulation was not generated for introduces additional complexity to the signal modelling for the resonant search channels. The signal models for these hypotheses, denoted as intermediate mass points, are derived by interpolating both the shape and normalization from the neighbouring mass points in which simulation samples were generated, denoted as nominal mass points. For the $\text{X}^{(0)} \to \text{HH}$ and $\text{X}^{(2)} \to \text{HH}$ searches,

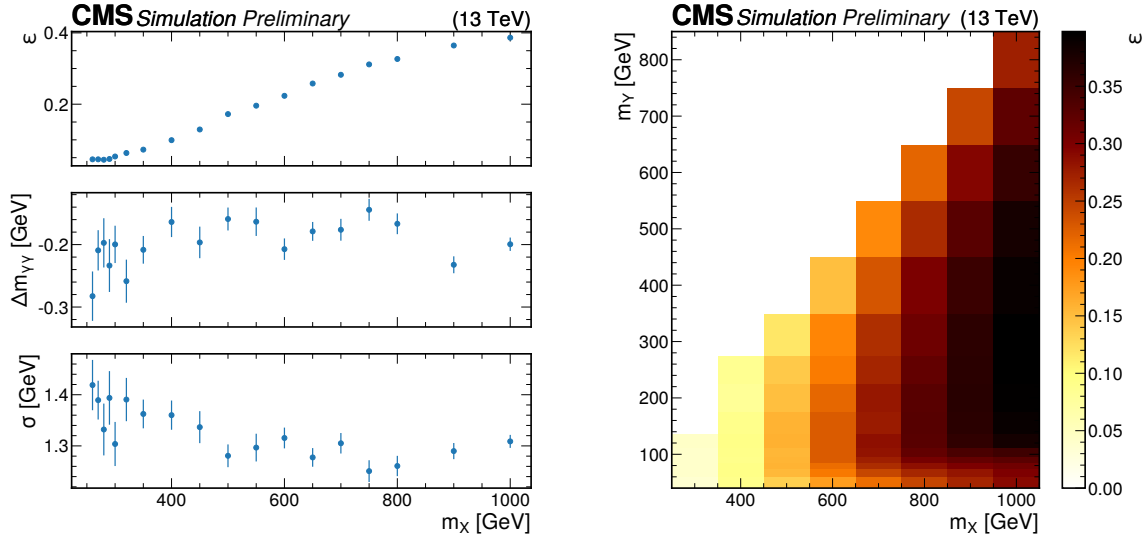

Figure 7: Signal efficiency, ϵ, and interpolated DCB shape parameters, $\Delta m_{\gamma\gamma}$ and σ, for the highest purity analysis category in the $X^{(2)} \to HH$ search, as functions of m_X (left). The first shape parameter, $\Delta m_{\gamma\gamma}$, is defined as $\overline{m}_{\gamma\gamma} - 125$. Signal efficiency in the (m_X,m_Y) plane for the highest purity analysis category in the $X \to Y(\tau\tau)H(\gamma\gamma)$ search (right).

the interpolation of parameters is performed in one dimension, m_X, whilst the $X \to YH$ searches require interpolation in two dimensions, m_X and m_Y.

As in the nonresonant search, DCB functions are used to model the signal. For the nominal mass points, the DCB shape parameters are extracted in a χ^2 fit to the simulated data points. For the intermediate mass points, $\overline{m}_{\gamma\gamma}$ and σ are determined by interpolating the values at the nominal mass points with a linear spline. The β_l, β_r, n_l and n_r parameters do not change significantly as a function of m_X and m_Y and therefore the values are simply taken from the closest nominal mass point.

In the $X^{(0)} \to HH$, $X^{(2)} \to HH$ and $X \to Y(\tau\tau)H(\gamma\gamma)$ search channels, a set of analysis categories are defined for each considered mass hypothesis, (m_X,m_Y). Signal events produced under the targeted (m_X,m_Y) hypothesis landing in analysis category, i, are normalized according to,

$$N_i(m_X, m_Y) = \sigma(pp \to X \to HH/YH \to \gamma\gamma\tau\tau) \cdot \epsilon_i(m_X, m_Y) \cdot \mathcal{L} \cdot \mathcal{C}(\vec{\theta}). \quad (6)$$

where the quantity $\sigma(pp \to X \to HH/YH \to \gamma\gamma\tau\tau)$ is the BSM cross section on which an upper limit is placed. The key difference with respect to equation 5 is that the signal efficiency, ϵ_i, becomes a function of the resonant particle masses, m_X and m_Y.

The signal efficiencies ϵ_i for the nominal mass points are computed directly from the simulation as the fraction of signal events landing in analysis category i. The functional form of $\epsilon_i(m_X, m_Y)$ is derived by creating a spline between the nominal mass point values. The signal efficiency at an intermediate mass point is then determined by interpolating along the spline. Both linear and cubic splines are considered, where the type of spline used at a given mass point is the one that leads to the lowest interpolation systematic uncertainty (see Section 9). In Fig. 7 (left), the DCB shape parameters and signal efficiency for the highest purity analysis category are shown as functions of m_X, for the $X^{(2)} \to HH$ search. Figure 7 (right) shows the signal efficiency in the (m_X,m_Y) plane, for the highest purity analysis category in the $X \to Y(\tau\tau)H(\gamma\gamma)$ search.

In the $X \to Y(\gamma\gamma)H(\tau\tau)$ search, the DCB mean is instead parametrized by $\Delta m = \overline{m}_{\gamma\gamma} - m_Y$ to account for the changing peak position as a function of m_Y. Additionally, the spacing in m_Y points in which a limit is calculated is driven by the experimental $m_{\gamma\gamma}$ resolution ($\approx$1.5 GeV at $m_Y = 125$ GeV). Given the large m_Y range considered in this search channel, it is not feasible to define optimized analysis categories at every m_Y point. It is also unnecessary since the kinematical distributions of the signal, which the pNN mainly uses to discriminate with, do not change significantly over the scale dictated by the resolution. Therefore, analysis categories are constructed for a less granular set of (m_X,m_Y) hypotheses, and when placing a limit at a particular m_Y, the closest set of analysis categories is used. This feature introduces an additional complication to the signal model, as it requires interpolating the signal model parameters for nearby values of m_Y, for which the analysis category is not optimized. These secondary interpolations are derived for the signal efficiency, Δm, and σ by propagating the neighbouring nominal mass point simulation samples into the analysis categories optimized for m_Y. The secondary interpolations for the $X \to Y(\gamma\gamma)H(\tau\tau)$ search are in general a smaller effect than the primary interpolation described above.

8.3 Background modeling

Background events are comprised of two contributions: the continuum background which forms a smoothly falling distribution in $m_{\gamma\gamma}$, and events from single H production in the diphoton final state $H \to \gamma\gamma$, which form a peak around M_H. The $H \to \gamma\gamma$ contribution is modelled using the same procedure used for signal events and is normalized according to the best available cross section and branching fraction recommendations from the LHC Higgs Working Group [14], within appropriate theory uncertainties. The remainder of this section details the modeling of the continuum background component. In the final maximum likelihood fit, the total background pdf is taken as the sum of the continuum and $H \to \gamma\gamma$ components.

The continuum background model is extracted directly from data using the discrete profiling method [68]. With this technique, the systematic uncertainty associated with choosing a particular analytic function to fit the $m_{\gamma\gamma}$ distribution is accounted for. This works by introducing a discrete nuisance parameter into the maximum likelihood fit which describes the choice of background function in an analysis category.

A large set of candidate function families is considered, including exponential functions, Bernstein polynomials, Laurent series and power law functions. For each family of functions, an $\mathcal{F}$-test [69] is performed to determine the maximum order of parameters to be used, whilst the minimum order is determined by placing a requirement on the goodness-of-fit to data. In the $\mathcal{F}$-test, each function is fit to the $m_{\gamma\gamma}$ distribution by minimizing twice the negative logarithm of the likelihood ($-2\Delta \ln L$). For the nonresonant HH, $X^{(0)} \to HH$, $X^{(2)} \to HH$ and $X \to Y(\tau\tau)H(\gamma\gamma)$ searches, the fits are performed only in the mass sidebands, defined as $m_{\gamma\gamma} \in [100, 115] \cup [135, 180]$ GeV, to avoid introducing a bias from the signal region.

The fitting strategy for the $X \to Y(\gamma\gamma)H(\tau\tau)$ searches is slightly different as it involves searching for a resonance Y over a large range in $m_{\gamma\gamma}$. To account for this, the mass sidebands used to construct the continuum background model are defined by a sliding blinded region in $m_{\gamma\gamma}$,

$$m_Y \pm 10\,\text{GeV} \cdot \frac{m_Y}{125.38\,\text{GeV}}, \tag{7}$$

where the size of the blinded region increases as a function of m_Y to account for the increase in the resonance width. This formula holds for all considered m_Y, with the exception of $70 < m_Y < 72$ GeV in the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ search. For these m_Y points, data with

$65 < m_{\gamma\gamma} < 68\,\text{GeV}$ is kept to form a low mass sideband; the fraction of signal events in this region from a $m_{\text{Y}} = 70\,\text{GeV}$ resonance is around 4%, and therefore will have a negligible impact on the construction of the background model.

In the resonant searches, the Bernstein polynomials were removed from the envelope and replaced by an exponential of a polynomial function. In the high-mass $\text{X} \to \text{Y}(\gamma\gamma)\text{H}(\tau\tau)$ search, the so-called "Dijet" functions [70] are also added to the continuum background model.

When extracting the results in the final maximum likelihood fits, the discrete profiling method minimizes the overall $-2\Delta \ln L$ considering all allowed functions for each analysis category. Both the normalization and the shape parameters of these functions are included as free parameters in the fit. A penalty term is added to $-2\Delta \ln L$ to take into account the number of floating parameters in each candidate function, thus penalizing functions with high complexity. Checks are performed to ensure that modeling the continuum background $m_{\gamma\gamma}$ distribution using the discrete profiling method introduces negligible bias to the final results.

8.4 Drell-Yan modelling

In the low-mass $\text{X} \to \text{Y}(\gamma\gamma)\text{H}(\tau\tau)$ search, there is an additional background from Drell-Yan (DY) where two electrons from a Z decay can be misidentified as photons, leading to a peak in the $m_{\gamma\gamma}$ spectrum close to the Z mass. To model this background contribution, a data-driven ABCD method [43] is employed. Three control regions are defined by inverting the electron veto (region B) or inverting the pNN selection (region C) or inverting both (region A). This leaves the signal region, labelled as D. By inverting the electron veto, regions A and B are enriched in DY. These regions constrain the shape of the DY distribution in $m_{\gamma\gamma}$. In regions A and C, the yields for DY are extracted, constraining the electron veto efficiency.

Since kinematical variables are input to the pNN, and the shape of the diphoton mass peak is dependent mainly on the kinematical variables of the two photons, the shape of the DY process is assumed to be the same in regions with the same pNN selection (in A and C, and in B and D). For regions A and C, the shape is modelled by a Gaussian except in cases where a poor goodness-of-fit is found and a DCB is used instead. When extracting the final results, the shape parameters are allowed to float. Similarly, the rates of DY in regions A, B and C are allowed to float. The rate of DY in region D is given by the rate in region B, multiplied by the electron veto efficiency (rate in region C divided by rate in region A).

Finally, given that there are 7 pNN regions, this method is applied 7 times, where there are 7 independent B and D regions, and the A and C regions remain the same.

9 Systematic uncertainties

As described in Section 8, the systematic uncertainty associated with the nonresonant background estimation is calculated with the discrete profiling method. Systematic uncertainties affecting the signal models and/or the resonant background models are implemented in one of two ways. First, uncertainties that modify the shape of the $m_{\gamma\gamma}$ distribution are incorporated as nuisance parameters affecting the mean and width of the DCB functions that comprise the signal model. Second, uncertainties that do not affect the shape of the $m_{\gamma\gamma}$ distribution are implemented as log-normal uncertainties in the event yield. Those of the first type are typically experimental sources of uncertainty that impact the energy of individual photons, while those of the second type include both theoretical and experimental sources.

9.1 Theoretical Uncertainties

Sources of theoretical uncertainties considered in this analysis are those that affect the $H \to \gamma\gamma$ branching fraction, the uncertainty for which is currently estimated to be around 3% [14], and those that affect the inclusive cross sections of ggF HH production and single H production.

Sources that affect the cross sections include the parton distribution function (PDF), the strong coupling constant (α_s), and variations of the renormalization (μ_R) and factorization (μ_F) scales which account for missing higher-order terms in the perturbative calculations of the cross sections. Additionally, for the ggF HH production cross section, there is a source of uncertainty related to the scheme and scale choice of the top mass [11]. These sources may also affect the kinematical variables of simulated events, thereby changing signal efficiencies. However, this effect has been found to be significantly smaller than the cross section uncertainty in recent $H \to \gamma\gamma$ measurements [71] and is therefore not considered.

The magnitudes of the uncertainties in the inclusive cross sections from these sources are taken from Refs. [4–10] for ggF HH production and from Refs. [14, 72, 73] for single H production. The total effect of uncertainties on HH production is found to be $^{+6.7\%}_{-23.2\%}$ for $\kappa_\lambda = 1$, while the effect on resonant background production is between 2% and 11%, depending on production mode.

9.2 Experimental Uncertainties

As described previously, the sources of experimental uncertainty are split into two categories: those modifying the shape of the $m_{\gamma\gamma}$ distribution (incorporated as nuisance parameters) and those affecting only the normalization (incorporated as log-normal uncertainties). Sources of experimental uncertainty that modify the shape of the $m_{\gamma\gamma}$ distribution are related to the corrections applied to the photon energy scale and resolution in simulation. These corrections, and associated uncertainties, are computed using $Z \to ee$ events [30]. The most dominant sources of experimental uncertainty which affect only the normalization of signal and resonant background processes are:

- *Integrated luminosity*: The integrated luminosities for the 2016, 2017, and 2018 data-taking years have individual uncertainties of 1.2–2.5% [44–46], while the overall uncertainty in the measurement for the 2016–2018 period is 1.6%. The possible correlations from different common sources of luminosity measurements are taken into account.
- *Trigger efficiency*: This efficiency is calculated with the "tag-and-probe" method using $Z \to ee$ events in data [74]. The size of uncertainty in the efficiency is between 0.1 and 3.5% depending on the signal process and category. An additional uncertainty is also added to include the trigger inefficiency arising from a gradual shift in the timing of the inputs of the ECAL L1 trigger for $|\eta| > 2.0$ in 2016–2017 data-taking years [27]. The size of this uncertainty is between 2 and 8%.
- *Diphoton preselection*: The efficiency after photon preselections is measured with simulated events and data, also using the "tag-and-probe" method with $Z \to ee$ events [74]. The ratio of the efficiencies from data and MC simulations is used to derive the correction factor and corresponding uncertainties. The size of the uncertainty in the efficiency is between 4 and 5%.
- *b jet tagging*: The DEEPJET discriminant distributions in MC simulations are reshaped by data-to-simulation scale factors [75]. The impact of different systematic uncertainty sources on the reshaping is studied. The most important effects lead to

efficiency uncertainties up to 10%.

- *Hadronic tau identification*: Selections are made using the DEEPTAU algorithm to distinguish genuine τ_h from jets, muons, and electrons. Measurements of the selection efficiency are made in data (see Section 2) and are used to correct simulation. Uncertainties associated with those corrections are propagated into an uncertainty in the signal efficiency, which is found to be up to 7%.
- *Signal model interpolation*: There is an uncertainty related to the interpolation required to calculate signal efficiencies for the intermediate mass points. For each spline used, an alternative spline is considered where the nominal mass point closest to the intermediate mass point is removed. The relative difference between the prediction of the nominal and alternative spline is taken as an uncertainty in the efficiency of the signal within that category.

Further sources of systematic uncertainty are included in the final fits to data but are negligible compared to the dominant sources described above. For every expected and observed limit reported in Section 10, the combined impact of all systematic uncertainties is less than 1%.

10 Results

For each search channel, a simultaneous maximum likelihood fit is performed to the $m_{\gamma\gamma}$ distributions in data in the corresponding analysis categories. In the resonant channels, a separate fit is performed for each considered mass hypothesis. The best-fit signal-plus-background pdfs are shown for the sum of analysis categories in each search channel in Fig. 8. The contributions from the individual analysis categories are weighted in the sum according to the ratio of S/S+B, where S and B are the expected signal and background yield estimates, respectively, in $\pm 1\sigma_{\text{eff}}$ of the signal peak. This is done such that the absolute signal yield is kept constant. The figures for the resonant search channels show the analysis categories for the m_X and (m_X,m_Y) points where the largest excess with respect to the background-only hypothesis is observed.

10.1 Results from the nonresonant search channel

In the nonresonant HH channel, an upper limit at the 95% CL is set on the production cross section of a pair of Higgs bosons, $\sigma(\text{pp} \to \text{HH})$, using the CL_s modified frequentist criterion [76], taking the LHC profile likelihood ratio as the test statistic [77], in the asymptotic approximation [76, 78, 79]. The observed (expected) 95% CL upper limit amounts to 930 (740) fb, which corresponds to 33 (26) times the SM prediction.

Limits are also derived for the nonresonant HH channels as a function of κ_λ, assuming that all other couplings of the Higgs boson have values as predicted by the SM. The result is shown in Figure 9. The variation in the excluded cross section as a function of κ_λ arises due to the different kinematical properties of nonresonant HH production for different values of κ_λ. At the 95% CL, the observed (expected) constraint on κ_λ rejects values outside of the interval $[-13, 18]$ ($[-11, 16]$). In this fit, the κ_λ-dependent NLO corrections in electroweak theory to single H production [80] are neglected in this result. Upper limits at the 95% CL are also set for the thirteen BSM benchmark scenarios listed in Table 1. These limits are shown in Fig. 10.

The results in the nonresonant search channel are not as sensitive as those found with alternative final states studied by CMS. For example, an analogous search in a final state with two bottom quarks and two photons found an observed upper limit of 7.7 times the SM prediction [81]. However, the results presented in this analysis are complementary and can be used in a combination of all final states studied by CMS.

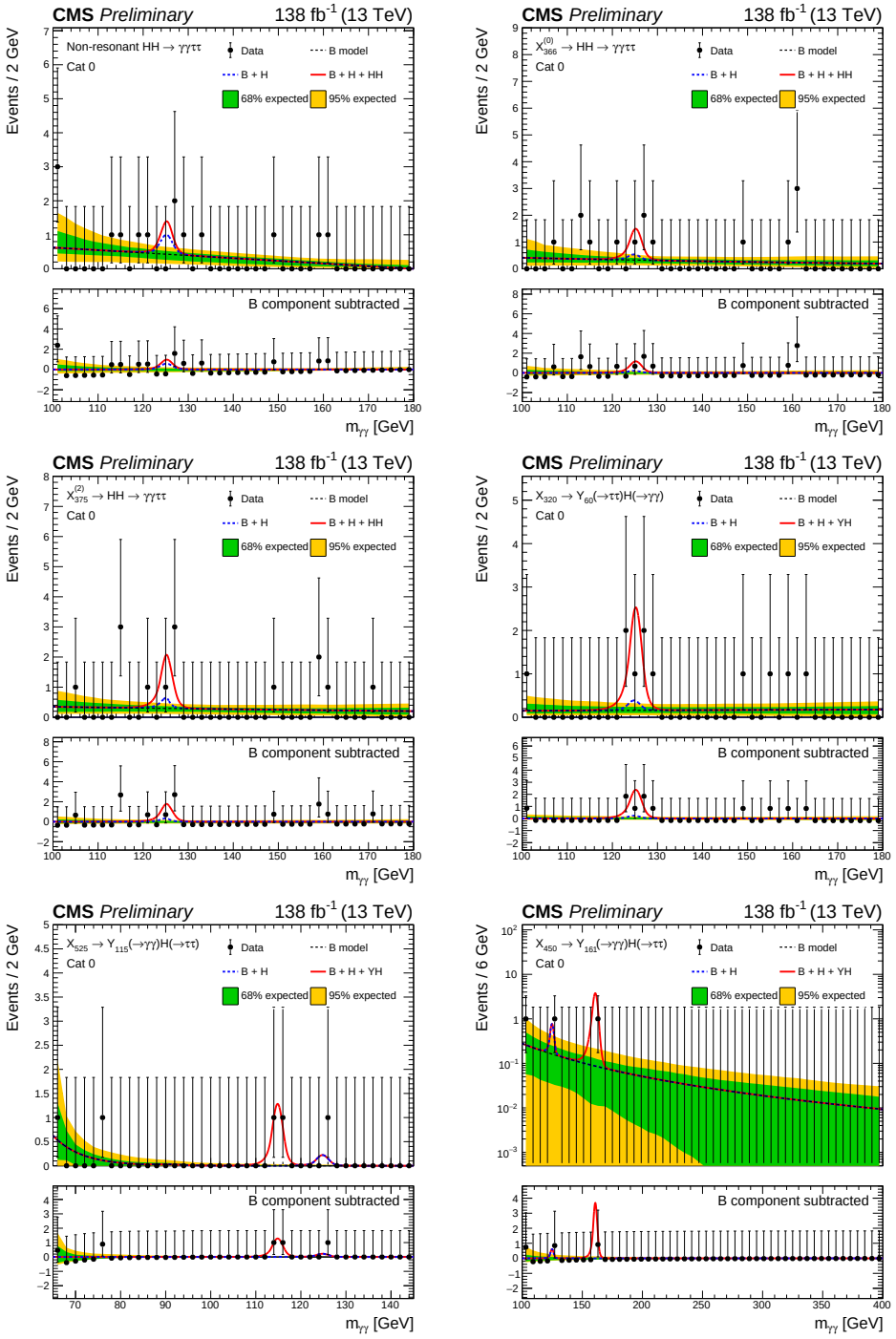

Figure 8: Data points (black) and signal-plus-background models for the most sensitive analysis category in each search channel are shown: nonresonant (top left), $X^{(0)} \rightarrow HH$ (top right), $X^{(2)} \rightarrow HH$ (middle left), $X \rightarrow Y(\tau\tau)H(\gamma\gamma)$ (middle right), low-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom left) and the high-mass $X \rightarrow Y(\gamma\gamma)H(\tau\tau)$ (bottom right). The analysis categories for the resonant search channels correspond to the mass hypotheses where the largest excesses with respect to the background-only hypothesis are observed. The one (green) standard deviation and two (yellow) standard deviation bands show the uncertainties in the background component of the fit. The solid red line shows the sum of the fitted signal and background components, the solid blue line shows the continuum background and the background from single H production together, and the dashed black line shows only the continuum background component. The lower panel in each plot shows the residual signal yield after subtraction of the background.

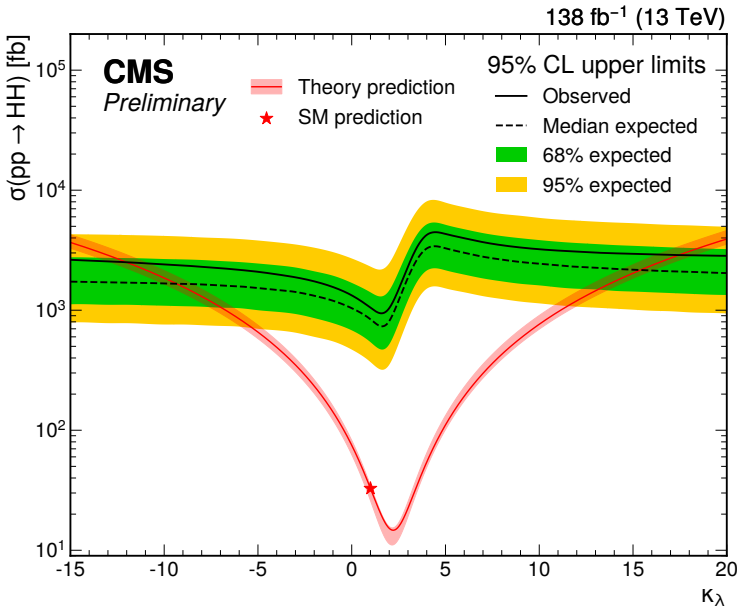

Figure 9: Expected and observed upper limits on the nonresonant HH production cross section at the 95% CL, obtained for different values of κ_λ. The green and yellow bands represent the one and two standard deviations for the expected limit, respectively. The theoretical prediction with the uncertainty of the cross section as a function of κ_λ is shown by the red band.

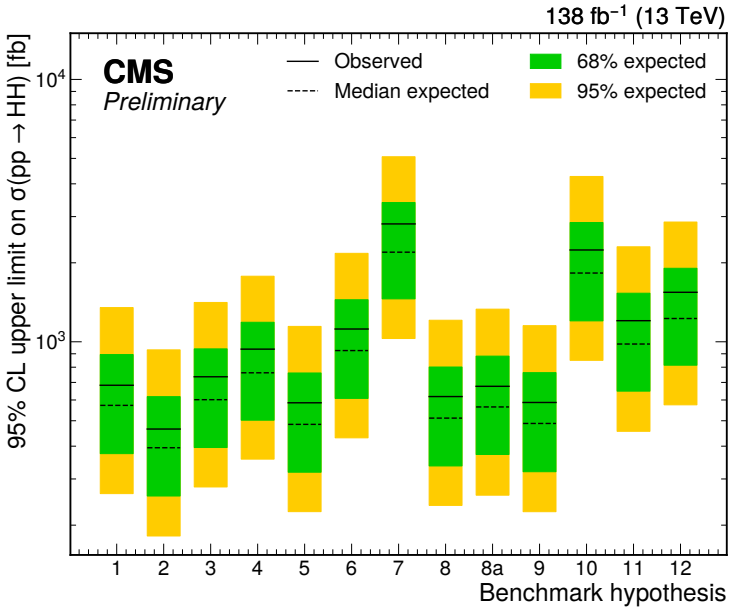

Figure 10: Expected and observed upper limits on the nonresonant HH production cross section at the 95% CL, for different thirteen BSM benchmark scenarios which consider different values of the couplings, κ_λ, κ_t, c_{2g}, c_g and c_2 (defined in Table 1). The green and yellow bands represent the one and two standard deviations for the expected limits, respectively.

10.2 Results from the resonant search channels

Figure 11 shows the upper limits on the resonant production cross section, $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{HH})$, as a function of the resonant mass m_X, for the $\mathrm{X}^{(0)} \to \mathrm{HH}$ (left) and $\mathrm{X}^{(2)} \to \mathrm{HH}$ (right) searches. The expected limits are shown to decrease as a function of m_X, which follows from the improved signal *vs.* background discrimination for high m_X values. The observed (expected) upper limits vary between 140 - 2200 fb (230 - 1700 fb) and 170 - 1800 fb (200 - 1700 fb) for the $\mathrm{X}^{(0)} \to \mathrm{HH}$ and $\mathrm{X}^{(2)} \to \mathrm{HH}$ searches, respectively. In the context of the RS bulk model, these results exclude m_X values up to 900 GeV for the spin-0 radion resonance at $\Lambda_R = 2\,\mathrm{TeV}$, and between 310 and 700 GeV for the spin-2 KK graviton resonance with coupling factor $\kappa/\overline{M_{pl}} = 1$.

The upper limits on $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{YH} \to \gamma\gamma\tau\tau)$ for the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$ search are shown in Fig. 12 and upper limits on $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{YH})B(\mathrm{Y} \to \gamma\gamma)$ for the $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ searches are shown in Figs. 13 and 14. In the top plots, the limits are shown as a function of m_Y for different m_X hypotheses (left) or as a function of m_X for different m_Y hypotheses (right). In each $\mathrm{X} \to \mathrm{YH}$ search channel, the expected limits are found to have a weak dependence on m_Y for all m_X values. The bottom plots show the observed upper limits in the 2D (m_X,m_Y) planes.

For the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$ search channel, the observed (expected) upper limits vary between 0.054 - 1.2 fb (0.89 - 0.66 fb), depending on the m_X and m_Y values. In the low-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ search, the observed (expected) upper limits vary between 0.53 - 15 fb (0.65 - 8.0 fb), depending on the m_X and m_Y values. These limits can be compared to a set of maximally allowed cross sections in the NMSSM taken from Ref. [83] which are calculated using the NMSSMTOOLS v5.6.2 package [84–86]. We find a region in (m_X,m_Y) parameter space where the observed limits are below the maximally allowed, implying that these results can be used to provide tighter constraints on the NMSSM. This region is shown by the red dashed area in Fig. 13. For the high-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ search, the observed (expected) upper limits vary between 0.51 - 9.7 fb (0.60 - 7.3 fb), depending on the m_X and m_Y values.

In the $\mathrm{X} \to \mathrm{HH}$ searches, no excesses with local significances above 1.7 standard deviations are found. In the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$ search channel, the highest local significance seen is 2.6 standard deviations at (m_X, m_Y) = (320, 60) GeV. In the low-mass and high-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ searches, the highest local significances found are 3.4 standard deviations at (m_X, m_Y) = (525, 115) GeV, and 3.2 standard deviations at (m_X, m_Y) = (462, 161) GeV respectively. However, given the large number of mass hypotheses considered in these searches - 401, 2661, and 1852 for the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$, low-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$, and high-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ searches respectively - the look-elsewhere effect is significant and the corresponding global significances are 2.2, 0.1 and 0.3 standard deviations.

As mentioned in Section 1, there are a number of recent excesses reported by CMS consistent with resonances of masses 650 GeV and 90 - 100 GeV. No significant excess is seen in the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$ search at these masses. However, in the $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ search, local significances of 2.6 σ and 2.3 σ are found for $m_\mathrm{Y} = 95\,\mathrm{GeV}$ and $m_\mathrm{X} = 600\,\mathrm{GeV}$ and $m_\mathrm{X} = 650\,\mathrm{GeV}$ respectively. Furthermore, these excesses are localized to these mass points as can be seen in the top plots of Fig. 13.

11 Summary

A search for the production of two scalar bosons in the $\gamma\gamma\tau\tau$ final state is presented. The search uses data from proton-proton collisions collected by the CMS experiment at the LHC in 2016–2018 at a centre-of-mass energy of 13 TeV, corresponding to $138\,\mathrm{fb}^{-1}$ of integrated luminosity.

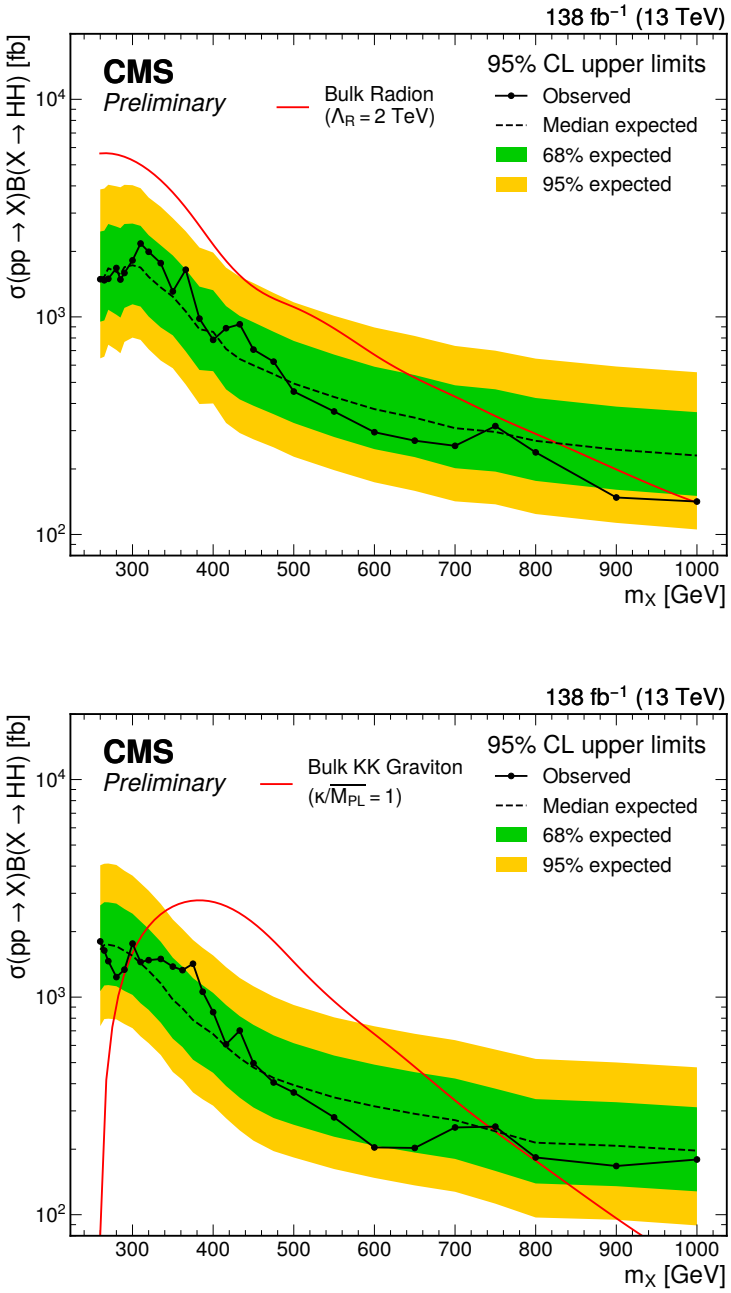

Figure 11: Expected and observed 95% CL upper limit on the resonant production cross section, $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{HH})$ for the spin-0 $\mathrm{X}^{(0)} \to \mathrm{HH}$ search (top) and spin-2 $\mathrm{X}^{(2)} \to \mathrm{HH}$ search (bottom). The dashed and solid black lines represent the expected and observed limits, respectively. The green and yellow bands represent the one and two standard deviations for the expected limit, respectively. The red lines show the theoretical predictions with different energy scales and couplings [82].

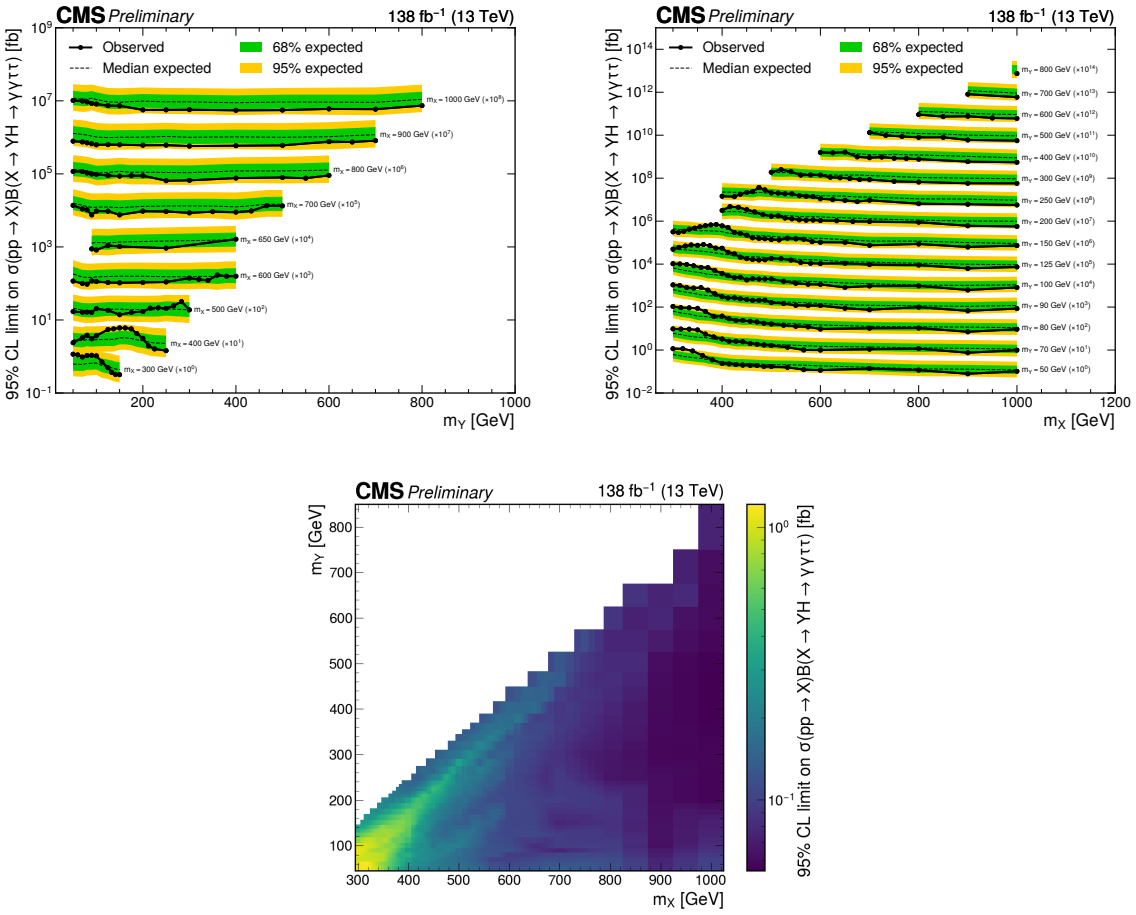

Figure 12: The top plot shows the expected and observed 95% CL upper limit on $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{YH} \to \gamma\gamma\tau\tau)$ for the $\mathrm{X} \to \mathrm{Y}(\tau\tau)\mathrm{H}(\gamma\gamma)$ search. The dashed and solid black lines represent the expected and observed limits, respectively. The green and yellow bands represent the one and two standard deviations for the expected limit, respectively. Limits are scaled by orders of 10, labeled in the plot, depending on m_X (left) or m_Y (right). The bottom plot shows the observed upper limits in the 2D (m_X,m_Y) plane. The values of the limits are shown by the colour scale.

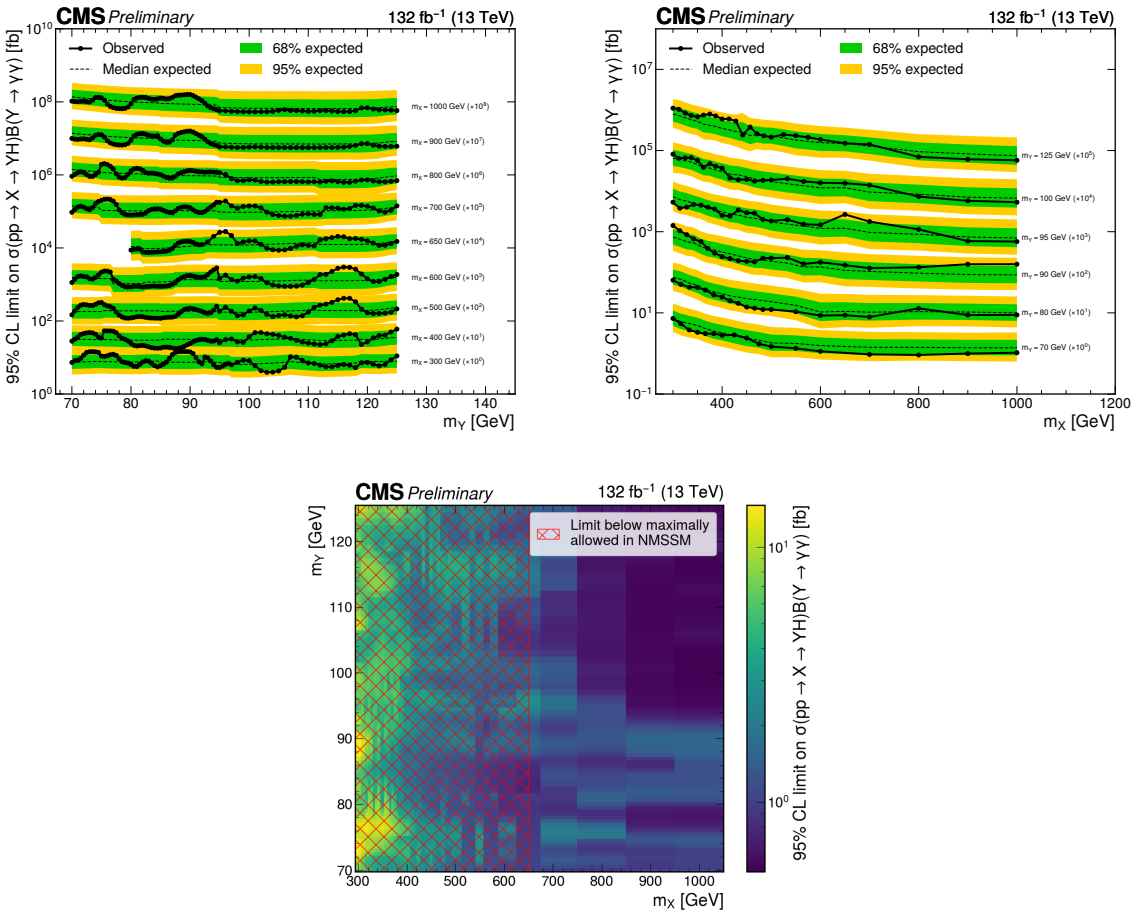

Figure 13: The top plot shows the expected and observed 95% CL upper limit on $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{YH})B(\mathrm{Y} \to \gamma\gamma)$ for the low-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ search. The dashed and solid black lines represent the expected and observed limits, respectively. The green and yellow bands represent the one and two standard deviations for the expected limit, respectively. Limits are scaled by orders of 10, labeled in the plot, depending on m_{X} (left) or m_{Y} (right). The bottom plot shows the observed upper limits in the 2D (m_{X},m_{Y}) plane. The values of the limits are shown by the colour scale. The red-hatched region indicates masses for which the observed limits are below the maximally allowed limits in the NMSSM taken from Ref. [83].

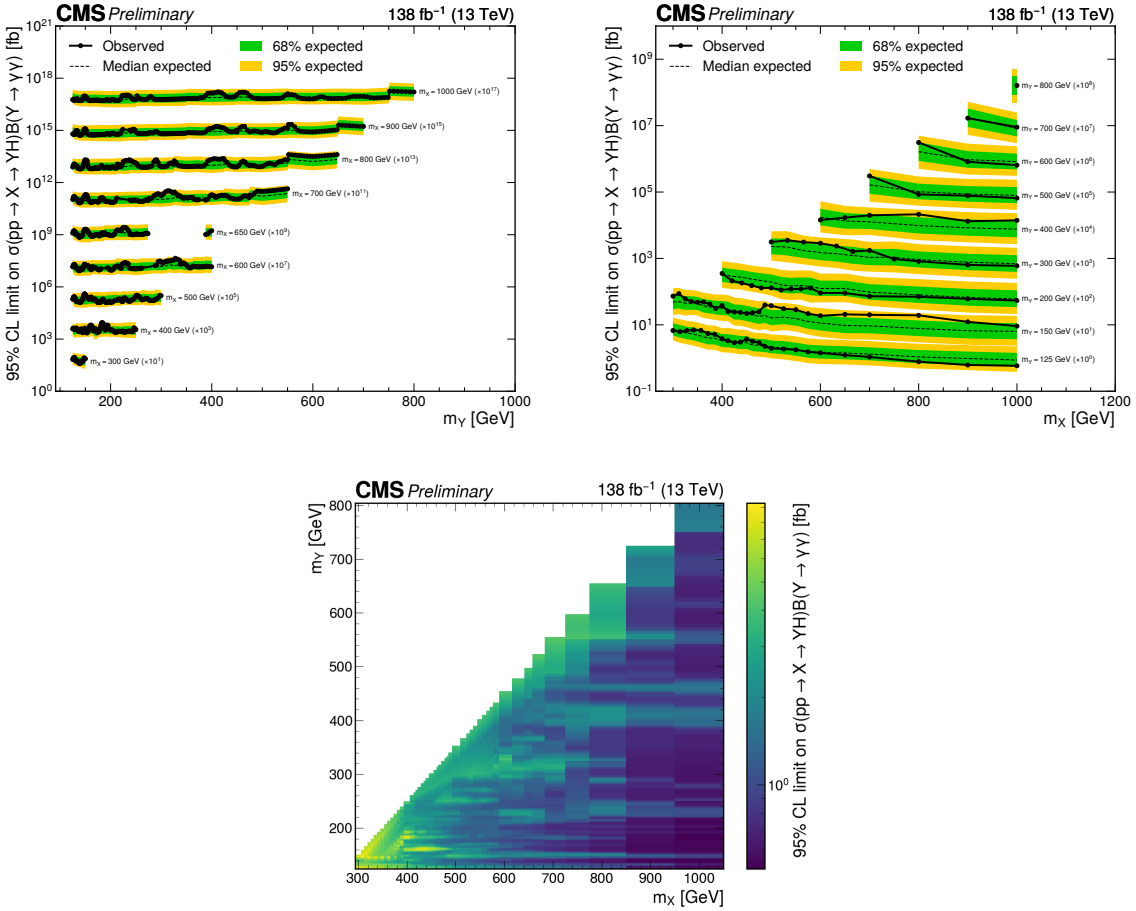

Figure 14: The top plot shows the expected and observed 95% CL upper limit on $\sigma(\mathrm{pp} \to \mathrm{X} \to \mathrm{YH})B(\mathrm{Y} \to \gamma\gamma)$ for the high-mass $\mathrm{X} \to \mathrm{Y}(\gamma\gamma)\mathrm{H}(\tau\tau)$ search. The dashed and solid black lines represent the expected and observed limits, respectively. The green and yellow bands represent the one and two standard deviations for the expected limit, respectively. Limits are scaled by orders of 100 (left) or 10 (right), labeled in the plot, depending on m_X (left) or m_Y (right). In the top-left plot, there is a discontinuity in the limits for $m_\mathrm{X} = 650\,\mathrm{GeV}$ which is due to the chosen limit granularity in m_X and m_Y. For the range $m_\mathrm{Y} \in [274, 388]\,\mathrm{GeV}$, there are limits placed at $m_\mathrm{X} = 600$, 633, 666 and 700 GeV instead of 600, 650 and 700 GeV. The bottom plot shows the observed upper limits in the 2D (m_X,m_Y) plane. The values of the limits are shown by the colour scale.

In total, five search channels are considered. One channel targets the nonresonant HH production via gluon-fusion, where no significant deviation from the background-only hypothesis is observed. Upper limits at the 95% CL on the HH production cross section are extracted for production in the SM and in several BSM scenarios. The observed upper limit for the SM production is found to be 930 fb, corresponding to 33 times the SM prediction, whilst the expected upper limit is 740 fb, corresponding to 26 times the SM prediction. The limit is also derived as a function of the Higgs boson self-coupling modifier, κ_λ, assuming all other Higgs boson couplings are as predicted in the SM. The self-coupling modifier, κ_λ, is constrained within the range $-13 < \kappa_\lambda < 18$ at the 95% CL. In addition, the limit is extracted for numerous BSM benchmark scenarios. The results are consistent with the SM predictions.

This analysis also targets the resonant production of two scalar bosons. Two channels are constructed to search for a resonance X decaying to a SM Higgs boson pair, $X \to HH$, for both the spin-0 resonance and spin-2 resonance scenarios. No significant deviation from the background-only hypothesis is observed. Furthermore, the analysis targets the $X \to YH$ process, where Y is an additional, lighter (than X) scalar particle. Three search channels are constructed, namely the $X \to Y(\tau\tau)H(\gamma\gamma)$, low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ and high-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ channels, which target the different decay chains in different mass regimes. The largest local significance across these channels is 3.4 standard deviations, which once considering the look-elsewhere effect, has a global significance of 0.1 standard deviations. Therefore, this analysis, as a standalone result, does not present any significant deviation from the standard model. However, when put into the context of recent excesses in CMS at resonance masses of 650 GeV and 95 GeV, the local significance of 2.3 at $m_X = 650$ GeV and $m_Y = 95$ GeV in the low-mass $X \to Y(\gamma\gamma)H(\tau\tau)$ analysis is interesting and warrants further measurements.

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] 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.

[5] S. Dawson, S. Dittmaier, and M. Spira, "Neutral Higgs boson pair production at hadron colliders: QCD corrections", *Phys. Rev. D* **58** (1998) 115012, doi:10.1103/PhysRevD.58.115012, arXiv:hep-ph/9805244.

[6] S. Borowka et al., "Higgs boson pair production in gluon fusion at next-to-leading order with full top-quark mass dependence", *Phys. Rev. Lett.* **117** (2016) 012001, doi:10.1103/PhysRevLett.117.079901, arXiv:1604.06447. [Erratum: doi:10.1103/PhysRevLett.117.079901].

[7] J. Baglio et al., "Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme", *Eur. Phys. J. C* **79** (2019) 459, doi:10.1140/epjc/s10052-019-6973-3, arXiv:1811.05692.

[8] D. de Florian and J. Mazzitelli, "Higgs boson pair production at next-to-next-to-leading order in QCD", *Phys. Rev. Lett.* **111** (2013) 201801, doi:10.1103/PhysRevLett.111.201801, arXiv:1309.6594.

[9] D. Y. Shao, C. S. Li, H. T. Li, and J. Wang, "Threshold resummation effects in Higgs boson pair production at the LHC", *JHEP* **07** (2013) 169, doi:10.1007/JHEP07(2013)169, arXiv:1301.1245.

[10] D. de Florian and J. Mazzitelli, "Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC", *JHEP* **09** (2015) 053, doi:10.1007/JHEP09(2015)053, arXiv:1505.07122.

[11] J. Baglio et al., "$gg \to hh$: Combined uncertainties", *Phys. Rev. D* **103** (2021) 056002, doi:10.1103/PhysRevD.103.056002, arXiv:2008.11626.

[12] CMS Collaboration, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery", *Nature* **607** (2022) 60, doi:10.1038/s41586-022-04892-x, arXiv:2207.00043.

[13] ATLAS Collaboration, "Constraints on the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at s=13 TeV", *Phys. Lett. B* **843** (2023) 137745, doi:10.1016/j.physletb.2023.137745, arXiv:2211.01216.

[14] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[15] F. Goertz, A. Papaefstathiou, L. L. Yang, and J. Zurita, "Higgs boson pair production in the D=6 extension of the SM", *JHEP* **04** (2015) 167, doi:10.1007/JHEP04(2015)167, arXiv:1410.3471.

[16] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370, doi:10.1103/PhysRevLett.83.3370, arXiv:hep-ph/9905221.

[17] W. D. Goldberger and M. B. Wise, "Modulus stabilization with bulk fields", *Phys. Rev. Lett.* **83** (1999) 4922, doi:10.1103/PhysRevLett.83.4922, arXiv:hep-ph/9907447.

[18] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall-Sundrum gauge hierarchy model", *Phys. Rev. Lett.* **84** (2000) 2080, doi:10.1103/PhysRevLett.84.2080, arXiv:hep-ph/9909255.

[19] A. Djouadi, "The anatomy of electro-weak symmetry breaking. II. the Higgs bosons in the minimal supersymmetric model", *Phys. Rept.* **459** (2008) 1, doi:10.1016/j.physrep.2007.10.005, arXiv:hep-ph/0503173.

[20] CMS Collaboration, "Search for a new resonance decaying into two spin-0 bosons in a final state with two photons and two bottom quarks in proton-proton collisions at $\sqrt{s}$ = 13 TeV", CMS Physics Analysis Summary CMS-HIG-21-011, CERN-EP-2023-132, 2023.

[21] CMS Collaboration, "Search for high mass resonances decaying into W^+W^- in the dileptonic final state with 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-20-016, 2022.

[22] CMS Collaboration, "Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **07** (2023) 073, doi:10.1007/JHEP07(2023)073, arXiv:2208.02717.

[23] CMS Collaboration, "Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-20-002, 2023.

[24] 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.

[25] 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.

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

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

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

[29] 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.

[30] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[31] CMS Collaboration, "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at sqrt(s) = 8 TeV", *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.

[32] 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.

[33] CMS Collaboration, "ECAL 2016 refined calibration and Run2 summary plots", CMS Detector Performance Summary CMS-DP-2020-021, 2020.

[34] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[35] 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.

[36] 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.

[37] 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.

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

[39] CMS Collaboration, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P10005, doi:10.1088/1748-0221/13/10/P10005, arXiv:1809.02816.

[40] CMS Collaboration, "Identification of hadronic tau lepton decays using a deep neural network", *JINST* **17** (2022) P07023, doi:10.1088/1748-0221/17/07/P07023, arXiv:2201.08458.

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

[42] P. Baldi et al., "Parameterized neural networks for high-energy physics", *Eur. Phys. J. C* **76** (2016) 235, doi:10.1140/epjc/s10052-016-4099-4, arXiv:1601.07913.

[43] S. Choi and H. Oh, "Improved extrapolation methods of data-driven background estimations in high energy physics", *Eur. Phys. J. C* **81** (2021) 643, doi:10.1140/epjc/s10052-021-09404-1, arXiv:1906.10831.

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

[45] 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.

[46] 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.

[47] 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.

[48] G. Heinrich et al., "NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers", *JHEP* **08** (2017) 088, doi:10.1007/JHEP08(2017)088, arXiv:1703.09252.

[49] 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.

[50] S. Jones and S. Kuttimalai, "Parton shower and NLO-matching uncertainties in Higgs boson pair production", *JHEP* **02** (2018) 176, doi:10.1007/JHEP02(2018)176, arXiv:1711.03319.

[51] G. Heinrich, S. P. Jones, M. Kerner, and L. Scyboz, "A non-linear EFT description of $gg \to hh$ at NLO interfaced to POWHEG", *JHEP* **10** (2020) 021, doi:10.1007/JHEP10(2020)021, arXiv:2006.16877.

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] E. Bothmann et al., "Event generation with SHERPA 2.2", *SciPost Phys.* **7** (2019) 34, doi:10.21468/SciPostPhys.7.3.034, arXiv:1905.09127.

[57] 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.

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

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

[60] 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.

[61] 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.

[62] 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.

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

[64] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

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

[66] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9 fb^{-1} of data from proton-proton collisions at 13 TeV with Phase-1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, 2018.

[67] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken, "Reconstruction of the Higgs mass in $h \to \tau\tau$ events by dynamical likelihood techniques", *J. Phys. Conf. Ser.* **513** (2014) 022035, doi:10.1088/1742-6596/513/2/022035.

[68] 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.

[69] R. A. Fisher, "On the interpretation of χ^2 from contingency tables, and the calculation of p", *Journal of the Royal Statistical Society* **85** (1922) 87, doi:10.2307/2340521.

[70] CMS Collaboration, "Search for narrow resonances in the b-tagged dijet mass spectrum in proton-proton collisions at $\sqrt{s} = 13$TeV", *Phys. Rev. D* **108** (2023) 012009, doi:10.1103/PhysRevD.108.012009, arXiv:2205.01835.

[71] CMS Collaboration, "Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{\text{s}}$ = 13 TeV", *JHEP* **07** (2021) 027, doi:10.1007/JHEP07(2021)027, arXiv:2103.06956.

[72] 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.

[73] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs cross sections: 3. Higgs properties", CERN Report CERN-2013-004, 2013. doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[74] 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.

[75] 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.

[76] 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.

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

[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] 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.

[81] CMS Collaboration, "Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2021) 257, doi:10.1007/JHEP03(2021)257, arXiv:2011.12373.

[82] A. Carvalho, "Gravity particles from warped extra dimensions, predictions for LHC", technical report, 2014. doi:10.48550/arXiv.1404.0102, arXiv:1404.0102.

[83] U. Ellwanger and C. Hugonie, "Benchmark planes for Higgs-to-Higgs decays in the NMSSM", *Eur. Phys. J. C* **82** (2022) 406, doi:10.1140/epjc/s10052-022-10364-3, arXiv:2203.05049.

[84] U. Ellwanger, J. F. Gunion, and C. Hugonie, "NMHDECAY: A Fortran code for the Higgs masses, couplings and decay widths in the NMSSM", *JHEP* **02** (2005) 066, doi:10.1088/1126-6708/2005/02/066, arXiv:hep-ph/0406215.

[85] U. Ellwanger and C. Hugonie, "NMHDECAY 2.1: An updated program for sparticle masses, Higgs masses, couplings and decay widths in the NMSSM", *Computer Physics Communications* **175** (2006) 290, doi:10.1016/j.cpc.2006.04.004.

[86] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, "micrOMEGAs_3: A program for calculating dark matter observables", *Comput. Phys. Commun.* **185** (2014) 960, doi:10.1016/j.cpc.2013.10.016, arXiv:1305.0237.