

---

# CMS Physics Analysis Summary

---

Contact: cms-pag-conveners-higgs@cern.ch

2019/05/20

Search for 2HDM neutral Higgs bosons through the  
 $H \rightarrow ZA \rightarrow \ell^+ \ell^- b\bar{b}$  process in proton-proton collisions at  
 $\sqrt{s} = 13 \text{ TeV}$

The CMS Collaboration

## Abstract

This note reports on a search for an extended scalar sector of the standard model, where a new CP-even (odd) scalar decays to a Z boson and a lighter CP-odd (even) scalar, which further decays to  $b\bar{b}$ . The Z boson is reconstructed via its decays to leptons. The analysed data were recorded in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ , collected by the CMS experiment at the LHC, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . Data and predictions from the standard model are in agreement within uncertainties. Upper limits at 95% confidence level are set on the production cross section times branching ratio, with masses of the resonances ranging up to 1000 GeV. The results are interpreted in the context of the Two-Higgs Doublet Model.



## 1 Introduction

The CMS and ATLAS experimental programmes are focusing efforts on the measurement of the properties of the newly discovered Higgs boson [1–3], which has a mass 125 GeV. Although all measurements, to date, are consistent with the standard model (SM) within the experimental uncertainties, additional Higgs bosons are predicted in several extensions of the SM. Examples of these extensions are the minimal supersymmetric extension of the SM (MSSM [4]) and the Two-Higgs-doublet model (2HDM) [5], whose phenomenology is based on the presence of an additional scalar Higgs doublet. The generic 2HDM potential has a large number of degrees of freedom. However, under the assumption that the electromagnetic gauge symmetry is preserved and that the bosonic sector of the theory is CP-invariant, by choosing the custodial phase and by suppressing tree-level flavor-changing neutral currents, the number of free parameters is reduced to six. If, in addition, the twisted custodial symmetry [5] is considered, then the free parameters for the theory are the mass of the Higgs boson,  $m_h$ ; the mass of the pseudoscalar  $A$ ,  $m_A$ ; the soft breaking term  $m_{12}^2$ ; the ratio of the vacuum expectation values,  $\tan\beta$ ; and the mixing angle  $\alpha$  between the two CP-even eigenstates. In the case of a non-twisted custodial symmetry scenario, characteristic of the MSSM scalar sector, the  $A$  and  $H$  bosons are exchanged.

Different models and assumptions alter the mass hierarchies and branching fractions, as shown in Fig. 1. In the region where  $\cos(\beta - \alpha)$  approaches 0, the light CP-even  $h$  boson has properties indistinguishable from a SM Higgs boson with the same mass. In this region, known as the alignment limit, the branching fraction of the heavier neutral boson into  $Z A$  is among the largest. The branching fractions of the  $H$  and  $A$  bosons for  $\cos(\beta - \alpha) = 0.01$ ,  $m_H = 300$  GeV and  $m_A = 200$  GeV are shown in Fig. 2, along with the cross section of the  $pp \rightarrow H \rightarrow ZA \rightarrow \ell\ell b\bar{b}$  process.

This note reports on a search for a new CP-even neutral Higgs boson,  $H$ , decaying into  $Z$  and a lighter CP-odd Higgs boson,  $A$ , where the  $Z$  decays into  $\ell^+\ell^-$ , and the  $A$  into  $b\bar{b}$ . The search is based on LHC proton-proton collision data at  $\sqrt{s} = 13$  TeV collected by the CMS experiment, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The analysis focuses on the invariant mass distributions of the  $\ell\ell b\bar{b}$  and  $b\bar{b}$  systems, searching for a resonant-like excess compatible with the  $H$  and  $A$  masses.

Searches for  $H \rightarrow ZA$  in the same final state have been performed at 8 TeV [6] and 13 TeV [7] by the ATLAS experiment, and at 8 TeV [8] by the CMS experiment. The search for  $H \rightarrow Zh$ , where  $h$  is the SM Higgs boson, has been also performed by the CMS experiment [9, 10].

## 2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionisation chambers embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [12].

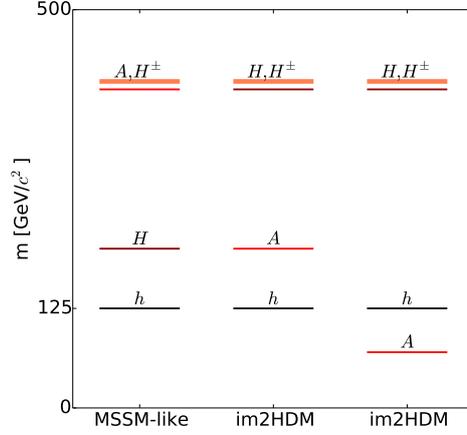


Figure 1: 2HDM mass hierarchies considered in the study: MSSM-like, where a heavy  $A$  is degenerate in mass with the charged scalars, and im2HDM [11], where the hierarchy is inverted making  $H$  the heaviest neutral scalar and degenerate in mass with charged scalars.

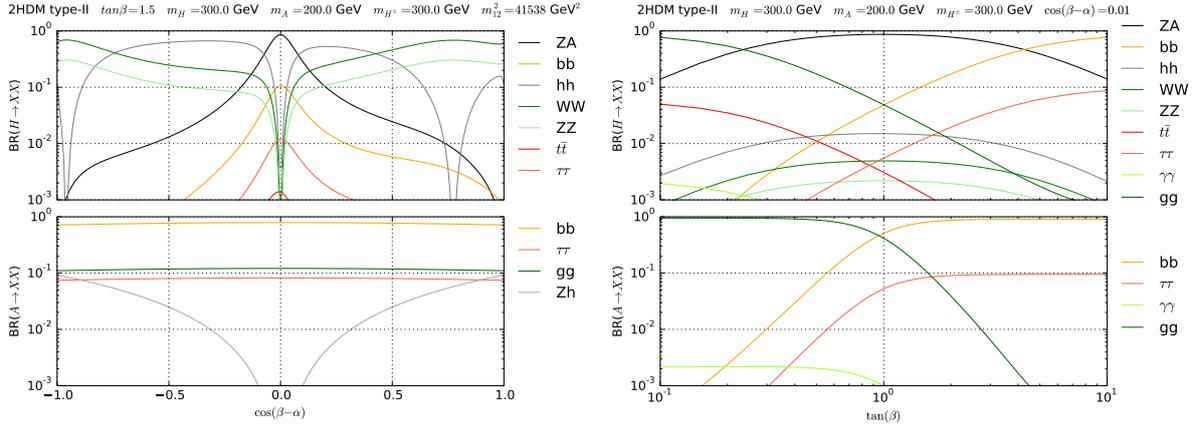


Figure 2:  $H$  and  $A$  branching fractions as a function of  $\cos(\beta - \alpha)$  for the following set of parameters:  $\tan \beta = 1.5$ ,  $m_H = 300 \text{ GeV}$ ,  $m_A = 200 \text{ GeV}$  (right).  $H$  and  $A$  branching fractions as a function of  $\tan \beta$  for the following set of parameters:  $\cos(\beta - \alpha) = 0.01$ ,  $m_H = 300 \text{ GeV}$  and  $m_A = 200 \text{ GeV}$  (left).

### 3 Event simulation and background predictions

The main background processes, in decreasing order of importance, are Drell–Yan (DY) and fully leptonic  $t\bar{t}$ . Single top quark, diboson, triboson,  $t\bar{t}V$ ,  $W$ +jets and SM single Higgs boson production are also considered. QCD multijet events with jets misidentified as leptons are negligible due to the dilepton selection described in Section 4. The dominant contribution arises from DY + two  $b$  quark jets production yielding the same final state as the signal process (two  $b$  quark jets and two leptons) when the  $Z$  boson decays leptonically.

Background simulation samples have been generated at next-to-leading order (NLO) in QCD using POWHEG 2 [13–17], and MADGRAPH5\_aMC@NLO versions 2.2.2.0 and 2.3.2.2 [18] with FxFx [19] for NLO jet merging and MADSPIN [20] to properly propagate spin information in

the matrix element of the process.

We produce 207 signal simulation samples of different mass hypotheses, with  $m_A$  and  $m_H$  ranging from 30 to 1000 GeV and from  $m_Z$  to 1000 GeV respectively. The choice of the mass hypotheses is strongly motivated by the need of achieving a complete coverage of the parameter space. The spacing between two adjacent mass hypotheses is chosen so as to take into account the worsening of the signal resolution as the mass increases. These samples are produced using MADGRAPH5\_aMC@NLO version 2.3.2.2 [18] interfaced with the parton shower generator PYTHIA 8.212 [21]. The parton distribution function (PDF) set is NNPDF 3.0 [22] at leading order (LO) in the four-flavour scheme, and the factorisation and renormalisation scales are estimated dynamically. The shower tune is CUEPT8M1 [23], derived from the MONASH tune [24].

For all processes, the detector response is simulated using a detailed description of the CMS apparatus, based on the GEANT4 package [25]. Additional pp interactions in the same and in the neighbouring bunch crossings (pileup) are generated with PYTHIA 8.212 [21], and overlapped with the simulated events of interest in order to reproduce the pileup measured in data.

All background processes are initially normalised to their most accurate theoretical cross sections. The  $t\bar{t}$ , DY, single top quark, and  $W^+W^-$  samples are normalised to next-to-next-to-leading order (NNLO) precision in QCD corrections [26–29], while the remaining diboson, triboson and  $t\bar{t}V$  processes are normalised to NLO precision in QCD corrections [18, 30]. The single Higgs boson production cross section is computed at NNLO precision in QCD corrections and NLO precision in electroweak corrections [31]. We indicate the single Higgs boson and the  $t\bar{t}V$  backgrounds with *Others* in the figures.

A control region consisting of events passing the selection as described in Section 4 is built, but with the requirement of containing two different flavour leptons. These events mostly arise from the  $t\bar{t}$  contribution, and this additional region is used in the final template fit to get a data-driven estimate of this background.

All the other backgrounds are fully estimated from simulation.

## 4 Event reconstruction and selection

Recorded events are collected with a set of dilepton triggers. The  $p_T$  thresholds applied by the triggers are asymmetric and channel-dependent, and vary from 17 to 23 GeV for the leading lepton and from 8 to 12 GeV for the subleading lepton. Trigger efficiencies are measured with a “tag-and-probe” technique [32] as a function of lepton  $p_T$  and  $\eta$  in a data control region consisting of  $Z \rightarrow \ell\ell$  events. Events with two oppositely charged leptons ( $e^+e^-$ ,  $\mu^+\mu^-$ ) are selected using asymmetric  $p_T$  requirements, chosen to be above the corresponding trigger thresholds, for leading and subleading leptons of 25 GeV and 15 GeV, respectively, for  $ee$  events; and of 20 GeV and 10 GeV for  $\mu\mu$  events. Electrons in the pseudorapidity range  $|\eta| < 2.5$  and muons in the range  $|\eta| < 2.4$  are considered. Events with different flavour leptons ( $e^\pm\mu^\mp$ ) are selected in order to constraint the top quark pair contribution in the analysis. The  $p_T$  requirements for leading and subleading leptons are 25 GeV and 15 GeV, respectively, for  $\mu e$  events; and 25 GeV and 10 GeV for  $e\mu$  events.

Electrons, reconstructed by associating tracks with ECAL clusters, are identified by a sequential selection using information on the cluster shape in the ECAL, track quality, and the matching between the track and the ECAL cluster. Additionally, electrons from photon conversions are

rejected [33]. Muons are reconstructed from tracks found in the muon system, associated with tracks in the silicon tracking detectors. They are identified based on the quality of the track fit and the number of associated hits in the different tracking detectors [34]. The lepton isolation, defined as the scalar  $p_T$  sum of all particle candidates in a cone around the lepton, excluding the lepton itself, divided by the lepton  $p_T$ , is required to be  $<0.06$  for electrons (with a cone of radius  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$ ) and  $<0.15$  for muons (with a cone of radius  $\Delta R = 0.4$ ). Lepton identification and isolation efficiencies in the simulation are corrected for residual differences with respect to data. These corrections are measured in a data sample, enriched in  $Z \rightarrow \ell\ell$  events, using a “tag-and-probe” method and are parameterised as a function of lepton  $p_T$  and  $\eta$ .

Jets are reconstructed using a particle flow (PF) technique [35]. PF candidates are clustered to form jets using the anti- $k_T$  clustering algorithm [36] with a distance parameter of 0.4, implemented in the FASTJET package [37]. Jet energies are corrected for residual nonuniformity and nonlinearity of the detector response [38]. Jets are required to have  $p_T > 20$  GeV,  $|\eta| < 2.4$ , and be separated from identified leptons by a distance of  $\Delta R > 0.3$ . The missing transverse momentum vector, defined as the projection onto the transverse plane relative to the beam axis, of the negative vector sum of the momenta of all PF candidates, is referred to as  $\vec{p}_T^{\text{miss}}$  [39, 40]. Its magnitude is denoted by  $p_T^{\text{miss}}$ . Corrections to the jet energies are propagated to  $\vec{p}_T^{\text{miss}}$ .

The reconstructed vertex with the largest value of summed physics-object  $p_T^2$  is taken to be the primary pp interaction vertex, considering the objects returned by a clustering algorithm applied to all charged tracks associated with the vertex, plus the corresponding associated  $\vec{p}_T^{\text{miss}}$ .

The DeepCSV algorithm [41, 42] is used to identify jets originating from b quarks. Jets are considered as b-tagged if they pass the medium working point of the algorithm, which provides around 70% efficiency with a mistag rate less than 1%. Correction factors are applied in the simulation to the selected jets to account for the different response of the DeepCSV algorithm between data and simulation [42]. Among all possible dijet combinations fulfilling the previous criteria, we select the two jets with the highest DeepCSV algorithm outputs.

After the object selection, a requirement of  $m_{\ell\ell} > 12$  GeV is applied to suppress quarkonia resonances and jets misidentified as leptons. The final object selection consists of two opposite-sign leptons and two b-tagged jets, after which a requirement of  $70 < m_{\ell\ell} < 110$  GeV is applied to enhance the presence of  $Z \rightarrow \ell\ell$  events. In addition, the events are required to have a  $p_T^{\text{miss}} < 80$  GeV in order to reduce the background contributions from processes with real  $p_T^{\text{miss}}$ , such as  $t\bar{t}$  production. Both requirements have negligible impacts on the signal efficiency.

The SM background processes contributing to this search after final selection are, in order of importance: DY production,  $t\bar{t}$  production, and other minor backgrounds such as single top quark and diboson production.

## 5 Signal extraction

We search for the process  $H \rightarrow ZA \rightarrow \ell\ell b\bar{b}$  via fully reconstructing its final state particles and applying selection cuts in order to remove as much background contribution as possible, as explained in Section 4. From the reconstructed objects, one is able to measure the mass resonances that we search for. Specifically, from the bjet pair one is able to reconstruct the invariant mass of the A boson; and from the bjet pair and the lepton pair, the invariant mass distribution of the H boson is obtained. Two categories are defined based on the lepton flavours

considered:  $e^+e^-$  and  $\mu^+\mu^-$ . The Z boson mass, reconstructed from two opposite-sign leptons, is used in the selection criteria described in Section 4 since it is common to any signal. The masses of the other two particles, H and A, vary according to the signal scenarios considered. Therefore, a simple and effective approach to isolate the signal is to search for an excess in the reconstructed  $m_{jj}$  and  $m_{ljj}$  distributions centered around the H and A candidate mass for each signal hypothesis.

The  $m_{jj}$  and  $m_{ljj}$  distributions are positively correlated for the signal hypothesis, and the correlation depends on the energy resolution of the b jets, which is measured less precisely than that of the leptons. Therefore, we choose an elliptical signal region, specific to each signal hypothesis, in order to optimize the sensitivity of the search. Figure 3 (left) shows the reconstructed mass distributions for three different signals in the parameter space along with its defined elliptical signal region. Because the shape of the signal is driven by the energy resolution of the final-state objects, ellipses take different sizes and tilt angles, depending on the masses being considered. A parametrisation is therefore performed such to guarantee a good description of the signal shape for each signal hypothesis, and for each ellipse it provides the center, the major and minor semi-axis, and the tilt angle. Since each ellipse must be well-centered around the maximum of the two-dimensional (2D) mass distribution, the reconstructed center is extracted from a one-dimensional (1D) Gaussian fit in both  $m_{ljj}$  and  $m_{jj}$ . The diagonalisation of the covariance matrix of the 2D distribution provides the axes of the ellipse and its tilt angle.

Elliptically shaped bins containing a fraction of the signal events that correspond to the standard deviations (s.d.) in a 2D Gaussian distribution are defined using a parameter called  $\rho$ . Since the shape of the signal is not exactly Gaussian, the size of the bins are varied such that ellipses with  $\rho = 1$  (2, 3, ...) contain 1 s.d. (2 s.d, 3 s.d., ...) of the signal events. Hence, the bulk of the signal is located at small values of  $\rho$ . Figure 3 (right) shows the variation of the ellipse size as a function of  $\rho$  in steps of 0.5, from 0 to 3. The  $m_{jj}$  and  $m_{ljj}$  distributions for  $e^+e^-$  and  $\mu^+\mu^-$  events are shown in Fig. 4.

Selected events in the  $m_{ljj}$ - $m_{jj}$  plane are classified in six regions around the center of the ellipse defined for each signal point. The regions are built in  $\rho$  steps as illustrated in Fig. 3 (right), and lead to a template containing six bins used to perform the statistical analysis.

The yield in data and the expected yields in background simulation are reported in Table 1 for each elliptical bin under the mass hypothesis  $m_H = 500$  GeV and  $m_A = 400$  GeV. The  $e^+e^-$  and  $\mu^+\mu^-$  categories are summed together.

Table 1: Expected and observed event yields prior to the fit in the signal region with  $m_H = 500$  GeV and  $m_A = 200$  GeV for each elliptical bin. The  $e^+e^-$  and  $\mu^+\mu^-$  categories are summed together.

Process	Yield					
	$0 \leq \rho < 0.5$	$0.5 \leq \rho < 1$	$1 \leq \rho < 1.5$	$1.5 \leq \rho < 2$	$2 \leq \rho < 2.5$	$2.5 \leq \rho < 3$
DY	$181 \pm 14$	$438 \pm 22$	$607 \pm 27$	$987 \pm 34$	$1440 \pm 42$	$2273 \pm 53$
$t\bar{t}$	$166 \pm 2$	$420 \pm 4$	$603 \pm 5$	$826 \pm 5$	$1165 \pm 6$	$1597 \pm 8$
Single top	$2.2 \pm 0.5$	$6.2 \pm 0.8$	$9 \pm 1$	$17 \pm 1$	$25.5 \pm 1.7$	$38 \pm 2$
VV(V)	$0.6 \pm 0.1$	$1.9 \pm 0.2$	$2.5 \pm 0.2$	$3.9 \pm 0.5$	$5.2 \pm 0.4$	$9.1 \pm 0.4$
Others	$0.9 \pm 0.2$	$3.7 \pm 0.3$	$5.1 \pm 0.3$	$8.4 \pm 0.4$	$11.7 \pm 0.5$	$18.1 \pm 0.6$
Total bkg.	$351 \pm 14$	$870 \pm 22$	$1227 \pm 27$	$1842 \pm 34$	$2647 \pm 42$	$3935 \pm 54$
Data	365	854	1231	1834	2608	3906

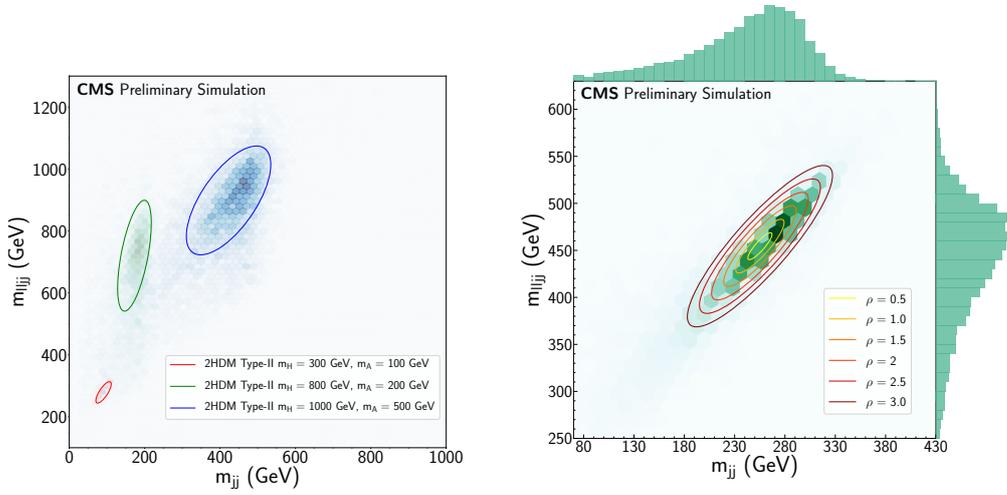


Figure 3: The  $m_{ljj}$ - $m_{jj}$  plane for signal samples under three different mass hypotheses, on which the parametrised ellipse is shown (left). A signal hypothesis with  $m_H = 500$  GeV and  $m_A = 300$  GeV is shown in the  $m_{ljj}$ - $m_{jj}$  plane (right). The different ellipses show the variation of the  $\rho$  parameter in steps of 0.5, from 0 to 3.

## 6 Systematic uncertainties

We investigate sources of systematic uncertainties and their impact on the statistical interpretation of the results by considering both uncertainties in the normalisation of the various processes in the analysis, as well as those affecting the shapes of the distributions.

Theoretical uncertainties in the cross sections of the backgrounds estimated using simulation are considered as systematic uncertainties in the yield predictions. The uncertainty in the total integrated luminosity is determined to be 2.5% [43].

The signal region contains events that have at least two b-tagged jets. One can build a control region by requiring events to pass the selection as described in Section 4 but with no b tag requirement for the jets. A discrepancy of up to 10% is observed across the reconstructed mass range, hinting at a mismodeling of the DY background in some specific regions of the parameter space. To correct for this discrepancy, additional shape uncertainties to this background are applied as follows. In the above-mentioned control region, the discrepancy between data and MC in the distributions of the reconstructed masses is quantified and used to reweigh each DY event. An uncertainty equal to 100% of the observed discrepancy is then applied. By construction, the severity of such mismodeling depends on the region of the parameter space being considered. This implicates that when regions with very different values of this uncertainty fall into the same elliptical bin, they will constrain each other in the maximum likelihood fit thus making the method lose its original purpose of properly accounting for the mismodeling in the final result. In order to mitigate as much as possible this effect, the parameter space is sampled in 42 regions of approximately  $150 \times 150 \text{ GeV}^2$ , and each region is then assigned the corresponding uncertainty. This procedure guarantees that such shape uncertainty is well uncorrelated across the whole parameter space.

The following sources of systematic uncertainties that affect the normalisation and shape of the templates used in the statistical evaluation are considered:

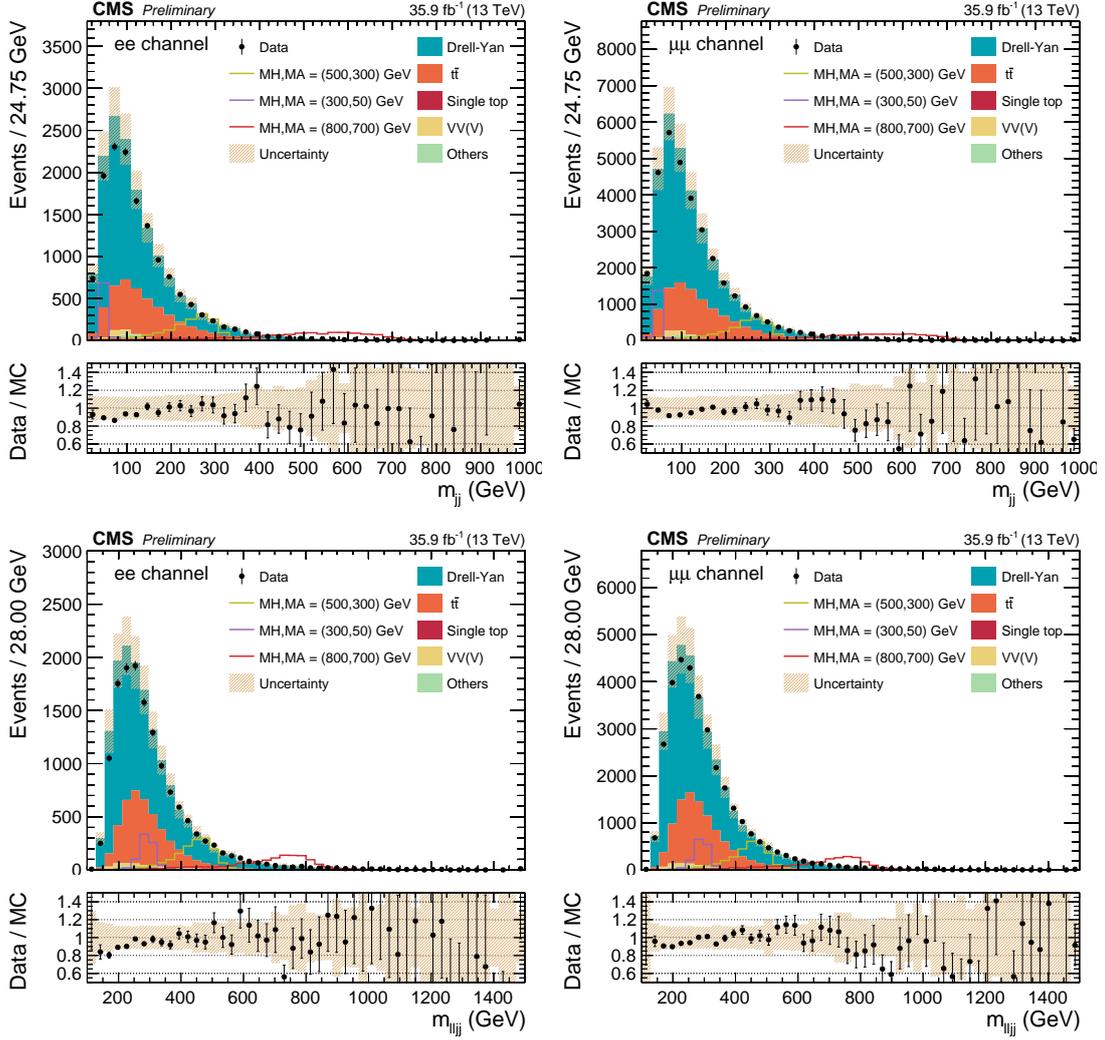


Figure 4: The  $m_{jj}$  and  $m_{ljj}$  distributions in data and simulated background events after requiring all the analysis cuts, for  $e^+e^-$  (left), and  $\mu^+\mu^-$  (right) events. The various signal hypotheses displayed have been scaled to a cross section of 1 pb for display purposes. Error bars indicate statistical uncertainties, while shaded bands show systematic uncertainties prior to the fit.

- **Trigger efficiency, lepton identification and isolation:** uncertainties in the measurement of trigger efficiencies as well as electron and muon isolation and identification efficiencies, are considered as sources of systematic uncertainties. These are evaluated as a function of lepton  $p_T$  and  $\eta$ , and their effect on the analysis is estimated by varying the corrections to the efficiencies by  $\pm 1$  standard deviation.
- **Jet energy scale and resolution:** uncertainties in the jet energy scale are of the order of a few percent and are estimated as a function of jet  $p_T$  and  $\eta$  [38]. A difference in the jet energy resolution of about 10% between data and simulation is accounted for by worsening the jet energy resolution in simulation by  $\eta$ -dependent factors. The uncertainty due to these corrections is estimated by a variation of the factors applied by  $\pm 1$  standard deviation. Variations of jet energies are propagated to  $\vec{p}_T^{\text{miss}}$ .
- **b tagging:** b tagging efficiency and light-flavour mistag rate corrections and associated uncertainties are determined as a function of the jet  $p_T$  [44]. Their effect on the analysis is estimated by varying these corrections by  $\pm 1$  standard deviation.

- **Pileup:** the measured total inelastic cross section is varied by  $\pm 4.6\%$  [45] to produce different expected pileup distributions.
- **Renormalisation and factorisation scale uncertainty:** this uncertainty is estimated by varying the renormalisation ( $\mu_R$ ) and the factorisation ( $\mu_F$ ) scales used during the generation of the simulated samples independently by factors of 0.5, 1, or 2. Unphysical cases, where the two scales are at opposite extremes, are not considered. An envelope is built from the 6 possible combinations by keeping maximum and minimum variations for each bin of the distributions, and is used as an estimate of the scale uncertainties for all the background and signal samples.
- **PDF uncertainty:** the magnitudes of the uncertainties related to the PDFs and the variation of the strong coupling constant for each simulated background and signal process are obtained using variations of the NNPDF 3.0 set [22], following the PDF4LHC prescriptions [28, 46].
- **Drell-Yan shape uncertainty:** 42 additional shape uncertainties are applied to DY events to correct for mismodeling of this background. Their values range up to 10% depending on the region of the parameter space. In Table 2 we show the uncertainties that contribute to a yield variation bigger than 1%, combined into a single uncertainty for convenience.
- **Simulated sample size:** the finite nature of simulated samples is considered as an additional source of systematic uncertainty. For each bin of the distributions, one additional uncertainty is added, where only the considered bin is altered by  $\pm 1$  standard deviation, keeping the others at their nominal value.

The effects of these uncertainties on the total yields in the analysis selection are summarised in Table 2, where the  $e^+e^-$  and  $\mu^+\mu^-$  categories are combined together, yielding, for some uncertainties, a range of impacts.

## 7 Results

A binned maximum likelihood fit is performed in order to extract best-fit signal cross sections. The fit is performed using the six binned templates mentioned before. Another bin is included in the fit containing the  $e^\pm\mu^\mp$  selection to further constraint the  $t\bar{t}$  and single top quark processes. The likelihood function is the product of the Poisson likelihoods over all bins of the templates and is given by

$$L(\beta_{\text{signal}}, \beta_k | \text{data}) = \prod_{i=1}^{N_{\text{bins}}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!},$$

where  $n_i$  is the number of observed events in bin  $i$  and the Poisson mean for bin  $i$  is given by

$$\mu_i = \beta_{\text{signal}} S_i + \sum_k \beta_k T_{k,i},$$

where  $k$  denotes all of the considered background processes,  $T_{k,i}$  is the bin content of bin  $i$  of the template for process  $k$ , and  $S_i$  is the bin content of bin  $i$  of the signal template. The parameter  $\beta_k$  is the nuisance parameter for the normalisation of the process  $k$ , constrained by theoretical uncertainties with a log-normal prior, and  $\beta_{\text{signal}}$  is the signal strength, unconstrained. For each systematic uncertainty affecting the shape (normalisation) of the templates, a nuisance parameter is introduced with a Gaussian (log-normal) prior.

The best-fit values for all the nuisance parameters, as well as the corresponding uncertainties,

Table 2: Summary of the systematic uncertainties prior to the fit and their impact in percentages on the total event yields for background and for a particular signal hypothesis of  $m_H = 379$  GeV and  $m_A = 172$  GeV.

Source	Background yield variation	Signal yield variation
Electron identification and isolation	2.7%	2.6%
Integrated luminosity	2.5%	2.5%
Jet energy scale	2.1–2.4%	0.1–0.3%
Jet b tagging (heavy-flavour jets)	2.3%	2.0%
PDFs	1.0%	0.5%
Pileup	0.3–0.9%	0.7–1.3%
b tagging (light-flavour jets)	0.7–0.8%	< 0.1%
Muon identification	0.5%	0.4%
Jet energy resolution	0.2%	0.2%
Trigger efficiency	0.1–0.3%	0.1–0.3%
Muon isolation	0.2%	0.1%
Affecting only $t\bar{t}$ (31.8% of the total bkg.)		
$\mu_R$ and $\mu_F$ scales	12.2–12.3%	
$t\bar{t}$ cross section	5.3%	
Affecting only Drell-Yan (64.5% of the total bkg.)		
$\mu_R$ and $\mu_F$ scales	9.6%	
Drell-Yan cross-section	4.9%	
Drell-Yan shape uncertainty	2.1–2.2%	
Simulated sample size	0.5–1.3%	
Affecting only VV (1.1% of the total bkg.)		
$\mu_R$ and $\mu_F$ scales	4.3–4.8%	
Affecting only signal		
$\mu_R$ and $\mu_F$ scales	1.8%	

are extracted by performing a binned maximum likelihood fit, in the background-only hypothesis to the data. Only nuisance parameters affecting the backgrounds are considered.

Figures 5 and 6 show final distributions under two different mass hypotheses for  $e^+e^-$ ,  $\mu^+\mu^-$ , and  $\mu^+e^- + \mu^-e^+$  events with all the nuisance parameters set to their best-fit values.

We set upper limits at 95% confidence level (CL) on the product of the production cross section and branching fraction for  $H \rightarrow ZA \rightarrow \ell\ell b\bar{b}$  using the asymptotic modified frequentist method (asymptotic  $CL_s$ ) [47–49] as a function of the A and H mass hypothesis. Figure 7 shows expected and observed (with  $\pm 1$ ,  $\pm 2$  standard deviation bands) 95% CL exclusion limits in the two dimensional  $m_{Hjj}$ - $m_{Ajj}$  plane under the 2HDM benchmark  $\tan\beta = 1.5$  and  $\cos(\beta - \alpha) = 0.01$ , while Figure 8 shows exclusion limits in the  $\tan\beta$ - $\cos(\beta - \alpha)$  plane for  $m_H = 379$  GeV and  $m_A = 172$  GeV.

The highest asymptotic local significance observed corresponds to  $3.9\sigma$  for the signal hypothesis with  $m_H = 627$  GeV and  $m_A = 162$  GeV, which globally becomes  $1.3\sigma$  once accounting for the look-elsewhere-effect [50].

## 8 Summary

This note reports on a search for a new CP-even neutral Higgs boson, H, decaying into a Z boson and a lighter CP-odd Higgs boson, A, where the Z decays into  $\ell^+\ell^-$ , and the A into  $b\bar{b}$ . The search is based on LHC proton-proton collision data at  $\sqrt{s} = 13$  TeV collected by the CMS experiment, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . We consider decays such that  $H \rightarrow ZA \rightarrow \ell\ell b\bar{b}$ , where H and A are additional neutral Higgs bosons with masses ranging from  $m_Z$  to 1000 GeV for H and 30 to 1000 GeV for A. Deviations from the SM expectations are observed with a global significance of  $1.3\sigma$  and so upper limits on the product of cross section and branching ratio are set. Limits are also set on the parameters of the type-II 2HDM model. Within this theoretical framework, H and A are the CP-even and CP-odd scalar bosons, respectively. The specific benchmark scenario corresponding to  $\tan\beta = 1.5$  and  $\cos(\beta - \alpha) = 0.01$  is excluded for  $m_H$  in the range 150–700 GeV and  $m_A$  in the range 30–295 GeV with  $m_H > m_A$ , or alternatively, for  $m_H$  in the range 125–280 GeV and  $m_A$  in the range 200–700 GeV with  $m_H < m_A$ . Results are also interpreted in the benchmark scenario where  $m_H = 379$  GeV and  $m_A = 172$  GeV. In this context, the region with  $\cos(\beta - \alpha)$  in the range -0.9–0.3 and  $\tan\beta$  in the range 0.5–7 is excluded. A larger region of the Type-II 2HDM parameter space is excluded with respect to previous searches.

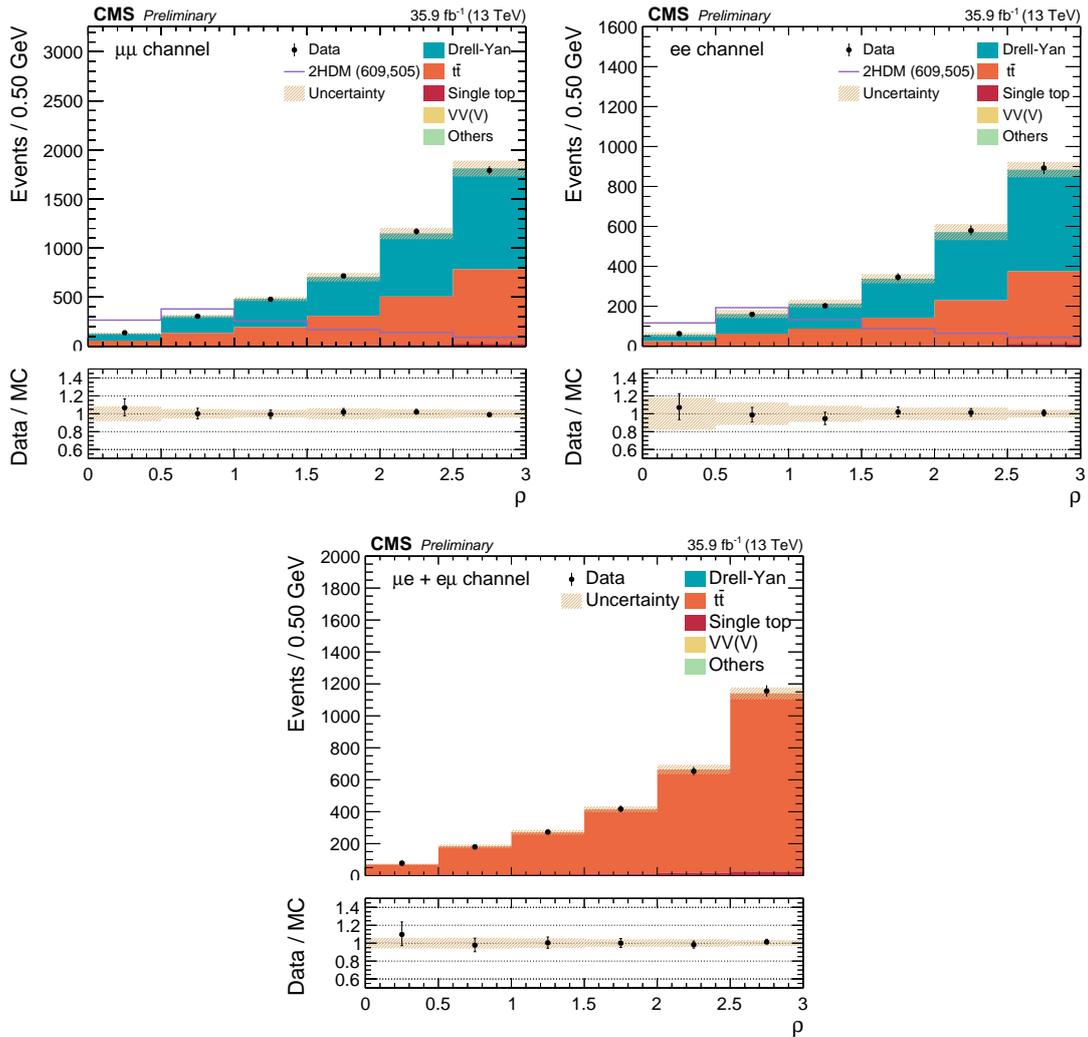


Figure 5:  $\rho$  distributions for  $\mu^+\mu^-$  (top left),  $e^+e^-$  (top right), and  $e^\pm\mu^\mp$  (bottom) events corresponding to a particle signal hypothesis with  $m_H = 609$  GeV and  $m_A = 505$  GeV. The signal is normalised to 20 nb. The mixed-flavour lepton category is a control region used in the template fit to further constrain the  $t\bar{t}$  background. The bins are defined by steps of 0.5 in  $\rho$ . Error bars indicate statistical uncertainties, while shaded bands show systematic uncertainties.

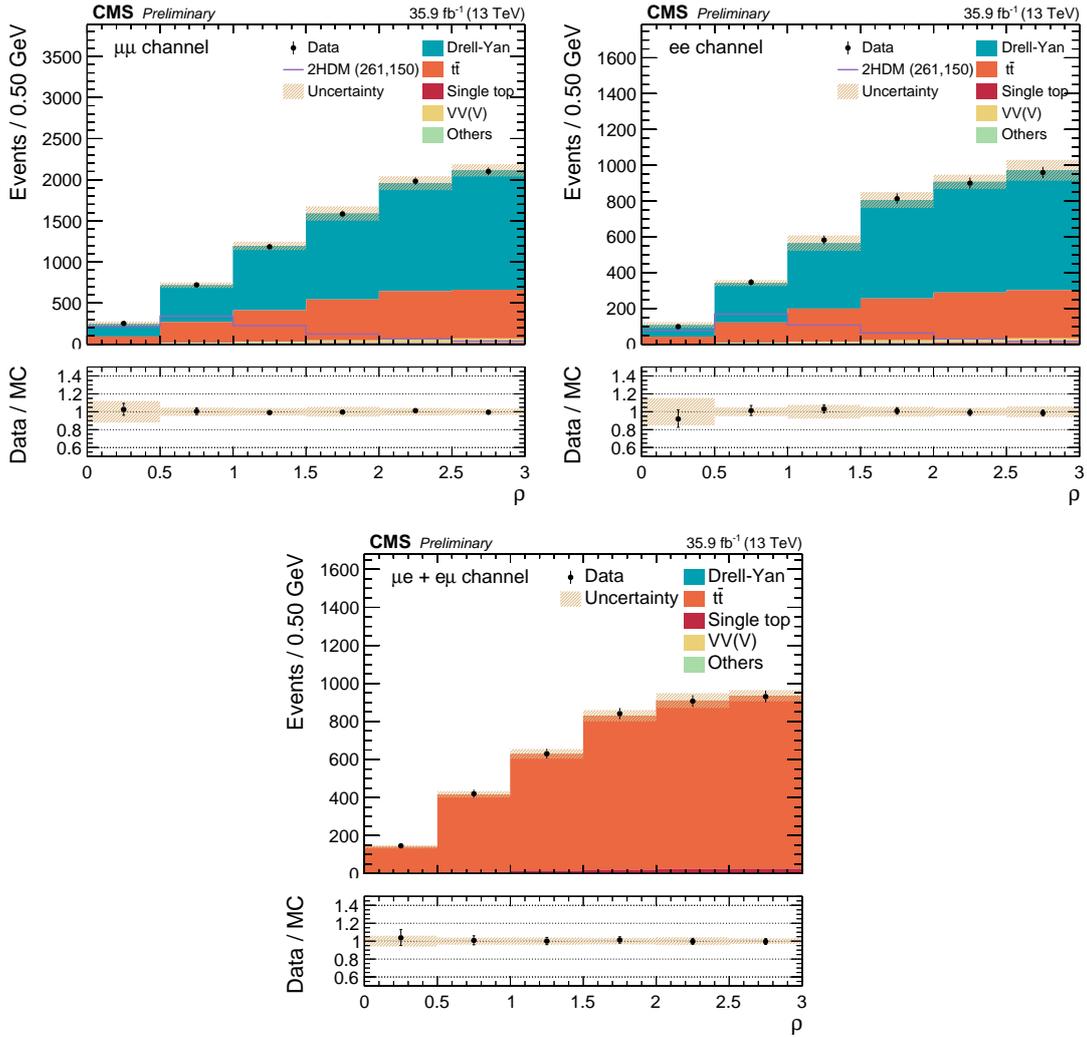


Figure 6:  $\rho$  distributions for  $\mu^+\mu^-$  (top left),  $e^+e^-$  (top right), and  $e^\pm\mu^\mp$  (bottom) events corresponding to a particle signal hypothesis with  $m_H = 261$  GeV and  $m_A = 150$  GeV. The signal is normalised to 20 nb. The mixed-flavour lepton category is a control region used in the template fit to further constrain the  $t\bar{t}$  background. The bins are defined by steps of 0.5 in  $\rho$ . Error bars indicate statistical uncertainties, while shaded bands show systematic uncertainties.

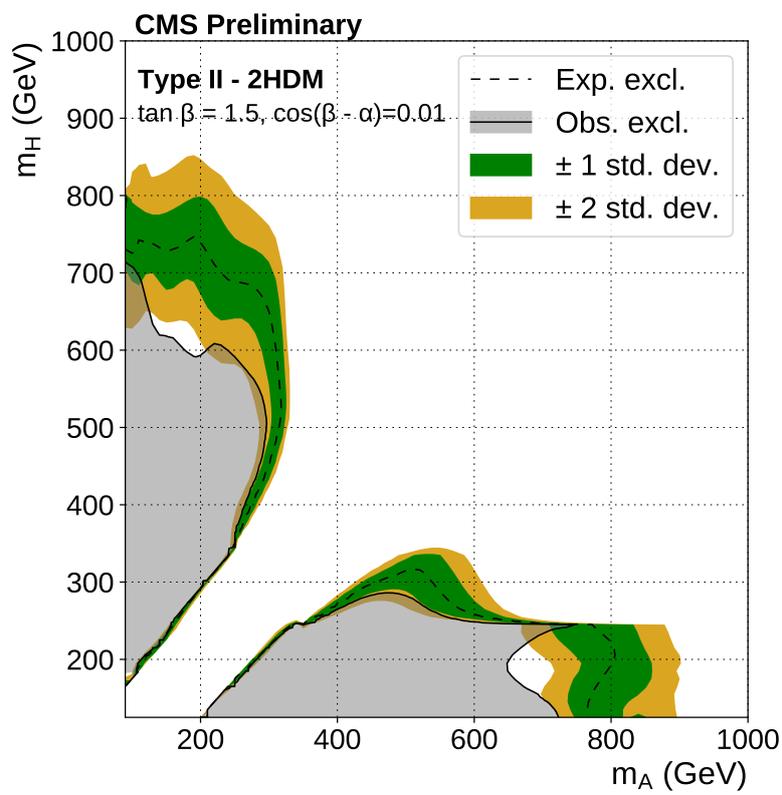


Figure 7: Expected (with  $\pm 1$ ,  $\pm 2$  standard deviation bands) and observed 95% CL exclusion limits for  $\tan\beta = 1.5$  and  $\cos(\beta - \alpha) = 0.01$  as a function of  $m_A$  and  $m_H$  in the Type-II 2HDM benchmark scenario. The limits are computed using the asymptotic  $CL_s$  method, combining the  $e^+e^-$  and  $\mu^+\mu^-$  channels.

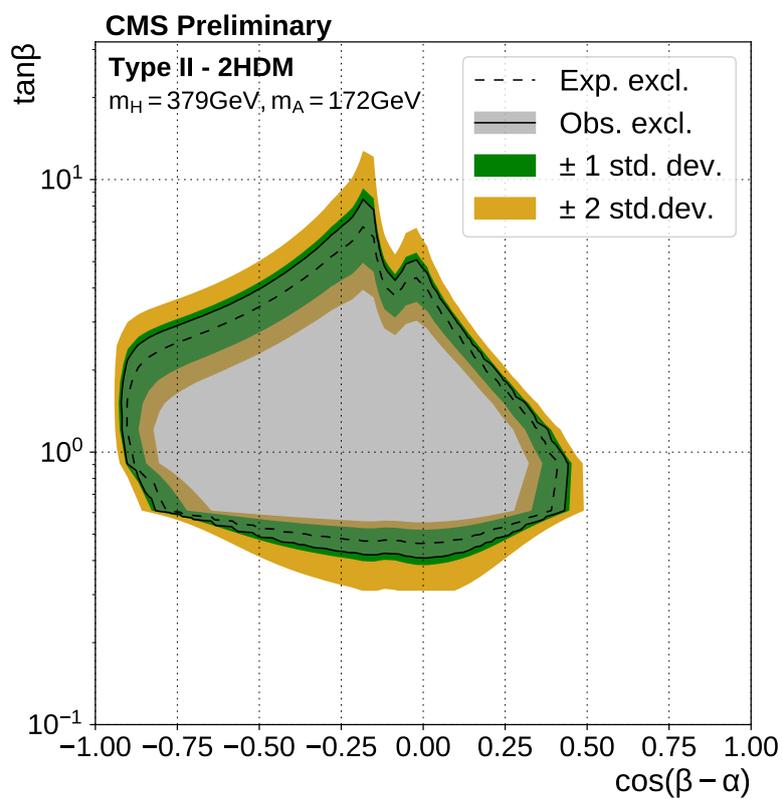


Figure 8: Expected (with  $\pm 1$ ,  $\pm 2$  standard deviation bands) and observed 95% CL exclusion limits for  $m_H = 379 \text{ GeV}$  and  $m_A = 172 \text{ GeV}$  as a function of  $\tan \beta$  and  $\cos(\beta - \alpha)$  in the Type-II 2HDM benchmark scenario. The limits are computed using the asymptotic  $\text{CL}_s$  method, combining the  $e^+e^-$  and  $\mu^+\mu^-$  channels.

## 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] C. Csaki, “The Minimal supersymmetric standard model (MSSM)”, *Mod. Phys. Lett. A* **11** (1996) 599, doi:10.1142/S021773239600062X, arXiv:hep-ph/9606414.
- [5] G. C. Branco et al., “Theory and phenomenology of two-Higgs-doublet models”, *Phys. Rept.* **516** (2012) 1, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.
- [6] ATLAS Collaboration, “Search for a CP-odd Higgs boson decaying to Zh in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector”, *Phys. Lett. B* **744** (2015) 163, doi:10.1016/j.physletb.2015.03.054, arXiv:1502.04478.
- [7] ATLAS Collaboration, “Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the  $\ell\ell b\bar{b}$  final state in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, arXiv:1804.01126.
- [8] CMS Collaboration, “Search for neutral resonances decaying into a Z boson and a pair of b jets or  $\tau$  leptons”, *Phys. Lett. B* **759** (2016) 369, doi:10.1016/j.physletb.2016.05.087, arXiv:1603.02991.
- [9] CMS Collaboration, “Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in  $\ell^+\ell^-\bar{b}b$  final states”, *Phys. Lett. B* **748** (2015) 221, doi:10.1016/j.physletb.2015.07.010, arXiv:1504.04710.
- [10] CMS Collaboration, “Search for a heavy pseudoscalar boson decaying to a Z boson and a Higgs boson at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis Summary CMS-PAS-HIG-18-005, 2018.
- [11] S. de Visscher et al., “Unconventional phenomenology of a minimal two-Higgs-doublet model”, *JHEP* **08** (2009) 042, doi:10.1088/1126-6708/2009/08/042, arXiv:0904.0705.
- [12] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **03** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [13] 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.
- [14] 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.

- 
- [15] 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.
- [16] S. Alioli, S. Moch, and P. Uwer, “Hadronic top-quark pair-production with one jet and parton showering”, *JHEP* **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.
- [17] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG:  $s$ - and  $t$ -channel contributions”, *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].
- [18] 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.
- [19] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [20] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”, *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.
- [21] T. Sjöstrand et al., “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [22] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [23] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2015) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [24] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J.* **C74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- [25] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [26] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”, *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.
- [27] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code”, *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.
- [28] M. Botje et al., “The PDF4LHC Working Group Interim Recommendations”, arXiv:1101.0538.
- [29] T. Gehrmann et al., “ $W^+W^-$  Production at Hadron Colliders in Next to Next to Leading Order QCD”, *Phys. Rev. Lett.* **113** (2014) 212001, doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.

- [30] J. M. Campbell, R. K. Ellis, and C. Williams, “Vector boson pair production at the LHC”, *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.
- [31] D. de Florian et al., “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”, doi:10.23731/CYRM-2017-002, arXiv:1610.07922.
- [32] CMS Collaboration, “Measurement of the inclusive W and Z production cross sections in pp collisions at  $\sqrt{s} = 7$  TeV with the CMS experiment”, *JHEP*. **10** (2011) 132, doi:10.1007/JHEP10(2011)132, arXiv:1107.4789.
- [33] 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.
- [34] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”, *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [35] 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.
- [36] 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.
- [37] 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.
- [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 the CMS missing transverse momentum reconstruction in pp data at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [40] CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector”, *Submitted to: JINST* (2019) arXiv:1903.06078.
- [41] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.
- [42] CMS Collaboration, “Heavy flavor identification at CMS with deep neural networks”, CMS Detector Performance Summary CMS-DP-2017-005, 2017.
- [43] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [44] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [45] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV”, *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.

- [46] S. Alekhin et al., “The PDF4LHC Working Group interim report”, 2011.  
arXiv:1101.0536.
- [47] 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.
- [48] A. L. Read, “Presentation of search results: the CL<sub>s</sub> technique”, *J. Phys. G* **28** (2002) 2693,  
doi:10.1088/0954-3899/28/10/313.
- [49] 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].
- [50] E. Gross and O. Vitells, “Trial factors for the look elsewhere effect in high energy  
physics”, *Eur. Phys. J.* **C70** (2010) 525–530,  
doi:10.1140/epjc/s10052-010-1470-8, arXiv:1005.1891.