I. INTRODUCTION

While ongoing analyses by both ATLAS and CMS show that the properties of the Higgs particle at $m_h \sim 125$ GeV are close to those expected for the Standard Model (SM) Higgs boson $h_{SM}$ [1–4], the complete nature of the scalar sector responsible for electroweak (EW) symmetry breaking remains to be determined. It is particularly interesting to ascertain whether the scalar sector consists of only one SU(2)$_L$ doublet or has a richer structure containing additional states. Addressing this question is a very important task for present and future studies at the LHC.

In this work, we concentrate on models with two-Higgs-doublet models (2HDM) (see Ref. [5] for a review), which appear in many extensions of the SM, such as the minimal supersymmetric SM or scenarios of viable electroweak baryogenesis [6–10]. In recent years, the region of the 2HDM parameter space allowed from Higgs coupling measurements by ATLAS and CMS [11,12] has been widely studied in the literature (see e.g. Refs. [13–32]). Various works have also discussed the constraints coming from LHC searches for neutral and charged scalars $A_0$, $H_0$, and $H^\pm$ via $A_0 \rightarrow Zh$, $A_0/H_0 \rightarrow \tau\tau$, $A_0/H_0 \rightarrow \gamma\gamma$, $H_0 \rightarrow ZZ/WW$, $H_0 \rightarrow hh$, $H^\pm \rightarrow tb$, and $H^\pm \rightarrow \tau\nu$ (see e.g. Refs. [30,31] for recent analyses). Interpreting these constraints assuming that the new scalars can only decay into SM states requires $H_0$, $A_0$, and $H^\pm$ to be relatively close in mass. In the following, we refer to this scenario as the degenerate 2HDM.

On the other hand, it has recently been shown [10] that sizable mass splittings between the 2HDM scalars (in particular a large $m_{A_0} - m_{H_0}$) favor a strong EW phase transition that could lead to baryogenesis. This provides an important physical motivation for 2HDM scenarios in which new decay channels for the heavier scalars are kinematically allowed (e.g. $A_0 \rightarrow Zh_0$), a situation which has not received enough attention in the literature (see, however, Refs. [33–36]). In the following, we refer to this scenario as the hierarchical 2HDM. It is the purpose of this work to fill this gap, providing a detailed discussion of the constraints on the 2HDM parameter space from 7 and 8 TeV LHC run 1 data, comparing the degenerate and hierarchical 2HDM scenarios. We will show that, besides significantly weakening the bounds from searches for these new scalars into SM states, the sizable mass splittings provide possibilities for novel searches (see e.g. Refs. [37,38]) which can yield complementary limits on the 2HDM parameter space. We assess the interplay between these searches, the standard searches for new scalars decaying directly into SM particles, and the measurement of 125 GeV Higgs signal strengths in constraining 2HDM scenarios. Furthermore, being at the onset of LHC run 2, we also outline the upcoming prospects for direct searches of the neutral scalars $H_0$ and $A_0$ in the hierarchical 2HDM at the 13 TeV run of LHC, through the discussion of benchmark plane scenarios.

After a review of the 2HDM in Sec. II, we discuss the measurements of Higgs signal strengths in the context of the 2HDM in Sec. III A. We then demonstrate the impact of the mass spectrum on LHC searches for $A_0/H_0$ in Secs. III B and III C as well as briefly commenting on $H^\pm$ searches in Sec. III D. In Sec. III E, we analyze the constraints that can be derived from the recent dedicated search of 2HDM neutral scalars with a sizable splitting by the CMS Collaboration [37,38], highlighting the strong complementarity with the standard searches and analyzing the interplay between these and Higgs measurements discussed in Sec. III A. Finally, in Sec. IV, we present benchmark plane scenarios for searches of these new scalars at LHC run 2.

II. (BRIEF) REVIEW OF THE 2HDM

In this section, we discuss the aspects of the 2HDM relevant to our analysis, defining at the same time our
notation. We consider a general 2HDM scalar potential with a softly broken $\mathbb{Z}_2$ symmetry in the absence of CP violation, which reads

$$V(H_1, H_2) = \mu^2 [H_1^\dagger H_1] + \mu^2 [H_2^\dagger H_2] - \mu^2 [H_1^\dagger H_2 + \text{h.c.}]$$

$$+ \frac{\lambda_1}{2} [H_1^\dagger H_1]^2 + \frac{\lambda_2}{2} [H_2^\dagger H_2]^2 + \frac{\lambda_3}{2} [H_1^\dagger H_2]^2$$

$$+ \lambda_4 [H_1^\dagger H_2]^2 + \frac{\lambda_5}{2} [(H_1^\dagger H_2)^2 + \text{h.c.}],$$

where the two scalar $SU(2)_L$ doublets $H_j$ ($j = 1, 2$) may be written as

$$H_j = (\phi_j^+, (v_j + h_j + i\eta_j)/\sqrt{2})^T.$$  (1)

The physical scalar sector of a 2HDM is comprised of two CP-even neutral scalars $h$ and $H_0$ (with $m_{H_0} \approx m_h$), a neutral CP-odd scalar $A_0$, and a charged scalar $H^\pm$. Throughout this work, we mostly assume that $h$ is the 125 GeV Higgs boson (the case of $H_0$ as the 125 GeV Higgs boson—with $h$ a lighter, undetected state—has been recently discussed in detail in Ref. [32]) and focus on the case where the other new states are heavier than $h$ (the possibility of $A_0$ being lighter than $m_h = 125$ GeV, although more experimentally constrained, has been explored recently in Ref. [39]). Apart from $m_h$ and $v = 246$ GeV, the scalar potential (1) may be parametrized in terms of the scalar masses $m_{H_0}$, $m_{A_0}$, and $m_{H^\pm}$; the squared mass scale $\mu^2$; and two angles $\beta$ and $\alpha$, the former being related to the ratio of vacuum expectation values of the two scalar doublets, $v_{1,2}$, via $\tan\beta \equiv v_2/v_1$ (with $v_1^2 + v_2^2 = v^2$) and the latter parametrizing the mixing between the CP-even states. The relation between the physical states $h$, $H_0$, $A_0$, and $H^\pm$ and the states $h_j$, $\eta_j$, and $\phi_j^+$ is given by

$$H^\pm = -s_\beta \phi_1^+ + c_\beta \phi_2^+ \quad A_0 = -s_\beta \rho_1 + c_\beta \rho_2$$

$$h = -s_\beta h_1 + c_\beta h_2 \quad H_0 = -c_\beta h_1 - s_\beta h_2$$

with $s_\beta$, $c_\beta$, $s_\alpha$, $c_\alpha \equiv \sin \beta$, $\cos \beta$, $\sin \alpha$, $\cos \alpha$, respectively. Regarding the couplings of the two doublets $H_{1,2}$ to fermions, the $\mathbb{Z}_2$ symmetry in (1), even when softly broken by $\mu^2$, may be used to forbid potentially dangerous tree-level flavor changing neutral currents (FCNCs) by requiring that each fermion type couple to one doublet only [40]. By convention, up-type quarks couple to $H_2$. In type I 2HDM, all the other fermions also couple to $H_2$, while for type II, down-type quarks and leptons couple to $H_1$. There are two more possibilities (depending on the $\mathbb{Z}_2$ parity assignment for leptons with respect to down-type quarks), but we focus here on types I and II, as they encode the relevant physics of 2HDMs with no tree-level FCNCs.

The parameters $t_\beta \equiv \tan \beta$ and $c_\beta \equiv \cos(\beta - \alpha)$ control the strength of the couplings of $h$, $H_0$, $A_0$, and $H^\pm$ to gauge bosons and fermions. Focusing on the neutral scalars, we denote the couplings normalized to the SM values (of $h_{\text{SM}}$) by $\kappa$-factors ($\kappa_V$ for gauge bosons, $\kappa_d$ for down-type quarks, and $\kappa_u$ for up-type quarks, which read

$$\begin{align*}
\kappa_V^h &= s_\beta - \alpha \\
\kappa_V^h &= c_\beta - \alpha \\
\kappa_V^{A_0} &= c_\beta \alpha \\
\kappa_V^{H_0} &= -s_\beta - \alpha \\
\kappa_V^{H_0} &= -c_\beta - \alpha \\
\kappa_V^{A_0} &= -s_\beta \alpha \\
\kappa_V^{H_0} &= -c_\beta \alpha \\
\kappa_V^\eta &= t_\beta \\
\kappa_V^\eta &= t_\beta
\end{align*}$$

Type I:  (3)

$$\begin{align*}
\kappa_V^h &= s_\beta - \alpha \\
\kappa_V^h &= c_\beta - \alpha \\
\kappa_V^{A_0} &= c_\beta \alpha \\
\kappa_V^{H_0} &= -s_\beta - \alpha \\
\kappa_V^{H_0} &= -c_\beta - \alpha \\
\kappa_V^{A_0} &= -s_\beta \alpha \\
\kappa_V^{H_0} &= -c_\beta \alpha \\
\kappa_V^\eta &= t_\beta \\
\kappa_V^\eta &= t_\beta
\end{align*}$$

Type II:  (4)

For $c_\beta - \alpha \to 0$, commonly referred to as the 2HDM alignment limit [41], $h$ has SM-like couplings to gauge bosons and fermions ($\kappa_V^h \to 1$, yielding $h \to h_{\text{SM}}$), while the coupling $H_0VV$ of $H_0$ to gauge bosons $V = W^\pm, Z$ vanishes ($\kappa_V^{H_0} \to 0$).

In order to obtain a viable 2HDM scenario, theoretical constraints from unitarity, perturbativity, and stability/ boundedness from below of the scalar potential (1) need to be satisfied. Tree-level stability of the potential $V(H_1, H_2)$ requires $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > -\sqrt{\lambda_1 \lambda_2}$, and $\lambda_3 + \lambda_4 - |\lambda_5| > -\sqrt{\lambda_1 \lambda_2}$ (see e.g. Ref. [42]). At the same time, tree-level unitarity (for a recent one-loop analysis, leading to slightly more stringent bounds; see Ref. [43]) imposes bounds on the size of various combinations of the quartic couplings $\lambda_j$ [44], like $|\lambda_3 + \lambda_4| < 8\pi, |\lambda_3 + \lambda_5| < 8\pi, |\lambda_3 + 2\lambda_4 + 3\lambda_5| < 8\pi$, and $|\lambda_1 + \lambda_2 + \sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_4^2}| < 16\pi$. Similar (although generically less stringent) bounds on $\lambda_j$ may be obtained from perturbativity arguments. We may express $\lambda_j$ in terms of the physical scalar masses, the mixing angles $\alpha, \beta$, and $\mu^2$:

$$\begin{align*}
\lambda_1 &= \frac{1}{v^2 \sin^2 \beta} (-\mu^2 t_\beta + m_{h_0}^2 s_\alpha^2 + m_{H_0}^2 c_\alpha^2), \\
\lambda_2 &= \frac{1}{v^2 \sin^2 \beta} (-\mu^2 t_\beta^2 + m_{H_0}^2 c_\alpha^2 + m_{H_0}^2 s_\alpha^2)
\end{align*}$$

(5)  (6)
As seen from (5)–(9), for a given set of values for $m_{H_0}, m_{A_0}, m_{A}^2, t_{\beta},$ and $c_{\beta - \alpha},$ only a certain range for $\mu^2$ is allowed by the combination of these theoretical constraints. In particular, $\lambda_{1,2} > 0$ directly imply an upper bound on $\mu^2$ from (5) and (6). It is, however, possible that no value of $\mu^2$ can satisfy all three theoretical requirements simultaneously, in which case such a set of values for the scalar masses and mixing angles would not be viable. If an allowed $\mu^2$ range exists, the size of trilinear scalar couplings such as $\lambda_{H_0 hh}$ and $\lambda_{H_0 A_0 A_0}$ (which control the partial widths $\Gamma_{H_0 \rightarrow hh}, \Gamma_{H_0 \rightarrow A_0 A_0}$ when these decays are kinematically allowed) or $\lambda_{H^+ H^-}$ (which controls the size of the charged scalar loop contribution to the $h \rightarrow \gamma \gamma$ decay amplitude, given by $\Delta_t^\pm$) depend on the value of $\mu^2$. Indeed, the trilinear couplings $\lambda_{H_0 hh}$ and $\lambda_{H_0 A_0 A_0}$ are given by

$$v\lambda_{H_0 hh} = \frac{2c_{\beta - \alpha}}{s_{2\beta}} \left[ \left( 1 - \frac{3s_{2\alpha}}{s_{2\beta}} \right) \mu^2 + \left( 2m_h^2 + m_{H_0}^2 \right) \frac{s_{2\alpha}}{2} \right]$$

$$v\lambda_{H_0 A_0 A_0} = 2 \left[ c_{\beta - \alpha} (2m_{A_0} - m_{H_0}^2) \right.$$ 

$$- 2 \left( s_{\beta - \alpha} \frac{c_{2\beta}}{s_{2\beta}} - c_{\beta - \alpha} \right) \left( m_{H_0}^2 - \frac{\mu^2}{s_{\beta - \alpha} c_{\beta - \alpha}} \right) \right].$$

Apart from vanishing in the alignment limit, if $s_{2\beta} - 3s_{2\alpha} \neq 0$, the coupling $\lambda_{H_0 hh}$ also vanishes for $\mu^2 = \left( 2m_h^2 + m_{H_0}^2 \right)(s_{2\alpha}s_{2\beta})/(6s_{2\alpha} - 2s_{2\beta})$, if such value of $\mu^2$ lies within the allowed range. Similarly, in the alignment limit, $\lambda_{H_0 A_0 A_0}$ vanishes for $t_{\beta} = 1$ or $\mu^2 = m_h^2 s_{\beta} c_{\beta}$. The trilinear coupling $\lambda_{H^+ H^-}$ reads

$$v\lambda_{H^+ H^-} = \left[ s_{\beta - \alpha} (m_h^2 - 2m_{H^+}^2) \right.$$ 

$$- 2 \left( c_{\beta - \alpha} \frac{c_{2\beta}}{s_{2\beta}} + s_{\beta - \alpha} \right) \left( m_h^2 - \frac{\mu^2}{s_{\beta - \alpha} c_{\beta - \alpha}} \right) \right].$$

so that $\Delta_t^\pm$ inherits a dependence on $\mu^2$ and other 2HDM parameters besides $m_{H^+}^2$ through $\lambda_{H^+ H^-}$. These trilinear couplings illustrate the phenomenological impact of the soft $Z_2$-breaking parameter in the 2HDM, which will be analyzed in more detail in Sec. III.

III. HIERARCHICAL VS DEGENERATE 2HDM: THE LHC RUN 1 LEGACY

Let us now concentrate on the mass spectrum of the 2HDM. We first note that constraints from measurements of EW precision observables (EWPO), in particular of the $T$-parameter, generically require $H^ \pm$ to be relatively degenerate with either $A_0$ or $H_0$ [45,46]. From a phenomenological perspective, we can then distinguish between a degenerate spectrum where all mass splittings among the new scalar states are small, $|m_{A_0} - m_{H_0}| \ll m_Z$, and a hierarchical spectrum for which the mass splitting among the new neutral scalars is sizable, $|m_{A_0} - m_{H_0}| \gtrsim m_Z$.

The main phenomenological feature of a hierarchical 2HDM spectrum is that the decays $q_i \rightarrow q_j V$, with $q_{i,j} = H_0, A_0$, $H^\pm$ ($i \neq j$), and $V = W^\pm, Z$ become kinematically allowed and generically yield the dominant branching fraction, with the decays into SM states comparatively suppressed. These considerations motivate performing a comparison of the allowed 2HDM parameter space for both types of spectra, assessing the impact of sizeable mass splitting(s).

In this respect, key probes of 2HDM scenarios are CMS and ATLAS measurements of Higgs signal strengths and searches for new scalar states at the LHC, which we analyze in detail in the next sections. When performing the analysis of LHC constraints (as well as the aforementioned theoretical constraints from vacuum stability, perturbativity, and unitarity), we consider either $m_{H^\pm} = m_{A_0}$ or $m_{H^\pm} = m_{H_0}$ which are simplifying limits motivated by EWPO (see also the discussion in Sec. III D).

A. Higgs signal strengths in the 2HDM

The values for the Higgs signal strengths measured by the ATLAS and CMS experiments during the 7 and 8 TeV LHC runs set an important constraint on the 2HDM parameter space [11] (see also Refs. [13,14,16,18–20,26,27,30]). Since, in the 2HDM, the Higgs couplings are simply rescaled with respect to the SM values, without modifying the kinematics of production or decay, the model prediction for the signal strength in a final state or experimental category $\gamma \gamma$ can be written as

$$\mu^\text{2HDM}_{\gamma \gamma} = \sum_i e_i \times \mu^i_{\gamma \gamma}.$$

Here, $e_i$ corresponds to the relative contribution of a particular Higgs production mode, $i$, to the observable, and $\mu^i_{\gamma \gamma}$ is the 2HDM signal strength for that production mode.

$$\mu_{\gamma \gamma} = \frac{\left[ \sigma_i(pp \rightarrow h) \times \text{BR}(h \rightarrow \gamma \gamma) \right]_{2\text{HDM}}}{\left[ \sigma_i(pp \rightarrow h) \times \text{BR}(h \rightarrow \gamma \gamma) \right]_{2\text{SM}}}.$$  

This is to be compared with the values obtained by ATLAS and CMS analyses in the relevant detection channels,
namely $h \to WW^*$ [47–51], $h \to ZZ^*$ [48,52,53], $h \to \gamma\gamma$ [54,55], $h \to \bar{b}b$ [56,57], and $h \to \tau\tau$ [58,59], with possibly numerous experimental categories therein. The various $\mu_{\ell\ell}$ are directly obtained from the $\kappa_{\ell\ell}$ factors in (3) and (4) for the corresponding final state and depend only on $c_{\beta-\alpha}$ and $t_\beta$. The sole exceptions are experimental categories within a $\gamma\gamma$ final state, since BR($h \to \gamma\gamma$) also involves the contribution to the $h \to \gamma\gamma$ decay amplitude from the charged scalar loop, $\Delta^\pm_H$, which introduces a dependence on $\mu^2$ and other physical parameters via the trilinear coupling, $\lambda_{hH^+H^-}$. As a result, a fully self-consistent comparison to the experimental data would require a generalized $\Delta^2$ likelihood fit in a multidimensional parameter space subject to the theoretical constraints on $\mu^2$ discussed above. These constraints lead to an indirect dependence on the full 2HDM parameter space, meaning that all possible experimental constraints on the model such as direct heavy Higgs searches and EW precision observables would have to be included in the general case. Recent studies [60–62] have shown that the charged scalar loop can give a sizable (5%–10%) contribution to the $h \to \gamma\gamma$ decay amplitude, in particular due to its potential nondecoupling behavior. However, its impact on the global Higgs signal strength fit is expected to be at most mild, and so we adopt a simplified approach of neglecting this term in the global fit by setting $\lambda_{hH^+H^-} = 0$.

The Higgs signal strength constraints can then be obtained by performing a $\Delta^2$ likelihood fit to the 2HDM parameters $c_{\beta-\alpha}$ and $t_\beta$, for which we use the public codes LILITH [63] and HiggsSignals [64,65]. The values of $e_i$ in (14) may be obtained from the experimental analyses and are provided in both these programs (e.g. for HiggsSignals, they may be found in Appendix A of Ref. [65]). The results are shown in Fig. 1 for type I (left) and type II (right) 2HDM. The green areas correspond to the 95% C.L. allowed region from LILITH, while the hatched-purple ones are those from HiggsSignals. Both show good agreement with analogous fits performed by CMS and ATLAS [11,12], with the fit from LILITH being slightly more constraining than both HiggsSignals and ATLAS/CMS. In type I, a sizable departure from alignment is allowed as soon as $t_\beta \gtrsim 1$, and the limit on $c_{\beta-\alpha}$ becomes both independent of $t_\beta$ and symmetric around $c_{\beta-\alpha} = 0$ for $t_\beta \gg 1$, which can be understood from (3). For type II, there are two distinct allowed regions: (i) the region close to the alignment limit $|c_{\beta-\alpha}| \ll 1$ corresponding to a SM-like Higgs $h$, with a mild preference for $c_{\beta-\alpha} > 0$ and $t_\beta \sim 1$, and (ii) the wrong-sign scenario $s_{\beta-\alpha} \sim 1$, for which $\kappa_{\beta-\alpha}^H < 0$, and $|1 + \kappa_{\beta-\alpha}^H| \ll 1$ (see e.g. Ref. [62] for a detailed discussion of this limit, possible only in type II).

**B. LHC searches for $A_0$ into SM states**

We discuss now the limits on the 2HDM parameter space from ATLAS and CMS searches of $A_0$ decaying via $A_0 \to Z h$ ($h \to \bar{b}b$, $h \to \tau\tau$) [66,67], via $A_0 \to \gamma\gamma$ [68,69] and $A_0 \to \tau\tau$ [70,71]. For the $A_0 \to \tau\tau$ searches, the production of $A_0$ in association with a $\bar{b}b$ pair is taken into account by the ATLAS/CMS experimental analyses in addition to production through gluon fusion, the former being important for type II at large values of $t_\beta$. Furthermore, we stress that, while the search via $A_0 \to Z h$ vetoes any b-tagged jets beyond those from $h \to \bar{b}b$ (see e.g. Ref. [66]), the $b$-jets resulting from the $pp \to \bar{b}bA_0$ process generically have large rapidity values and consequently yield a very low b-tagging efficiency [72]. Thus, we also consider $\bar{b}b$-associated production of $A_0$ in the $A_0 \to Z h$ searches and do not implement a b-tagging efficiency suppression in this case.

**FIG. 1.** 95% C.L. regions from a likelihood fit to Higgs signal strengths in the $(c_{\beta-\alpha}, t_\beta)$ plane, for type I (left) and type II (right) 2HDM, using LILITH (solid-green region) and HiggsSignals (hatched-purple region). See the text for details on the fit.
In order to derive the bounds on the 2HDM parameter space, we compute the $0$ production cross section in gluon fusion and in association with $bb$ at next-to-leading order (NNLO) in QCD with StSuHi [73] for types I and II as a function of $t_\beta$ and $m_{\tilde A_0}$ and then use 2HDMC [74] to compute the branching fractions for $A_0 \to \tau\tau$, $A_0 \to \gamma\gamma$, $A_0 \to Zh$, and $h \to bb$, $\tau\tau$ as a function of $t_\beta$, $c_{\beta-\alpha}$, $m_{\tilde A_0}$, and $m_{\tilde H_0}$. The $95\%$ C.L. exclusion region in the ($c_{\beta-\alpha}$, $t_\beta$) plane resulting from these searches is shown in Fig. 2 for different values of $m_{\tilde A_0}$ and $m_{\tilde A_0} - m_{\tilde H_0}$ and discussed below. The progressive weakening of each search with increasing mass splitting is a ubiquitous feature of our results (see also Figs. 4 and 5), such that the limit corresponding to a spectrum with a greater mass splitting is always contained within the limit of a less split scenario. We therefore permit ourselves to not distinguish the mass scenarios for a given final state in each figure since it is implicit that the best limit always corresponds to the degenerate case and the worst to the most hierarchical. In any case, the important information lies in the union of the limits for each mass scenario under which we shade in a particular color.

Let us consider first a high mass scenario for $A_0$, above the $t\bar t$ threshold: The exclusion region for $m_{\tilde A_0} = 500$ GeV in types I and II is shown in gray, respectively, in Fig. 2’s top left and top right. The only sensitive channel above the $t\bar t$ threshold is $A_0 \to Zh$, and for type II also $A_0 \to \tau\tau$ in $bb$-associated production. Nevertheless, we see that for low/moderate $t_\beta$, these searches only constrain values of $|c_{\beta-\alpha}| \gtrsim 0.15$. The green region corresponds to the exclusion for $m_{\tilde H_0} = 500$ GeV, when $A_0$ only decays into SM states. As $m_{\tilde H_0}$ decreases and the decay $A_0 \to H_0Z$ becomes kinematically allowed, the current limits from searches of SM decay channels weaken significantly, as the orange and purple regions in Fig. 2 (top) show, respectively, for $m_{\tilde H_0} = 300$ GeV and $m_{\tilde H_0} = 150$ GeV.

The impact of a sizable $m_{\tilde A_0} - m_{\tilde H_0}$ splitting is even more important for $m_{\tilde A_0} \lesssim 300$ GeV in types I and II, as shown in Fig. 2 (middle) for type I (left) and type II (right), both in the degenerate scenario $m_{\tilde H_0} = 300$ GeV (green region) and for a hierarchical scenario with $m_{\tilde H_0} = 150$ GeV (purple region). In the former, the limits from $A_0 \to Zh$ searches are stringent, ruling out $|c_{\beta-\alpha}| \gtrsim 0.02$ for $t_\beta < 6$ in type I. Even for $c_{\beta-\alpha} \to 0$, $A_0 \to \gamma\gamma$, and $A_0 \to \tau\tau$, searches constrain the regions of $t_\beta \lesssim 2$ and $t_\beta \lesssim 3$, respectively, for types I and II. In contrast, for the hierarchical scenario, the $A_0 \to \gamma\gamma$ and $A_0 \to \tau\tau$ searches only constrain values of $t_\beta \lesssim 0.5$, while the sensitivity of the $A_0 \to Zh$ searches also reduces drastically.

Finally, we also present the limits for a light $A_0$, with $m_{\tilde A_0} = 150$ GeV in Fig. 2 (bottom). In this case, we do not consider a hierarchical 2HDM scenario (with $A_0$ being the heavier state), as it would require $c_{\beta-\alpha} \to 0$ to avoid nonobservation of $H_0$ at LEP (we will, however, briefly discuss this region of parameter space in Sec. III E). Both for type I (left) and type II (right), the $A_0 \to \tau\tau$ and $A_0 \to \gamma\gamma$ searches yield the constraint $t_\beta \gtrsim 1.5$, while for type II the searches for $A_0 \to \tau\tau$ in $bb$-associated production also yield a limit $t_\beta < 10$.

The above discussion highlights the fact that, in the presence of a sizable mass splitting $m_{\tilde A_0} - m_{\tilde H_0}$, the sensitivity of searches for $A_0$ decaying into SM final states becomes highly suppressed, with the sole exception of $A_0 \to \tau\tau$ searches in $bb$-associated production at high $t_\beta$ in type II 2HDM. Let us, however, emphasize that the lighter state in the hierarchical 2HDM scenario, in this case $H_0$, would decay solely into SM states. Thus, the constraints on the parameter space from searches of $H_0$ into SM final states would fully apply for a hierarchical scenario with $m_{\tilde A_0} - m_{\tilde H_0} > 0$. LHC searches for $H_0$ will be analyzed in Sec. III C.

Before moving on to the next section, let us discuss the impact of theoretical constraints from unitarity, perturbativity, and stability of the 2HDM scalar potential on Fig. 2, where the theoretically excluded regions are shown in gray. Focusing on the case $m_{H^\pm} = m_{\tilde A_0}$, and defining $m^2_{\tilde A_0} - m^2_{\tilde H_0} \equiv \Delta^2 \geq 0$, Fig. 2 shows that the exclusion becomes more important as $m_{\tilde H_0}$ increases, particularly for $t_\beta \gg 1$. The departure from alignment also has a strong impact on the theoretically viable parameter space, especially for $c_{\beta-\alpha} < 0$. These features may be understood from the interplay of $\lambda_1 > 0$ and various unitarity limits. Writing $\lambda_1$ as

$$
\lambda_1 v^2 = m^2_{\tilde A_0} - t_\beta(1 + t^2_\beta)\Omega^2 - (m^2_{\tilde H_0} - m^2_{\tilde A_0}) [c_{\beta-\alpha}(t_\beta - 1) - 2t_\beta s_{\beta-\alpha} c_{\beta-\alpha}] \tag{15}
$$

with $\Omega^2 \equiv m^2 - m^2_{\tilde H_0}s_{\beta-\alpha} c_{\beta}$, we see that for $m^2_{\tilde H_0} \gg m^2_{\tilde A_0}$ (neglecting $m^2_{\tilde A_0}$ in (15)) and $t_\beta > 1$, $\Omega^2 < 0$ is required to satisfy $\lambda_1 > 0$ for either $c_{\beta-\alpha} < 0$ or $c_{\beta-\alpha} t_\beta \gg 1$. This in turn impacts the unitarity requirements, e.g.

$$
|\lambda_3 + \lambda_4| \sim \frac{\Delta^2}{v^2} + \frac{m^2_{\tilde H_0} c_{\beta-\alpha}}{v^2} [s_{\beta-\alpha}(t_\beta - t^{-1}_\beta) - 2c_{\beta-\alpha}] < 8\pi|\lambda_3 + 2\lambda_4 + 3\lambda_5| \sim \frac{3\Delta^2}{v^2} \frac{4}{s_{\beta-\alpha} c_{\beta}} \frac{\Omega^2}{v^2} + m^2_{\tilde A_0} c_{\beta-\alpha} [s_{\beta-\alpha}(t_\beta - t^{-1}_\beta) - 2c_{\beta-\alpha}] \quad < 8\pi.
$$

which are then violated for $t_\beta \gg 1$ and/or $m^2_{\tilde H_0} \gg v^2$, as no cancellation among terms is possible in both $|\lambda_3 + \lambda_4|$ and $|\lambda_3 + 2\lambda_4 + 3\lambda_5|$ (since $\Delta^2 \geq 0$, $\Omega^2 < 0$).
FIG. 2. Current 95% C.L. excluded region by ATLAS (solid lines) and CMS (dashed lines), respectively, for $m_{A_0} = 500, 300, 150$ GeV (top/middle/bottom) and for 2HDM types I/II (left/right), coming from searches of $gg/b\bar{b} \rightarrow A_0 \rightarrow Zh (\ell \rightarrow b\bar{b})$ (black lines), $gg \rightarrow A_0 \rightarrow \gamma\gamma$ (yellow lines), $gg \rightarrow A_0 \rightarrow \tau\tau$ (red lines), and $b\bar{b} (A_0 \rightarrow \tau\tau)$ (light-blue lines). In each case, the limits in the degenerate scenario $m_{H_0} = m_{A_0}$ are shown in green, while those for the hierarchical scenario(s) are shown in purple/orange. The various gray regions correspond to the theoretically excluded regions for the degenerate and hierarchical 2HDM scenarios (darker gray as $m_{H_0}$ is lower). As discussed in the text, limits from a given search get progressively weaker with increasing mass splitting and are always nested within each other, starting from the degenerate scenario and moving toward increasingly hierarchical mass spectra.
FIG. 3. Current 95% C.L. excluded region (in green) by ATLAS (solid lines) and CMS (dashed lines), respectively, for $m_{H_0} = 150$ GeV and 2HDM types I/II (left/right), coming from searches of $gg/b\bar{b} \rightarrow H_0 \rightarrow WW/ZZ$ (black lines), $gg \rightarrow A_0 \rightarrow \tau\tau$ (red lines), and $b\bar{b} \rightarrow A_0 \rightarrow \tau\tau$ (light-blue lines), and $gg \rightarrow A_0 \rightarrow \gamma\gamma$, both for $\mu^2 = \mu_{\text{min}}^2/\mu^2 = \mu_{\text{max}}^2$ (yellow/pink lines). The gray region is theoretically excluded.

We note that the above requirement $\Omega^2 < 0$ to satisfy $\lambda_1 > 0$ may be avoided for $c_{\beta-\alpha}(t_\beta^2 - 1) - 2t_\beta t_{\beta-\alpha} \sim 0$, for which the last term in (15) vanishes. This cancellation, which happens for $c_{\beta-\alpha} = 2t_\beta/(1 + t_\beta^2)$, is observed for $m_{H_0} = 500$ GeV and $c_{\beta-\alpha} > 0$ in Fig. 2 (top). We also note that in exact alignment $c_{\beta-\alpha} = 0$, $\Omega^2 = 0$ automatically yields $\lambda_1 > 0$, and all other boundedness-from-below requirements are also trivially satisfied for $\Delta^2 \geq 0$. The unitarity constraints are then only violated for $\Delta^2 \gg v^2$, and thus $c_{\beta-\alpha} = 0$ is always allowed in Fig. 2.

C. LHC searches for $H_0$ into SM states

We turn now to analyzing the constraints from LHC searches for $H_0$. The relevant searches to be considered are $H_0 \rightarrow ZZ \rightarrow \ell\ell\ell\ell$ [75] (and in the low mass region also $H_0 \rightarrow WW$ [76]), $H_0 \rightarrow hh \rightarrow b\bar{b}\gamma\gamma$ [77] by ATLAS and $H_0 \rightarrow WW$, $ZZ$ [78] (both low and high mass regions), $H_0 \rightarrow hh \rightarrow b\bar{b}\gamma\gamma$ [79], and $H_0 \rightarrow hh \rightarrow b\bar{b}b\bar{b}$ [80] by CMS. In all these searches, $b\bar{b}$-associated production of $H_0$ is implicitly included\(^1\) together with gluon fusion. In addition, the ATLAS/CMS searches via $A_0/H_0 \rightarrow \gamma\gamma$ [68,69] and via $A_0/H_0 \rightarrow \tau\tau$ [70,71] discussed in the previous section also apply in this case.

As in the previous section, we use SUSHi to compute the gluon fusion and $b\bar{b}$-associated $H_0$ production cross sections at NNLO in QCD for types I and II as a function of $c_{\beta-\alpha}, t_\beta$, and $m_{H_0}$ and then use 2HDMC to compute the branching fractions for $H_0 \rightarrow \tau\tau$, $H_0 \rightarrow \gamma\gamma$, $H_0 \rightarrow ZZ$, $H_0 \rightarrow hh$, and $h \rightarrow b\bar{b}, \gamma\gamma$ as a function of $c_{\beta-\alpha}, t_\beta, m_{H_0}, m_{A_0}$, and $\mu^2$. We stress that, contrary to the $A_0$ case, the value of $\mu^2$ has a significant impact on the $H_0$ branching fractions via the modification of the trilinear coupling $\lambda_{H_0 hh}$, which changes the $H_0 \rightarrow hh$ partial width (recall the discussion at the end of Sec. II). In order to account for the dependence of $\mu^2$ on the 95% C.L. limits, we compute the theoretically viable $\mu^2$ range as a function of $c_{\beta-\alpha}, t_\beta, m_{H_0}$, and $m_{A_0}$ and derive the bounds on the values of $\mu^2$ that respectively minimize ($\mu_{\text{min}}^2$) and maximize ($\mu_{\text{max}}^2$) the $H_0 \rightarrow hh$ branching fraction within the allowed $\mu^2$ range.

We begin now by discussing the scenario with a light $H_0$ and consider the 95% C.L. exclusion region for $m_{H_0} = 150$ GeV in the degenerate scenario, as shown in Fig. 3. Due to the absence of the $H_0 \rightarrow hh$ decay in this case, the $H_0 \rightarrow ZZ$ and $H_0 \rightarrow \tau\tau$ branching fractions are not sensitive to the value of $\mu^2$, and only $H_0 \rightarrow \gamma\gamma$ is mildly dependent via the $H^\pm$ loop contribution. Nevertheless, Fig. 3 shows that the important limits in the $c_{\beta-\alpha}, t_\beta$ plane are given by $H_0 \rightarrow ZZ$ and $H_0 \rightarrow \tau\tau$ searches, with $H_0 \rightarrow \gamma\gamma$ less sensitive. As has been emphasized in Sec. III B, the present limits are complementary to those from $A_0$ searches in the hierarchical 2HDM, e.g. for the $(m_{A_0}, m_{H_0}) = (300, 150)$ GeV and $(500, 150)$ GeV benchmarks considered in Fig. 2.

In Fig. 4, we show the limits from $H_0$ searches for $m_{H_0} = 300$ GeV, for types I/II (left/right). Here, the presence of the decay $H_0 \rightarrow hh$ requires us to take into account the $\mu^2$ dependence in the limit extraction, and we show the limits for $\mu^2 = \mu_{\text{min}}^2$ (top) and $\mu^2 = \mu_{\text{max}}^2$ (bottom).

\(^1\)For $H_0 \rightarrow WW, ZZ$ searches, $b\bar{b}$-associated production generally fails the vector boson fusion and $V$-associated production analysis tags and so is included in the gluon fusion category. For $H_0 \rightarrow hh$ searches, the analysis is inclusive with respect to $H_0$ production.
In the former case, the strongest limits come from $H_0 \to ZZ$ searches, with $H_0 \to hh$ playing no relevant role because of its suppressed branching fraction. Moreover, in this case, the presence of a sizable $m_{H_0} - m_{A_0}$ splitting does lead to a significant reduction of the limits on the 2HDM parameter space from these searches. In contrast, for $\mu^2 = \mu^2_{\text{max}}$, the $H_0 \to hh$ searches provide the dominant constraint for low and moderate $t_\beta$, and these limits do not change significantly in a hierarchical 2HDM scenario, as the branching fraction of $H_0 \to hh$ is still the dominant one in this case. A similar situation occurs for $m_{H_0} = 500$ GeV, as shown in Fig. 5. Again, for $\mu^2 = \mu^2_{\text{min}}$ (top), $H_0 \to ZZ$ searches provide the only meaningful constraint, which gets significantly weakened in the hierarchical scenario $m_{H_0} - m_{A_0} \gg m_Z$. For $\mu^2 = \mu^2_{\text{max}}$, the LHC searches for $H_0 \to hh$ in $b\bar{b}b\bar{b}$ and $b\bar{b}\gamma\gamma$ are the most constraining, being particularly sensitive around $t_\beta \sim 1$, and the limits only get mildly weakened in the hierarchical 2HDM scenario. In addition, for $m_{H_0} = 500$ GeV, there is no appreciable difference between types I and II for low and moderate $t_\beta$, with $b\bar{b}$-associated production of $H_0$ in $H_0 \to \tau\tau$ searches constraining the $t_\beta \gg 1$ region in type II.

Contrary to Figs. 2, 4 and 5 do not show the would-be limits on the $(c_\beta, t_\beta)$ plane from searches of $H_0$ in regions which are not viable theoretically; these limits depend crucially on the value of $\mu^2$ (in contrast with the situation for $A_0$ searches discussed in Sec. III B), and the theoretical bounds correspond precisely to the absence of an allowed $\mu^2$ range.

Figures 4 and 5 highlight that, for $\Delta^2 < 0$ (and $m_{H_0}^2 = m_{A_0}^2$), the theoretical bounds from stability,

---

**FIG. 4.** Current 95% C.L. excluded region by ATLAS (solid lines) and CMS (dashed lines) for $m_{H_0} = 300$ GeV and respectively for 2HDM types I/II (left/right) in the case $\mu^2 = \mu^2_{\text{min}}/\mu^2 = \mu^2_{\text{max}}$ (top/bottom). The limits come from searches of $gg \to b\bar{b} \to H_0 \to WW/ZZ$ (black lines), $gg \to A_0 \to \tau\tau$ (red lines), $b\bar{b} \to A_0 \to \tau\tau$ (light-blue lines), $gg \to A_0 \to \gamma\gamma$ (yellow lines), $gg/b\bar{b} \to H_0 \to hh \to b\bar{b}\gamma\gamma$ (pink lines). The two scenarios considered are $m_{H_0} = m_{A_0}$ (degenerate; green exclusion region) and $m_{H_0} = 150$ GeV (hierarchical; purple exclusion region). The light/dark gray areas correspond to the theoretically excluded regions for the degenerate/hierarchical 2HDM scenarios. As discussed in the text, limits from a given search get progressively weaker with increasing mass splitting and are always nested within each other, starting from the degenerate scenario and moving toward increasingly hierarchical mass spectra.

---
unitarity, and perturbativity are significantly more important than for the previously discussed $\Delta^2 > 0$ case and in particular constrain the alignment limit $c_{\beta - \alpha} = 0$. The stability conditions for $c_{\beta - \alpha} = 0$ read

$$
\lambda_1 = \frac{m_h^2}{v^2} - t_\beta (1 + t_\beta^2) \frac{\Omega^2}{v^2} > 0 \\
\lambda_2 = \frac{m_h^2}{v^2} \frac{(1 + t_\beta^2) \Omega^2}{v^2} > 0 \\
\lambda_3 = \frac{m_h^2}{v^2} \frac{1}{s_\beta c_\beta} \frac{\Omega^2}{v^2} > -\sqrt{\lambda_1 \lambda_2} \\
\lambda_4 + |\lambda_5| = \frac{m_h^2}{v^2} + \frac{\Delta^2}{v^2} - \frac{1}{s_\beta c_\beta} \frac{\Omega^2}{v^2} - \frac{\Delta^2}{v^2} > -\sqrt{\lambda_1 \lambda_2}.
$$

(17)

The first three inequalities in (17) are trivially satisfied for $\Omega^2 \leq 0$. For $|\Delta^2| \gg v^2$, the last one, however, requires $\Omega^2 \sim s_\beta c_\beta \Delta^2$, and this affects the unitarity bounds which depend on $\lambda_1 + \lambda_2$, e.g.,

$$
\left| \lambda_1 + \lambda_2 + \sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_3^2} \right| \sim 2 \frac{|\Delta^2| t_\beta}{v^2} (t_\beta \gg 1) \\
\sim 2 \frac{|\Delta^2|}{v^2} t_\beta (t_\beta \ll 1) \quad (18)
$$

such that only values $t_\beta \sim 1$ are allowed if $|\Delta^2| \gg v^2$.

D. Overview of constraints on $H^\pm$

Before we comment on the limits from direct searches of $H^\pm$, let us emphasize that there are two other important sources of constraints on the mass of $H^\pm$ in this case: (i) Flavor physics yields important bounds on $m_{H^\pm}$, the most stringent one coming from the $H^\pm$ contribution to the flavor violating decay $b \to s\gamma$. For type II 2HDM, this leads to a lower bound $m_{H^\pm} > 480$ GeV at 95% C.L. [81], while
for type I the bound is milder and depends on $t_\beta$ [82].
(ii) EWPOs strongly prefer $m_{H^\pm} \sim m_{A_0}$ or $m_{H^\pm} \sim m_{H_0}$, (this last condition is mildly modified away from the alignment limit $c_{\beta-\alpha} = 0$), as a splitting between the charged and neutral components of the doublet breaks custodial symmetry. While some degree of splitting is allowed by EWPOs, it cannot be sizable (see e.g. the analysis of Ref. [31]). In the present work, we have chosen for simplicity to make $H^\pm$ degenerate with the heavier of the two neutral scalars $H_0$, $A_0$, as neglecting small mass splittings between the charged and neutral scalars does not have an appreciable impact on the theoretical constraints on the model, nor on the phenomenological analysis, and satisfies EWPOs. Regarding the bounds from flavor physics, while particularly for type II they motivate our choice of pairing $H^\pm$ with the heavier state among $H_0$ and $A_0$, we do not consider them as limits stricto sensu, meaning that for type II we still discuss scenarios in which both $m_{A_0}$ and $m_{H_0}$ are below 480 GeV. We also stress that, since for type I the $b \to s\gamma$ bound is not as severe, it could be possible for $H^\pm$ to pair with the lighter state among $H_0$ and $A_0$. For a hierarchical 2HDM scenario, this would also open either the decay $A_0 \to W^+H^\mp$ or $H_0 \to W^\pm H^\mp$ and would make the LHC limits from searches of $A_0$ and $H_0$ into SM states even weaker, opening at the same time further opportunities for direct searches of these new states (see e.g. Refs. [35,36]).

We now briefly discuss the current bounds from searches of $H^\pm$ by ATLAS and CMS. For a light $H^\pm$, $m_{H^\pm} < m_t = 173$ GeV, ATLAS searches for $t \to H^\pm b$ in top quark pair production with the full data set of run 1 [83,84] set a 95% C.L. bound on the branching fraction $\text{BR}(t \to H^\pm b) \times \text{BR}(H^\pm \to \nu\tau) < [0.0023, 0.013]$ in the mass range $m_{H^\pm} \in [80 \text{ GeV}, 160 \text{ GeV}]$. For $m_{H^\pm} > m_{A_0}$, ATLAS searches for $H^\pm$ produced in association with a top quark [84] yield the bound $\sigma(pp \to tH^\pm + X) \times \text{BR}(H^\pm \to \nu\tau) < [0.76 \text{ pb}, 4.5 \text{ fb}]$ in the range $m_{H^\pm} \in [180 \text{ GeV}, 1000 \text{ GeV}]$. We, however, note that these bounds do not result generically in meaningful constraints, since $\text{BR}(H^\pm \to \nu\tau) \ll 1$ when the decay $H^\pm \to tb$ is open. Moreover, in the hierarchical scenario, $\text{BR}(H^\pm \to \nu\tau)$ may be further suppressed by the presence of either $H^\pm \to W^\pm A_0$ or $H^\pm \to W^\pm H_0$ decays.

**E. Filling the gaps: $A_0 \to ZH_0/H_0 \to ZA_0$ searches**

Our previous analysis highlights that, while direct searches for heavy neutral Higgs bosons at the LHC may provide a wide coverage across the 2HDM parameter space, and complementary to measurements of signal strengths, bounds from searches assuming direct decays of the neutral scalars into SM particles become much weaker in a hierarchical 2HDM scenario, and new searches are needed to fill in the gaps. It is also clear that the new searches capable of probing a hierarchical 2HDM are precisely those which exploit the sizable mass splittings among the neutral scalars, namely $^2$ $H_0 \to ZA_0$ or $A_0 \to ZH_0$ (we note that if $H_0$ is the SM-like Higgs and $m_{H_0} = 125 \text{ GeV} > m_b$, then $A_0 \to Zh$ should be considered instead). For $H_0 \to ZA_0$, the dominant final state is generically $\ell\ell b\bar{b}$. In the latter case, the relevant final state to search for would depend on the dominant decay mode of $H_0$ [10]. In alignment, $H_0 \to b\bar{b}$ (eventually, $H_0 \to t\bar{t}$ if $m_{H_0} > 340 \text{ GeV}$) would dominate. For a sizable departure from alignment, the dominant decay mode would be $H_0 \to W^+W^-$, yielding $A_0 \to ZH_0 \to \ell\ell W^+W^- (W^+W^- \to \ell\ell\ell\ell)$ as the most sensitive final state.

In Refs. [37,38], the CMS Collaboration has performed the first analysis of such signatures, with an integrated luminosity of $L = 19.8 \text{ fb}^{-1}$ at 7 TeV, in the $\ell\ell b\bar{b}$ and $\ell\ell\tau\tau$ final states. We discuss here the limits on the 2HDM parameter space that may be derived from that search in the alignment limit, having checked that in this case the $\ell\ell b\bar{b}$ final state is significantly more sensitive than $\ell\ell\tau\tau$ across the whole parameter space. We stress that from the point of view of the CMS analysis, the limits on the production cross sections for $A_0 \to ZH_0$ and $H_0 \to ZA_0$ are identical for the same kinematical mass point. However, the translation between these limits and the constraints on the 2HDM parameter space is quite different in the two cases.

We first concentrate on $A_0 \to ZH_0$ ($A_0 \to Zh$ for $m_{H_0} = 125 \text{ GeV}$) and show in Fig. 6 the lower bounds on $t_\beta$ in the $(m_{A_0}, m_{H_0})$ plane for type I (top left) and type II (top right) from gluon fusion production of $A_0$. We assume the alignment limit, that is $c_{\beta-\alpha} = 0$ if $h$ is the SM-like Higgs and $s_{\beta-\alpha} = 0$ if $H_0$ to is instead the SM-like Higgs. The search constrains up to $t_\beta \sim 5$ around $m_{A_0} = 380 \text{ GeV}$ and additionally yields the limit $t_\beta \gtrsim 2$ for $m_t < 80 \text{ GeV}$ and $m_{A_0} < 600 \text{ GeV}$. Similarly, for type II, we can derive upper bounds on $t_\beta$ from the $b\bar{b}$-associated production of $A_0$, shown in Fig. 6 (bottom). We, however, expect a weakening of all these limits once there is departure from the alignment limit, and we emphasize that searches for the $H_0 \to W^+W^-$ and ZZ decay modes are very much needed in this region (note that for $t_\beta^1 s_{\beta-\alpha} - c_{\beta-\alpha} \sim 0$, direct searches for $H_0 \to W^+W^-$ and ZZ which assume gluon fusion production will not be sensitive to $H_0$).

In order to illustrate the complementarity between the above limits from $A_0 \to ZH_0$ searches and those from the most sensitive ATLAS/CMS searches for $A_0$ and $H_0$ decaying directly into SM states analyzed in Secs. III B and III C, as well as their interplay with measurements of

---

$^2$Other decay modes could also be promising, like $A_0/H_0 \to W^+H^- \nu\tau$ or $H^\pm \to W^\pm A_0/H_0$, depending on $m_{H^\pm}$ [35,36].

$^3$Other competitive final states are $H_0 \to W^+W^- (W^+W^- \to \ell\ell\ell\ell)$ and $H_0 \to ZZ$ yielding $A_0 \to ZH_0 \to \ell\ell\ell\ell\ell\ell$ [34].
Higgs signal strengths from Sec. III A (we take here the limits obtained with HIGGSIGNALS), we present a summary of the various bounds on the \((c_β−α, t_β)\) plane in Figs. 7 and 8 for types I/II (left/right): Fig. 7 (top) shows the combined limits for \(m_{A_0} = m_{H_0} = 150 \text{ GeV}\) (only the degenerate scenario is considered in this case). Focusing then on \(m_{A_0} = 300 \text{ GeV}\), Fig. 7 (bottom) highlights the fact that close to \(c_β−α = 0\) the CMS search for \(A_0 \rightarrow ZH_0 \rightarrow \ell\ell b\bar{b}\) in the hierarchical 2HDM scenario yields a superior sensitivity to the one obtained in the degenerate 2HDM scenario via the union of limits from \(A_0\) and \(H_0\) searches for low \(t_β\) (for high \(t_β\) in type II, the \(A_0/\ H_0 \rightarrow \tau\tau\) searches always have a superior sensitivity to \(A_0 \rightarrow ZH_0\), and we choose not to show the bounds from the latter in Figs. 7 and 8). It is also interesting to note that, while in the degenerate scenario the combination of \(A_0\) and \(H_0\) searches exclude the type II wrong-sign region allowed by Higgs signal strength measurements, in the hierarchical scenario, the wrong-sign region is allowed by direct searches.

We note that for \(m_{A_0} = 150 \text{ GeV}\) and \(m_{A_0} = 300 \text{ GeV}\), the choice between \(\mu^2 = \mu^2_{\text{min}}\) and \(\mu^2 = \mu^2_{\text{max}}\) does not impact the limits shown in Fig. 7 since the di-Higgs searches are not the most constraining in this plane. For \(m_{A_0} = 500 \text{ GeV}\), the situation is different, as shown in Fig. 8. Here, the degenerate case is the least constrained, cutting into the edges of the type I light Higgs limits and not significantly affecting the type II exclusions near alignment. In the \(\mu^2_{\text{max}}\) scenario, the di-Higgs searches improve the limits toward alignment around \(t_β \sim 1\). As one decreases the \(H_0\) mass, the picture changes considerably, with the direct \(H_0\) searches proving particularly effective for \(m_{H_0} = 300 \text{ GeV}\), even near alignment.

FIG. 6. Bounds on \(t_β\) in the \((m_{A_0}, m_{H_0})\) plane from the search for \(gg/ b\bar{b} \rightarrow A_0 \rightarrow Z\phi_0 \rightarrow \ell\ell b\bar{b}\) performed in Refs. [37,38], for 2HDM of type I (left) and type II (right). In each case, \(\phi_0 = H_0\) for \(m_{H_0} > m_h = 125 \text{ GeV}\), while \(\phi_0 = h\) for \(m_{H_0} = 125 \text{ GeV} > m_h\). The alignment limit is assumed. The gray regions are theoretically excluded.
The analysis of the signal strength measurements of Higgs signals, obtained with HiggsSignals (dotted black lines; see Sec. III A) and from the most sensitive ATLAS/CMS searches for $A_0$ and $H_0$ decaying directly into SM states: the green region corresponds to the exclusion in the degenerate scenario $m_{H_0} = m_{A_0}$; purple regions correspond to the exclusion in the hierarchical scenario (see Secs. III B and III C). The dashed blue region corresponds to the exclusion from the CMS $A_0 \to ZH_0 \to \ell\ell b \bar{b}$ search [37,38] in the hierarchical scenario. The gray regions are theoretically excluded.

Turning to $H_0 \to ZA_0$, we show the limits on $t\beta$ in the $(m_{A_0}, m_{H_0})$ plane in Fig. 9, for type I (left) and type II (right) in the alignment limit. A few comments are in order. First, the limits are expected to be weaker than for $A_0 \to ZH_0$, as the production cross section for $A_0$ is larger than that for $H_0$ for the same mass. More importantly, when $m_{H_0} > 2m_{A_0}$, the decay $H_0 \to A_0A_0$ becomes kinematically possible, which weakens the bounds from $H_0 \to ZA_0$ and also makes them dependent on $\mu^2$, since $\text{BR}(H_0 \to A_0A_0)$ does depend on this parameter. Figure 9 (top) shows the limits for $\mu^2 = \mu^2_{\text{min}}$, while Fig. 9 (bottom) shows the limits for $\mu^2 = \mu^2_{\text{max}}$ which are identical for $m_{H_0} < 2m_{A_0}$ but much weaker for $m_{H_0} > 2m_{A_0}$ as expected. Note also that for $t\beta = 1$, the limits are identical in both mass regions, since $\lambda_{H_0A_0A_0} = 0$ and the $\mu^2$ dependence therefore disappears. Additionally, a small allowed region appears around masses of $m_{A_0} = 100$ and $m_{H_0} = 260$ GeV. In the region $m_{H_0} > 2m_{A_0}$, the width of $H_0$ very quickly reaches $\Gamma_{H_0}/m_{H_0} > 0.15$, for which the bounds from the analysis [37,38] are no longer robust (these regions are shaded in Fig. 9). This is in contrast with $A_0 \to ZH_0$ bounds, for which $\Gamma_{A_0}/m_{A_0} < 0.15$ throughout the whole allowed parameter space. From the comparison of Figs. 6 and 9, it is also apparent that for a sizable splitting $m_{H_0} - m_{A_0} > 0$, the 2HDM parameter space is much more theoretically constrained than for a splitting.
\(m_{A_0} - m_{H_0} > 0\) of the same magnitude, and the constraints become more stringent as \(t_\beta\) increases (recall the discussion in Sec. III C). For example, for \(t_\beta > 2\ m_{H_0} \lesssim 500\) GeV is required, as shown in Fig. 9.

IV. HIERARCHICAL 2HDM AND LHC RUN II

LHC run 2 at 13 TeV represents a great opportunity to dig further into the parameter space of hierarchical 2HDM scenarios, since the sensitivity of the searches described in the previous section is limited mainly by small cross section values at the 8 TeV run of the LHC. While a detailed analysis of the LHC run 2 prospects for the hierarchical scenario of the 2HDM is beyond the scope of this work, we present in this section benchmark planes in \((m_{A_0}, m_{H_0})\) for \(A_0 \rightarrow ZH_0\) searches, classified according to the 2HDM type (I/II) and the proximity to the alignment limit.

In Fig. 10 (top), we provide \(\sigma(gg \rightarrow A_0 \rightarrow ZH_0) \times \text{BR}(H_0 \rightarrow X)\) for types I/II (left/right) and a reference value \(t_\beta = 3\), with \(X\) being the relevant decay mode of \(H_0\) in each case. In alignment \(c_{\beta-\alpha} = 0\) (Fig. 10, top), \(X = \bar{b}b\), and we choose \(c_{\beta-\alpha} = 0.55\) for type II and \(c_{\beta-\alpha} = 0.3\) for type I.

4We note that, while this benchmark point is excluded by the LILITH likelihood fit to Higgs signal strengths, it is allowed by the corresponding fit by HiggsSignals (see Fig. 1) as well as the ATLAS/CMS fits [11,12], and so we choose to consider it.
We show in each case the constraint from the LHC run 1 $A_0 \to ZH_0 \to Z\bar{b}b$ CMS search [37,38], noting that, besides providing useful limits in alignment (recall Fig. 6), it can also constrain the $(m_{A_0}, m_{H_0})$ plane away from alignment. This is most relevant in type II, where $\kappa_{H_0} = 1$ increases with $t_\beta$, and for $m_{H_0} \lesssim 180$ GeV as shown in Fig. 10 (see also Figs. 7 and 8). In contrast, for the benchmarks chosen away from alignment, there are no limits from $g g / b \bar{b} \to H_0 \to Z A_0 \to \ell \ell b \bar{b}$ performed in Refs. [37,38], for 2HDM of type I (left) and type II (right) and for $\mu_2 = \mu_2^{\text{min}} / \mu_2^{\text{max}}$ (top/bottom, respectively). The gray regions are theoretically excluded, while the shaded regions correspond to $\Gamma_{H_0} / m_{H_0} > 0.15$, which goes beyond the narrow width assumption made in the analysis. Excluded regions are therefore the areas between the colored contours which lie outside of the corresponding shaded area delimited by the dot-dashed line of the same color. The dotted-black line corresponds to $m_{H_0} = 2m_{A_0}$, above which the decay $H_0 \to A_0 A_0$ becomes kinematically possible.

![FIG. 9. Bounds on $t_\beta$ in the $(m_{A_0}, m_{H_0})$ plane for $c_{\beta-\alpha} = 0$, from the search for $g g / b \bar{b} \to H_0 \to Z A_0 \to \ell \ell b \bar{b}$ performed in Refs. [37,38], for 2HDM of type I (left) and type II (right) and for $\mu_2 = \mu_2^{\text{min}} / \mu_2^{\text{max}}$ (top/bottom, respectively). The gray regions are theoretically excluded, while the shaded regions correspond to $\Gamma_{H_0} / m_{H_0} > 0.15$, which goes beyond the narrow width assumption made in the analysis. Excluded regions are therefore the areas between the colored contours which lie outside of the corresponding shaded area delimited by the dot-dashed line of the same color. The dotted-black line corresponds to $m_{H_0} = 2m_{A_0}$, above which the decay $H_0 \to A_0 A_0$ becomes kinematically possible.](image-url)
V. DISCUSSION AND OUTLOOK

Uncovering the full structure of the SM scalar sector and its possible extensions will be a central task for the LHC in the coming years. The results will have important implications not only for our understanding of the mechanism of electroweak symmetry breaking but also for the origin of visible matter and the nature of dark matter. Extensions of the SM scalar sector that address one or both of these open questions may yield distinctive signatures at the LHC via modifications of the SM Higgs boson properties and/or the observation of new states.

In this work, we have investigated the constraints on the parameter space of CP-conserving two-Higgs-doublet models of types I/II in light of the ATLAS/CMS results from LHC run 1. A key difference from the many similar analyses already existing in the literature is that the latter generally assume a nearly degenerate 2HDM spectrum for the new scalar states, which can then only decay into SM particles. While the properties of the observed 125 GeV Higgs are not affected by the mass spectrum of the new scalars (as discussed in Sec. III A), a large mass splitting between two or more of the new scalar states,
e.g. $m_{A_0} - m_{H_0} \gtrsim m_Z$, causes new decay channels of the heavier scalars to open and become dominant. For such a hierarchical 2HDM, we show that the constraints usually obtained in the literature are significantly weakened. On the other hand, the new decay channels constitute novel ways of searching for these scalar states, e.g. $A_0 \rightarrow ZH_0$, and we show how they can be used to fill in the gaps left by previous analyses. We also highlight the importance of the $\mu^2$ parameter, through its impact on the phenomenology of the heavier $CP$-even scalar $H_0$ and its sensitivity to unitarity and stability constraints.

ACKNOWLEDGMENTS

We thank Jeremy Bernon for useful discussions. The work of S.H. and K.M. is supported by the Science Technology and Facilities Council under Grant No. ST/L000504/1. J. M. N. is supported by the People Programme (Marie Curie Actions) of the European Union Seventh Framework Programme (FP7/2007-2013) under REA Grant No. PIF-EA-2013-625809. G. C. D. is supported by the German Science Foundation under the Collaborative Research Center (SFB) 676 Particles, Strings and the Early Universe.
HIERARCHICAL VERSUS DEGENERATE 2HDM: THE LHC …