Theory and phenomenology of two-Higgs-doublet models

G.C. Branco [a], P.M. Ferreira [b,c,*], L. Lavoura [a], M.N. Rebelo [a], Marc Sher [d], João P. Silva [b,e]

[a] Departamento de Física and Centro de Física Teórica de Partículas, Instituto Superior Técnico, Technical University of Lisbon, 1049-001 Lisboa, Portugal

[b] Instituto Superior de Engenharia de Lisboa, 1959-007 Lisboa, Portugal

[c] Centro de Física Teórica e Computacional, University of Lisbon, 1649-003 Lisboa, Portugal

[d] High Energy Theory Group, College of William and Mary, Williamsburg, VA 23187, USA

[e] Centro de Física Teórica de Partículas, Instituto Superior Técnico, Technical University of Lisbon, 1049-001 Lisboa, Portugal

ARTICLE INFO

ABSTRACT

We discuss theoretical and phenomenological aspects of two-Higgs-doublet extensions of the Standard Model. In general, these extensions have scalar mediated flavour changing neutral currents which are strongly constrained by experiment. Various strategies are discussed to control these flavour changing scalar currents and their phenomenological consequences are analysed. In particular, scenarios with natural flavour conservation are investigated, including the so-called type I and type II models as well as lepton-specific and inert models. Type III models are then discussed, where scalar flavour changing neutral currents are present at tree level, but are suppressed by either a specific ansatz for the Yukawa couplings or by the introduction of family symmetries leading to a natural suppression mechanism. We also consider the phenomenology of charged scalars in these models. Next we turn to the role of symmetries in the scalar sector. We discuss the six symmetry-constrained scalar potentials and their extension into the fermion sector. The vacuum structure of the scalar potential is analysed, including a study of the vacuum stability conditions on the potential and the renormalization-group improvement of these conditions is also presented. The stability of the tree level minimum of the scalar potential in connection with electric charge conservation and its behaviour under CP is analysed. The question of CP violation is addressed in detail, including the cases of explicit CP violation and spontaneous CP violation. We present a detailed study of weak basis invariants which are odd under CP. These invariants allow for the possibility of studying the CP properties of any two-Higgs-doublet model in an arbitrary Higgs basis. A careful study of spontaneous CP violation is presented, including an analysis of the conditions which have to be satisfied in order for a vacuum to violate CP. We present minimal models of CP violation where the vacuum phase is sufficient to generate a complex CKM matrix, which is at present a requirement for any realistic model of spontaneous CP violation.

© 2012 Elsevier B.V. All rights reserved.

Contents

---

* Corresponding author at: Centro de Física Teórica e Computacional, University of Lisbon, 1649-003 Lisboa, Portugal.
  E-mail addresses: gbranco@ist.utl.pt (G.C. Branco), pedrommferreira@gmail.com, ferreira@cii.fc.ul.pt (P.M. Ferreira), balio@cftp.ist.utl.pt (L. Lavoura), rebelo@ist.utl.pt (M.N. Rebelo), mtsher@wm.edu (M. Sher), jpsilva@cftp.ist.utl.pt (J.P. Silva).

1. Introduction

The gauge boson and fermion sectors of the Standard Model of the electroweak interactions have been extremely well probed phenomenologically; yet, its scalar sector has not yet been directly explored. In the Standard Model (SM) the simplest possible scalar structure – just one $SU(2)$ doublet – is assumed [1–5]; on the contrary, the fermion structure, with more than one family and with family mixing, is not simple at all.

One critical piece of evidence about the scalar structure is the parameter ρ. In the $SU(2) \times U(1)$ gauge theory, if there are n scalar multiplets ϕ_i, with weak isospin I_i, weak hypercharge Y_i, and vacuum expectation value (vev) of the neutral components v_i, then the parameter ρ is, at tree level [6],

$$\rho = \frac{\sum\limits_{i=1}^{n} \left[I_i \left(I_i + 1 \right) - \frac{1}{4} Y_i^2 \right] v_i}{\sum\limits_{i=1}^{n} \frac{1}{2} Y_i^2 v_i}. \tag{1}$$

Experimentally [7] ρ is very close to one. According to Eq. (1), both $SU(2)$ singlets with $Y = 0$ and $SU(2)$ doublets with $Y = \pm 1$ give $\rho = 1$, since they both have $I(I+1) = \frac{3}{4} Y^2$. Other scalars with vevs in much larger $SU(2)$ multiplets, scalars with small or null vevs, and models with triplets and a custodial $SU(2)$ global symmetry [8], are compatible with $\rho = 1$; but such scalar sectors tend to be large and complex—the simplest extension of the SM consists in simply adding scalar doublets and singlets.

In this review we focus on one of the simplest possible extensions of the SM—the two-Higgs-doublet model (2HDM) [9]. There are many motivations for 2HDMs. The best known motivation is supersymmetry [10]. In supersymmetric theories the scalars belong to chiral multiplets and their complex conjugates belong to multiplets of the opposite chirality; since multiplets of different chiralities cannot couple together in the Lagrangian, a single Higgs doublet is unable to give mass simultaneously to the charge 2/3 and charge −1/3 quarks. Moreover, since scalars sit in chiral multiplets together with chiral spin-1/2 fields, the cancellation of anomalies also requires that an additional doublet be added. Thus, the Minimal Supersymmetric Standard Model (MSSM) contains two Higgs doublets.

Another motivation for 2HDMs comes from axion models [11]. Peccei and Quinn [12] noted that a possible CP-violating term in the QCD Lagrangian, which is phenomenologically known to be very small, can be rotated away if the Lagrangian contains a global $U(1)$ symmetry. However, imposing this symmetry is only possible if there are two Higgs doublets. While the simplest versions of the Peccei–Quinn model (in which all the New Physics was at the TeV scale) are experimentally ruled out, there are variations with singlets at a higher scale that are acceptable, and the effective low-energy theory for those models still requires two Higgs doublets [11].

Still another motivation for 2HDMs is the fact that the SM is unable [13] to generate a baryon asymmetry of the Universe of sufficient size. Two-Higgs-doublet models can do so, due to the flexibility of their scalar mass spectrum [13] and the existence of additional sources of CP violation. There have been many works on baryogenesis in the 2HDM [14–23]. Exciting new possibilities for explicit or spontaneous CP violation constitute one of the attractive features of 2HDMs.

With the Large Hadron Collider (LHC) starting to produce data, time seems appropriate for a review of 2HDMs. The Higgs sector of the Standard Model is very predictive, with the Higgs mass being the only free parameter, and it will be tested at the LHC over the entire theoretically preferred mass ranges within the next few months. In contrast, due to the larger number of free parameters in the 2HDM, it will take much longer to probe the entire parameter space of the various models. Should the Higgs not be seen at the LHC in the next few months, the 2HDM will be one of the simplest alternatives. With charged Higgs bosons, pseudoscalars and different decay modes and branching ratios, the experimental challenges will be quite different than in the Standard Model. While it may not be possible to completely probe the entire parameter space of the various 2HDMs at the LHC, most of the parameter space can be probed, and this is further incentive for a review of the various forms of the 2HDM and their experimental signatures.

We shall explicitly exclude supersymmetric models from this review. The Higgs sector of supersymmetric models is extremely well-studied and Djouadi [24] has written a very comprehensive review of it. We shall also not include models with scalar $SU(2)$ singlets in addition to the two doublets, since those models usually include many additional parameters.

In general, the vacuum structure of 2HDMs is very rich. The most general scalar potential contains 14 parameters and can have CP-conserving, CP-violating, and charge-violating minima. In writing that potential one must be careful in defining the

various bases and in distinguishing parameters which can be rotated away from those which have physical implications. However, most phenomenological studies of 2HDMs make several simplifying assumptions. It is usually assumed that CP is conserved in the Higgs sector (only then can one distinguish between scalars and pseudoscalars), that CP is not spontaneously broken, and that discrete symmetries eliminate from the potential all quartic terms odd in either of the doublets; however, usually one considers all possible real quadratic coefficients, including a term which softly breaks these symmetries. We shall also make those assumptions in the early chapters of this report but will subsequently discuss relaxing them. Under those assumptions, the most general scalar potential for two doublets Φ_1 and Φ_2 with hypercharge $+1$ is

$$V = m_{11}^2 \, \Phi_1^\dagger \Phi_1 + m_{22}^2 \, \Phi_2^\dagger \Phi_2 - m_{12}^2 \left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) + \frac{\lambda_1}{2} \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{\lambda_2}{2} \left(\Phi_2^\dagger \Phi_2 \right)^2$$

$$+ \lambda_3 \, \Phi_1^\dagger \Phi_1 \, \Phi_2^\dagger \Phi_2 + \lambda_4 \, \Phi_1^\dagger \Phi_2 \, \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} \left[\left(\Phi_1^\dagger \Phi_2 \right)^2 + \left(\Phi_2^\dagger \Phi_1 \right)^2 \right], \tag{2}$$

where all the parameters are real. For a region of parameter space, the minimization of this potential gives

$$\langle \Phi_1 \rangle_0 = \begin{pmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{pmatrix}, \qquad \langle \Phi_2 \rangle_0 = \begin{pmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{pmatrix}. \tag{3}$$

With two complex scalar $SU(2)$ doublets there are eight fields:

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, \quad a = 1, 2. \tag{4}$$

Three of those get 'eaten' to give mass to the $W^\pm$ and Z^0 gauge bosons; the remaining five are physical scalar ('Higgs') fields. There is a charged scalar, two neutral scalars, and one pseudoscalar. With the above minimum, the mass terms for the charged scalars are given by

$$\mathcal{L}_{\phi^\pm \text{ mass}} = \left[m_{12}^2 - (\lambda_4 + \lambda_5)v_1 v_2 \right] \left(\phi_1^-, \quad \phi_2^- \right) \begin{pmatrix} \frac{v_2}{v_1} & -1 \\ -1 & \frac{v_1}{v_2} \end{pmatrix} \begin{pmatrix} \phi_1^+ \\ \phi_2^+ \end{pmatrix}. \tag{5}$$

There is a zero eigenvalue corresponding to the charged Goldstone boson $G^\pm$ which gets eaten by the $W^\pm$. The mass-squared of the 'charged Higgs' is $m_+^2 = \left[m_{12}^2/(v_1 v_2) - \lambda_4 - \lambda_5 \right] \left(v_1^2 + v_2^2 \right)$. The mass terms for the pseudoscalars are given by

$$\mathcal{L}_{\eta \text{ mass}} = \frac{m_A^2}{v_1^2 + v_2^2} \left(\eta_1, \quad \eta_2 \right) \begin{pmatrix} v_2^2 & -v_1 v_2 \\ -v_1 v_2 & v_1^2 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}. \tag{6}$$

This gives a pseudoscalar Goldstone mode together with the mass-squared of the physical pseudoscalar, $m_A^2 = \left[m_{12}^2/(v_1 v_2) - 2\lambda_5 \right](v_1^2 + v_2^2)$. Note that, when $m_{12}^2 = 0$ and $\lambda_5 = 0$, the pseudoscalar becomes massless. This is due to the existence, in that limit, of an additional global $U(1)$ symmetry which is spontaneously broken. Finally, the mass terms for the scalars are given by

$$\mathcal{L}_{\rho \text{ mass}} = -\left(\rho_1, \quad \rho_2 \right) \begin{pmatrix} m_{12}^2 \frac{v_2}{v_1} + \lambda_1 v_1^2 & -m_{12}^2 + \lambda_{345} v_1 v_2 \\ -m_{12}^2 + \lambda_{345} v_1 v_2 & m_{12}^2 \frac{v_1}{v_2} + \lambda_2 v_2^2 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \tag{7}$$

with $\lambda_{345} = \lambda_3 + \lambda_4 + \lambda_5$. The mass-squared matrix of the scalars can be diagonalized and the angle α is defined to be the rotation angle that performs that diagonalization.

Perhaps the single most important parameter in studies of 2HDMs is

$$\tan \beta \equiv \frac{v_2}{v_1}. \tag{8}$$

The angle β is the rotation angle which diagonalizes the mass-squared matrices of the charged scalars and of the pseudoscalars. If one redefines the doublets as $H_1 = \cos \beta \, \Phi_1 + \sin \beta \, \Phi_2$ and $H_2 = -\sin \beta \, \Phi_1 + \cos \beta \, \Phi_2$, one finds that the lower component of H_1 has a (real and positive) vev $v/\sqrt{2}$, where $v \equiv \left(v_1^2 + v_2^2 \right)^{1/2}$, while H_2 has null vev. The two parameters α and β determine the interactions of the various Higgs fields with the vector bosons and (given the fermion masses) with the fermions; they are thus crucial in discussing phenomenology. Still, one should keep in mind the assumptions that were made in defining them.

In this review, we shall begin by discussing the phenomenology of the above restricted version of the 2HDM. A feature of general 2HDMs is the existence of tree-level flavour-changing neutral currents (FCNC). One can avoid these potentially dangerous interactions by imposing discrete symmetries in several possible ways. In Section 2, the phenomenological analyses of 2HDMs without tree-level FCNCs are presented, including decays and production of neutral scalars and pseudoscalars, bounds from LEP and the Tevatron, and expectations for the LHC.

If one does not impose discrete symmetries, then there are tree-level FCNCs in the 2HDM. A prototype model of this kind, the so-called type III 2HDM, is discussed in Section 3, along with other such models, such as the Branco–Grimus–Lavoura (BGL) model and models of Minimal Flavour Violation.

All 2HDMs have charged scalar bosons ('charged Higgses'). An analysis of charged-Higgs production and decay, for all the 2HDMs of Section 2, is presented in Section 4; this is followed by a discussion of charged-Higgs phenomenology in models with tree-level FCNC.

In Section 5 we relax the rather strict assumptions made in this introduction and in Sections 2–4. An analysis of the full scalar potential, including the vacuum structure with or without CP violation and the possible symmetries one can impose on the potential, is presented.

As noted earlier, the 2HDM offers possibilities for new sources of CP violation. This is analysed in some detail in Section 6. In particular we discuss the weak-basis invariant conditions for the 2HDM Lagrangian to be CP invariant and we present a minimal realistic extension of the Standard Model, with spontaneous CP violation.

In Section 8 we briefly summarize some of our conclusions. Some isolated topics that the reader may want to consult separately – constraints on the parameters of the potential from unitarity, renormalization-group running of the parameters of 2HDMs, and contributions to the oblique parameters from the scalar sector of 2HDMs – are left to appendices.

2. Models with natural flavour conservation

The most serious potential problem facing all 2HDMs[1] is the possibility of tree level flavour-changing neutral currents(FCNC). For example, the Yukawa couplings of the $Q = -1/3$ quarks will, in general, be

$$\mathcal{L}_Y = y_{ij}^1 \bar{\psi}_i \psi_j \Phi_1 + y_{ij}^2 \bar{\psi}_i \psi_j \Phi_2, \tag{9}$$

where i, j are generation indices. The mass matrix is then

$$M_{ij} = y_{ij}^1 \frac{v_1}{\sqrt{2}} + y_{ij}^2 \frac{v_2}{\sqrt{2}}. \tag{10}$$

In the Standard Model, diagonalizing the mass matrix automatically diagonalizes the Yukawa interactions, therefore there are no tree-level FCNC. In 2HDMs, however, in general y^1 and y^2 will not be simultaneously diagonalizable, and thus the Yukawa couplings will not be flavour diagonal. Neutral Higgs scalars ϕ will mediate FCNC of the form, for example, $\bar{d}s\phi$.

These FCNC can cause severe phenomenological difficulties. The $\bar{d}s\phi$ interaction, for example, will lead to K–$\bar{K}$ mixing at tree level. If the coupling is as large as the b-quark Yukawa coupling, the mass of the exchanged scalar would have to exceed 10 TeV [25,26]. Nonetheless, under reasonable assumptions, models with these FCNC may still be viable. They will be discussed in the next chapter. In this chapter, however, we will assume that tree level FCNC are completely absent, due to a discrete or continuous symmetry.

It is easy to see that if all fermions with the same quantum numbers (which are thus capable of mixing) couple to the same Higgs multiplet, then FCNC will be absent. This was formalized by the Paschos–Glashow–Weinberg theorem [27,28] which states that a necessary and sufficient condition for the absence of FCNC at tree level is that all fermions of a given charge and helicity transform according to the same irreducible representation of $SU(2)$, correspond to the same eigenvalue of T_3 and that a basis exists in which they receive their contributions in the mass matrix from a single source. In the Standard Model with left-handed doublets and right-handed singlets, this theorem implies that all right-handed quarks of a given charge must couple to a single Higgs multiplet. In the 2HDM, this can only be ensured by the introduction of discrete or continuous symmetries.

Looking at the quark sector of the 2HDM, there are only two possibilities. In the type I 2HDM, all quarks couple to just one of the Higgs doublets (conventionally chosen to be Φ_2). In the type II 2HDM, the $Q = 2/3$ right-handed (RH) quarks couple to one Higgs doublet (conventionally chosen to be Φ_2) and the $Q = -1/3$ RH quarks couple to the other (Φ_1). The type I 2HDM can be enforced with a simple $\Phi_1 \to -\Phi_1$ discrete symmetry, whereas the type II 2HDM is enforced with a $\Phi_1 \to -\Phi_1, d_R^i \to -d_R^i$ discrete symmetry. Note that the original Peccei–Quinn models as well as supersymmetric models give the same Yukawa couplings as in a type II 2HDM, but do it by using continuous symmetries.

We will in this chapter consider that there is no CP violation in the vacuum expectation values (vevs) of the scalar doublets $\Phi_{1,2}$. This means that $v_{1,2}$ will be assumed to be both real and (without loss of generality) non-negative. Thus

$$\Phi_j = \begin{pmatrix} \phi_j^+ \\ \left(v_j + \rho_j + i\eta_j\right)/\sqrt{2} \end{pmatrix}, \tag{11}$$

with $v_1 = v \cos \beta$ and $v_2 = v \sin \beta$. Then, the neutral Goldstone boson is $G^0 = \eta_1 \cos \beta + \eta_2 \sin \beta$. The linear combination of the η_j orthogonal to G^0 is the physical pseudoscalar

$$A = \eta_1 \sin \beta - \eta_2 \cos \beta. \tag{12}$$

---

[1] All multi-Higgs-doublet models in general face this potential problem.

Table 1
Models which lead to natural flavour conservation. The superscript
i is a generation index. By convention, the u_R^i always couple to Φ_2.

Model	u_R^i	d_R^i	e_R^i
Type I	Φ_2	Φ_2	Φ_2
Type II	Φ_2	Φ_1	Φ_1
Lepton-specific	Φ_2	Φ_2	Φ_1
Flipped	Φ_2	Φ_1	Φ_2

The physical scalars are a lighter h and a heavier H, which are orthogonal combinations of ρ_1 and ρ_2:

$$h = \rho_1 \sin\alpha - \rho_2 \cos\alpha, \tag{13}$$
$$H = -\rho_1 \cos\alpha - \rho_2 \sin\alpha. \tag{14}$$

Notice that the Standard-Model Higgs boson would be

$$\begin{aligned} H^{\text{SM}} &= \rho_1 \cos\beta + \rho_2 \sin\beta \\ &= h \sin(\alpha - \beta) - H \cos(\alpha - \beta). \end{aligned} \tag{15}$$

As shown by Carena and Haber [29], one can, without loss of generality, assume that β is in the first quadrant, *i.e.* that both v_1 and v_2 are non-negative real; also, one can add π to α, *i.e.* invert the sign of both the h and H fields, without affecting any physics. In the tree-level MSSM, α is in the fourth quadrant, but this is not the case in the general 2HDM, therefore we will choose α to be either in the first or the fourth quadrant. We will choose our independent variables to be $\tan\beta$ and α, which are single valued over the allowed range.

It is conventionally assumed, in discussions of type I and type II 2HDMs, that the right-handed leptons satisfy the same discrete symmetry as the d_R^i and thus the leptons couple to the same Higgs boson as the $Q = -1/3$ quarks. However, the Glashow–Weinberg theorem does not require this, and there are two other possibilities. In the "lepton-specific" model, the RH quarks all couple to Φ_2 and the RH leptons couple to Φ_1. In the "flipped" model, one has the $Q = 2/3$ RH quarks coupling to Φ_2 and the $Q = -1/3$ RH quarks coupling to Φ_1, as in the type II 2HDM, but now the RH leptons couple to Φ_2. The phenomenology of these models is, as we will see, quite different. In one of the earliest papers [30], the names "Model III" and "Model IV" were used for the flipped and lepton-specific models, respectively. The term "Model III", however, has become associated with the 2HDM with tree-level FCNCs (the subject of Chapter III). In other early papers [31–33], the terms "Model I" and "Model II" were used for the lepton-specific and flipped models respectively, and in even earlier works [34,35], the terms IIA and IIB were used. More recently [36], the terms type X and type Y were used for the lepton-specific and flipped models. The four models which lead to natural flavour conservation are presented in Table 1. It is straightforward to find a Z_2 symmetry which will ensure that only these interactions exist.

In a somewhat related work, Pich and Tuzón [37,38] simply assumed that the Yukawa coupling matrices of Φ_1 and Φ_2 in flavour space are proportional. This then eliminates all tree-level FCNC, and gives three arbitrary proportionality constants. Note that this assumption is *ad hoc* and, in general, is not radiatively stable [39]—one would obtain FCNC couplings being generated radiatively, as was analysed recently in Ref. [40]. However, Serôdio has recently proposed a UV completion of the Pich–Tuzón model [41]. Varzielas [42] has studied how family symmetries in multi-Higgs doublet models may give a justification for the alignment hypothesis. Each of the four models (as well as the Inert Doublet model discussed later) then arises as a specific choice of the proportionality constant (and only these choices allow for a symmetry [39]). Another recent, very general, formulation in which the various models are special cases is shown in Ref. [43]. One should keep in mind that even if a 2HDM without FCNC is correct, it will take some time to determine all of the couplings to determine which 2HDM it is, and the Pich–Tuzon parametrization might be a valuable guide for phenomenologists. In addition, the Pich–Tuzon parametrization might arise in other models; for example, the three doublet model of Cree and Logan [44] reproduces the Pich–Tuzon model in its charged Higgs Yukawa couplings. Of particular interest is the fact that if the proportionality constants are complex, one has CP violating effects. It has been noted [38,45] that loop corrections induce flavour changing currents of the Minimal Flavour Violation form, and bounds on the charged-Higgs mass were discussed. A similar approach was recently used by Mahmoudi and Stal [46], who studied the constraints on the charged-Higgs mass from meson decays and FCNC transitions, using a more general model-independent approach, getting results in the four models as special cases.

The Yukawa couplings can now be determined. In the Standard Model, the coupling of the fermion f to the Higgs boson is m_f/v. Following the notation of Aoki et al. [36], we define the parameters ξ_h^f, ξ_H^f, ξ_A^f through the Yukawa Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^{\text{2HDM}} = &- \sum_{f=u,d,\ell} \frac{m_f}{v} \left(\xi_h^f \bar{f} f h + \xi_H^f \bar{f} f H - i\xi_A^f \bar{f} \gamma_5 f A \right) \\ &- \left\{ \frac{\sqrt{2} V_{ud}}{v} \bar{u} \left(m_u \xi_A^u P_L + m_d \xi_A^d P_R \right) d H^+ + \frac{\sqrt{2} m_\ell \xi_A^\ell}{v} \overline{\nu_L} \ell_R H^+ + \text{H.c.} \right\} \end{aligned} \tag{16}$$

where $P_{L/R}$ are projection operators for left-/right-handed fermions, and the factors ξ are presented in Table 2.

Table 2
Yukawa couplings of u, d, ℓ to the neutral Higgs bosons h, H, A in the four different models. The couplings to the charged Higgs bosons follow Eq. (16).

	Type I	Type II	Lepton-specific	Flipped
ξ_h^u	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
ξ_h^d	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_h^ℓ	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$
ξ_H^u	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
ξ_H^d	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
ξ_H^ℓ	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$
ξ_A^u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
ξ_A^d	$-\cot\beta$	$\tan\beta$	$-\cot\beta$	$\tan\beta$
ξ_A^ℓ	$-\cot\beta$	$\tan\beta$	$\tan\beta$	$-\cot\beta$

In all models, the coupling of the neutral Higgs bosons to the W and Z are the same: the coupling of the light Higgs, h, to either WW or ZZ is the same as the Standard-Model coupling times $\sin(\beta-\alpha)$ and the coupling of the heavier Higgs, H, is the same as the Standard-Model coupling times $\cos(\alpha-\beta)$. The coupling of the pseudoscalar, A, to vector bosons vanishes.

In this section, we will summarize some of the work done on these four models, and will follow with a more detailed discussion in the following sections.

There are relatively few studies which directly compare all four models. One of the earliest papers to mention all four models was by Barger et al. [30], who studied the charged-Higgs phenomenology but assumed fairly light top quarks. The famous Higgs Hunter's Guide [47] mentions all four, but concentrates only on the type I and type II 2HDMs. Grossman [31] also discusses all four models, but focuses on models with more than two doublets, and concentrates on the charged Higgs sector. Akeroyd has several papers in which all four models are discussed. In an early paper with Stirling [32], the phenomenology of the charged Higgs boson at LEP2 was analysed in each model, and this was followed [33] by a study of the neutral sector at LEP2. In addition, he looked [49] at LHC phenomenology in all four models, focusing in particular on the Higgs branching ratios to $\gamma\gamma$ and $\tau\tau$. More recently, Barger et al. [50] performed a comprehensive analysis of the couplings in all models with natural flavour conservation, including doublets and singlets; the four models appear as special cases.

There are two recent papers comparing Higgs decays in all four models. Aoki et al. [36] study the decays of the Higgs bosons in each model, summarize current phenomenological constraints and look at methods of distinguishing the models at colliders, although they focus on the type II and lepton-specific models and assume that the heavy Higgs bosons are not too heavy (typically with masses below 200 GeV). Arhrib et al. [51] study the decays of the light Higgs in each model, although the main point of their work concerns double-Higgs production at the LHC.

Recently, a new computer code was written by Eriksson et al. [52]. The code allows one to input any of the different Z_2 symmetries, or even more general couplings, and calculates all two-body and some three-body Higgs boson decays, and the oblique parameters S, T and U and other collider constraints.

The least studied model is the flipped model (the word was coined in Ref. [50]); even works that discuss all four models generally focus less on this structure than the others. The only paper dedicated entirely to the flipped model was the very recent article of Logan and MacLennan [53]. They studied the charged-Higgs phenomenology in that model, including branching ratios and indirect constraints and analyse prospects at the LHC.

The lepton-specific model was first discussed in two papers by Barnett et al. [34,35] in the context of extremely light Higgs scalars. The model was recently analysed extensively by Su and Thomas [54]. They studied theoretical and experimental constraints on the model and showed that there can be substantial enhancement of the couplings between the charged leptons and the neutral Higgs scalar. Logan and MacLennan [55] considered the constraints on the charged-Higgs mass, with bounds arising from lepton flavour universality and direct searches, and discuss prospects at the LHC. Goh et al. [56] discussed whether a lepton-specific model could explain the leptonic cosmic-ray signals seen by PAMELA and ATIC, and studied the implications for the LHC. A very recent analysis of this possibility, including analyses of astrophysical results and direct dark matter detection, can be found in the work of Boucenna and Profumo [57]. In another analysis, Cao et al. [58] assumed that the 3σ discrepancy [59] between theory and experiment in the $g-2$ of the muon is primarily due to the lepton-specific model—this requirement substantially reduces the available parameter space, forcing the model to have a very light pseudoscalar and very large values of $\tan\beta$, and they analysed this parameter space. Finally, Aoki et al. [60] looked at neutrino masses and dark matter in the lepton-specific model, but did add singlets.

The type I 2HDM [61] is the second most studied.[2] In the quark sector, it is identical to the lepton-specific model, thus many results from studies of the type I 2HDM apply to the lepton-specific as well. A special limit of the type I 2HDM

---

[2] Recent developments in string phenomenology [62] suggest that a type I 2HDM is generic among the vacua of the heterotic string, providing new motivation for study of this model.

is $\alpha = \pi/2$, in which case the fermions all completely decouple from the lightest Higgs; this limit is referred to as the fermiophobic limit. Note that even in this limit, the coupling does reappear at the one-loop level, but it will in any event be very small. Later in this chapter, it will be shown that in the inert doublet models, this limit can be obtained exactly. The earliest discussions of the fermiophobic limit in the context of imminent Tevatron data were those of Stange et al. [63], of Diaz and Weiler [64] and of Barger et al. [65], who looked at Higgs production and decay through photon loops. Shortly thereafter, Akeroyd [66] studied the other phenomenological implications during the early Tevatron runs. Brücher and Santos [67] mentioned all four models, but then focused on the fermiophobic limit of the type I 2HDM, studying the decays of the various Higgs bosons and the constraints on the model from LEP 2. The same authors together with Barroso [68,69] used the possibility of a fermiophobic Higgs to look at $h \rightarrow \gamma\gamma$ in two different versions of the type I 2HDM, showing that this decay could distinguish between them.

Moving away from the fermiophobic limit, there are many papers looking at the type I 2HDM. A very early discussion of the fermiophobic, gauge-phobic and fermiophilic limits was given by Pois et al. [70], who studied top production below the $t\bar{t}$ threshold in Higgs decays. The Higgs Hunter's Guide [47] contains some analysis (although they focus on the type II 2HDM and the MSSM), but this was before the top mass was known to be very heavy. In a series of papers, Akeroyd and collaborators [48,71–74] considered Higgs decays into lighter Higgs bosons, charged-Higgs decays into a W and a pseudoscalar, double-Higgs production, Higgs decays to $\gamma\gamma$, $\tau^+\tau^-$ at the LHC, and the possibility of a very light Higgs, respectively, all in the context of the type I 2HDM. There have also been studies of the contribution to the anomalous magnetic moment of the muon in the type I 2HDM [75,76].

The type II 2HDM is by far the most studied, since it is the structure present in supersymmetric models. A voluminous Physics Reports review article in 2008 by Djouadi [24] analyses the Higgs bosons of the MSSM in great detail. We will not discuss the MSSM Higgs structure and phenomenology in this work, and refer the reader to Djouadi's review article. Here, we will only focus on the differences between the general type II 2HDM and the MSSM.

The most crucial difference is that the general type II 2HDM does not have a strict upper bound on the mass of the lightest Higgs boson, which is an important characteristic of the MSSM. In addition, the scalar self-couplings are now arbitrary. Another important difference is that the mixing parameter α, which in the MSSM is given in terms of $\tan\beta$ and the scalar and pseudoscalar masses, is now arbitrary. Finally, in the MSSM the charged-scalar and pseudoscalar masses are so close that the decay of the charged Higgs into a pseudoscalar and a real W is kinematically forbidden, while it is generally allowed in the type II 2HDM (although see Refs. [77,78] for possible exceptions).

The Higgs Hunter's Guide has numerous phenomenological bounds on the type II 2HDM, but they have become quite outdated. There have been several more recent works [79–81] on $B \rightarrow X_s\gamma$; the work by Misiak [81] would also apply to the lepton-specific model. A recent summary of bounds on the charged-Higgs mass in the type II 2HDM is by Krawczyk and Sokolowska [82], and detailed analyses of charged-Higgs production at hadron colliders can be found in Refs. [83–85]. Asakawa et al. [86] compare associated W plus charged Higgs production in the type II 2HDM with that in the MSSM. As noted above, the decay of the charged Higgs into a pseudoscalar and a W is forbidden in the MSSM, but not in the more general type II 2HDM, and a detailed study of that decay is found in Ref. [72], which also discusses the type I 2HDM. Krawczyk and Temes [87] studied constraints from leptonic tau decays. Finally, a very recent study by Kaffas et al. [88] looked at all these (and other) processes in a very comprehensive analysis of the type II 2HDM parameter space. There are many additional papers on the type II 2HDM; their results will be discussed in subsequent sections.

In addition, there is another model which has natural flavour conservation, in which the quarks and charged leptons all couple to Φ_2, but the right-handed neutrino couples to Φ_1. In this model, there are Dirac neutrino masses, and the vacuum expectation value of Φ_1 must be O(eV). The only way to have such a small vev is through an approximate symmetry. The model originally used a Z_2 symmetry which was softly [89] or spontaneously [90] broken, but this allows for right-handed neutrino masses. Extending the symmetry to a $U(1)$ and breaking it softly (to avoid a Goldstone boson) gives the model of Davidson and Logan [91]. An interesting question concerns the effects of quantum corrections to the small vacuum expectation value and that is discussed in Ref. [92]. The phenomenology of this model (which is basically the type I 2HDM with a right-handed neutrino added) is quite interesting, especially in the charged-Higgs sector, and will be discussed subsequently.

In this chapter, we will first discuss the decays of the neutral Higgs bosons in the various models, followed by an analysis of the production of the neutral Higgs. Then the constraints due to both collider bounds as well as lower energy processes (such as B decays, b production, the anomalous magnetic moment of the muon, etc.) will be presented. Finally, the "inert" doublet model will be discussed.

2.1. Higgs decays

The key methods of distinguishing the various 2HDMs from each other and from the Standard Model involve the branching ratios in Higgs decays. In this section, we will focus on the decays of the neutral scalars (h, H, A) of the 2HDM. In the Standard Model, the Tevatron is sensitive to the $\bar{b}b$, WW, ZZ Higgs decays, whereas at the LHC the decays $\gamma\gamma$, WW, ZZ are more important. The branching ratios and total width of the Standard Model Higgs are very well studied; the most comprehensive analysis is found in the recent review of Djouadi [93]. The results are in Fig. 1. One can see that the WW and ZZ decays are dominant for Higgs masses above 160 GeV; they would provide the best signature at either the Tevatron
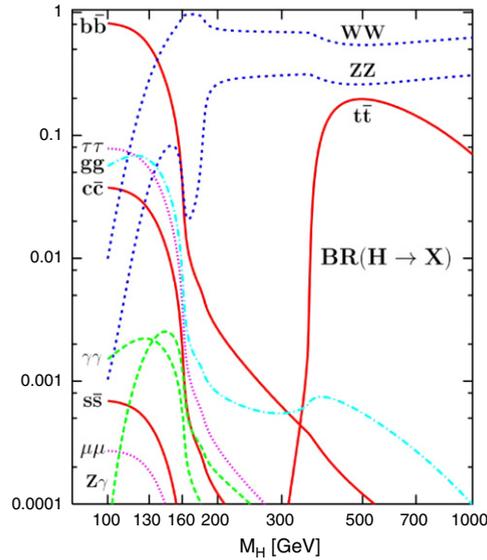
Fig. 1. The branching ratios for the decay of the SM Higgs boson as a function of its mass.

or the LHC. Below 160 GeV, the $b\bar{b}$ branching ratio is more important at the Tevatron, but that decay mode is swamped by large backgrounds at the LHC, which must rely on the $\gamma\gamma$ mode in the low-mass range.

We now must look at branching ratios in 2HDMs. As shown in the last section, the branching ratios will not depend exclusively on the masses, but also on α and β. This makes plots of the various branching ratios necessarily incomplete. Ignoring, for now, the possibility of heavy Higgs bosons decaying into lighter ones, one has branching ratios depending on m_H, α and β for each Higgs (h, H, A) and for each model. This makes a comprehensive analysis difficult.

How have other authors dealt with this previously? In the comprehensive review of all four models by Aoki et al. [36], they plotted the branching ratios of H and A for each of the four models as a function of $\tan\beta$, setting the mass to 150 GeV and $\cos(\alpha - \beta) = 0$. In this limit, the decays of the H and A to WW and ZZ vanish. While certainly an interesting and important limit, this does correspond to a single point in the two-dimensional mass$-\cos(\alpha - \beta)$ parameter space. Akeroyd [49] plotted the branching ratio of the light Higgs as a function of $\tan\beta$, but also chose only a few points in the parameter space. In their analysis of the lepton-specific model, Su and Thomas chose a specific benchmark point for α and β and plotted the branching ratios as a function of the mass. These are just some examples.

In order to cut down on the complexity of the graphs, we will not display decay modes which can never be seen at the LHC and Tevatron (at least within the next decade), although they will be included in the computation of the branching ratios. Long before these modes could be measured, many of the parameters of the 2HDM will be known, and more comprehensive analyses will be done. Thus, we will only consider decays into $t\bar{t}$, $b\bar{b}$, WW, $\tau\tau$ and $\gamma\gamma$ (and possibly into other Higgs bosons). In fact,

- In the type I and type II 2HDMs the ratio of the $\tau\tau$ to $b\bar{b}$ modes is fixed at $m_\tau^2/3m_b^2$. Using the fact that one must use the running b-quark mass at the 100 GeV scale (which is approximately 3.0 GeV) and including radiative corrections, this ratio is fixed at approximately 10%, and thus we will not explicitly include the $\tau\tau$ mode in the figures.
- The decay width into ZZ is related to the one into WW by only the weak mixing angle and phase–space factors, and thus will have the same ratio as in the Standard Model. As a result, the branching ratio into ZZ will not be directly displayed.
- The decay into gluons (while important in production) cannot be directly measured due to very large backgrounds.
- Other decay modes, such as $c\bar{c}$, $Z\gamma$ and $\mu\mu$, have either too small branching ratios or too large backgrounds.

Again, only neutral Higgs fields will be discussed in this chapter.

In all four of the 2HDMs, a lower bound on $\tan\beta$ of roughly 0.3 can be obtained from the requirement that the top-quark Yukawa coupling be perturbative. Since large Yukawa couplings have positive beta functions, if they start out large then they will exceed the perturbative limit (as well as unitarity) at a relatively low scale (see Section 5.7 and Appendix A). It is hard to see how a $\tan\beta$ near or below 0.3 can be accommodated. At the other extreme, in the type II 2HDM, the bottom-quark Yukawa coupling will be non-perturbative if $\tan\beta$ exceeds roughly 100. In the context of the MSSM, Barger et al. [94] and Carena et al. [95] showed that perturbative unification could be achieved for $\tan\beta < 60$, but the MSSM has beta functions very different from those of the 2HDM. Kanemura and collaborators [96,97] looked at bounds on $\tan\beta$ from perturbative unification in the non-supersymmetric 2HDM, and Akeroyd et al. [98,99] looked for violations of tree-level unitarity in a large number of processes and concluded that values of $\tan\beta$ greater than 30 are disfavoured, although there are some regions of parameter space in which they are allowed. More recently, Arhrib et al. [51] argued that perturbative and unitarity constraints, in all four models, require that $\tan\beta < 6$ for all but a very small region of parameter space, and this is confirmed
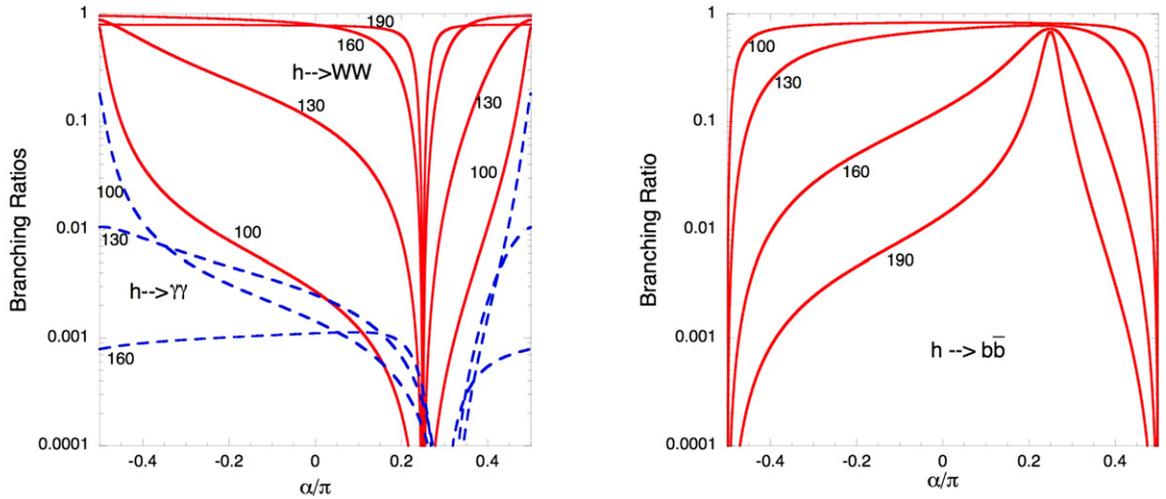
Fig. 2. The type-I 2HDM light-Higgs branching ratios into W pairs, diphotons and $b\bar{b}$ are plotted as a function of α for $\tan\beta = 1$ and for various values of the Higgs mass (in GeV). In the left figure, the solid lines correspond to $h \to WW$ and the dashed lines to $h \to \gamma\gamma$.

in the work of Kaffas et al. [88]. Thus, we will focus on values of $\tan\beta$ between 1 and 6, but will mention the effects of larger $\tan\beta$ in a few instances.

One region of interest is the decoupling region. This is the region of parameter space in which the H, A and charged Higgs all are much heavier than the h, and it is thus possible to integrate out the heavy fields. The resulting effective theory is then like the Standard-Model Higgs sector, with corrections to the various couplings due to the heavy sector. This was discussed in detail by Gunion and Haber [100]. Later, Mantry et al. [101] and Randall [102] examined Higgs decays in the decoupling region and showed how one could obtain a measurable sensitivity to the high scale even if the heavy scalars are inaccessible at the LHC, pointing out the importance of accurate measurements of the branching fractions.

2.1.1. Higgs decays in the type I 2HDM

The type I 2HDM has the simplest discrete symmetry and its couplings can be easily described. The coupling of the light neutral Higgs, h, to fermions is the same as in the Standard Model but multiplied by $\cos\alpha / \sin\beta$ while its couplings to WW and ZZ are multiplied by $\sin(\alpha - \beta)$. For the heavy neutral Higgs, H, these factors are $\sin\alpha / \sin\beta$ and $\cos(\alpha - \beta)$, respectively. Thus one can determine the widths of the various decays by simple multiplication, with the exception of the $\gamma\gamma$ decay, in which the contribution of the W loop and that of the fermion loops are multiplied by their respective factors.

There are a few interesting limits. If $\sin(\alpha - \beta)$ $(\cos(\alpha - \beta))$ vanishes, then the h (H) field is gauge-phobic, *i.e.* it does not couple to WW and ZZ, radically changing the phenomenology of Higgs decays. If $\cos\alpha$ $(\sin\alpha)$ vanishes, then the h (H) is fermiophobic. This is a particularly interesting limit, since then the $\gamma\gamma$ decay can become dominant well below the WW threshold. Although there is no symmetry that can enforce this limit (and a non-zero $\sin(\alpha - \beta)$ would be generated at one-loop if it is set equal to zero at tree level), the phenomenological implications are so dramatic that study of the limit is warranted.[3]

For $\tan\beta = 1$, the branching ratios of the light Higgs have been plotted in Fig. 2 for various values of the Higgs mass. One can clearly see that $\alpha = \pm\pi/2$ is the fermiophobic limit, where the branching ratio to fermions vanishes, and that $\alpha = \beta$ is the gauge-phobic limit, where the branching ratios to WW and to ZZ vanish. Note that the branching ratio to $\gamma\gamma$ at the gauge-phobic point does not quite vanish since there is a small contribution from top-quark loops.

For the light Higgs h we have considered masses ranging from 100 GeV to 190 GeV. An h heavier than 190 GeV may have problems with electroweak precision results [103], although in the 2HDM, such a heavy Higgs could be made compatible with these results, as discussed in Ref. [104]. An h lighter than 100 GeV might be allowed by the LEP data [105] if its coupling to ZZ were suppressed by a sufficiently small value of $\sin(\alpha - \beta)$; for instance, a 70 GeV h would be allowed [106] if $\sin^2(\alpha - \beta)$ were no larger than 3%. For a recent discussion, see the work of Gupta and Wells [107]. Since phenomenological consistency is obtained only for a narrow range of α around the gauge-phobic point, we will not allow for $m_h < 100$ GeV in the plots, but any interesting physics of a lighter h will be discussed in the text.

In the Standard Model, for a Higgs mass of 100 GeV, the decay to WW is heavily suppressed since at least one of the Ws must be far off-shell, and the $b\bar{b}$ mode is dominant. Thus, as seen in Fig. 2, one must be very close to the fermiophobic limit in order for the $b\bar{b}$ branching ratio to be small. As the Higgs becomes heavier, the WW decay mode is less suppressed and eventually dominates; for larger Higgs masses, the decay into $b\bar{b}$ is very small except very close to the gauge-phobic point.

---

[3] Notice, though, that this limit corresponds to the inert model, in which only one of the doublets gains a vev.

Note that the $b\bar{b}$ branching ratio never reaches unity, even at the gauge-phobic point, due to the contribution of the gluonic decay (which depends on the coupling of the Higgs to the top quark). Comparing with the Standard Model decays, one can see that the relative branching ratios can differ substantially.

For larger values of $\tan\beta$, the results look very similar, with slightly different slopes. The main difference is that the gauge-phobic point is at $\alpha = \beta$ and thus moves to the right. In the case of $\tan\beta = 30$, the gauge-phobic and fermiophobic points are only two degrees apart, leading to very steep slopes in the curves (since the gauge-phobic and fermiophobic points cannot coincide, the Higgs will certainly decay, but in this case its width will be quite narrow).

The diphoton decay, $h \to \gamma\gamma$, has contributions from both W loops and t loops, therefore it does not vanish in either the gauge-phobic or fermiophobic limit (the contribution of W loops in the Standard Model is substantially larger than that of top-quark loops, so the gauge-phobic limit does cause a suppression). An analysis of this mode in the type-I 2HDM was carried out by Posch [108], who found that an enhancement over the Standard Model branching ratio of as much as 70% was possible.

The fermiophobic limit is of special interest. Although the mass limit on the Standard-Model Higgs boson is 114.4 GeV [109], this assumes that the coupling of the Higgs boson to ZZ is not suppressed, and the bound can be much weaker if one is fairly close to the gauge-phobic limit. Many studies have been done concerning the possibility of a Higgs much lighter than 100 GeV in the fermiophobic limit. Note that the decay into WW drops off dramatically as the Higgs mass is below the W mass, therefore the fermiophobic limit leads to a dominant decay into $\gamma\gamma$ for these masses. CDF [110] and D0 [111] have published bounds on a fermiophobic Higgs (by looking for the diphoton mode) which are slightly over 100 GeV, but they still assume that the coupling to ZZ is not suppressed, which may not be the case (note that if $\alpha = \pm\pi/2$, then $\sin(\alpha - \beta) = \pm\cos\beta$ which can be quite small in the large $\tan\beta$ limit).

In a comprehensive analysis, Akeroyd and Diaz [74] noted that if the coupling to gauge bosons is suppressed, then the coupling of the light Higgs to a charged Higgs and a W is not suppressed (one scales as $\sin(\alpha - \beta)$ and the other as $\cos(\alpha - \beta)$) thus one can pair-produce the charged Higgs and from there a pair of light Higgs bosons. D0 [112] has looked at this and found a lower bound on the light-Higgs mass, which is approximately 80 GeV for a charged-Higgs mass below 100 GeV and approximately 50 GeV for a charged-Higgs mass below 150 GeV. This is based on less than a single fb^{-1}, and will be improved substantially in the near future. Subsequently, Akeroyd et al. [73] looked at this process at the LHC, which would be particularly relevant if the charged Higgs boson were heavier. Finally, it was shown [67,68,113] that charged-Higgs loops can substantially alter the diphoton branching ratio in 2HDMs, although this depends on unknown scalar self-couplings, *viz.* on the coupling hH^+H^-. Since one can see from Fig. 2 that the dominance of the diphoton decay mode (especially for a very light Higgs) will only occur if α is extremely close to $\pm\pi/2$, the region of parameter space is very small, but the signature is sufficiently dramatic that searches should continue. Note that the fermiophobic limit is only relevant for the type I 2HDM.

Throughout the above we have neglected the possibility that the light Higgs can decay into other Higgs. In fact, in both the type I and type II 2HDMs there is a decay that could be important for a range of parameter space, and yet it has not, to our knowledge, been explored substantially. The range of parameter space will occur if the pseudoscalar mass M_A is less then $m_h - m_Z$. Specifically, one can look for $h \to ZA$. This could occur if the pseudoscalar is very light (a few GeV) and the Higgs is near the current bound (this would fit electroweak precision tests better), or else if the pseudoscalar mass is comparable to the gauge-boson masses and the lightest Higgs scalar is considerably heavier. The $h \to ZA$ decay has generally not been mentioned, and we know of no experimental searches for this mode (although it has been mentioned by DELPHI [114]). The primary reason for this lack of attention is that the rate is proportional to $\cos^2(\alpha - \beta)$, which is very small in the MSSM. But in the general 2HDM there is no reason that $\cos(\alpha - \beta)$ is small. The width is given by[4]

$$\Gamma(h \to ZA) = \frac{g^2 m_h^3 \cos^2(\alpha - \beta)}{64\pi m_W^2} \lambda^{3/2} \tag{17}$$

where $\lambda = (1 - (m_Z^2 - m_A^2)/m_h^2)^2 - 4m_Z^2 m_A^2/m_h^4$. One can compare this with the more well-known decay of the light Higgs to WW, which is (for decays into real gauge bosons)

$$\Gamma(h \to WW) = \frac{g^2 m_h^3 \sin^2(\alpha - \beta)}{64\pi M_W^2} \lambda' \tag{18}$$

where $\lambda' = \sqrt{1 - 4x}(1 - 4x + 12x^2)$, with $x = M_W^2/m_h^2$. For $\cos^2(\alpha - \beta) > 1/2$, $h \to ZA$ will actually *dominate* Higgs decays. The branching ratio is plotted in Fig. 3, for a range of h masses, choosing $\tan\beta = 1$ and $m_A = 5$ GeV. One can see the dip at $\cos(\alpha - \beta) = 0$; this is the region expected in supersymmetric models. At $\alpha = \beta$, one has the gauge-phobic limit, and thus the $h \to ZA$ decay dominates (except at a Higgs mass of 100 GeV, where the $b\bar{b}$ decay is still substantial). For larger $\tan\beta$, as before, the maxima and minima simply shift to the right. For larger m_A, the only change will be in λ, but as long as the light Higgs is reasonably heavier than $m_Z + m_A$, there will be a range of parameters where this decay is large.

One sees that there is a substantial range of parameters for which this relatively unstudied decay mode could be substantial, and even dominate. The only discussion of it that we are aware of is the analysis, in the type I 2HDM, by

---

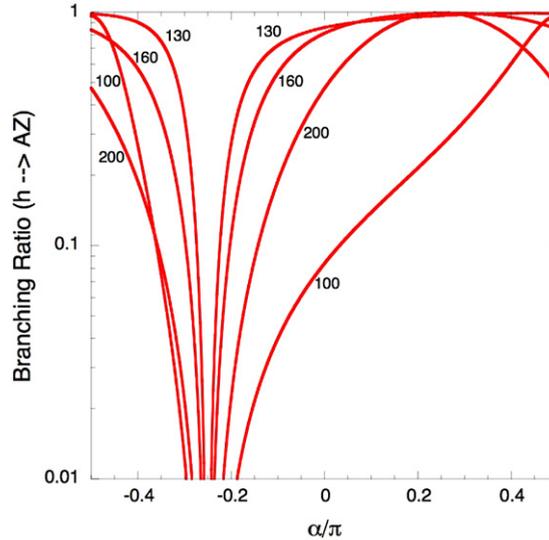[4] Throughout this work, π refers to the mathematical constant [115], not the pion field.

Fig. 3. Branching ratios of the light Higgs boson *h* into a *Z* and a pseudoscalar, for various values of the mass (in GeV) of *h*. The value of $\tan\beta$ is chosen to be 1 and the mass of the pseudoscalar is chosen to be 5 GeV.

Akeroyd [71], who looked at the decay of a neutral scalar into a *virtual Z* and a pseudoscalar. There has been no analysis of detector capabilities. Even if the *Z* is real, the decay may be challenging [116]. At a hadron collider, the *h* longitudinal boost will not be known, so the *Z* will not be monoenergetic; moreover, the decays of the pseudoscalar may not be easy to observe (especially if it decays into *b* quarks). But if this is the dominant decay, further analysis is needed. As we will see below, the decay involving the heavier of the neutral scalars, $H \rightarrow ZA$ has been discussed in detail, since this can also be substantial in supersymmetric models.

What about Higgs decays into pairs of scalars? By definition, the *H* is heavier, but the pseudoscalar could be lighter, leading to $h \rightarrow AA$. There has been substantial discussion of the possibility of a very light *A*, especially in a series of papers by Dermisek and Gunion [117–121]. One motivation is that precision electroweak fits prefer the light Higgs to be lighter than the 114 GeV bound, and this could be allowed if the $h \rightarrow AA \rightarrow 4\tau$ or 4 jet signature exists. LEP has published [122] a bound of 82 GeV for the light Higgs, independent of decay modes (assuming one is not near the gauge-phobic limit), but there is still a substantial allowed window. If the *A* is heavier than twice the *b* quark mass, then $b\bar{b}$ decays will dominate. So bounds generally can be found for *A* masses between 4 and 10 GeV and Higgs masses between 80 and 115 GeV. The bounds are generally expressed in terms of ξ, where

$$\xi^2 = \left(\frac{g_{hVV}}{g_{hVV}^{SM}} \right)^2 \text{BR}\left(h \rightarrow AA \right) \left[\text{BR}\left(A \rightarrow \tau^+\tau^- \right) \right]^2, \tag{19}$$

and in the bounds the range [0.1, 1] is typically [123] considered for ξ^2. In the type I 2HDM, however, ξ^2 will generally be smaller. The reason is that the *A* can, in the type I 2HDM, decay into $c\bar{c}$. Since the running *c*-quark mass at the 10 GeV scale is not that much smaller than the τ mass, the rates of $c\bar{c}$ and $\tau^+\tau^-$ will be similar. Thus BR$\left(A \rightarrow \tau^+\tau^- \right)$ and, consequently, ξ will be smaller. This will not be the case in the type II 2HDM. We will not have much to say about this mode here, since it will depend on the completely arbitrary quartic couplings of the scalar potential (unlike the MSSM or NMSSM, where those couplings are specified). The experimental possibility of measuring these quartic couplings in the context of the four models discussed in this section was considered by Arhrib et al. [51]. The possibility that this decay mode could suppress the other ones should be kept in mind.

We now look at the other neutral Higgs bosons. The heavier neutral scalar, *H*, has a coupling to fermions proportional to $\sin\alpha/\sin\beta$ and a coupling to vector bosons proportional to $\cos(\alpha - \beta)$. These are *identical* to the couplings of the light Higgs, *h*, if one shifts α to $\alpha - \pi/2$. As a result, for the decay modes of *H* one may still use Fig. 2 after shifting the graph to the left by $\pi/2$ (recall that a shift by π in α does not affect any physics). So the fermiophobic point will be $\alpha = 0$ and, for $\tan\beta = 1$, the gauge-phobic point will be at $\alpha = -\pi/4$. As $\tan\beta$ increases, just as in the case for *h*, the gauge-phobic point moves towards the fermiophobic point. It should be emphasized that if there is a detection of a single scalar field with a mass of, say, 190 GeV, then it will be impossible to determine if it is an *h* or an *H* by looking at the decays, since they are identical after a redefinition of α.

Unlike the light Higgs field, however, the *H* can be substantially heavier. As can be seen from Fig. 2, the branching ratio into *WW* will be fairly constant for heavier masses, with a very sharp dip at the gauge-phobic point (which is now at $\alpha = -\pi/4$ for $\tan\beta = 1$). But, the branching ratio into fermions will change as the top-quark threshold is reached. In Fig. 4 we have plotted the branching ratio of the *H* into $b\bar{b}$ and $t\bar{t}$. Note that the curve for *H* decays into $b\bar{b}$ at a Higgs mass at 200 GeV is
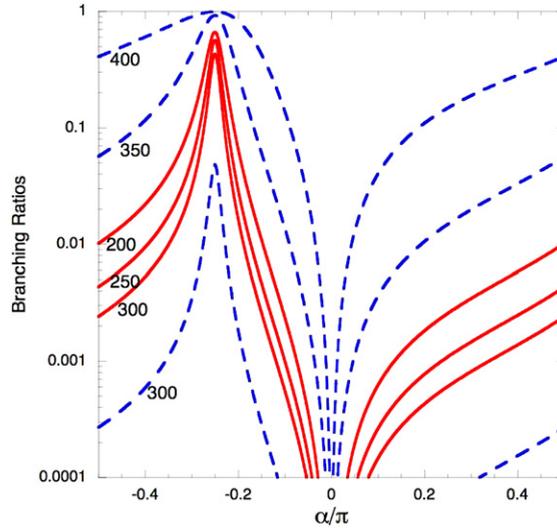
Fig. 4. Branching ratios of the heavy Higgs boson into $b\bar{b}$ (solid lines) and $t\bar{t}$ (dashed lines), for various values of the heavy-Higgs mass. We have chosen $\tan\beta = 1$.

identical to that for h decays in Fig. 2, shifted by $\pi/2$. As the Higgs mass increases, the branching ratio to $b\bar{b}$ decreases (except at the gauge phobic point) until the top-quark threshold is reached. As the threshold is reached, top quarks dominate (with a branching ratio of nearly 100% at the gauge-phobic point).

Our previous discussion about Higgs bosons decaying into other Higgs bosons still applies. In this case, one can also have $H \rightarrow hh$ if h is light enough. Since this depends on unknown scalar self-couplings, we will not have much to say about this mode. The same is true for the $H \rightarrow AA$ mode if the pseudoscalar is very light.

It was noted above that the decay of the light Higgs into a Z and a light pseudoscalar has not been studied, primarily because it is small in the MSSM. However, the decay of the heavy Higgs into ZA has been discussed in detail by Dermisek and Gunion [124] and by Kao et al. [125], who note that it can, for a reasonable range of parameters, dominate the decays of the H.

An interesting scenario that has, to our knowledge, not been studied is that the h and H could be very close in mass (a few GeV mass difference would be within the LHC experimental resolution). While fine-tuned, this possibility leads to the question of whether this could be distinguished from the Standard Model Higgs. If the decays into WW or ZZ are considered, then the h (H) will have a $\cos^2(\alpha - \beta)$ ($\sin^2(\alpha - \beta)$) factor in the decay rate, leading to a Standard Model decay rate if both are added. However, the production rate will depend on the top-quark coupling (since gluon fusion dominates), as will the $\gamma\gamma$ decay rate, leading to a more complicated and model dependent picture. This possibility deserves further analysis.

Finally, one can look at the decays of the pseudoscalar. Here, all fermion couplings are multiplied by $\cot\beta$, and there are no couplings to a pair of vector bosons. Thus, the branching ratios will be independent of α and β, and are just given by fermion mass ratios and phase–space. It should be remembered that all fermions are accessible (in the type I 2HDM) and thus the branching ratio into $\tau^+\tau^-$ will always be similar (within a factor of two) to the one into $c\bar{c}$. If the A is heavier than 10 GeV, then decays into $b\bar{b}$ will dominate (until the top-quark threshold is reached). Since it is difficult to detect $b\bar{b}$ at a hadron collider, one can look at other decays, such as $\tau^+\tau^-$, $W^\pm H^\mp$, Zh, ZH, gg, $\gamma\gamma$, $Z\gamma$. A very comprehensive analysis of all these decays in the type I 2HDM can be found in the article by Kominis [126].

If the mass of the pseudoscalar is above $m_h + M_Z$, but below 350 GeV (when the decays into $t\bar{t}$ will start to dominate), a possible decay mode is $A \rightarrow hZ$. Given that there are no decays to a pair of vector bosons, and the decay into bottom quarks is suppressed by m_b^2, this decay could be dominant over the region [126–128]. The decay width is given in the Higgs Hunter's Guide [47]. It does depend on $\cos(\alpha - \beta)^2$ and one would expect the decay to dominate if it is kinematically accessible and if that quantity is not small. A recent discussion about detection at the LHC is in Ref. [107]. The other decays mentioned above are loop effects and will generally be small, but may be experimentally easier to detect.

Kominis [126] has studied the diphoton decay channel and the Zh channel in the type I and type II 2HDMs and compared it with the results of the MSSM. He found that in the lower end of the mass range, due to the difficulty in seeing $b\bar{b}$ pairs at a hadron collider, the diphoton mode would be the most promising one. The process $A \rightarrow hZ$ provides a very clear signature in the intermediate mass range in which the h is between 40 and 160 GeV. These two modes are thus complementary. This work was done in 1994, and an updated analysis would be welcome.

Recently, Bar-Shalom et al. [129] studied 2HDMs in which dynamical electroweak symmetry breaking is triggered by condensation of 4th generation fermions. Their models have one "heavy" doublet with a large vev and a "light" doublet with a much smaller vev. The heavy doublet couples to the 4th generation fermions and, in one of the models to the 3rd generation of fermions, in another to the top quark only, and in another to only the 4th generation fermions. The phenomenology is
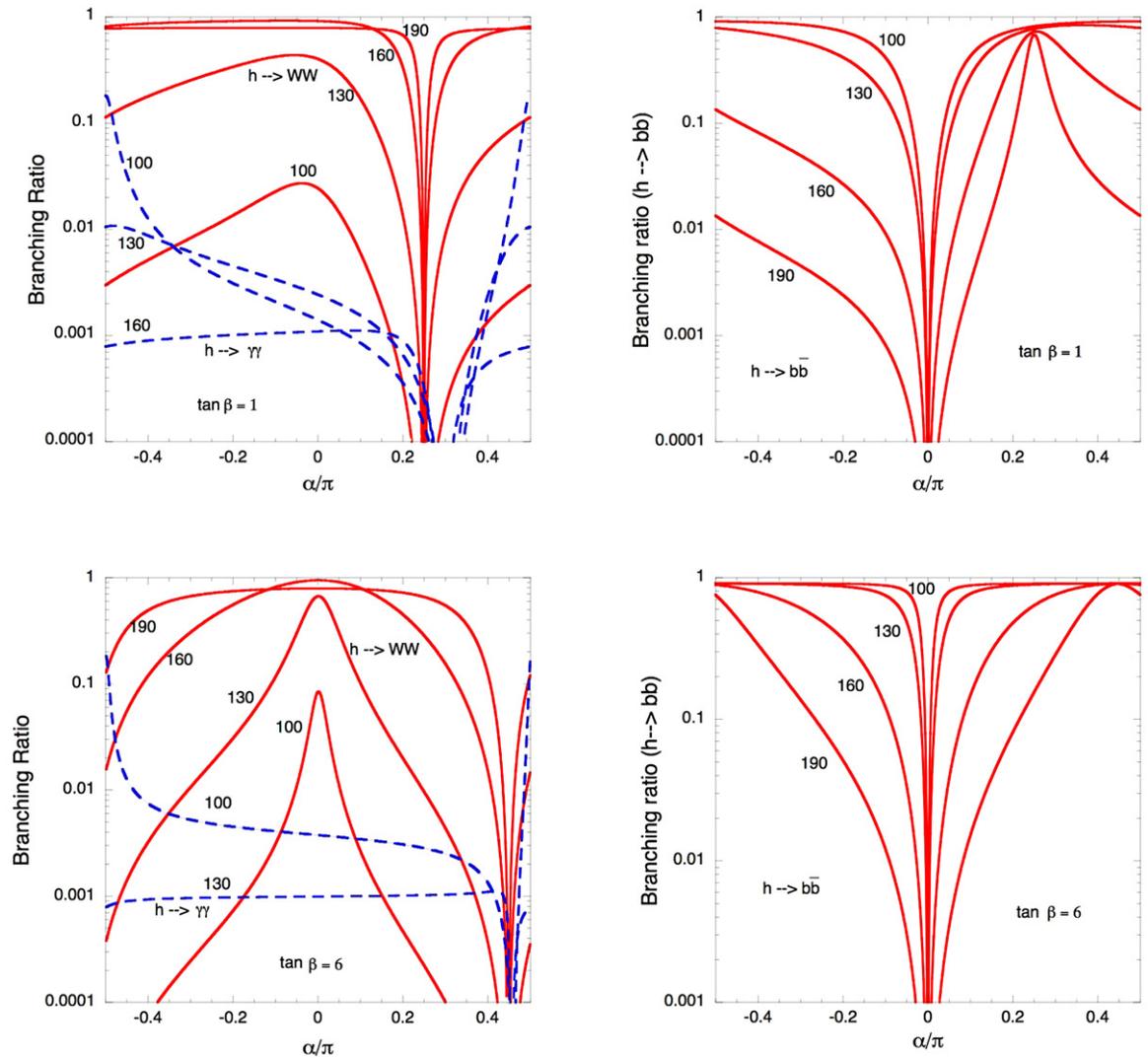
Fig. 5. The type II 2HDM light-Higgs branching ratios into W pairs, diphotons and $b\bar{b}$ are plotted as a function of α for $\tan\beta = 1$ and $\tan\beta = 6$ and for various values of the Higgs mass (in GeV). In the left-hand figures, the solid lines correspond to $h \to WW$ and the dashed lines to $h \to \gamma\gamma$. The branching ratio into Z pairs has the same ratio to the one into W pairs as in the Standard Model.

similar to the type I 2HDM. They study the phenomenology of the models, including precision electroweak measurements, rare decays and collider implications.

2.1.2. Higgs decays in the type II 2HDM

The type II 2HDM is the most studied one, since the couplings of the MSSM are a subset of the couplings of the type II 2HDM. The coupling of the light neutral Higgs, h, to fermions depends on the fermion charge. The coupling of the $Q = 2/3$ quarks is the same as in the type I 2HDM, *i.e.* it is the Standard-Model coupling times $\cos\alpha/\sin\beta$. On the other hand, the coupling of the $Q = -1/3$ quarks and of the leptons is the Standard-Model coupling times $-\sin\alpha/\cos\beta$. In the large $\tan\beta$ scenario, this means that the couplings of the $Q = -1/3$ quarks and of the leptons are much larger than in the type I 2HDM. In fact, one can see that the ratio of the bottom quark Yukawa coupling to that of the top quark is approximately $\tan\alpha\tan\beta$ times the same ratio for the type I 2HDM, and this can drastically affect the phenomenology. The couplings to gauge bosons are the same as in the type I 2HDM. Therefore, one still has the gauge-phobic limit for h if $\sin(\alpha - \beta) = 0$, but there is no fully fermiophobic limit. If $\alpha = \pm\pi/2$ ($\alpha = 0$), then the h will not couple to the $Q = 2/3$ quarks ($Q = -1/3$ quarks and leptons), but no choice of α will eliminate all couplings to fermions.

The branching ratios of the light neutral Higgs are plotted in Fig. 5 for $\tan\beta = 1$ and $\tan\beta = 6$. Note that, as expected, the curves for $\tan\beta = 1$ are very similar to those in Fig. 2 if one shifts α by $\pi/2$ and flips the sign. The b-phobic value of α is zero in the type II 2HDM. At the b-phobic point, say for a Higgs mass of 100 GeV, neither the $b\bar{b}$ nor the WW, ZZ modes are substantial. In this case, the $c\bar{c}$ and gluon–gluon decays dominate. Since neither can be seen at the LHC, we have not

Fig. 6. Branching ratios of the heavy Higgs boson into fermions for $\tan \beta = 6, 30$. The solid (dashed) lines are the branching ratios into $b\bar{b}$ ($t\bar{t}$) for various values of the heavy-scalar mass.

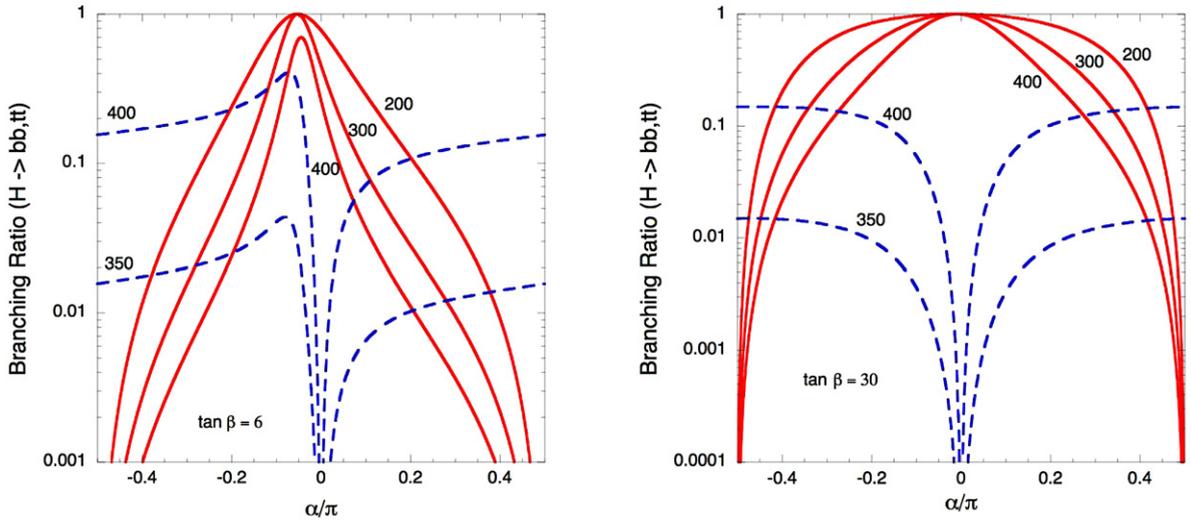included them here, although a very similar figure with these modes included appears in Arhrib et al. [51] for $m_h = 110$ GeV and $\tan \beta = 1$.

Unlike the type I 2HDM, however, there is in this case a strong dependence on $\tan \beta$. In Fig. 5 we also show the decays for $\tan \beta = 6$. One sees that, as expected, the gauge-phobic point $\alpha = \beta$ moves to the right, and the b-phobic point remains at $\alpha = 0$. For a relatively light h, with a mass of 100 GeV, the enhanced coupling of the b quark when $\tan \beta = 6$ causes the phase–space suppression of the WW decay mode to be quite dramatic, with virtually no WW pairs except at the b-phobic point. Even at that point, the WW branching ratio is only around 10%, since the gluon–gluon decay goes through a top-quark loop which is not suppressed at the b-phobic point, and thus becomes dominant. Still, slightly away from the b-phobic point, the $b\bar{b}$ mode already dominates again although, for a light h, the diphoton mode is not negligible. It is interesting to note that this model contains a sizeable region of parameter space in which a light Higgs with a mass of 160 GeV will still predominantly decay into b quarks rather than W pairs.

For $\tan \beta = 30$, the gauge-phobic point is extremely close to $\alpha = \pi/2$, and the trend becomes even more dramatic, with the $m_h = 190$ GeV curve becoming close to the $m_h = 130$ GeV curve of the $\tan \beta = 6$ plot. At this point, even a relatively heavy h will have negligible couplings to WW or ZZ for most values of α.

It should be kept in mind that the branching ratio into $\tau^+\tau^-$ is approximately 10% of the one into $b\bar{b}$. The backgrounds for the $b\bar{b}$ mode are huge at the LHC, rendering that decay mode very difficult to observe, while the backgrounds for $\tau^+\tau^-$ are much less severe and therefore, even if the $b\bar{b}$ mode cannot be seen at the LHC, h could still be detected through its decay to $\tau^+\tau^-$.

As noted earlier, the recent review article by Djouadi [24] studies the Higgs phenomenology of the MSSM in great detail, and that is a special case of the type II 2HDM. Much of the analysis in that article will be relevant here. In the general case, there are some differences. In the MSSM, the light Higgs cannot be much heavier than 130 GeV, as it can here. In addition, α is determined in terms of other Higgs masses and β and is thus much more restricted. Nonetheless, the basic features shown in Fig. 5 are, for h masses at 130 GeV and below, also present in the MSSM. For large $\tan \beta$, the $b\bar{b}$ decay mode dominates for most of parameter space. However, the MSSM (at tree level) does not allow $\alpha = 0$, and thus the b-phobic region does not occur. The major differences, then, are that the general type II 2HDM allows for a heavier h and that it also allows for the possibility that a fairly light h could still predominantly decay into W pairs.

In the discussion of the type I 2HDM, we also discussed the decays of the light Higgs into a Z plus a pseudoscalar and of a light Higgs into two pseudoscalars. Nothing much changes in the type II 2HDM. One still has a region of parameter space in which $h \rightarrow ZA$ can be substantial, and $h \rightarrow AA$ can be significant if A is light. The only difference is that if the A has a mass between 4 and 10 GeV, then its branching ratio into $\tau^+\tau^-$ will be larger for large $\tan \beta$, facilitating detection.

Turning to the other neutral Higgs bosons, one can see that the couplings of the heavier scalar H are identical to those of the h if one replaces α by $\alpha - \pi/2$. Thus the decay modes of the H, if it is lighter than 200 GeV, will be identical to those in Fig. 5 after α is shifted to the left by $\pi/2$. For heavier H fields, however, there is a substantial difference between the type I and type II 2HDMs. This is because in the type II 2HDM, the ratio of the branching fraction into bottom quarks compared to the one into top quarks varies as $\tan^2 \beta \tan^2 \alpha$ compared with the type I 2HDM. In addition, the b-phobic and t-phobic points do not coincide. For $\tan \beta = 1$, the results for the $t\bar{t}$ mode are identical to those in Fig. 4, and the results for the $b\bar{b}$ are the same but with α shifted by $\pi/2$ (so the b-phobic points are at $\alpha = \pm\pi/2$). In Fig. 6, we plot the branching ratios of the heavy Higgs into fermions for $\tan \beta = 6$ and $\tan \beta = 30$. Again, one sees that the t-phobic and b-phobic points are now different. There is here a striking difference between the type I and type II 2HDMs.
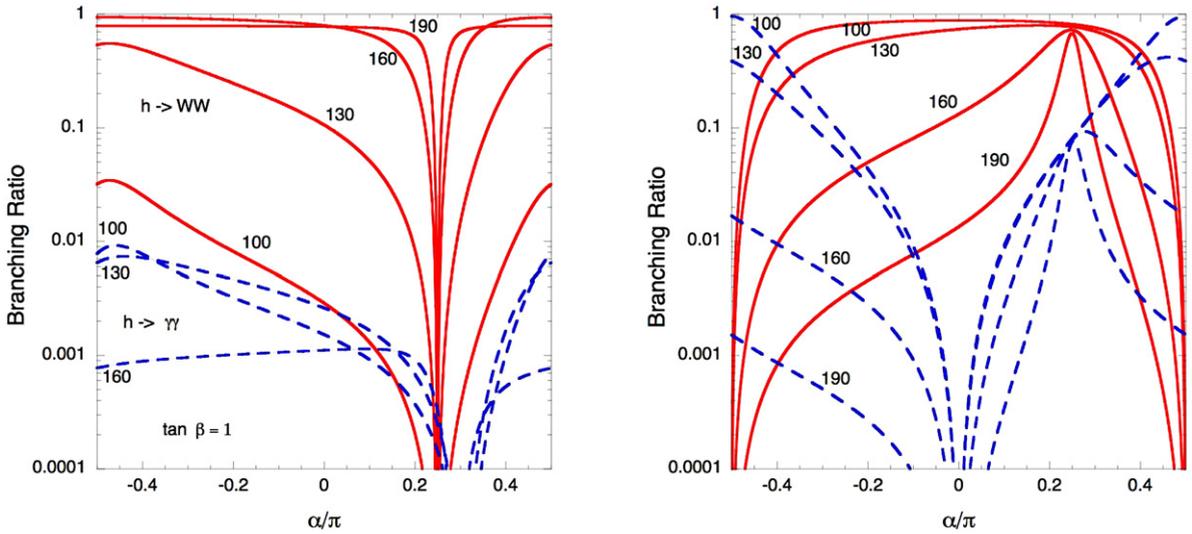
Fig. 7. Branching ratios of the light Higgs boson in the lepton-specific 2HDM. We have taken $\tan \beta = 1$. In the left figure, the solid (dashed) lines are the branching ratios into WW ($\gamma\gamma$); in the right figure, the solid (dashed) lines are the branching ratios into $b\bar{b}$ ($\tau^+\tau^-$) for various values of the Higgs-boson mass.

The pseudoscalar does not decay into gauge bosons, and thus it will decay into the heaviest fermions accessible, as in the type I 2HDM. However, now the decay to $Q = -1/3$ quarks ($Q = 2/3$ quarks) will be multiplied by $\tan^2 \beta$ ($\cot^2 \beta$), and for large $\tan \beta$ this will strongly suppress the coupling to top quarks. In fact, for $\tan \beta$ greater than about 6–7, the decay into bottom quarks will always exceed that into top quarks.

Thus, for $\tan \beta$ near 1, the results for pseudoscalar decay are identical to those of the type I 2HDM. For a pseudoscalar mass in between 4 and 10 GeV, the main decay mode is $\tau^+\tau^-$, with a comparable branching ratio into $c\bar{c}$. Above 10 GeV, the principal decay is into $b\bar{b}$ (with $\tau^+\tau^-$ being approximately 10% of $b\bar{b}$). As discussed in detail by Kominis [126], it might be easier, in this mass range, to look for the one-loop diphoton mode. Once the mass exceeds 200 GeV, $A \rightarrow hZ$ becomes possible, and then above 350 GeV, decays into top quarks dominate.

For $\tan \beta$ much larger, say larger than 10, the decay of a 4–10 GeV pseudoscalar is almost entirely into $\tau^+\tau^-$, due to the absence of charm decays. Above 10 GeV, $b\bar{b}$ dominates, and this domination continues for all masses; since the $b\bar{b}$ mode is difficult to observe at the LHC, it may be necessary to look at the diphoton mode to find the pseudoscalar. Kominis [126] discusses these possibilities in detail.

It is straightforward to see (as discussed above) how the results of this section can be expanded to larger values of $\tan \beta$— the results do not change substantially, except for the ratio of h decays into bottom quarks relative to top quarks. For a much more recent discussion with many references, which look at both the CP-conserving and CP-violating models, see the papers of Kaffas et al. [88,130].

There is a class of models in which the electroweak symmetry is broken by the condensation of a strongly coupled fermion sector. Although this sector could come from a fourth generation, it need not. As noted originally by Luty [131], if this strongly interacting sector respects isospin invariance, then the resulting low energy theory is a two-Higgs doublet model. Using an RG-improved Nambu–Jona-Lasinio model, Burdman and Haluch [132] studied the effective low-energy scalar sector. They found that, not surprisingly, the scalars are all fairly heavy (in the 600–800 GeV region), but also found that the pseudoscalar is light, with a mass ranging from 10 to 120 GeV. They discuss the phenomenology of the model, including precision electroweak fits, and find it similar to a type II 2HDM with an unusual mass spectrum and with $\tan \beta \sim 1$.

2.1.3. Higgs decays in the lepton-specific 2HDM

The couplings of the quarks to the Higgs bosons in the lepton-specific (LS) 2HDM are identical to those in the type I 2HDM, but the couplings of the leptons are quite different. In the previous models, the branching ratio into $\tau^+\tau^-$ was roughly 10% of the branching ratio into $b\bar{b}$ for all values of the parameters. This is not the case in the LS 2HDM. There are two major differences between the LS 2HDM and the type I 2HDM. Firstly, the branching ratio of the h into $\tau^+\tau^-$ can be much larger and, compared with the branching ratio into $b\bar{b}$, grows as $\tan^2 \beta \cot^2 \alpha$. Secondly, the b-phobic value of α is $\pm\pi/2$, whereas in the LS 2HDM, the τ-phobic value of α is 0. This can dramatically affect the phenomenology—the $\tau^+\tau^-$ branching ratio can exceed that of $b\bar{b}$ even for values of $\tan \beta$ near unity.

As in the type I 2HDM, the h field (H field) will be gauge-phobic if $\sin(\alpha - \beta)$ ($\cos(\alpha - \beta)$) vanishes.

In Fig. 7, we have plotted the branching ratios of the light Higgs, h, into WW, $\gamma\gamma$, $b\bar{b}$ and $\tau^+\tau^-$ for various h masses, assuming $\tan \beta = 1$. With such a low $\tan \beta$, one might expect that there would be relatively little contribution from the $\tau^+\tau^-$ mode. However, since the τ-phobic and b-phobic points are different, one can see that near the b-phobic point, the
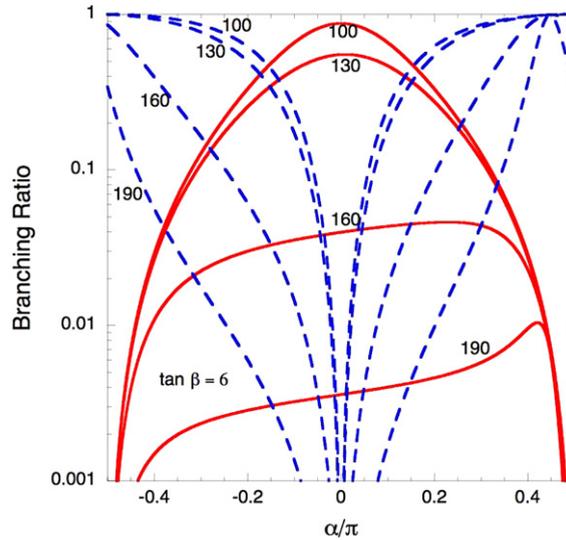
Fig. 8. Branching ratios of the light Higgs boson into fermions in the LS 2HDM with $\tan \beta = 6$. The solid (dashed) lines are the branching ratios into $b\bar{b}$ ($\tau^+\tau^-$) for various values of the Higgs-boson mass.

$\tau^+\tau^-$ mode dominates the decays of the h for $m_h = 100$ GeV. This is understandable, since the WW mode is phase–space suppressed and one is near the b-phobic point. Thus, near $\alpha = \pm\pi/2$, the $\tau^+\tau^-$ mode supersedes the other branching ratios. Note that away from this point, the curves look virtually identical to those in the type I 2HDM, Fig. 2, as expected. Su and Thomas [54] have pointed out that the region of parameter space near the b-phobic point will violate perturbativity or vacuum stability in the limit where the H and A fields are quite heavy. They studied the LHC discovery potential of the LS 2HDM in the case where the low-energy spectrum only contained one light Higgs boson.

One can also look at larger values of $\tan \beta$, where the coupling to τs will increase substantially. We have plotted the branching ratios into fermions for $\tan \beta = 6$ in Fig. 8. One sees a dramatic increase in the $\tau^+\tau^-$ mode away from the τ-phobic point. In fact, even for a Higgs mass of 160 GeV, there is a range of α [33] for which the $\tau^+\tau^-$ mode is the *dominant* decay. The fact that the curves do not add up to 100% is due to our not having included the gluon–gluon and $c\bar{c}$ decay modes in the figure; these cannot be measured easily at the LHC. The paper of Arhrib et al. [51] has very similar figures which (for a specific h mass of 110 GeV) show these decays as well. By comparing this figure with Fig. 7, one can see the pattern for very large $\tan \beta$. For $\tan \beta = 100$, for example, which is still allowed by perturbation theory (although, as noted earlier, may have difficulties with tree-level unitarity), the $\tau^+\tau^-$ mode would completely dominate the h decays for all masses of h, except very close to $\alpha = 0$. A comprehensive analysis of the $\tau^+\tau^-$ decay mode at the LHC can be found in the papers of Belyaev et al. [133,134]. They note that if decays to gauge bosons are kinematically forbidden, and decays to other Higgs bosons are not allowed, then one can write a simple formula:

$$\text{BR}\left(h \rightarrow \tau^+\tau^-\right) = \frac{\frac{\sin^2 \alpha}{\cos^2 \beta} \, \text{BR}\left(h_{\text{SM}} \rightarrow \tau^+\tau^-\right)}{\left(\frac{\sin^2 \alpha}{\cos^2 \beta} - \frac{\cos^2 \alpha}{\sin^2 \beta}\right) \text{BR}\left(h_{\text{SM}} \rightarrow \tau^+\tau^-\right) + \frac{\cos^2 \alpha}{\sin^2 \beta}}. \tag{20}$$

This equation works well for Higgs masses below 130 GeV. Above that mass, WW^* decays can become important. A recent study of multi-tau-lepton signatures can be found in the work of Kanemura, et al. [135].

A decay mode of the h that may be important and is one of the easiest to detect is $h \rightarrow \mu^+\mu^-$. The branching ratio is 0.0035 times that of the $\tau^+\tau^-$ mode. If the latter does dominate, then the decay into muons should be clearly detectable. Discussions of this mode can be found in Refs. [136–138].

The previous discussion of $h \rightarrow ZA$ and $h \rightarrow AA$ in the type I 2HDM also applies here. If the coupling to τs is large, this will suppress the branching ratios somewhat from the discussion there, but the general character of the analysis will not change.

What about the other neutral Higgs? Just as in the type I 2HDM, the couplings of the heavier neutral scalar, H, are identical to that of the lighter Higgs, h, if one shifts α to $\alpha - \pi/2$. In Fig. 9, we have plotted the decays of H into top quarks, bottom quarks, and tau leptons for $\tan \beta = 6$. We see that for H masses below the top threshold, the decays into $\tau^+\tau^-$ dominate as soon as one moves away from the τ-phobic point (now at $\alpha = \pm\pi/2$ for the H coupling). For heavier masses, the $t\bar{t}$ mode dominates the $b\bar{b}$ mode. For larger $\tan \beta$, the curves for the $\tau^+\tau^-$ mode widen and can eventually dominate for most values of α.

The discussion of other decays, such as $H \rightarrow ZA$ or $H \rightarrow AA$, is the same as for the type I 2HDM.

We finally turn to the decays of the pseudoscalar A. Here, there are no decays into WW or ZZ, and thus the pseudoscalar decays primarily into fermions. This provides a remarkable opportunity for discovery of the pseudoscalar. The ratio of the
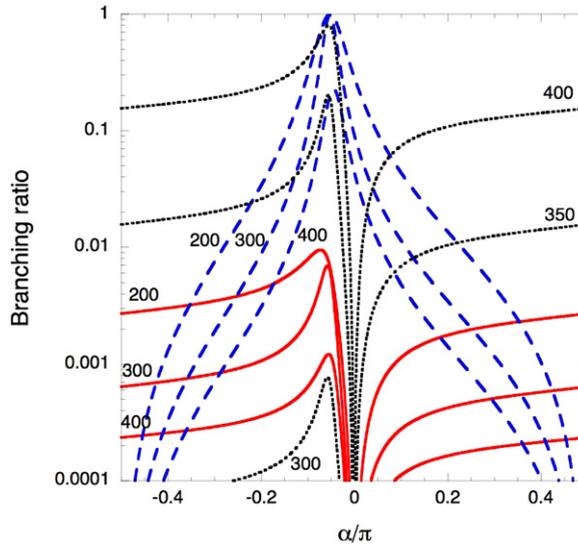
Fig. 9. Branching ratios of the heavy Higgs boson into fermions for $\tan \beta = 6$. The solid (dashed, dotted) lines, which are red (blue, black) are the branching ratios into $b\bar{b}$ ($\tau^+\tau^-$, $t\bar{t}$) for various values of the Higgs-boson mass.

branching fraction into $\tau^+\tau^-$ to the one into $b\bar{b}$ is proportional to $\tan^4 \beta$, which can easily exceed unity. In fact, including the mass effects, one can show that

$$\frac{\Gamma \left(A \rightarrow \tau^+\tau^- \right)}{\Gamma \left(A \rightarrow b\bar{b} \right)} = \left(\frac{\tan \beta}{1.76} \right)^4, \tag{21}$$

and therefore, even for relatively small $\tan \beta$, the $\tau^+\tau^-$ mode will dominate. In fact, for $\tan \beta > 3$, the branching ratio exceeds 90%. This is independent of the A mass, as long as it is below the $t\bar{t}$ threshold. As before, the branching ratio into $\mu^+\mu^-$ is 0.0035 times the one into $\tau^+\tau^-$, and if the latter dominates, this gives a sizeable branching ratio into an easily observed signature.

2.1.4. Higgs decays in the flipped 2HDM

In the flipped 2HDM, the RH leptons couple to the same Higgs doublet as the RH up quarks. As in the lepton-specific model, the τ-phobic point is different from the b-phobic point, leading to a region of parameter space in which the $\tau^+\tau^-$ branching ratio can exceed the $b\bar{b}$ branching ratio. But unlike the lepton-specific model, the flipped model cannot have a huge enhancement of the τ coupling to any of the scalars. This is because any enhancement of the τ coupling would also enhance the top-quark coupling, and a large enhancement of the latter would cause serious problems with perturbation theory and unitarity. One does not expect an enhancement of the $\tau^+\tau^-$ mode in the flipped model as one changes $\tan \beta$, like the one observed in the lepton-specific model, cf. Figs. 7 and 8.

The branching ratios of the light scalar h are presented in Fig. 10 for $\tan \beta = 1$. One sees a region of parameter space in which the $\tau^+\tau^-$ mode dominates $b\bar{b}$ decays, but it never reaches branching ratios as high as in the lepton-specific case (the remainder of the decays in the $\alpha = 0$ region is a mix of gg and $c\bar{c}$). All other branching ratios are very similar to those from the type II 2HDM. If $\tan \beta$ is increased, the region of $\tau^+\tau^-$ dominance narrows, and the region in which the $b\bar{b}$ mode dominates grows, as in the larger $\tan \beta$ curves in the type II 2HDM. The α range for which the $\tau^+\tau^-$ decay is dominant can grow if one takes $\tan \beta < 1$, but perturbation theory breaks down for $\tan \beta < 0.3$.

We have plotted the branching ratios of the decays into fermions for $\tan \beta = 0.3$ and $\tan \beta = 6$ and $m_h = 100$ GeV in Fig. 11 (for such a low m_h the decays into vector bosons are very small). As expected, for $\tan \beta = 6$, the $b\bar{b}$ mode completely dominates except for a very narrow region around the b-phobic point. On the other hand, for $\tan \beta = 0.3$, $\tau^+\tau^-$ decays dominate over $b\bar{b}$ decays for most values of α (although they never exceed a 40% branching ratio due to $c\bar{c}$ and gluon–gluon decays). The fact that in this model there is a region of parameters in which the $c\bar{c}$, gluon–gluon and $\tau^+\tau^-$ decay modes all contribute equally was first noted in Ref. [33]. Although charm pairs will be very difficult to observe at either the LHC or Tevatron, one can see plots of the branching ratios into charm pairs for this model in Ref. [51]. For the other modes, the discussion is not much different from that of the type II 2HDM. In the region where the $\tau^+\tau^-$ decay dominates one can look at $\mu^+\mu^-$ decays, which occur at a rate 0.0035 times the one of $\tau^+\tau^-$.

The couplings of the heavy Higgs, H, are identical to those of the light Higgs, when α is shifted by $\pi/2$. The difference is that the H can, if heavy enough, decay into top quarks. For $\tan \beta = 0.3$, the decays into $t\bar{t}$ and $\tau^+\tau^-$ are both enhanced, whereas for $\tan \beta = 6$, the decay into $b\bar{b}$ is substantially enhanced, similar to Fig. 6. A discussion for fairly light H and A bosons can be found in Ref. [36].
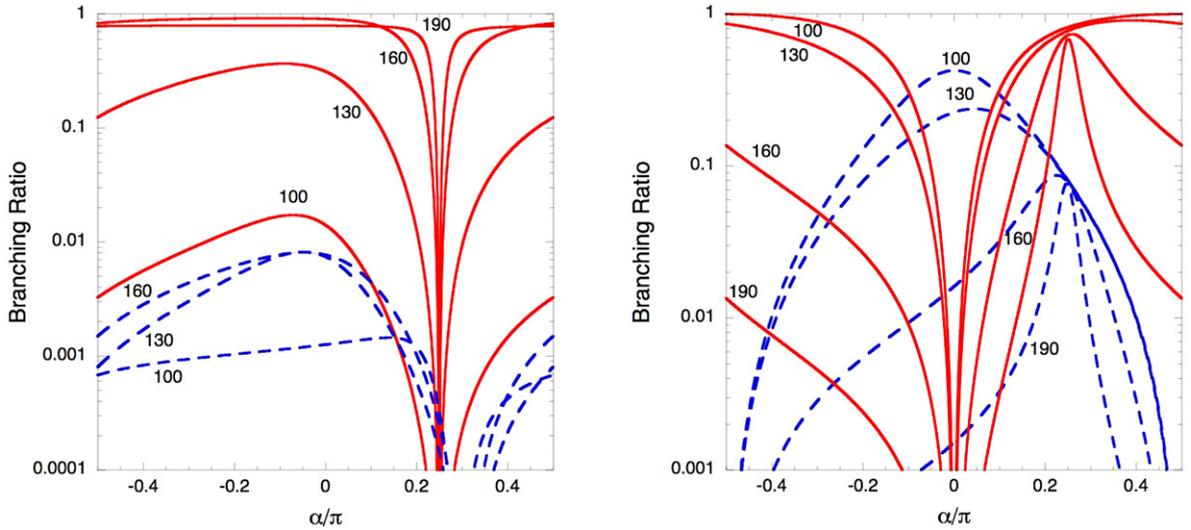
Fig. 10. Branching ratios of the light Higgs boson in the flipped model. We have taken $\tan\beta = 1$. In the left figure, the solid (dashed) lines are the branching ratios into WW, $(\gamma\gamma)$; in the right figure, the solid (dashed) lines are the branching ratios into $b\bar{b}$ $(\tau^+\tau^-)$ for various values of the Higgs-boson mass.
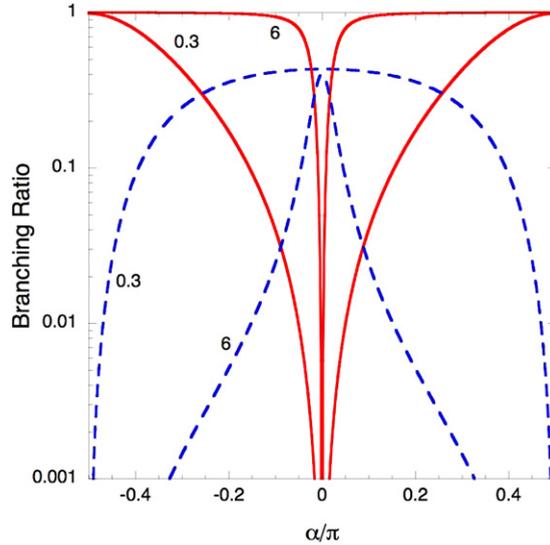
Fig. 11. Branching ratios of a Higgs boson with mass 100 GeV into fermions for $\tan\beta = 0.3$ and $\tan\beta = 6$ in the flipped model. The solid (dashed) lines, which are red (blue) are the branching ratios into $b\bar{b}$ $(\tau^+\tau^-)$.

The pseudoscalar Higgs does not decay into vector bosons, therefore its primary decays are into $\tau^+\tau^-$ and $b\bar{b}$. One has, in the flipped 2HDM,

$$\frac{\Gamma\left(A \to \tau^+\tau^-\right)}{\Gamma\left(A \to b\bar{b}\right)} = \left(\frac{\cot\beta}{1.76}\right)^4; \tag{22}$$

for $\tan\beta \geq 1$, this is 10% or less. For $\tan\beta = 0.3$, this would be roughly 10, leading to dominance of the $\tau^+\tau^-$ decay, as in the lepton-specific model, but for a different region of parameter space.

2.1.5. Higgs decays in the neutrino-specific 2HDM

In the neutrino-specific (NS) 2HDM the RH quarks and charged leptons all couple to Φ_2, but the right-handed neutrino couples to Φ_1. In this model, there are Dirac neutrino masses, and the vacuum expectation value of Φ_1 must be O(eV). As a result, the Yukawa coupling of Φ_1 to neutrinos can be O(1). The model originally used a Z_2 symmetry which was softly [89] or spontaneously [90,139] broken, in order to have such a small vev, but this allows for right-handed neutrino masses. The latter model, by Gabriel et al. [90] and Wang [139], also has a very light scalar, with mass of O(eV). Extending the symmetry
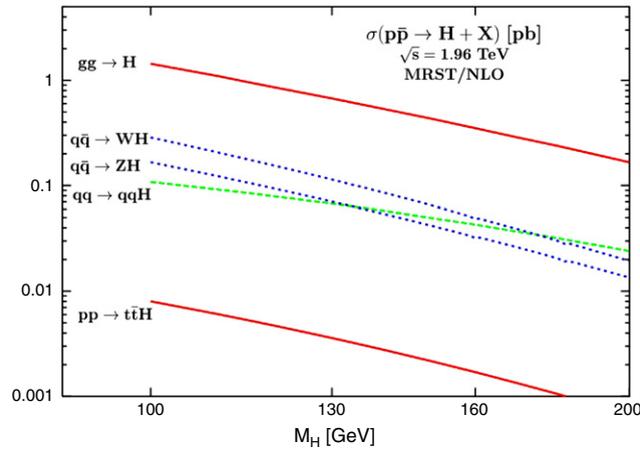
Fig. 12. Production cross sections for the Standard-Model Higgs at the Tevatron.
Source: This figure is from Djouadi's review article [24].

to a $U(1)$ and breaking it softly (to avoid a Goldstone boson) gives the model of Davidson and Logan [91]. This model is much less fine-tuned, since it requires that the soft $U(1)$ breaking term be of the same order as the electron mass (the small O(eV) vev arises in a see-saw like pattern). Thus, there are several versions of the model—those with a Φ_1 vev of O(eV) and a very light scalar, those with a similar vev and no light scalar, and those that use a see-saw type mechanism and have a Φ_1 vev of O(MeV).

The first of these models, with a light scalar, which was the original proposal of Gabriel et al. [90] and Wang, et al. [139], has very recently been excluded. Following work on the astrophysics of the model [140], Zhou [141] pointed out that there are serious problems with the model. In particular, the neutrinos emitted by SN1987a would, if there is a light scalar, interact strongly with the relic neutrino background and would not reach Earth. In addition, the effects of the neutrinos in the early universe would cause problems with the WMAP data. Thus, these models are excluded.

Because of the small vev, mixing between the Higgs doublets is negligible, and thus one scalar behaves just as the Standard Model Higgs. In the spontaneously broken version, they can also decay into the very light scalars, but these are invisible and thus the Standard Model Higgs would decay invisibly. In the Davidson–Logan model [91], the light Higgs decays as in the Standard Model. The phenomenology of the charged Higgs, which decays into charged leptons and a right-handed neutrino, is interesting and discussed there. The phenomenology of the new neutral scalars in the model, however, is much less interesting, since they will only decay into neutrinos, and thus appear as invisible Higgs decays.

In another version [142,143], there is relatively little fine-tuning in the potential, and the $\Phi_1^\dagger \Phi_2$ term has a coefficient as large as $10\,\text{GeV}^2$. Through a see-saw type mechanism, the vev of Φ_2 is O(MeV), and a further see-saw gives the light neutrino masses. The vacuum stability of the model is discussed by Haba and Horita in Ref. [144]. A recent comprehensive analysis of this model, including precision constraints, can be found in Ref. [145]. There are also interesting leptogenesis effects, which can be found in works of Haba and Seto [146,147].

2.2. Higgs production

The production of the Higgs boson of the Standard Model at hadron colliders is very well studied. Calculations of two-loop corrections, including next-to-next-to-leading order calculations and soft-gluon resummations, have all been carried out in detail. A very extensive discussion, with hundreds of references, is in the review by Djouadi [93].

The leading channel at the Tevatron and at the LHC is gluon fusion, in which two gluons and the Higgs are at the vertices of a triangle with a top quark going round the loop. The bottom quark plays little role since its coupling to the Higgs is so small (note that this may not hold in a 2HDM).

The next leading channel, at the LHC, is W fusion or Z fusion. At the Tevatron, Higgsstrahlung off a W or a Z is the second leading channel. Higgsstrahlung off a top quark might also be observable.

The production cross sections, from Djouadi's review, are in Figs. 12 and 13. Although gluon fusion dominates, it may not always provide the best signature. For example, in order to detect a Higgs H of mass 120 GeV at the Tevatron. One must look for decays $H \to b\bar{b}$. If the Higgs is produced via gluon fusion, then the background of $b\bar{b}$ will be much too large. Instead, Higgsstrahlung from a W or Z will help tag the $b\bar{b}$ with the decay of the W or Z. Thus, the searches for a low-mass Higgs at the Tevatron focus on WH and ZH production, while the searches for a high-mass Higgs focus on gluon fusion. In practice, of course, all production modes must be considered.

In the following, we will look at each of these processes, and discuss how they are modified in the 2HDMs explored in this chapter. In most cases we will determine the ratio of the production cross section in the 2HDM at hand to the one in the Standard Model, and then one can use Djouadi's figures to determine the absolute cross section.
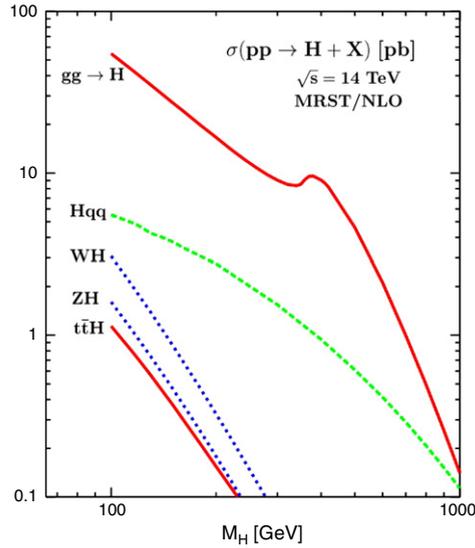
Fig. 13. Production cross sections for the Standard-Model Higgs at the LHC.
Source: This figure is from Djouadi's review article [24].

2.2.1. Gluon fusion

In the Standard Model, at the parton level, the cross section for gluon fusion to a Higgs of mass m_h is $m_h^2 \delta \left(\hat{s} - m_h^2 \right) \sigma_o$, where

$$\sigma_o = \frac{G_\mu \alpha_s^2}{512 \sqrt{2} \pi} \left| \sum_q A_{1/2}^h \left(\tau_q \right) \right|^2. \tag{23}$$

Here, G_μ is the Fermi constant from muon decay, α_s is the strong coupling constant, $\tau_q = m_h^2 / \left(4 m_q^2 \right)$, and $A_{1/2}^h \left(\tau \right) = 2 \left[\tau + (\tau - 1) f \left(\tau \right) \right] / \tau^2$, where

$$f \left(\tau \right) = \begin{cases} \arcsin^2 \left(\sqrt{\tau} \right) & \Leftarrow \quad \tau \leq 1, \\ -\frac{1}{4} \left(\log \frac{1 + \sqrt{1 - 1/\tau}}{1 - \sqrt{1 - 1/\tau}} - i\pi \right)^2 & \Leftarrow \quad \tau > 1. \end{cases} \tag{24}$$

This can be substantially simplified in the limits of large or small τ, which apply in most cases that we will consider. In the limit $m_q \gg m_h$, i.e. when $\tau_q \ll 1$, $A_{1/2}^h \left(\tau_q \right) \to 4/3$, whereas in the limit $m_q \ll m_h$, i.e. when $\tau_q \gg 1$, $A_{1/2}^h \left(\tau_q \right) \to - \left[\log \left(4 \tau_q \right) - i\pi \right]^2 / \left(2 \tau_q \right)$.

In the 2HDM, one's first thought is that the change in the cross section would be trivial. If only top-quark loops contribute, then the only difference is in the coupling of the Higgs to the top quark, and so one would multiply the SM cross section by $(\cos \alpha / \sin \beta)^2$ in the case of h production, or $(\sin \alpha / \sin \beta)^2$ in the case of H production, in all four 2HDMs. In the case of gluon-fusion production of the pseudoscalar A, the form factor is instead $A_{1/2}^A \left(\tau_q \right) = 2 f \left(\tau_q \right) / \tau_q$, and moreover one must multiply the whole cross section by $\cot^2 \beta$.

While the reasoning above indeed applies in models without a substantially enhanced b-quark Yukawa coupling, such as the type I 2HDM or the lepton-specific 2HDM, it does not necessarily apply in the cases of the type II and flipped 2HDMs (the behaviour of the lepton couplings has no effect on gluon fusion). This is because the b-quark loops can become crucial, and for large $\tan \beta$ can actually dominate the cross section.

In the type I or lepton-specific 2HDM, the production cross section of a light Higgs, h, through gluon fusion is simply multiplied by the factor $(\cos \alpha / \sin \beta)^2$. In the decoupling limit, i.e. when $\cos(\alpha - \beta)$ is small, this is near unity, but in general it is quite smaller. This factor is somewhat discouraging for the $\tau^+ \tau^-$ signature of the lepton-specific model. As one can see from Fig. 7, $h \to \tau^+ \tau^-$ is dominant (when $\tan \beta = 1$) near $\alpha = \pm \pi / 2$, precisely where the factor $(\cos \alpha / \sin \beta)^2$ strongly suppresses the production rate. For larger $\tan \beta$, however, as seen in Fig. 8, the region in which the $\tau^+ \tau^-$ decays dominate does extend a substantial distance away from $\alpha = \pm \pi / 2$, and then the suppression would not be that severe.

In the type II or flipped 2HDM, the contribution of the top quark is also multiplied by $(\cos \alpha / \sin \beta)^2$, but now the diagram with a b-quark loop can contribute for large $\tan \beta$, since then the b Yukawa coupling becomes large. The b-loop contribution to the amplitude is multiplied by $- \tan \alpha \tan \beta$ relative to the t-loop contribution. Using the limits of small τ_q (when $q = t$) and large τ_q (when $q = b$) above, we find that the cross section increases over the type I 2HDM cross section by a factor of $|1 + (5 - 8i) \tan \alpha \tan \beta / 100|^2$ for $m_h = 100$ GeV, and that this correction roughly scales as m_h^{-2} for other values of m_h.

Note that $\tan\alpha$ can be of either sign in the type II 2HDM; in the MSSM it is always negative. Numerically, for positive α, this is 1.1 (1.4, 2.9, 12.0) for $\tan\alpha\tan\beta = 1$ (3, 10, 30). For negative α, this is 0.9 (0.8, 0.9, 6.0) for $\tan\alpha\tan\beta = -1$ ($-3, -10, -30$). We see that for large $\tan\beta$, one can get an increase of almost an order of magnitude in the production rate relative to the type I 2HDM. However, one should keep in mind the discussion of the last section in which it is noted that values of $\tan\beta$ larger than 6 are allowed (by unitarity and perturbation theory) only for a very small region of parameter space, so very large enhancements are unlikely.

For the heavy neutral Higgs, H, the results are similar. In the type I or lepton-specific 2HDM, one must multiply the SM production cross section by the factor $(\sin\alpha/\sin\beta)^2$. As before, this is small in the region where the $\tau^+\tau^-$ branching ratio is large. In the type II or flipped 2HDM, the factor $-\tan\alpha\tan\beta$ in the previous paragraph becomes $\cot\alpha\tan\beta$. The results of the previous paragraph are thus qualitatively unchanged. Note that when the mass of the H mass rises beyond the mass of the top quark, the rate of gg-fusion production of H drops quickly, just as in the Standard Model.

For gg-fusion production of the pseudoscalar A, the form factor $A^A_{1/2}$ is simply $2f(\tau)/\tau$, which in the limit $m_t \gg m_A$ is 2 instead of 4/3. There is also a factor $\xi^u_A = \cot\beta$ in the amplitude. This yields a total factor $(9/4)\cot^2\beta$ in the production cross section of an A in the type I 2HDM compared to the production cross section of an SM Higgs of identical mass. This will be substantial for small $\tan\beta$. For the type II or flipped 2HDM, the ratio of production cross section relative to the type I 2HDM is, for $m_A = 100$ GeV, $\left|1 - (3.5 - 4i)\tan^2\beta/100\right|^2$. Numerically, this is 0.93 (0.82, 0.59, 1.44) for $\tan^2\beta = 1$ (3, 10, 30). The difference of production cross sections between the various 2HDMs is thus much less dramatic.

2.2.2. WH and ZH production, WW and ZZ fusion, and $b\bar{b}H$ and $t\bar{t}H$ production

It is straightforward to see how these production rates are affected in 2HDMs. First consider the production processes involving vector bosons ($W^* \to WH$, $Z^* \to ZH$, WW to$H$ or $ZZ \to H$ where H is a neutral Higgs). In all four 2HDMs, for the light Higgs h the SM rate is multiplied by $\sin^2(\alpha - \beta)$, and for the heavy Higgs H it is multiplied by $\cos^2(\alpha - \beta)$. The pseudoscalar A cannot be produced via this mode, since there are no W^+W^-A and ZZA vertices.

For $t\bar{t}H$ production, one simply multiplies the SM rate by the appropriate coupling-constant factor. In all four models, this is $(\cos\alpha/\sin\beta)^2$ for the light scalar h, $(\sin\alpha/\sin\beta)^2$ for the heavy scalar H, and $\cot^2\beta$ for the pseudoscalar.

Contrary to what happens in the Standard Model, there is in some 2HDMs the possibility of substantial $b\bar{b}H$ production. The rate, well above the threshold, is, compared to the SM $t\bar{t}H$ production cross section, negligible for the type I and lepton-specific 2HDMs, but for the type II and flipped 2HDMs it is $(\sin\alpha/\cos\beta)^2 (m_b/m_t)^2$ for the light h. A detailed study, with numerous references, of this mode in the MSSM by Schaarschmidt [148–150], showed that an h in the 120–200 GeV mass range can be discovered at the LHC through this process, if $\tan\beta > 6$, by looking for the signature $h \to \tau^+\tau^-$. In fact, for a large $\tan\beta$, which is not unnatural in the MSSM, this can become the primary discovery mode, especially for a relatively light Higgs h. One would expect a similar result in the type II 2HDM (the flipped model would have less enhancement in the $h \to \tau^+\tau^-$ decay mode), although, as noted earlier, large values of $\tan\beta$ are not favoured in the type II 2HDM.

2.2.3. Other production mechanisms

With several neutral scalars, another production mechanism can become important. One can produce two Higgs bosons (either the same one or different ones). (This can also occur in the Standard Model, of course.) In any diagram in which a scalar is produced, that scalar can then convert itself into two scalars via a trilinear coupling. In addition, one can have double-Higgsstrahlung off a W or Z, or one can have a gluon fusion into two scalars via a box diagram with a heavy quark in the loop; the latter mechanisms are independent of the trilinear coupling. An analysis of all these mechanisms was done in Ref. [151]. Total cross sections for the LHC were given in Djouadi's review [93], and are dominated by gluon fusion, giving cross sections of order 10 fb over the low-mass range for the scalars. Although detection is much more difficult due to large backgrounds [152,153], analyses [154] indicate that it might be possible to distinguish the trilinear coupling from zero with an integrated luminosity of 300 fb^{-1}, for scalar masses in the 150–200 GeV range; a higher luminosity is needed for lighter scalars.

In the MSSM there are many more possibilities involving neutral scalars. One can have $q\bar{q} \to Z^* \to hA$, HA or, through triangle and box graphs, $gg \to hh$, HH, hH, AA, hA, HA. In addition, one can have double Higgsstrahlung. These are discussed in Refs. [155,156] and summarized in Djouadi's review [24]. For the quark annihilation process, cross sections range from 10 to 100 fb as m_A varies from 100 to 170 GeV, and for gluon fusion the cross sections (for $\tan\beta = 30$) are a little more than an order of magnitude larger. Detection is difficult, but possible, and is summarized in Ref. [24].

In the general 2HDM, one has the same processes as in the MSSM, but there is now a much larger parameter space. The quark-annihilation process has precisely the same form as in the MSSM, but now one need no longer have a small $\cos(\alpha - \beta)$, and thus the ZhA coupling can be larger than in the MSSM. This is encouraging and leads to some interesting possibilities. For example, in the lepton-specific model and for $\tan\beta > 2$, the dominant decay of the A is into $\tau^+\tau^-$, and for much of parameter space the decay of the light Higgs is also into $\tau^+\tau^-$. Thus one might have four-τ events with branching ratios as high as tens of femtobarns. This signature needs further investigation. A study of pair production of the lightest Higgs bosons in the type II model was carried out in Refs. [157–159]. They showed that while pair production in the Standard Model is very difficult to observe at the LHC, it can be bigger in the type II model, and they also show that there can be sensitivity to the quartic couplings, which could help distinguish the model from the MSSM.

For the gluon-initiated process, triangle diagrams produce a single $h/H/A$, real or virtual, which then converts into a pair of scalars. Alas, this process is proportional to trilinear scalar couplings and, while these are known in the MSSM, they are unknown in the general 2HDM. The box diagrams which give gluon fusion into two scalars will be similar to those of the MSSM. Thus all one can really say is that the rate could be substantially larger than in the MSSM, but accurate predictions are impossible.

2.3. The inert Higgs model

The inert Higgs model is a 2HDM with an unbroken Z_2 symmetry under which one of the doublets transforms non-trivially, *viz.* $\Phi_2 \to -\Phi_2$, and all other SM fields are invariant. This 'parity' imposes natural flavour conservation. Initially a similar model [160] was introduced to explain neutrino masses. More recently such a model was proposed in the context of radiative neutrino masses [161] and also to attack the naturalness problem of the SM by allowing for a larger mass (between 400 and 600 GeV) for the SM Higgs while keeping full consistency with electroweak precision tests [162], thus solving the 'little hierarchy' problem [163]. Even more recently, an inert doublet was introduced to allow for the possibility of several mirror families of fermions [164].

In the inert Higgs model the Higgs doublet Φ_2 – the inert doublet – does not couple to matter and acquires no vacuum expectation value, leaving the Z_2 symmetry unbroken. The scalar spectrum consists of the SM-like Higgs obtained from Φ_1 and one charged and two neutral states from Φ_2. Since the Z_2 is unbroken the lightest inert particle will be stable and will contribute to the dark matter density [161,162]. This possibility has been analysed by several authors [165–172]. The early cosmological evolution of the model has been discussed by Ginzburg et al. in [173].

The scalar potential is the one in Eq. (2) but with $m_{12}^2 = 0$. The asymmetric phase, where

$$\langle \phi_1^0 \rangle = \frac{v}{\sqrt{2}} \quad \text{and} \quad \langle \phi_2^0 \rangle = 0, \tag{25}$$

corresponds to a sizeable region of parameter space [161,162] and the scalar masses are given by

$$
\begin{aligned}
m_h^2 &= \lambda_1 v^2, & m_S^2 &= m_{22}^2 + \frac{(\lambda_3 + \lambda_4 + \lambda_5)\, v^2}{2}, \\
m_+^2 &= m_{22}^2 + \frac{\lambda_3 v^2}{2}, & m_A^2 &= m_{22}^2 + \frac{(\lambda_3 + \lambda_4 - \lambda_5)\, v^2}{2},
\end{aligned}
\tag{26}
$$

where h is the usual Higgs boson obtained from Φ_1, $(S + iA)/\sqrt{2}$. is the neutral component of the inert Higgs doublet, and H^+ is the charged inert Higgs field. The doublet Φ_1 gives mass to the gauge bosons $W^\pm$ and Z^0. In the limit of Peccei–Quinn symmetry, $\lambda_5 \to 0$, the neutral inert scalars become degenerate. Direct detection of halo dark matter places a limit on this degeneracy [174].

The inert scalars can be produced at colliders through their couplings to the electroweak gauge bosons subject to the constraint of the Z_2 symmetry [175,176]. In addition, they also participate in cubic and quartic Higgs couplings:

$$
\begin{aligned}
V_{\text{interactions}} &= \frac{\lambda_2}{2} \left(H^+ H^- + \frac{S^2 + A^2}{2} \right)^2 + \lambda_3 \left(vh + \frac{h^2}{2} \right) \left(H^+ H^- + \frac{S^2 + A^2}{2} \right) \\
&\quad + \frac{\lambda_4 + \lambda_5}{2} \left(vh + \frac{h^2}{2} \right) S^2 + \frac{\lambda_4 - \lambda_5}{2} \left(vh + \frac{h^2}{2} \right) A^2.
\end{aligned}
\tag{27}
$$

As pointed out in Ref. [175], assuming the mass hierarchy $m_+^2 > m_A^2 > m_S^2$, the stable scalar S appears as missing energy in the decays of H^+ and A. Since there is no linear term in A or S in Eq. (27), the decay of A must occur through the gauge interaction

$$\frac{g}{2c_W} Z_\mu \left(S\partial^\mu A - A\partial^\mu S \right). \tag{28}$$

Hence [175] the dominant decay of A is into $Sf\bar{f}$, where f is a fermion, either a lepton or a quark, and S appears as missing energy. Concerning $H^\pm$, the gauge interactions with S and A are given by:

$$\frac{ig}{2} W_\mu^- \left(S\partial^\mu H^+ - H^+\partial^\mu S \right) + \frac{g}{2} W_\mu^- \left(A\partial^\mu H^+ - H^+\partial^\mu A \right) + \text{H.c.} \tag{29}$$

Hence the dominant decays of $H^\pm$ are into $W^\pm S$ and $W^\pm A$, with, in the second case, subsequent decay of A into ZS.

There are also trilinear gauge interactions among $H^\pm$, Z, and γ, as well as the quadrilinear terms required by gauge invariance.

The LEP I data on the width of the Z^0 gauge boson force the sum of the masses of the S and A to be larger than the mass of the Z^0 [169,175], thus preventing $Z^0 \to AS$. The electroweak precision tests put constraints on the inert-scalar mass splittings as a function of the h mass [162]. If the Z_2-odd scalars are much lighter than the SM Higgs boson, they may have

a great impact on the direct search for the latter, because $h \rightarrow SS$ and $h \rightarrow AA$ may become the dominant decay channels. In particular, the $h \rightarrow SS$ channel is invisible. The lower limit $m_h > 114.4$ GeV, obtained by LEP II, was based on a direct search via $h \rightarrow b\bar{b}$, which is dominant for a light Higgs boson in the SM. In the present case this limit can be relaxed down to about 106 GeV, assuming the mass splitting of the new neutral scalars to be $m_A - m_S = 10$ GeV, due to the fact that in this model the invisible decay of the Higgs boson may be dominant and, as a result, the decay branching ratio into $b\bar{b}$ is highly suppressed.

A scan over the parameter space was performed in [175], assuming the charged-Higgs mass to be much larger than m_h, the mass splitting of the new neutral scalars to be $m_A - m_S = 10$ GeV, and also taking into account relevant constraints including the Z^0-decay constraint mentioned above and the vacuum stability bound. The invisible decay mode $h \rightarrow SS$ is found to dominate in the light- and intermediate-mass region of m_h, i.e. when $100 < m_h < 160$ GeV. The 10 GeV mass gap between A and S implies that the contribution of the SS mode to the decay of the SM-like Higgs boson h is much larger than that of the AA mode. In particular, the branching ratio for the invisible mode is 50%–65% in the mass region $100 < m_h < 150$ GeV and for $m_S \sim 40$–60 GeV. As a result, the usual decay modes of the SM Higgs boson are highly suppressed. This fact, together with the strong suppression of the $\gamma\gamma$ mode, will make it very difficult, in this framework, to use this mode to detect the SM-like Higgs boson at the LHC in the mass range $100 < m_h < 150$ GeV [175]. However, it was also shown there that it is very promising to look for the SM-like Higgs boson through its invisible decay in the so-called weak-boson-fusion process [177]. This process is of the form $q\bar{q} \rightarrow q'q'VV \rightarrow q'q'h$ (V denotes a gauge boson), with the subsequent decay of h to undetectable particles. The decay mode $h \rightarrow W^+W^-$ dominates over the invisible one for h heavier than 160 GeV, where it starts behaving as in the SM. The decay pattern and decay branching ratios of the new neutral scalars have also been examined in Ref. [175]. They conclude that an S with mass ~ 50 GeV should be observable at the LHC and, at the same time, it would constitute a good dark-matter candidate. A complementary analysis of collider phenomenology of the inert Higgs model was performed in Ref. [176].

The possibility of extending the inert Higgs model in order to introduce CP-violation in the scalar sector was considered in Refs. [178,179]. In another extension of the inert doublet model, Barr and Kephart [180] pointed out that a difficulty with generic multi-doublet models is that many fine-tunings are necessary to alleviate the hierarchy problem, and thus, if one believes that the hierarchy problem of the Standard Model is solved via anthropic fine-tuning, then having more than one Higgs doublet becomes exceedingly unlikely. They looked for the conditions under which an N doublet model only has a single fine-tuning, and found that if the N doublets formed a representation of a global symmetry group, then only one fine-tuning is necessary, and classified all possible symmetry groups with $N \leq 6$. They found that in some models, such as $SU(N)$ with an N-plet, the extra Higgs doublets become completely inert, and the scalars and pseudoscalars are all degenerate in mass. Different mass relations arise from different representations.

A different version of the inert Higgs model, which has recently received increasing attention, is the Lee–Wick Standard Model (LWSM) [181–183], whose neutral sector is similar to the inert Higgs model but whose charged sector is similar to a type II model with $\tan\beta = 1$. In this model, for every SM field, a higher-derivative kinetic term is introduced; for a Higgs scalar, this term is quartic in the derivatives. These terms lead to the presence of additional poles in the propagators. By introducing auxiliary fields, the model can be written as one without higher-derivative terms, but with additional fields corresponding to Lee–Wick partners (one for each field of the SM). The signs of the kinetic and mass terms of these partners are opposite to those of normal particles, i.e. the states have negative norm.

A major attractiveness of the LWSM is the elimination of quadratic divergences in the Higgs sector. In supersymmetry, the minus sign which cancels those divergences arises from the different statistics of the partners; in the LWSM, it arises from a sign difference in the propagator due to the negative-norm states. If those states are not stable, then they do not appear as 'out' states in the S matrix and unitarity is preserved. While one does have microcausality violation, no logical paradoxes arise macroscopically [182,184]. There is now an extensive literature on the LWSM [185–197].

A comprehensive analysis of the Higgs sector of the LWSM was recently carried out by Carone and Primulando [198]. They were motivated by the realization [189] that electroweak precision constraints do not severely constrain the Lee–Wick mass scale for the Higgs sector. The Higgs doublet and its Lee–Wick partner form a 2HDM, but the mixing between the neutral states is symplectic rather than orthogonal. The neutral-scalar sector is very similar to the one of the inert Higgs model, but with some differences in signs. The primary focus of Ref. [198] was on the charged Higgs sector (which will be discussed in Section 4), but they did find bounds from LEP, showed that constraints in the charged sector also constrain the neutral sector, and have plotted bounds in the neutral-Higgs-boson plane. The effects of the model on electroweak parameters and on the $Zb\bar{b}$ coupling were studied by Chivukula et al. [199] A more detailed analysis of the neutral sector, including bounds from direct searches at LEP and the Tevatron, as well as prospects at the LHC, was very recently published by Alvarez et al. [200]; they find interesting differences in the usual sum rules in the neutral sector; for example, the usual relation $g^2_{h1-V-V} + g^2_{h2-V-V} = 1$ has a different sign between the two terms. They used the HiggsBounds code to study implications at colliders.

Another 2HDM, called the "quasi-inert" model, was recently introduced by Cao et al. [201]. In that model, motivated by the possible observation of W plus dijets at the Tevatron [202], there is a second doublet whose tree-level vev vanishes, and which couples primarily to the first generation of quarks. They set to zero the $\Phi_1^\dagger \Phi_2 +$ H.c. term in the potential, which can be accomplished by a Z_2 symmetry in which both Φ_2 and u_R change sign. This symmetry must be weakly broken to allow the up quark to get a mass, and this results in a small vev for Φ_2. Cao et al. calculate electroweak precision effects and flavour constraints on their model. The primary motivation for the model is that the Tevatron can produce the charged

Higgs ($p\bar{p} \to H^\pm$), which subsequently decays into $H^0 W^\pm$ or $A^0 W^\pm$, leading to $\ell^\pm \nu jj$ or to $W^\pm$ plus dijets. Since the charged Higgs is produced resonantly, the signal can be large. Cao et al. show that a reasonable region of parameter space exists which can explain the recent observation of W plus dijets at the Tevatron [202]. Even if this observation is not confirmed, the model is interesting in its own right and leads to other unique signatures. Another scenario by Chen et al. [203] switched the role of the neutral and charged scalars, and then the dijet comes directly from the charged scalar. This can be produced non-resonantly by $W^\pm H^\mp$ production, and can also explain the non-observation of a resonance in $\ell \nu jj$ by CDF (although the bounds on that are much weaker and a resonance could still exist). Similar models, which focus more on explaining the $B_s \to \mu\mu$ rate but also discuss the W plus dijet signature, can be found in Refs. [203,204].

3. Models with tree-level flavour-changing neutral currents

3.1. The type III 2HDM

In the previous chapter it was shown that one can eliminate the potentially dangerous tree-level FCNC through a discrete symmetry. Suppose, however, that we reject any such symmetry. The tree-level FCNC can certainly be suppressed by making the neutral scalars extremely heavy, but scalar masses in the multi-TeV range (or higher) seem unnatural. In this section, we examine the constraints from FCNC and show that a reasonable *Ansatz* for the neutral flavour-changing couplings allows for scalar masses well below the TeV scale.

It is easiest to discuss the tree-level FCNC in the Higgs basis described in Section 5. In that basis, the scalar doublets are rotated so that the vev is entirely in the first doublet, while the second doublet has zero vev. The general Yukawa couplings can be written as

$$\mathcal{L}_{\text{Yukawa}} = \eta^U_{ij} \bar{Q}_{iL} \tilde{H}_1 U_{jR} + \eta^D_{ij} \bar{Q}_{iL} H_1 D_{jR} + \eta^L_{ij} \bar{L}_{iL} H_1 E_{jR} + \hat{\xi}^U_{ij} \bar{Q}_{iL} \tilde{H}_2 U_{jR} + \hat{\xi}^D_{ij} \bar{Q}_{iL} H_2 D_{jR} + \hat{\xi}^L_{ij} \bar{L}_{iL} H_2 E_{jR} + \text{H.c.}, \tag{30}$$

where H_1 and H_2 are the two scalar doublets. In the Higgs basis those doublets have been rotated so that only H_1 has a vev, *i.e.*

$$\langle H_1 \rangle_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}, \qquad \langle H_2 \rangle_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{31}$$

where v is real. In this basis, only the Yukawa couplings of the doublet H_1, *viz.* the η_{ij}, generate fermion masses; those η_{ij} may be bi-diagonalized and do not lead to tree-level FCNC. When that bi-diagonalization is performed, the neutral flavour-changing couplings become

$$\mathcal{L}_{\text{FCNC}} = \xi^U_{ij} \bar{U}_{iL} H^{0*}_2 U_{jR} + \xi^D_{ij} \bar{D}_{iL} H^0_2 D_{jR} + \xi^L_{ij} \bar{L}_{iL} H^0_2 L_{jR}, \tag{32}$$

where

$$\xi^{U,D,L} = V^{U,D,L\dagger}_L \hat{\xi}^{U,D,L} V^{U,D,L}_R. \tag{33}$$

Since V_R is completely unknown and the $\hat{\xi}$ are arbitrary, these $\xi^{U,D,L}$ coefficients are arbitrary; in order to look at specific processes, some assumptions must be made about their magnitudes.

One of the earliest papers discussing tree-level FCNC was the one of Bjorken and Weinberg [205], who considered radiative muon decay and chose $\xi^L_{\mu e}$ to be the Yukawa coupling of the muon. Later, in 1980, McWilliams and Li [25] and Shanker [26] considered K–$\bar{K}$ mixing, as well as many processes involving kaon and muon decays. They argued that the heaviest fermion sets the scale for the entire Yukawa-coupling matrix. The flavour-changing vertex should be the product of the largest Yukawa coupling and a mixing angle factor. Since they did not know the mixing angle factors, they set them equal to one. Thus, the $\xi^{U,D,L}$ were set equal to the top, bottom, and tau Yukawa couplings, respectively. The most stringent bound came from K–$\bar{K}$ mixing and led to a lower bound of 150 TeV on the mass of H^0_2. For most of the 1980's, this led most authors to assume that there must be a discrete symmetry which prohibits the FCNC, and attention focused on the type I and type II 2HDMs.

Cheng and Sher [206] argued that this estimate of the lower bound is not reasonable. They argued that the most conspicuous feature of the fermion mass structure is its hierarchical structure and that, therefore, setting all the flavour-changing couplings to be equal to the heaviest-fermion Yukawa coupling was not reliable. They proposed what has since become known as the Cheng–Sher *Ansatz*: that the flavour-changing couplings should be of the order of the geometric mean of the Yukawa couplings of the two fermions. In other words,

$$\xi_{ij} = \lambda_{ij} \sqrt{m_i m_j} \frac{\sqrt{2}}{v}, \tag{34}$$

where the λ_{ij} are of order one. Since the most severe bounds on FCNC arise from the first two generations and this *Ansatz* especially suppresses the Yukawa couplings of those generations, it will reduce the lower bound on the Higgs mass.

More specifically, Cheng and Sher's argument was as follows. Consider a model with n Higgs doublets Φ_i ($i = 1 \ldots n$) and call λ'_i the matrix of Yukawa couplings to, say, the charge $-1/3$ quarks. First suppose that the fermion mass matrix is of

the Fritzsch form[5]:

$$
M = \begin{pmatrix} 0 & A & 0 \\ A & 0 & B \\ 0 & B & C \end{pmatrix}.
\tag{35}
$$

In this case, calling the eigenvalues m_1, $-m_2$, and m_3, one has $A \simeq \sqrt{m_1 m_2}$, $B \simeq \sqrt{m_2 m_3}$, and $C \simeq m_3$. Cheng and Sher then simply assumed that the Yukawa-coupling matrices were given by

$$
\lambda'_i = \frac{\sqrt{2}}{v_i} \begin{pmatrix} 0 & A_i & 0 \\ A_i & 0 & B_i \\ 0 & B_i & C_i \end{pmatrix},
\tag{36}
$$

and that their matrix elements had the same structure as the full mass matrix, namely $A_i = a_i \sqrt{m_1 m_2}$, $B_i = b_i \sqrt{m_2 m_3}$, and $C_i = c_i m_3$, with coefficients a_i, b_i, c_i of order unity. In other words, the requirements, obtained by comparing the λ'_i with M,

$$
\sum_i a_i = \sum_i b_i = \sum_i c_i = 1
\tag{37}
$$

are not satisfied through any fine-tuned cancellations among the different couplings. Essentially, the *Ansatz* states that mass matrix zeros are not obtained through any cancellations among non-zero matrix elements, and that non-zero mass matrix elements are also not obtained through precise cancellations among larger terms. Cheng and Sher then pointed out that this argument does not apply only to mass matrices with the Fritzsch structure, but to any other structure in which one requires that the hierarchy of eigenvalues does not arise through delicate cancellations. It was later noted [208] that, if the hierarchical structure is due to approximate flavour symmetries, then the Cheng–Sher *Ansatz* will be satisfied. The *Ansatz* is thus quite general.

In the type III[6] model the Cheng–Sher *Ansatz* is assumed and its implications are explored. Many papers focused on a few specific processes, including Δm_B [210], $t \rightarrow ch$, and $h \rightarrow \bar{t}c + \bar{c}t$ [209], rare μ, τ, and B decays (emphasizing $B \rightarrow K\mu\tau$) [211], $\mu \rightarrow e\gamma$ (at two-loop level) [212], $t \rightarrow c\gamma$ and $t \rightarrow cZ^0$ [213], muon–electron conversion [214], and $b \rightarrow s\gamma$ [215]. In 1996, an extremely comprehensive analysis of the model by Atwood et al. [216] looked at many of these processes as well as at implications for $Z^0 \rightarrow b\bar{b}$ and $Z^0 \rightarrow b\bar{s} + s\bar{b}$.

Over the years, the various bounds have steadily improved. If the type III model is correct, then one would expect the λ_{ij} in Eq. (34) to be all of order unity, but this is a fairly loose requirement since there are unknown mixing angles. Scalar masses will enter in all specific processes. In the Higgs basis, the imaginary part of H_2^0 is the usual pseudoscalar, A, whereas the real part of H_2^0 is a linear combination of h and H. Since H_2^0 has no vev, its coupling to W^+W^- vanishes, and thus the relevant mixing angle which rotates (H_2^0, H_1^0) into (h, H) is $\alpha - \beta$. In a model with tree-level scalar exchange one has an effective mass in the matrix element given by

$$
\frac{1}{m_{\text{eff}}^2} = \frac{c^2}{m_h^2} + \frac{s^2}{m_H^2},
\tag{38}
$$

where $s \equiv \sin(\alpha - \beta)$ and $c \equiv \cos(\alpha - \beta)$. Henceforth, we shall refer to the scalar mass m_{eff} and to the pseudoscalar mass m_A as free parameters. Some processes, such as $B_s \rightarrow \mu\tau$, proceed only through pseudoscalar exchange, while other processes, such as $B_s \rightarrow K\mu\tau$, proceed only through scalar exchange. The most stringent bounds in the quark sector come from meson–antimeson mixing, to which both scalars and pseudoscalar contribute.

The most recent analysis of F^0–$\bar{F}^0$ mixing, where $F = K$, D, B_d, or B_s, was by Gupta and Wells [107], who considered the tree-level exchange of scalars and pseudoscalar and found that

$$
\Delta m_F = \frac{\xi_{ij}^2}{m_F} \left(\frac{S_F}{m_{\text{eff}}^2} + \frac{P_F}{m_A^2} \right),
\tag{39}
$$

$$
S_F = \frac{B_F f_F^2 m_F^2}{6} \left[1 + \frac{m_F^2}{(m_i + m_j)^2} \right],
\tag{40}
$$

$$
P_F = \frac{B_F f_F^2 m_F^2}{6} \left[1 + \frac{11 m_F^2}{(m_i + m_j)^2} \right].
\tag{41}
$$

In these expressions, m_F is the meson mass, f_F is the pseudoscalar decay constant, and B_F is the vacuum insertion parameter defined in Ref. [216]. Adding in quadrature the theoretical and experimental errors to the SM prediction, Gupta and Wells

---

[5] At the time of Cheng and Sher's paper the Fritzsch *Ansatz* gave acceptable mixing angles. A generalization of it was proposed, and its phenomenology discussed in detail, in Ref. [207].

[6] The model was first referred to as type III in Ref. [209].

demanded that the sum of the SM value and the new contribution not exceed the experimental value by more than two standard deviations for the B_d and B_s systems; for the D and K systems, they demanded that the new contribution not exceed the experimental value by more than two standard deviations. The results, using the Cheng–Sher *Ansatz* and assuming that $m_{\text{eff}} = m_A = 120$ GeV, are

$$(\lambda_{ds}, \lambda_{uc}, \lambda_{bd}, \lambda_{bs}) \leq (0.1, 0.2, 0.06, 0.06). \tag{42}$$

This might seem problematic for the type III model, where one expects $\lambda_{ij} \sim 1$. But, the pseudoscalar contribution is typically a factor of 7–10 larger than the scalar contribution, therefore, increasing m_A to, say, 400 GeV increases all the bounds in Eq. (42) by more than a factor of three. This was illustrated more explicitly by Golowich et al. [217], who have shown that the bound is increased substantially as the pseudoscalar mass increases.[7] In addition, as pointed out in Ref. [216], there are additional diagrams, involving boxes and triangle graphs, some of them with a charged Higgs, and a mild cancellation involving these contributions would also weaken the bounds. Finally, it has been argued [218] that it might be more appropriate to use $\lambda_{ij}/\tan\beta$ rather than λ_{ij}, which for large $\tan\beta$ would also decrease the bounds.

Golowich et al. [217] also studied the effects of λ_{ct} and λ_{ut} on D–$\bar{D}$ mixing. For scalar masses between 100 and 400 GeV, they found that $\sqrt{\lambda_{ut}\lambda_{ct}}$ must be less than, approximately, 8–10. Thus, bounds on top-quark flavour-changing neutral currents are much weaker. Direct bounds on λ_{tc} can be obtained from $t \to ch$, which was first discussed by Hou [209] and much more recently in Refs. [219–221]. For a Higgs mass of 120 GeV, the branching ratio is roughly $0.005 \lambda_{ct}^2$. The fact that the SM decay $t \to bW$ fits its prediction implies that the branching ratio for $t \to ch$ cannot be very large, and thus one would expect λ_{ct} to be less than 10 or so, but we know of no current experimental bounds on this process. The signature for $t \to ch$ is quite different from that of $t \to H^+b$, as discussed in Section 4.

Other processes will bound products of the λ_{ij} instead of the individual λ_{ij}. For example, the processes $B_s \to \mu^+\mu^-$, with tree-level pseudoscalar exchange, and $B \to K\mu^+\mu^-$, with tree-level scalar exchange, will bound the product $\lambda_{bs}\lambda_{\mu\mu}$; it has been shown [107,222] that data will be obtained at the Tevatron and LHC which will bound the product to approximately 0.6 (for scalar or pseudoscalar masses of 120 GeV). One of the most interesting decays, since it vanishes in the SM, is $B \to K\mu\tau$. This process was searched for by BABAR [223] and, while the bounds only give $\sqrt{\lambda_{bs}\lambda_{\mu\tau}}$ to be less than O(10), the absence of an SM rate makes this promising for future B factories.

In the above we have focused on quarks. Bounds for leptons in the type III model have also been discussed extensively— with large mixing angles in the leptonic charged weak current, one imagines that mixing in the neutral current might be large. The anomalous magnetic moment of the muon, with neutral-scalar exchange, is proportional to $\lambda_{\mu\tau}^2$ and does not depend on any other couplings. This was first used in Ref. [224] to obtain a bound $\lambda_{\mu\tau} < 50$. Later, the experimental precision improved and there is now a significant discrepancy, of approximately 3σ, between that observable and its SM prediction (the precise discrepancy depends on the dataset used in determining hadronic corrections [225,226]). Diaz et al. [227,228] assumed that the discrepancy arises from flavour-changing neutral-Higgs exchange and obtained a *lower* bound on $\lambda_{\mu\tau}$, with $10 < \lambda_{\mu\tau} < 80$.

In the same work, Diaz et al. [227] considered $\mu \to e\gamma$, in which a scalar is exchanged and the internal fermion is a τ. Taking a range of values for the scalar mass, and a heavy pseudoscalar, they obtained $\lambda_{e\tau}\lambda_{\mu\tau} < 0.04$. This might be marginally acceptable for the type III model, similar to the bounds on F–$\bar{F}$ mixing discussed above. In this case, however, a very recent analysis by Hou et al. [229] has a different expression for the $\mu \to e\gamma$ width and they find that $\mu \to e\gamma$ is acceptable for $\lambda_{e\tau} = \lambda_{\mu\tau} = 1$ if the scalar mass is above 150 GeV. They also note that the contribution of the scalar and pseudoscalar to $\mu \to e\gamma$ have opposite signs, hence some cancellation might be possible. In any event, failure to observe a signal in $\mu \to e\gamma$ within the next three years at the MEG experiment – which will be sensitive to an additional two orders of magnitude in the rate – would be a serious problem for the type III model. Diaz et al. then used the lower bound on $\lambda_{\mu\tau}$ discussed in the previous paragraph to obtain $\lambda_{e\tau} < 10^{-3}$. Although their work did not explicitly use the *Ansatz* of the type III model, their conclusion – if the discrepancy in the anomalous magnetic moment of the muon is due to scalar exchange then there must be a substantial hierarchy between $\lambda_{\mu\tau}$ and $\lambda_{e\tau}$ – is robust, and the size of that hierarchy would rule out the type III model.

Other bounds may be obtained by mixing leptons and quarks. In Ref. [230] the process $\tau \to \mu P$, where P is a pseudoscalar meson, was studied. It was claimed that this process gives $\lambda_{\mu\tau} < 10^{-3}$, which is stronger than the previous bounds by several orders of magnitude. However, in Ref. [230] it was assumed that $\lambda_{uu}, \lambda_{dd}, \lambda_{ss}, \lambda_{bs}$, and λ_{bb} are all larger than 100, and that the light scalar and pseudoscalar have mass 120 GeV, leading to this stringent bound. Subsequently [231], in a study of $\tau \to \mu PP$, these assumptions were relaxed and $\lambda_{\mu\tau}\lambda_{ss} < 200$ was found.

Other bounds on the λ_{ij} do not involve flavour-changing neutral currents. If one assumes that only λ_{tt} and λ_{bb} are non-zero, then neutral scalars play no role, and one can bound λ_{tt} and λ_{bb} by considering the charged-Higgs contributions to $b \to s\gamma$ [232–234]; one obtains a bound on λ_{tt} of approximately 1.7, for a charged-Higgs mass below 300 GeV. More recently, Mahmoudi and Stal [46] analysed $b \to s\gamma$ and Δm_{B_d} as well as B, K, and D_s decays, and obtained bounds on all the second- and third-generation diagonal λ_{ij} (including for the leptons). Huang and Li [235] showed the Δm_{B_s} can also be used to bound λ_{cc} and λ_{ss}. Since we are focusing here on the FCNC, we shall not discuss these bounds further, but refer the reader to Ref. [46] for a detailed analysis. The type III model says nothing about the flavour hierarchy problem. Blechman

---

[7] The reader is cautioned that the parameter Δ_{ij} in Ref. [217] differs from our λ_{ij} by $\Delta_{ij} = \lambda_{ij}/\sqrt{2}$.

et al. [236], within a type III model scenario, looked for a basis independent constraint on the Yukawa couplings that would generate a mass hierarchy with the Yukawa couplings being of the same order. Their constraint was that the determinant of the Yukawa couplings of Φ_2 vanish. They require a fairly large, but not unreasonable, value of $\tan\beta$. The values of the λ_{ij} in their model depend on the pseudoscalar and heavy scalar masses. The phenomenological constraints are discussed in Ref. [236].

Over the past twenty years, bounds on the λ_{ij} have steadily improved, and the type III model is now facing challenges. In the leptonic sector, a negative result in the MEG experiment would essentially rule out the model (unless there is a cancellation between the scalar and pseudoscalar contributions). In the quark sector, a pseudoscalar in the 100–200 GeV mass range would cause problems for B–$\bar{B}$ and B_s–$\bar{B}_s$ mixing, and one might expect an observable signal in $B_s \to \mu^+\mu^-$. With the large top-quark sample at the LHC, the most promising signature is likely to be $t \to ch$; if h is light (as expected) and the branching ratio is below 10^{-3}, then the type III model would be in serious difficulty.

3.2. BGL models

All the tree-level flavour-changing transitions in the SM are mediated by the charged weak current, with the flavour mixing controlled by the CKM (quark mixing) matrix. Branco et al. [237] have built explicitly a class of 2HDMs (that we call BGL models) in which the tree-level flavour-changing couplings of the neutral scalars are related in an exact way to elements of the CKM matrix. In BGL models the required suppression of the scalar-mediated FCNC is obtained through relations involving the small off-diagonal elements of the CKM matrix V. Some variants of the BGL models fall into a wider category, which was coined later on [238] as models of Minimal Flavour Violation (MFV) [239–241].

The Standard Model with three families, consisting of $SU(2)$ doublets (Q_L and L_L) and $SU(2)$ singlets (U_R, D_R, and E_R) has a large flavour group of unitary transformations $G_F = U(3)^5$ which commutes with the gauge group. The gauge group G_F can be decomposed as

$$G_F \equiv SU(3)_q^3 \times SU(3)_l^2 \times U(1)_B \times U(1)_L \times U(1)_Y \times U(1)_{PQ} \times U(1)_{E_R}, \tag{43}$$

where

$$SU(3)_q^3 = SU(3)_{Q_L} \times SU(3)_{U_R} \times SU(3)_{D_R}, \tag{44}$$

$$SU(3)_l^2 = SU(3)_{L_L} \times SU(3)_{E_R}. \tag{45}$$

The notation is borrowed from Ref. [238]. The Yukawa couplings break the flavour group G_F. One can formally recover this flavour invariance by treating the Yukawa couplings as dimensionless fields (spurions), transforming in such a way that the Yukawa interactions become G_F-invariant. The MFV hypothesis consists of assuming that, even if New Physics exists, G_F is only broken by the Yukawa couplings, with dominance of the Yukawa coupling of the top quark. The most general scalar potential under the MFV hypothesis was recently derived [242].

The relevant question for BGL models is: under what conditions the neutral-scalar couplings in 2HDMs are only functions of the CKM matrix V? Namely, in BGL models these exact functions result from the imposition of discrete symmetries. Similar functions had been considered previously as an *ad hoc* assumption [208,243–245].[8]

Let us write down the Yukawa interactions:

$$\mathcal{L}_Y = -\overline{Q_L^0} \left(Y_1^d \Phi_1 + Y_2^d \Phi_2 \right) d_R^0 - \overline{Q_L^0} \left(Y_1^u \tilde{\Phi}_1 + Y_2^u \tilde{\Phi}_2 \right) u_R^0 + \text{H.c.} \tag{46}$$

After spontaneous symmetry breaking the quark mass matrices are

$$M_d = \frac{1}{\sqrt{2}} \left(v_1 Y_1^d + v_2 e^{i\alpha} Y_2^d \right), \qquad M_u = \frac{1}{\sqrt{2}} \left(v_1 Y_1^u + v_2 e^{-i\alpha} Y_2^u \right), \tag{47}$$

where α a general phase for the vev of Φ_2 and v_1 and v_2 are, without loss of generality, real. These matrices are bi-diagonalized as

$$U_{dL}^\dagger M_d U_{dR} = D_d \equiv \text{diag}\,(m_d, m_s, m_b), \tag{48}$$

$$U_{uL}^\dagger M_u U_{uR} = D_u \equiv \text{diag}\,(m_u, m_c, m_t). \tag{49}$$

In terms of the quark mass eigenstates u and d, the Yukawa couplings are:

$$\mathcal{L}_Y = \left[\frac{\sqrt{2}H^+}{v} \bar{u} \left(V N_d \gamma_R - N_u^\dagger V \gamma_L \right) d + \text{H.c.} \right] - \frac{H^0}{v} \left(\bar{u} D_u u + \bar{d} D_d d \right)$$

$$- \frac{R}{v} \left[\bar{u} \left(N_u \gamma_R + N_u^\dagger \gamma_L \right) u + \bar{d} \left(N_d \gamma_R + N_d^\dagger \gamma_L \right) d \right] + i \frac{I}{v} \left[\bar{u} \left(N_u \gamma_R - N_u^\dagger \gamma_L \right) u - \bar{d} \left(N_d \gamma_R - N_d^\dagger \gamma_L \right) d \right], \tag{50}$$

---

[8] Another proposal for the structure of the scalar couplings to fermions is the suggestion that the two Yukawa couplings are aligned in flavour space [38,246,247].

where $v \equiv \sqrt{v_1^2 + v_2^2} = (\sqrt{2} G_F)^{-1/2} \approx 246$ GeV (G_F is the Fermi constant), $\gamma_L = (1 - \gamma_5)/2$, $\gamma_R = (1 + \gamma_5)/2$, and

$$H^0 = \frac{1}{v} (v_1 \rho_1 + v_2 \rho_2), \tag{51}$$

$$R = \frac{1}{v} (v_2 \rho_1 - v_1 \rho_2), \tag{52}$$

$$I = \frac{1}{v} (v_2 \eta_1 - v_1 \eta_2). \tag{53}$$

The fields ρ_j, η_j ($j = 1, 2$) arise when one expands [9] the neutral scalar fields around their vevs: $\phi_j^0 = (e^{i\alpha_j}/\sqrt{2})(v_j + \rho_j + i\eta_j)$. The physical neutral-scalar fields are linear combinations of H^0, R, and I.

The flavour-changing neutral currents are controlled by the matrices N_d and N_u, which are given by:

$$N_d = \frac{1}{\sqrt{2}} U_{dL}^\dagger \left(v_2 Y_1^d - v_1 e^{i\alpha} Y_2^d \right) U_{dR}, \tag{54}$$

$$N_u = \frac{1}{\sqrt{2}} U_{uL}^\dagger \left(v_2 Y_1^u - v_1 e^{-i\alpha} Y_2^u \right) U_{uR}. \tag{55}$$

In general these matrices are not diagonal, hence there are FCNC. Let us consider, for example, N_d and rewrite it as [248]:

$$N_d = \frac{v_2}{v_1} D_d - \frac{v_2}{\sqrt{2}} \left(\frac{v_2}{v_1} + \frac{v_1}{v_2} \right) U_{dL}^\dagger e^{i\alpha} Y_2^d U_{dR}; \tag{56}$$

clearly, the first term in the right-hand side conserves flavour, but the second one leads, in general, to FCNC. The CKM matrix is given by $V = U_{uL}^\dagger U_{dL}$, therefore, if we want N_d to be entirely controlled by V, we need, on the one hand, to get rid of its dependence on U_{dR} and, on the other hand, to relate $U_{dL}^\dagger$ to V. A solution to these two requirements by means of symmetries was found in Ref. [237] and corresponds to the BGL models.

In Ref. [237] a flavour symmetry is imposed which constrains U_{uL} to be of the form

$$U_{uL} = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{57}$$

where $\times$ denotes an arbitrary entry. This leads, from the definition of V, to:

$$V_{3j} = (U_{dL})_{3j}. \tag{58}$$

In addition, the flavour symmetry imposes the following condition

$$Y_2^d U_{dR} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \times & \times & \times \end{pmatrix}. \tag{59}$$

This guarantees that only the third row of U_{dL} appears in N_d, as can be easily checked from Eq. (56). This row is exactly the one that coincides with the third row of V.

To get rid of U_{dR} in N_d, the flavour symmetry also enforces the relation

$$\frac{v_2 e^{i\alpha}}{\sqrt{2}} Y_2^d = P M_d, \tag{60}$$

where P is a projection matrix. In order to be consistent with Eq. (59), the matrix P must be given by:

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{61}$$

immediately leading to texture-zeros for Y_2^d such that Eq. (59) is verified. Furthermore, the appearance of M_d in Eq. (60) allows to absorb U_{dR}.

Branco, Grimus, and Lavoura have imposed the following symmetry S on the 2HDM:

$$Q_{L3}^0 \to e^{i\psi} Q_{L3}^0, \qquad u_{R3}^0 \to e^{2i\psi} u_{R3}^0, \qquad \Phi_2 \to e^{i\psi} \Phi_2, \tag{62}$$

where $\psi \neq 0, \pi$ and all other fields are invariant under S. The Yukawa couplings consistent with this symmetry have the structure

$$Y_1^d = \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2^d = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \times & \times & \times \end{pmatrix}, \qquad Y_1^u = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2^u = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \times \end{pmatrix}. \tag{63}$$

Notice that the matrices Y^u are block diagonal; this is crucial in order for U_{uL} to be as in Eq. (57). Furthermore, these four matrices satisfy

$$PY_1^d = 0, \qquad PY_2^d = Y_2^d, \qquad PY_1^u = 0, \qquad PY_2^u = Y_2^u. \tag{64}$$

The structure of zeros in the matrix Y_2^d leads to the important relation:

$$\left(U_{dL}^\dagger Y_2^d\right)_{ij} = \left(U_{dL}^\dagger\right)_{i3} \left(Y_2^d\right)_{3j} = V_{i3} \left(Y_2^d\right)_{3j}. \tag{65}$$

Together with Eq. (60) inserted in Eq. (56), one then obtains [237]

$$(N_d)_{ij} = \frac{v_2}{v_1} (D_d)_{ij} - \left(\frac{v_2}{v_1} + \frac{v_1}{v_2}\right) V_{i3}^\dagger V_{3j} (D_d)_{jj}, \tag{66}$$

whereas

$$N_u = -\frac{v_1}{v_2} \operatorname{diag}(0, 0, m_t) + \frac{v_2}{v_1} \operatorname{diag}(m_u, m_c, 0). \tag{67}$$

In this example there are scalar-mediated FCNC in the down sector but no FCNC in the up sector.

In general, BGL models are six different models, as was emphasized in [237]. Three of these models have FCNC only in the down sector, and are obtained from the three different projection matrices similar to the matrix P introduced above, but, in each case, with the diagonal unit entry in one of the other two alternative positions. Another three models are obtained by exchanging the patterns of zeros of the Y_i^d and Y_i^u matrices, *i.e.* by exchanging the up and down quarks. From Eq. (62) it is straightforward to write down the flavour symmetries corresponding to each of the six cases. The relations given in Eq. (64) result from the imposed symmetry. All BGL models obey relations of this type for the corresponding projection matrices. These relations guarantee that the scalar flavour-changing neutral couplings can be written in terms of quark masses and CKM-matrix entries [249]. The stability of these equations under renormalization is crucial; this feature was analysed in [250], with the help of the one-loop renormalization-group equations for the Yukawa couplings generalized from Ref. [39].

In each of the six models, one can study the phenomenological constraints by comparison with the type III model. For the specific model discussed in the last paragraph, for example, one can compare the FCNC coupling *bsH* in the two models directly:

$$\lambda_{bs} \frac{\sqrt{2m_b m_s}}{v} \leftrightarrow \frac{m_b}{v} \left(\frac{v_1}{v_2} + \frac{v_2}{v_1}\right) V_{ts}^* V_{tb} \tag{68}$$

which numerically gives $\lambda_{bs} = 0.14 \left(\frac{v_1}{v_2} + \frac{v_2}{v_1}\right)$, as compared to $\lambda_{bs} = 1$ in the type III model. Similar expressions can be found for the other FCNC couplings, as well as for the other five models. Thus, the bounds in the previous section can be used to find the phenomenological bounds in the BGL model as well.[9]

In the example given here the scalar-mediated FCNC are suppressed by the matrix elements of the third row of the CKM matrix; this is due to the fact that P is the projection matrix with unit (33) entry. In the alternative cases, corresponding to a different projection matrix P, the suppression is given by the row indicated by the non-zero entry in the corresponding matrix P. In these additional cases the suppression of the scalar-mediated FCNC is not as strong as in the example above. From the point of view of the authors of Ref. [238], these additional cases do not qualify as MFV models, since they do not comply with the ingredients imposed by their definition of MFV. In the next section we generalize BGL models and compare them with the MFV definition of Ref. [238].

3.3. MFV generalized

In Ref. [249] the question was addressed of how to find a general expansion for N_d^0 and N_u^0 which conforms with the requirement of having all the flavour-changing couplings of the neutral scalars related in an exact way to elements of the quark mixing matrix. Here, N_d^0 and N_u^0 denote N_d and N_u when still in a weak basis:

$$N_d^0 = U_{dL} N_d U_{dR}^\dagger = \frac{1}{\sqrt{2}} \left(v_2 Y_1^d - v_1 e^{i\alpha} Y_2^d\right), \tag{69}$$

$$N_u^0 = U_{uL} N_u U_{uR}^\dagger = \frac{1}{\sqrt{2}} \left(v_2 Y_1^u - v_1 e^{-i\alpha} Y_2^u\right). \tag{70}$$

---

[9] In Ref. [251] a specific analysis of BGL models, especially their consequences for neutral-meson–antimeson mixing, has been performed.

The strategy was imposing that N_d^0 and N_u^0 be only functions of M_d and M_u with no other flavour dependence. Furthermore, N_d^0 and N_u^0 should transform appropriately under weak-basis (WB) transformations.

Weak-basis transformations are defined by:

$$Q_L^0 \rightarrow W_L Q_L^0, \qquad d_R^0 \rightarrow W_R^d d_R^0, \qquad u_R^0 \rightarrow W_R^u u_R^0. \tag{71}$$

Under this transformation, the quark mass matrices M_d and M_u transform like

$$M_d \rightarrow W_L^\dagger M_d W_R^d. \qquad M_u \rightarrow W_L^\dagger M_u W_R^u, \tag{72}$$

and the matrices U_{dL}, U_{dR}, U_{uL}, and U_{uR} defined in Eq. (48) and (49) transform as:

$$U_{dL} \rightarrow W_L^\dagger U_{dL}, \qquad U_{uL} \rightarrow W_L^\dagger U_{uL}, \qquad U_{dR} \rightarrow W_R^{d\dagger} U_{dR}, \qquad U_{uR} \rightarrow W_R^{u\dagger} U_{uR}. \tag{73}$$

Under a WB transformation N_d^0 and N_u^0 transform in the same way as M_d and M_u, respectively. Furthermore, the Hermitian matrices $H_{d,u} \equiv M_{d,u} M_{d,u}^\dagger$ transform under a WB as:

$$H_d \rightarrow W_L^\dagger H_d W_L, \qquad H_u \rightarrow W_L^\dagger H_u W_L. \tag{74}$$

From Eq. (48) and (49) it follows that

$$U_{dL}^\dagger H_d U_{dL} = D_d^2, \qquad U_{uL}^\dagger H_u U_{uL} = D_u^2. \tag{75}$$

It is convenient, for our propose, to write H_d, H_u in terms of projection operators [252]:

$$H_d = \sum_i m_{di}^2 P_i^{dL}, \tag{76}$$

where

$$P_i^{dL} = U_{dL} P_i U_{dL}^\dagger, \qquad (P_i)_{jk} = \delta_{ij}\delta_{ik}. \tag{77}$$

Analogous expressions hold for H_u.

Based on the above considerations, the following expansion for N_d^0 and N_u^0, with the correct transformation properties under WB transformations, was proposed [249]:

$$N_d^0 = \lambda_1 M_d + \lambda_{2i} U_{dL} P_i U_{dL}^\dagger M_d + \lambda_{3i} U_{uL} P_i U_{uL}^\dagger M_d + \cdots, \tag{78}$$

$$N_u^0 = \tau_1 M_u + \tau_{2i} U_{uL} P_i U_{uL}^\dagger M_u + \tau_{3i} U_{dL} P_i U_{dL}^\dagger M_u + \cdots. \tag{79}$$

In the quark mass eigenstate basis these equations become

$$N_d = \lambda_1 D_d + \lambda_{2i} P_i D_d + \lambda_{3i} V^\dagger P_i V D_d + \cdots, \tag{80}$$

$$N_u = \tau_1 D_u + \tau_{2i} P_i D_u + \tau_{3i} V P_i V^\dagger D_u + \cdots. \tag{81}$$

This meets the requirement of having all the flavour-changing couplings related in an exact way to the quark mixing matrix, with no additional flavour dependence.

In Eq. (78) and (79) the lambda and tau coefficients are dimensionless. The building blocks of the expansion are given explicitly in Eq. (77) for the index d; they are of the same form for the index u. Notice that, in addition, there is the possibility of having different coefficients for difference values of the index i (ranging from 1 to 3), corresponding to different projectors P_i.

In Ref. [238] a different expansion is used. There, the building blocks are $Y_d Y_d^\dagger$ and $Y_u Y_u^\dagger$, where Y_d and Y_u are the Yukawa-coupling matrices. These matrices are also dimensionless. Such an expansion can be accommodated in Eqs. (78) and (79), however, but in that case the lambdas and taus would become dimensionless functions of the quark masses.

In theories where the MFV requirement results from the imposition of a symmetry there are constraints on the coefficients lambda and tau appearing in Eq. (80) and (81). This is the case for the BGL example presented in the previous subsection, which corresponds to the truncation

$$N_d^0 = \frac{v_2}{v_1} M_d - \left(\frac{v_2}{v_1} + \frac{v_1}{v_2}\right) U_{uL} P_3 U_{uL}^\dagger M_d, \tag{82}$$

$$N_u^0 = \frac{v_2}{v_1} M_u - \left(\frac{v_2}{v_1} + \frac{v_1}{v_2}\right) U_{uL} P_3 U_{uL}^\dagger M_u, \tag{83}$$

with lambdas and taus fully determined as functions of $\tan\beta \equiv v_2/v_1$.

More details about different aspects of this MFV generalization can be found in Ref. [249].

A study on the effectiveness of the two different hypothesis, Natural Flavour Conservation and Minimal Flavour Violation, in suppressing the strength of flavour-changing neutral currents in models with more than one Higgs doublet was performed in Ref. [253].

3.4. Two-Higgs leptonic minimal flavour violation

In order to study the phenomenological implications of models with an extended Higgs sector it is necessary to specify the lepton sector in addition to the quark sector. Furthermore, the analysis of stability under renormalization requires the entire set of renormalization-group equations both in the quark and lepton sectors. In Ref. [250] the extension to the lepton sector of models of BGL type was considered. In particular the minimal discrete symmetry required in order to implement the models in a natural way was given and stability was analysed. Different extensions of the MFV principle to the leptonic sector were considered previously in Refs. [254–256].

The case of Dirac-type neutrinos is straightforward. The Yukawa couplings for both the quark and the lepton sectors are given by:

$$\mathcal{L}_Y = -\overline{Q_L^0}\left(Y_1^d \Phi_1 + Y_2^d \Phi_2\right) d_R^0 - \overline{Q_L^0}\left(Y_1^u \tilde{\Phi}_1 + Y_2^u \tilde{\Phi}_2\right) u_R^0 - \overline{L_L^0}\left(Y_1^e \Phi_1 + Y_2^e \Phi_2\right) l_R^0 - \overline{L_L^0}\left(Y_1^\nu \tilde{\Phi}_1 + Y_2^\nu \tilde{\Phi}_2\right) \nu_R^0 + \text{H.c.,} \quad (84)$$

where Y_j^e and Y_j^ν denote the couplings of the left-handed leptonic doublets L_L^0 to the right-handed charged leptons l_R^0 and neutrinos ν_R^0. If there are no additional Majorana mass terms, the parallel between the quark and leptonic sectors allows to apply similar discussions to both sectors.

In the case of Majorana-type neutrinos there is no lepton-number conservation. In the seesaw framework with three right-handed neutrinos, an additional invariant mass term of Majorana type for right-handed neutrinos, *i.e.* $(1/2)\nu_R^{0T} C^{-1} M_R \nu_R^0$, must be included in the Lagrangian. As a result there will be three light neutrinos ν_i and three heavy neutrinos N_i. In order to obtain FCNC in the charged-lepton sector completely controlled by the lepton mixing (PMNS) matrix, together with no FCNC in the light-neutrino sector, a Z_4 symmetry is required. An example is [250]:

$$L_{L3}^0 \to e^{i\psi} L_{L3}^0, \qquad \nu_{R3}^0 \to e^{2i\psi} \nu_{R3}^0, \qquad \Phi_2 \to e^{i\psi} \Phi_2, \quad (85)$$

with $\psi = \pi/2$; all other fields are invariant under Z_4. The most general matrices Y_j^e, Y_j^ν, and M_R consistent with this Z_4 symmetry have the following structure:

$$Y_1^e = \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2^e = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \times & \times & \times \end{pmatrix}, \quad (86)$$

$$Y_1^\nu = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y_2^\nu = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \times \end{pmatrix}, \qquad M_R = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & \times \end{pmatrix}, \quad (87)$$

where $\times$ denotes an arbitrary entry while the zeros are enforced by the symmetry Z_4. Note that the choice of Z_4 is crucial in order to guarantee $(M_R)_{33} \neq 0$ and, thus, a non-vanishing det M_R.

The scalar couplings to the neutrinos are more involved than those to the charged leptons, since they include couplings to two light neutrinos, two heavy neutrinos, or one light and one heavy neutrino. They are explicitly given in Ref. [250].

As a result of the Z_4 symmetry, the scalar potential acquires an exact ungauged accidental continuous symmetry, which is not, though, a symmetry of the full Lagrangian. The simplest way of avoiding the ensuing pseudo-Goldstone boson after spontaneous symmetry breaking is through the introduction of a soft symmetry-breaking term of the form $m_{12}^2 \Phi_1^\dagger \Phi_2 + \text{H.c.}$

4. Charged Higgs bosons

One of the most important features of all 2HDMs is the existence of a charged scalar $H^\pm$. This state is orthogonal to the longitudinal component $G^\pm$ of the gauge boson $W^\pm$. The study of the properties of $H^\pm$ will be essential to understand which 2HDM – if any – has been chosen by nature. Most of the phenomenological studies of 2HDMs have indeed focused on the charged scalar. The $H^\pm$ can be readily pair-produced through Drell–Yan processes and – unlike the neutral scalars – can never decay "invisibly". In addition to direct production, it can have sizeable indirect effects in B physics and is a major topic of analysis in studies of rare B decays.

The first paper to include the words "charged Higgs"[10] in its title was the one by Tomozawa [258], in which the charged scalar in a vector-like model has been studied. In the context of a 2HDM, the phenomenology of charged scalars was first

---

[10] As emphasized by Georgi [257], the phrase "charged Higgs" is misleading. If the Higgs is defined to be the field that acquires a vev, then the Higgs cannot be charged since electromagnetism is unbroken. If the charged Higgs is defined to be the charged member of an $SU(2)$ doublet whose neutral component acquires a vev, then the charged Higgs is the longitudinal component $G^\pm$ of the $W^\pm$. The phrase "charged scalar" is thus preferable, yet the nomenclature "charged Higgs" is by now so common that we shall adopt it here.

Table 3
The parameters X, Y, and Z in Eq. (88) for the four models without FCNC.

	Type I	Type II	Lepton-specific	Flipped
X	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
Y	$\cot\beta$	$-\tan\beta$	$\cot\beta$	$-\tan\beta$
Z	$\cot\beta$	$-\tan\beta$	$-\tan\beta$	$\cot\beta$

discussed in detail by Donoghue and Li [259]. Following that there was an explosion of interest in the charged Higgs. This interest is sufficiently great that a series of three conferences studying the properties of charged scalars – CHarged 2006, CHarged 2008, and CHarged 2010 – have been held in Uppsala; the proceedings of those conferences are readily available [260–262] and provide an enormous source of information on both the theory and phenomenology of charged Higgs bosons.

For all four models without tree-level FCNC discussed in Section 3, the most general Yukawa couplings were written in [30] as

$$\mathcal{L}_{H^\pm} = -H^+ \left(\frac{\sqrt{2}\, V_{ud}}{v} \bar{u} \left(m_u X P_L + m_d Y P_R \right) d + \frac{\sqrt{2}\, m_\ell}{v} Z \bar{\nu}_L \ell_R \right) + \text{H.c.} \tag{88}$$

where V_{ud} is the element of the CKM matrix corresponding to the charge 2/3 quark u and the charge $-1/3$ quark d. The values of X, Y, and Z depend on the particular model and are given in Table 3. In the type I model the couplings to all fermions are suppressed if $\tan\beta \gg 1$, meaning a fermiophobic charged Higgs. In the same limit $\tan\beta \gg 1$ one has in the lepton-specific model a quark-phobic but leptophilic model, which could lead to a huge branching ratio for $H^\pm \to \tau^\pm \nu$. In both cases, the quark-phobic nature of the model eliminates constraints from rare B decays. The type II model is the most studied one; large contributions to rare B decays are possible in it. The flipped model has only recently been studied. In the next section we shall review the properties of the charged Higgs in each of these models.

In models with tree-level FCNC, discussed in Section 3, it is much more convenient to use the Higgs basis, in which one doublet has a vev and the other doublet is vev-less (see the discussion in Section 5). In the models without tree-level FCNC, this basis is a disadvantage, since the Yukawa couplings and discrete symmetry adequately specify the basis; with FCNC this disadvantage vanishes. This was pointed out by Atwood et al. [216]. As shown there, in the Higgs basis, the H_1 field (which acquires vev) has diagonal couplings to fermions, which are identical to those in the Standard Model, whereas the couplings of H_2 to quarks are given by

$$\mathcal{L} = \bar{Q}_{Li} \left(\hat{\xi}^U_{ij} \tilde{H}_2 u_{Rj} + \hat{\xi}^D_{ij} H_2 d_{Rj} \right) + \text{H.c.} \tag{89}$$

where the quark fields are mass eigenstates and the matrices $\hat{\xi}^{U,D}$ are in general not diagonal. Defining the rotation matrices $V_{L,R}^{U,D}$, the neutral flavour-changing couplings are related to the original ones by V_L and V_R. Since the definition of the $\xi_{ij}^{U,D}$ is arbitrary, one can, without loss of generality, replace the rotated couplings by the original ones. For the charged flavour changing couplings, we then have

$$\xi^U_{\text{charged}} = \xi^U V_{CKM}, \qquad \xi^D_{\text{charged}} = V_{CKM} \xi^D. \tag{90}$$

The Cheng–Sher Ansatz discussed in Section 3 gives the matrices $\xi_{ij}^{U,D}$ as a constant of O(1) times the geometric mean of the respective Yukawa couplings. Similar results are found for the leptonic sector, with V_{CKM} replaced by the V_{PMNS} matrix. In the third section of this Chapter we shall review the phenomenology of the charged Higgs in these models.

4.1. Models without tree-level FCNC

The differences among the various 2HDMs concern the Yukawa couplings to fermions. At the recent CHarged 2010 meeting there have been many discussions of benchmarks for these various 2HDMs [262]. They were summarized, in the context of LHC searches, by Guedes et al. [263]. More details, together with numerous plots, are in the talk by Santos at that meeting [262]. A large group has studied specific benchmarks for each of the various models; that work is summarized in the reports by Krawczyk et al. and by Osland et al. at the meeting. Some of the models, other than the type II model, have been discussed in the talk by Akeroyd [262]. A very comprehensive analysis of charged Higgs phenomenology for all four models can be found in the recent article of Aoki et al. [36]. Jung et al. [246] have considered many processes within their general "aligned two Higgs doublet model"—in which it is assumed that the Yukawa-coupling matrices are proportional, hence FCNC are absent at tree level. Since that model has many parameters, we shall not go through their analysis explicitly. However, a valuable feature of that analysis is that, from its results, one can explore the limits from indirect processes in all four models, which are all limiting cases of the aligned two Higgs doublet model. A discussion of the decay of the top quark into $H^+ b$ in a very general 2HDM can be found in Ref. [264].

We shall summarize in the following the production and decay of the charged Higgs in each of the four models, and discuss various constraints. In the type II model and in the flipped model there is a lower bound of about 300 GeV for the mass of $H^\pm$ stemming from $b \to s\gamma$; this bound comes from charged Higgs bosons in the loop. However, this bound is only
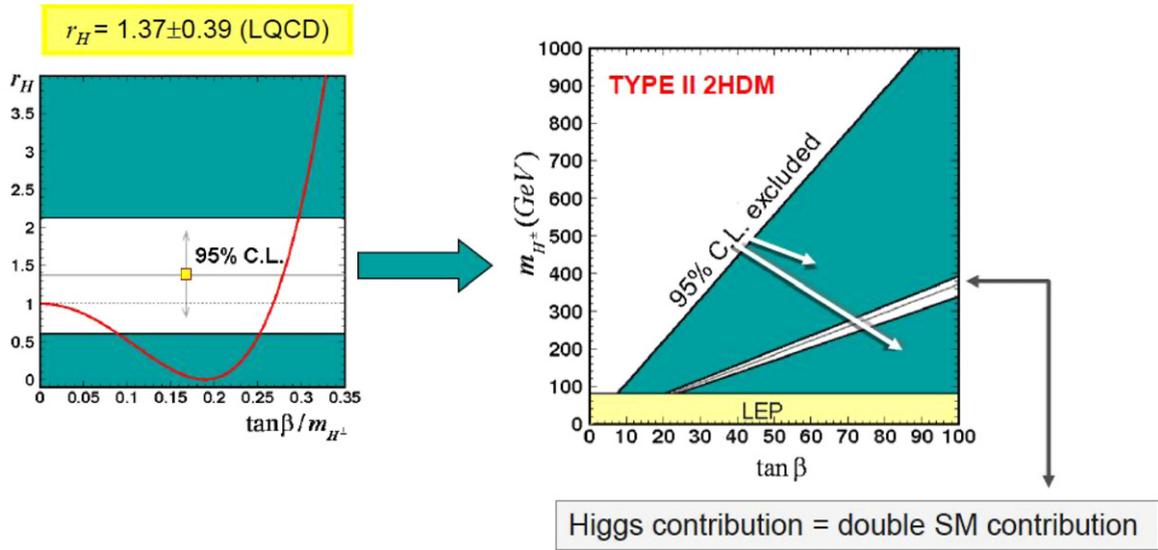
Fig. 14. On the left one has the prediction for r_H – the ratio between the value of BR $(B^+ \rightarrow \tau^+\nu_\tau)$ in the type II 2HDM to that in the Standard Model – as a function of $\tan\beta/m_{H^\pm}$. Given the experimental value, one has on the right the ensuing constraints in the $m_{H^\pm}$–$\tan\beta$ plane; the narrow window corresponds to the allowed region near $\tan\beta/m_{H^\pm} = 0.2$ and the yellow band corresponds to the region of $m_{H^\pm}$ already excluded by direct searches at LEP. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Figure from the talk by Rozanska at the CHarged 2010 workshop [262].

valid if there is no additional New Physics. If there is New Physics, even a relatively mild cancellation can weaken the bound substantially. Therefore, we shall also examine other, somewhat weaker bounds.

4.1.1. The type II model

Since supersymmetric and Peccei–Quinn models are all of type II, the type II model is the most studied one. The Yukawa couplings are given in Table 3; one can see that the coupling of the charged Higgs to the top and bottom quarks is governed by either the bottom-quark mass times $\tan\beta$, which may be large, or by the top-quark mass times $\cot\beta$. As a result, one expects potentially large virtual effects in b-quark decays and mixing. In fact, one of the major motivations for the B factories was the possibility of New Physics coupling strongly to the third generation; the type II 2HDM is the simplest example of this.

A very strong bound on the mass of the charged Higgs comes from studies of $\bar{B} \rightarrow X_s\gamma$. The charged Higgs appears in the loop. A nice review with a comprehensive list of references is in the article of El Kaffas et al. [88]. Early analyses can be found in Refs. [265–268]. An explosion of interest occurred after the realization that the top quark was heavy and after QCD corrections were considered [269–274]. These models were shown to be quite scale dependent [275,276] and this has led to a calculation of the next-to-leading (NLO) order, where the scale dependence is substantially reduced [277–280]. Discussions can be found in the reviews of Haisch [281] and of Hurth and Nakao [282], and more recent results on all B decays in the recent talks of Rozanska [262] and Hurth [283]. The effects of the charged Higgs in the loop add constructively [79,80, 284–286]. The most detailed calculation is the $O(\alpha_s^2)$ calculation by Misiak [81] and gives $m_{H^\pm} > 295$ GeV for virtually all values of $\tan\beta$ (the bound increases slightly for low values of $\tan\beta$, see [281]). A more recent measurement of the rate by BELLE [287] does not change this bound.

Other bounds are even more severe, but only at large $\tan\beta$. The process $B \rightarrow \tau\nu$ has been studied. This is not a loop process and proceeds instead through tree-level virtual $H^\pm$ exchange. The rate is given by [288,289]

$$\frac{\text{BR}(B^+ \rightarrow \ell^+\nu_\ell)}{\text{BR}(B^+ \rightarrow \ell^+\nu_\ell)_{\text{SM}}} = \left(1 - \frac{m_B^2 \tan^2\beta}{m_{H^\pm}^2}\right)^2. \tag{91}$$

Combining the BELLE [290,291] and BABAR [292] results, the Heavy flavour Averaging Group found [293] the branching fraction to be $(1.64 \pm 0.34) \times 10^{-4}$. The ratio of this measured value to the Standard Model prediction [294] is 1.37 ± 0.39. This gives a lower bound which, at 95% confidence level, excludes a region which rises from 300 GeV for $\tan\beta = 40$ to 1100 GeV for $\tan\beta = 100$. However, there is a small window in this region which is still allowed. The reason for that window is shown in Fig. 14, where one can see a narrow region of $\tan\beta$ which is still allowed.

That window can be closed by considering the process $B \rightarrow D^{(*)}\tau\nu_\tau$. In that process the CKM angles are better known; also, many of the experimental and theoretical uncertainties cancel out in the ratio [295,296] in which the τ is replaced by a lighter lepton—in which case the contribution from $H^\pm$ exchange is negligible. It is found (see Ref. [295] for details) that the window is completely closed by this process—which does have its own window, but that is closed by $B^+ \rightarrow \tau^+\nu_\tau$.
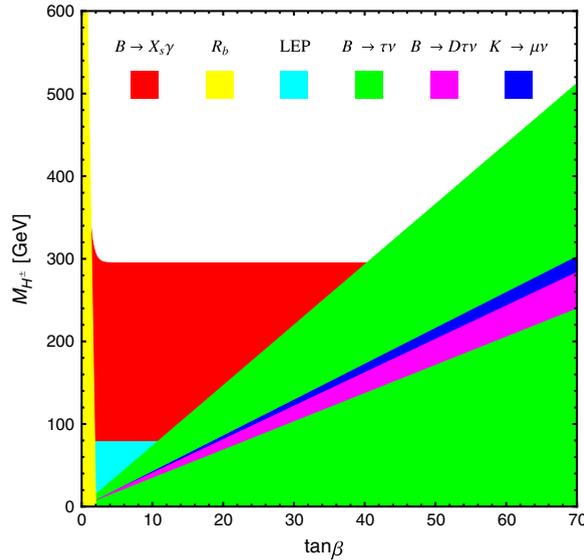
Fig. 15. Bounds in the $m_{H^\pm}$–tan β plane from various B-physics constraints.
Source: This figure was extracted from the article by Haisch [281].

For small values of tan β, the value of R_b, which is the ratio of $\Gamma\left(Z \to b\bar{b}\right)$ to the total hadronic width of the Z, can be affected at one loop through the exchange of $H^\pm$. Haber and Logan found [297] that the ensuing constraints will be more severe than those from $b \to s\gamma$ for values of tan $\beta < 1.4$; radiative corrections to their results are in Ref. [298]. Very similar bounds were obtained [299] by considering Δm_B and Δm_{B_s}. Other bounds, which tend to be weaker, can be found from rare K, D, and τ decays; a comprehensive analysis is found in the recent paper of Deschamps et al. [300]. Such a comprehensive analysis can be very valuable since it is possible that New Physics might weaken some, but not all, of the bounds.

The results are shown in Fig. 15. There one clearly sees the various contributions. Note that the direct bound from LEP is substantially weaker than the indirect bounds.

It is still important to look at bounds from direct production of $H^\pm$, because it is always possible that New Physics weakens the bounds from $b \to s\gamma$. In fact, that is precisely what occurs in supersymmetric models. It can be seen that a weakening of this bound would allow, for smaller values of tan β, charged-Higgs masses somewhat below 100 GeV.

To study direct production bounds, and to explore prospects for the LHC, the production rates and branching ratios of the charged Higgs are needed. The branching ratios have been studied quite thoroughly—Ref. [47] contains an extensive review of the literature. A recent analysis [53] used the FORTRAN code HDECAY [301]. The program includes final-state mass effects, full one-loop QCD corrections and running masses, and off-shell decays to $t\bar{b}$ below threshold. In principle, decays to $S^0 W^\pm$ are also possible, where S^0 is a neutral scalar. However, the masses of these scalars are unknown and, moreover, the branching ratios in most of the parameter space are much smaller than 1%; thus, they are neglected in much of this discussion. We shall, however, discuss these possibilities at the end, since there could be a substantial rate in the type II model – but not in the MSSM – for part of the parameter space, as emphasized by Kanemura et al. [302]. Note that there is no $Z^0 W^\pm H^\mp$ vertex in the 2HDM.

The results are shown in Fig. 16. We see that, above the kinematic limit, the decays into $t\bar{b}$ are completely dominant, but, for lower masses $m_{H^\pm}$, the decays into $\tau^+ \nu_\tau$ dominate. The decay into $c\bar{s}$ can also be significant for small values of tan β. Now, all the current direct searches are sensitive only to charged-Higgs masses below the kinematic threshold for decays into $t\bar{b}$. Therefore, the main decay modes in the present direct searches are into $\tau^+ \nu_\tau$ and $c\bar{s}$. The LEP combined limit on $m_{H^\pm}$ is 78.6 GeV and has been calculated [303] assuming the existence of only those two decay modes (in the flipped model the decay into $c\bar{b}$ is more significant than the decay into $\tau^+ \nu_\tau$ [53]).

Searches have also been performed at the Tevatron. The CDF Collaboration [304] has assumed that BR$\left(H^+ \to c\bar{s}\right) = 1$ and has therefrom found bounds on the branching ratio of $t \to H^+ b$. Logan and MacLennan [53] noted that the CDF analysis also applies to decays into $c\bar{b}$ and therefore also applies to the flipped model. However, in the type II model the CDF assumption is valid only for small (less than 1) values of tan β and only for charged-Higgs masses near the LEP bound. The D0 Collaboration [305], on the other hand, has considered both scenarios in which BR$\left(H^+ \to c\bar{s}\right) = 1$ and BR$\left(H^+ \to \tau^+ \nu_\tau\right) = 1$; the latter scenario is certainly valid at large tan β in the type II model. D0 [306] and CDF [307] have also looked for associated production of a charged Higgs with a W.

These papers give upper limits on the branching ratio of the top quark into a charged Higgs and a bottom quark. This can be converted into a bound in the m_{H^+}–tan β plane. We use the well-known expressions for the branching ratio (see the appendix of Ref. [53] for a simple expression) and find the results in Fig. 17. As expected, the present bounds only exclude regions of parameter space for either large or small tan β; a charged-Higgs mass of about 100 GeV is thus still allowed. Recall
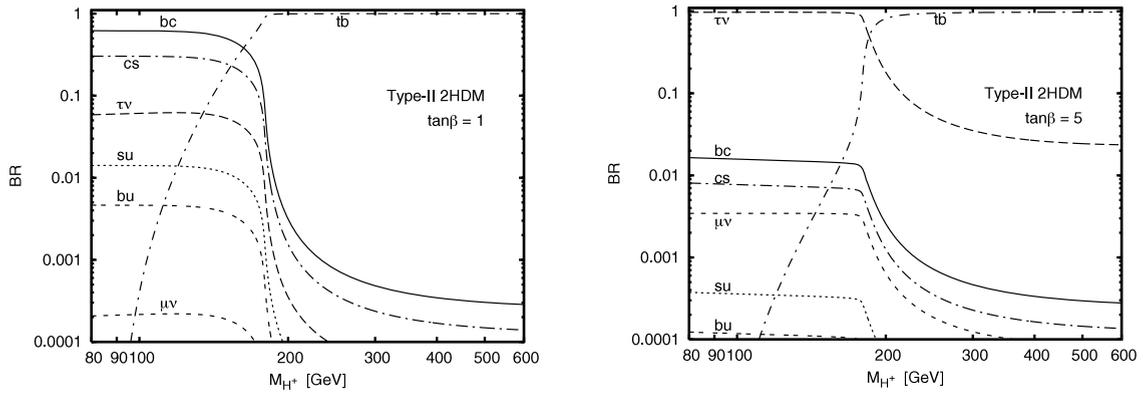
Fig. 16. The branching ratios of the charged Higgs in the type-II 2HDM. Decays into $S^0 W^\pm$ (S^0 is a neutral Higgs boson) are not included.
Source: This figure is from Ref. [53].
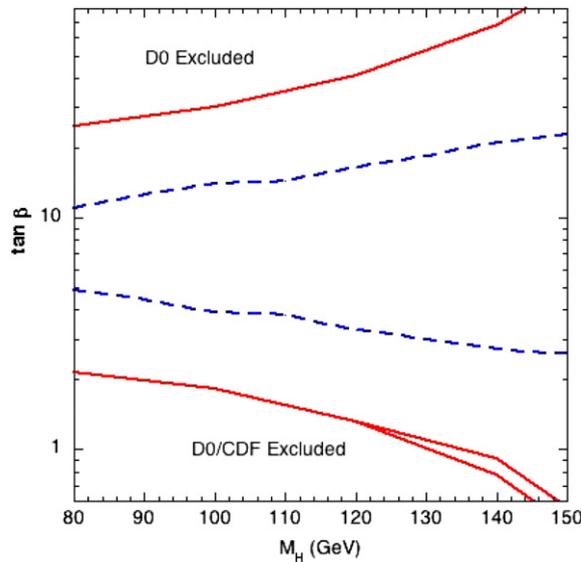
Fig. 17. Current Tevatron bounds on the mass M_H of the charged Higgs as a function of $\tan\beta$. The small sliver near the lower-right corner is excluded by CDF but not by D0. The dashed lines are the bounds which are projected for the ATLAS experiment with 30 fb^{-1} of data; the regions that would then be excluded are those outside the dashed lines.

that one is assuming that some New Physics cancels the contribution to $b \rightarrow s\gamma$, and it is possible that this New Physics could also lead to alternative production and decay mechanisms.

In the future much of this parameter space will be probed at LHC. The ATLAS [308] and CMS [309] Collaborations have studied top-quark pair production in which one of the tops decays into $H^+ b$. They have assumed that the charged Higgs H^+ subsequently decays into $\tau^+ \nu$, as is the case in the type II model except for very small $\tan\beta$ of O(1). Techniques to improve these studies have been discussed in Ref. [310]. Also shown in Fig. 17 is the expected reach of LHC, as found in the ATLAS analysis for a collision energy of 14 TeV and with 30 fb^{-1} of data; ATLAS gave the expected reach in the $t \rightarrow H^+ b$ branching ratio and we have converted this into a bound in the $m_{H^\pm}$–$\tan\beta$ plane. The CMS analysis was explicitly model dependent by using the MSSM, and in that particular case it could partially close the allowed window at intermediate $\tan\beta$. Again, for intermediate values of $\tan\beta$ there is in general a large window that can be closed neither at the Tevatron nor at the LHC. The ILC, on the other hand, will readily be able to produce charged-Higgs pairs and thus to cover the entire parameter space for Higgs masses lower the mass of the top. There has also been a discussion of the possibility of detecting the $H^\pm \rightarrow \mu^+ \nu_\mu$ after 300 fb^{-1} in Ref. [311].

For larger Higgs masses the decay of H^+ is overwhelmingly into $t\bar{b}$. This will be subject to very large backgrounds[11] and is thus quite challenging. For moderately large $\tan\beta$ the branching ratio of $H^+ \rightarrow \tau^+ \nu_\tau$ is approximately 10% and this might be sufficient to pick out the signal. In fact, most search strategies must study the latter decay mode.

---

[11] If the charged Higgs is produced in associated with a top quark, it has been argued [312] that one might be able to detect the decay into $\bar{t}b$.

The primary production mechanisms at the LHC are single charged-Higgs production and pair production. The single charged-Higgs production processes include gluon fusion [313], quark fusion [314–317], quark–gluon fusion [83,85], associated $W^\pm H^\mp$ production [318–324] and $AH^\pm$, $HH^\pm$, $hH^\pm$ production [325–329]. Most of these works have focused on the MSSM, although enhancement in the general type II model is referred to in Ref. [86]. Pair production has also been calculated: the Drell–Yan and $b\bar{b} \rightarrow H^+H^-$ processes most recently in [330], and weak-boson fusion in [331]. Another production mechanism is associated production of a charged Higgs with a top quark, which can lead to interesting consequences for the top quark polarization and the angular correlations of its decay products [332], and also in the left–right asymmetry in the polarized top quark production cross section [333]. Although strictly a type II model, associated $W^\pm H^\mp$ production in a model with spontaneous CP violation (i.e. a relative phase between the vevs) was studied in Ref. [334]. They found no observable effects from associated production, but did note that the model allows the charged Higgs mass to be substantially smaller than the conventional type II model.

The ATLAS [308] and CMS [335] Collaborations have used the production cross sections to explore the reach of the LHC. Their conclusions are that detection of the charged Higgs boson through its decay into $t\bar{b}$ will be swamped by large backgrounds, hence the decay mode $\tau^+\nu_\tau$ is the most promising one. Both analyses use various versions of the MSSM but one would not expect the results in the type-II 2HDM to differ substantially. In each case, the reach on the branching ratio is give, and we have converted this into a reach in the $m_{H^\pm}$–$\tan\beta$ plane. The ATLAS Collaboration finds a discovery (*i.e.*, 5σ) reach which ranges (for an integrated luminosity of 30 fb^{-1}) from $\tan\beta = 28$ for $m_{H^\pm} = 200$ GeV to $\tan\beta = 65$ for $m_{H^\pm} = 350$ GeV (these values of $\tan\beta$ are lower bounds). The CMS Collaboration, in a slightly different MSSM scenario, finds a discovery reach, for the same integrated luminosity, which varies from $\tan\beta = 28$ for $m_{H^\pm} = 200$ GeV to $\tan\beta = 65$ for $m_{H^\pm} = 450$ GeV. The exclusion bounds correspond to much lower values of $\tan\beta$, of course—ATLAS, for example, can exclude $\tan\beta > 12$ (50) for charged-Higgs masses of 200 (600) GeV. The main production mode is $gb \rightarrow tH^-$. For associated $W^\pm H^\mp$ production, a signal can be found [321] for large $\tan\beta$ in the 150–300 GeV mass region.

Since most of the above analyses were done in the context of the MSSM, it is important to focus on the differences in the type II 2HDM. As noted in Ref. [321], if the mass of a neutral scalar is larger than $m_{H^\pm} + m_W$, then one can resonantly produce this neutral scalar in the *s*-channel, leading to a huge enhancement in the cross section for associated production. It is difficult to make a plot of the expected reach, though, since the rate is very sensitive to the neutral-scalar couplings and mass, but one should keep in mind the possibility that the rate for associated $H^\pm W^\mp$ production might be considerably larger than expected.

A detailed analysis of the way in which the phenomenology of the type II 2HDM may differ from the one of the MSSM was recently carried out by Kanemura et al. [302]. Their main point is that the Higgs masses are very tightly constrained in the MSSM, therefore certain decays that would otherwise be allowed cannot occur there. For example, for much of parameter space in the MSSM the charged Higgs, the heavy neutral scalar H, and the pseudoscalar A are very close in mass, therefore decays such as $H^\pm \rightarrow W^\pm A$ or $H^\pm \rightarrow W^\pm H$ are kinematically forbidden, and $H^\pm \rightarrow W^\pm h$ is suppressed by phase–space. In the general type II 2HDM, on the contrary, those decays are allowed.

To be specific, Kanemura et al. [302] have considered $\tan\beta = 1$–3 and $m_{H^\pm} > 250$ GeV—as noted above, the LHC will be insensitive to such masses in all the MSSM scenarios. They also consider the case $m_A = m_{H^\pm}$, $m_H = 150$ GeV, and $m_h = 50$ GeV as an illustrative example. For $\sin(\beta - \alpha) = 0.1$, they find that, for $\tan\beta = 3$ the dominant decay mode (over 90% branching ratio) of the charged Higgs is into $W^\pm h$ and for $\tan\beta = 1$ the branching ratio of that mode rises from 40% to 80% as the charged-Higgs mass goes from 250 to 600 GeV. They obtained similar results for $\sin(\beta - \alpha) = 0.9$ (now with $m_h = 120$ GeV in order to avoid LEP bounds), except that the dominant decay mode then is $W^\pm H$. Note that, for the charged-Higgs mass region that they consider the dominant branching ratio is into $t\bar{b}$ in the MSSM.

This has a huge effect on the phenomenology. By the time the LHC has been running for a few years, the neutral Higgs bosons will presumably have been discovered and, once their masses are known, the decay $H^\pm \rightarrow W^\pm(h, H)$ should be quite straightforward to detect. Thus, for a substantial part of the parameter space, the charged Higgs will be easier to detect in the type II 2HDM than in the MSSM.

Borzumati and Djouadi [336] have studied the observation of the decays $H^\pm \rightarrow W^\pm(h, H)$ at LEP and the Tevatron. It should be noted that their rate can be subject to large radiative corrections [72,337]. Searches for the charged Higgs by using the decay mode $H^\pm \rightarrow b\bar{b}W^\pm$ have been performed at the Tevatron [338] and will be studied at the LHC [339,340].

The production and decay of the charged Higgs in the "Complex two Higgs doublet model" was studied in Ref. [334]. In this model, one violates CP invariance by making the m_{12} term in Eq. (2) complex. They calculate CP violating asymmetries for H^+ and H^- production and decays, including one loop effects. They modified the codes for FeynArts and FormCalc to incorporate this model. It is found that CP violating asymmetries are substantially smaller than in the MSSM, and do not exceed 3%.

4.1.2. The type I model

Since both supersymmetric models and Peccei–Quinn models require a type II 2HDM, there has been substantially less discussion of the type I model. Nonetheless, many of the early works on 2HDMs discuss that model and it is often considered in more general analyses. The seminal paper of Barger et al. [30] analysed numerous constraints on the charged Higgs in the type I 2HDM, but this was before LEP and a relatively light top quark was assumed. There have been numerous studies since then, as will be seen in this section.
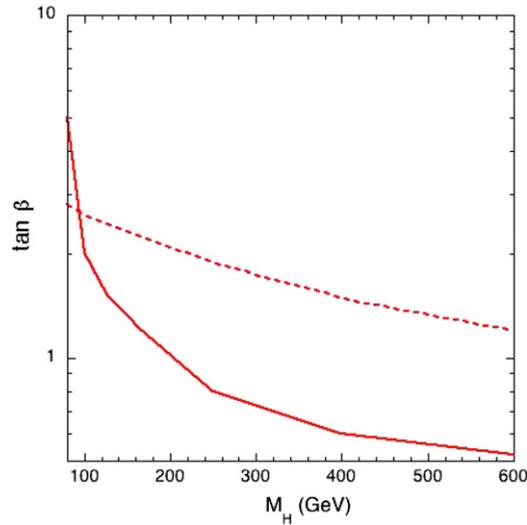
Fig. 18. Lower bounds on $\tan \beta$ in the type I 2HDM as a function of the charged-Higgs mass M_H. The solid line is the bound from Ref. [246] and comes from consideration of $Z \to b\bar{b}$, ϵ_K, and Δm_{B_s}. The dashed line is the bound in Refs. [46] and [107], which arises from $B \to X_s \gamma$ and is very sensitive to assumptions and to input parameters.

There are several key features of the type I model [341]. The relative branching ratios of the charged Higgs decaying into fermions are independent of $\tan \beta$. Assuming that there are no decays into lighter scalars – such as $H^+ \to AW^+$, hW^+, HW^+ – the decay into $\tau^+ \nu_\tau$ will have a branching ratio of 70% and the decay into $c\bar{s}$ will have branching ratio equal to 30%, when $m_{H^\pm}$ is below the top threshold. Above that threshold, H^+ decays almost always into $t\bar{b}$. Since the couplings to fermions are proportional to $\cot \beta$, the charged Higgs becomes fermiophobic in the large $\tan \beta$ limit. This might avoid constraints from flavour physics, allowing for the possibility of a charged Higgs in the 100 GeV mass range.

If kinematically possible, one may consider the decays $H^+ \to SW^+$, where S may be either h, H, or A. In this case, the W^+ may even be virtual, since in the large $\tan \beta$ (fermiophobic) limit a three-body decay can still dominate decays into two fermions [71]. In fact, it was pointed out in that paper that a three-body decay can dominate even if $\tan \beta$ is not so large, since the Yukawa couplings of the H^+ are small if its mass is below the one of the top quark.

As in the type-II 2HDM, we shall begin by considering the constraints from indirect processes. Since all of these involve Yukawa couplings, which vanish in the large $\tan \beta$ limit, they will provide bounds only in the low-$\tan \beta$ region. As noted earlier, a recent comprehensive review by Jung et al. [246] has considered many processes within their general "aligned 2HDM". One may use their results to explore the limits from indirect processes in all four 2HDMs, which are all limiting cases of the general approach of Pich et al.; they give bounds in each of the four 2HDMs that we discuss explicitly (the bounds for the type II model agree with those in the last subsection).

In the type I 2HDM the strongest bounds come from the process $Z \to b\bar{b}$, from ϵ_K, and from Δm_{B_s}. The value of R_b, the ratio of the width $Z \to b\bar{b}$ to the total hadronic width of the Z, has been calculated in 2HDMs [79,297,342], most recently in Ref. [298]. Comparing it with the experimental result one obtains bounds with vary from $\tan \beta > 2$ to $\tan \beta > 1$ when the charged-Higgs mass varies from 80 to 400 GeV. The value of the mass difference between the B_s and the $\bar{B}_s$ was calculated in [343] (some minor but not insignificant errors in the calculation were pointed out in [88]), and this also provides bounds. Recent calculations by Buras et al. [344,345] of ϵ_K also give constraints.

Putting all of these together leads to the excluded region in Fig. 18. There, the solid line corresponds to the bounds of Jung et al. [246] for the processes in the previous paragraph. As they note, other bounds from $B \to D\tau\nu$ followed by $\tau \to \ell\nu$ are weaker. In Ref. [246] the bound from $\bar{B} \to X_s \gamma$ is also weaker, but Ref. [46] – which is primarily concerned with models with tree-level FCNC and will be discussed in the next section – obtains a stronger bound from that process, shown as the dotted line in Fig. 18.[12] The precise value for radiative b decays is very sensitive to input parameters and to theoretical errors, so this is not necessarily a disagreement between [46] and [246] (one should also note that $\tan \beta$ is not directly measurable). In any event, one can see that there are no substantive bounds on $m_{H^\pm}$ for moderate and large $\tan \beta$, which is in sharp contrast to what happens in the type II 2HDM.

We next proceed to the direct bounds. For the type II 2HDM it was necessary to discuss the branching ratios as a function of $\tan \beta$. That is not needed for the type I 2HDM, since the relative fermionic branching ratios are in this case independent of $\tan \beta$ and are, approximately, 65% into $\tau^+ \nu_\tau$ and 35% into $c\bar{s}$ (below the top threshold, which is the relevant case for the current direct bounds). The combined LEP bound [303] is 78.7 GeV. However, this assumes that there are no non-fermionic decays such as $H^\pm \to AW^\pm$ (where the $W^\pm$ is virtual). The DELPHI Collaboration [346] also searched for a charged Higgs

---

[12] A separate analysis by Gupta and Wells [107] found results similar to those of Ref. [46].

Table 4
Branching ratio of the charged Higgs into $\tau \nu$ for various values of the charged Higgs mass (in GeV) and of tan β. For $m_{H^\pm} = 100$ GeV and tan $\beta = 1$ the remaining 30% branching ratio is into $c\bar{s}$; for all the other entries the remaining branching ratio is almost entirely into $t\bar{b}$.

Charged Higgs mass	tan $\beta = 1$	tan $\beta = 5$	tan $\beta = 10$	tan $\beta = 20$
100	0.70	0.95	0.99	1.00
200	0.05	0.20	0.80	0.97
300	0.00	0.05	0.40	0.92

that decays into either $\tau^+ \nu_\tau$, $c\bar{s}$, or AW^*; it found a bound of 76.7 GeV by assuming that the A was heavier than 12 GeV, similar results were also reported by the OPAL Collaboration [347].

As for the type II model, the Tevatron also has bounds from Drell–Yan pair production of a charged Higgs. The CDF bounds in Ref. [304] have assumed that BR $\left(H^+ \rightarrow c\bar{s}\right) = 1$, which is never the case in the type I 2HDM, so those bounds do not apply. The D0 bound [305] has considered the case in which the sum of BR $\left(H^+ \rightarrow c\bar{s}\right)$ and BR $\left(H^+ \rightarrow \tau^+ \nu_\tau\right)$ is 1, which applies here unless there are light scalars. The result is the lower line in Fig. 17. This bound, however, is slightly weaker than the one derived by Jung, Pich, and Tuzón from indirect processes. Thus, no stronger bounds, as of yet, have been obtained by the Tevatron on the type I 2HDM.

The ATLAS study of a light charged Higgs possibly being produced in top-quark decays [308] was explicitly focused on the type II 2HDM. However, the primary difference between the type I and type II models is in the coupling to $t\bar{b}$. In the type II model that coupling is large at large tan β, whereas it is always small for large tan β in the type I model; both models give similar results for small tan β. As a result, the discovery bound at the LHC would correspond to the lower dashed line in Fig. 17. For large tan β, the top-quark branching ratio into a charged Higgs is too small to detect in the type I 2HDM.

A study by Aoki et al. [348] of the light charged Higgs at the LHC explicitly focused on the type I and Lepton-Specific models has very recently appeared. They analyse production rates from top pair production (in which one top decays into $H^+ b$), single top production, and direct production through $cs \rightarrow H^+ +$ jet. They plot the reach of the LHC for 10 and 30 inverse femtobarns at $\sqrt{s} = 14$ TeV and find that over the range of charged Higgs masses up to 150 GeV, upper bounds on tan β between 6 and 10 can be obtained (recall that the very large tan β limit is fermiophobic in the type I model). Charged Higgs pair production is also considered, and it is shown that constraints are more parameter-dependent, but the process might also be detected at the LHC. An analysis at 7 TeV is currently under investigation by the same authors.

For larger charged-Higgs masses the decay into $t\bar{b}$ will be swamped by background. The $\tau^+ \nu_\tau$ decay, which has a branching ratio as high as 10% in the type II 2HDM, is much less important in the type I 2HDM and typically has a branching ratio of order $m_\tau^2 / m_t^2 \sim 10^{-4}$. This makes detection at the LHC impossible if the primary decay is fermionic.

However, there are also neutral scalars S (which may be h, H, or A in the 2HDM) and one may consider $H^\pm \rightarrow SW^\pm$. In the large tan β limit the charged Higgs is fermiophobic and $H^\pm \rightarrow SW^\pm$ would be the leading decay mode if kinematically accessible (and it might still be the leading decay mode even if the $W^\pm$ is virtual [336]). We thus see that in the large tan β limit the most promising decay modes are – unlike what occurs in the type II model – into a neutral scalar and a W. As noted in the last subsection, searches of a charged Higgs through the decay mode $H^\pm \rightarrow b\bar{b}W^\pm$ have been studied at the Tevatron [338] and LHC [339,340]. A plot of the results is premature since there are several additional parameters (such as α and the scalar mass) but the decay $H^\pm \rightarrow b\bar{b}W^\pm$ offers the best hope. An updated analysis of this mode would be welcome.

4.1.3. The lepton-specific model

In the lepton-specific (LS) 2HDM the same Higgs doublet couples to both the up-type and the down-type (right-handed) quarks, just as in the type I 2HDM, therefore the charged Higgs is quark-phobic for large tan β. Indirect bounds arising from hadronic decays will thus be identical to those in the type I model. But, in contrast to what happens in the type I 2HDM, at large tan β the charged Higgs becomes strongly leptophilic in the LS 2HDM, hence the $\tau^+ \nu_\tau$ decay mode of the H^+ can be dominant. In fact, as we shall see, that decay can be dominant *even above the $t\bar{b}$ threshold*, leading to quite dramatic experimental signatures.

Some early discussions of the charged Higgs in the LS 2HDM can be found in Refs. [30,32,349]. It has also been discussed recently in the context of dark-matter models [56] and of neutrino mass models [60]. The most recent comprehensive analyses of the phenomenology of the charged Higgs in the LS 2HDM are the articles by Su and Thomas [54], by Aoki et al. [36], and by Logan and MacLennan [55]; we shall follow those analyses closely. The branching ratio of the charged Higgs into $\tau^+ \nu_\tau$ is given in Table 4 (branching ratios below 5% are not shown). Decays into $W^\pm(h, H, A)$ are not included in the branching ratio computations; by using MSSM values for the mixing angles and for the light-scalar masses, Logan and MacLennan have shown that the branching ratios never exceed 10%, rendering irrelevant for the LS 2HDM some detection strategies involving decays into scalars which have been discussed for the type I and type II 2HDMs.

We see that, for tan $\beta = 1$, the branching ratios are identical to those in the type I model. But, the $\tau^+ \nu_\tau$ decay rapidly becomes dominant as tan β increases and is significant – and sometimes dominant – even above the $t\bar{b}$ threshold. Another important feature is that the total width of the charged Higgs is much lower than in other models, remaining below 1 GeV over the entire mass range for tan $\beta < 40$. Thus one would expect very monochromatic τs in the decay.

The indirect bounds from B decays are very similar in the LS and type I 2HDMs, since they are only relevant for $\tan \beta$ close to 1, where the Yukawa couplings are identical in both models [246]. The decay $B \to X_s \gamma$ gives precisely the same bound as shown in Fig. 18 for the type I 2HDM. The decays $B \to \ell^+ \ell^-$ and $b \to c \tau \nu$ do not give useful bounds [55]. One can get an effect from $B^+ \to \tau^+ \nu_\tau$, since the charged Higgs can mediate this decay, but that effect also turns out to be negligible for charged-Higgs masses which are not yet excluded by the direct searches. Logan and MacLennan [55] have also considered deviations from flavour universality in τ decays, finding an exclusion region for a fairly light charged Higgs that begins at $\tan \beta = 65$ for $m_{H^\pm} = 100$ GeV, and rises linearly to $\tan \beta = 200$ for $m_{H^\pm} = 250$ GeV.

The LEP direct-search limit of 78.6 GeV for the type I 2HDM still applies here. OPAL [350] has presented a stronger bound of 92.0 GeV under the assumption that the branching ratio into $\tau^+ \nu_\tau$ is 100%. This assumption is valid for large $\tan \beta$ and thus, over the entire range of $\tan \beta$, the lower bound will vary from 78.6 to 92.0 GeV. At the Tevatron, the D0 bound discussed for the type I 2HDM also applies in the LS 2HDM, but (as in the former case) it is not stronger than the bounds already imposed by indirect searches. Similarly, the ATLAS bound discussed in the previous section, for a charged Higgs to be produced in top decays, applies indifferently in the type I and LS 2HDMs.

There is an additional possibility at the LHC. In all other 2HDMs, detection of the charged Higgs above the $t\bar{b}$ threshold is very difficult due to the huge backgrounds; the analyses have had to rely on the small branching fraction of the H^+ into $\tau^+ \nu_\tau$. In the LS model, however, that branching fraction is very large and will (provided $\tan \beta$ is not too small) dominate. This sounds promising but, unfortunately, the production rate of H^+ through $gg \to \bar{t} b H^+$ scales like $\cot^2 \beta$ and that will suppress the result. Still, a sufficient enhancement at intermediate $\tan \beta$ may occur. As noted by Aoki et al. [36], one could readily distinguish the type II and LS 2HDMs through the decay rates into leptons as opposed to quark, if the charged Higgs could be detected.

One can also look for pair production of $H^+ H^-$ proceeding to $\tau^+ \nu_\tau \tau^- \bar{\nu}_\tau$. A similar phenomenon was studied by Davidson and Logan [91,351] in the neutrino-specific 2HDM, to be discussed shortly. They looked at charged-Higgs pair production at the LHC but, in their case, the decays of the charged Higgs studied were into $\mu\nu$ and $e\nu$. It is clear that $\tau\nu$ will be much more difficult and the missing energy in the τ decay will render irrelevant some of the cuts that Davidson and Logan used. The possibility of pair production of charged Higgs above the top threshold with the charged Higgs decaying entirely into $\tau\nu$ had not yet, to our knowledge, been studied.

One could look at associated production of the charged Higgs with a neutral Higgs, proceeding to three τs and missing energy. The only detailed analysis of this possibility that we are aware of is the one of Aoki et al. [36]; at the LHC with 300 fb^{-1} of accumulated luminosity, they have found – for scalar masses of 150 GeV – that the signal is comparable to the background – which originates primarily from $W^\pm Z$ – and they have discussed possible cuts. They note that τ misidentification (hadrons being identified as τs) may be a serious problem; realistic simulations are necessary. They also study $\mu\mu\tau\nu$ events; the rate is much smaller, since the branching ratio of neutral scalars into muons is much less than the one into τs, but the resolution of the muon pair is much better. If the neutral scalar has already been identified, this could provide a very useful tag for the charged Higgs. Once again, realistic simulations are needed.

Of course, at an eventual ILC, signatures will be very clean and backgrounds negligible, and all of the various 2HDMs will be readily distinguishable [36].

4.1.4. The flipped 2HDM

In the flipped 2HDM the Yukawa couplings to the quarks are the same as in the type II 2HDM but the Yukawa couplings to the leptons are proportional to $\cot \beta$ instead of $\tan \beta$. As a result, the various bounds from hadronic processes such as $B \to X_s \gamma$ are identical in the flipped and type II 2HDMs. Without a cancellation due to New Physics, the lower bound on the charged-Higgs mass is close to 300 GeV. Although the flipped model has been discussed in several recent general studies of the four 2HDMs [36,50,246,262], the only paper explicitly dedicated to the phenomenology of the charged Higgs in the flipped 2HDM that we are aware of is the one by Logan and MacLennan [53].

For moderately large $\tan \beta$, branching ratios in the flipped and type II 2HDMs are quite different. For $\tan \beta = 1$ the branching ratios are identical, see the left panel of Fig. 16. But as $\tan \beta$ increases the branching ratios to leptons become suppressed, as shown in Fig. 19. One can see that, below the $t\bar{b}$ threshold, the dominant decay of H^+ is into $\tau^+ \nu_\tau$ for $\tan \beta \sim 1$ and into $c\bar{b}$ or $c\bar{s}$ for larger $\tan \beta$.

As stated above, the indirect bounds on $m_{H^\pm}$ from $b \to s\gamma$, from Δm_B, and from R_b are identical in the flipped and type II 2HDMs. The first process yields a lower bound of 295 GeV on $m_{H^\pm}$ and the other two processes exclude the region $\tan \beta < 1$ for higher masses—see Jung et al. [246] for plots of these bounds. As noted earlier, though, it is possible that New Physics cancels out the effects of the charged Higgs in the loop—even a fairly mild cancellation would weaken the bounds.

The bound on $m_{H^\pm}$ from LEP [303] explicitly assumed that H^+ decays into either $\tau^+ \nu_\tau$ or $c\bar{s}$. While this is true in the type II 2HDM, it is not true for the flipped 2HDM, since the decay into $c\bar{b}$ may dominate for large or moderate $\tan \beta$. The ALEPH Collaboration [352] produced bounds independent of the quark flavours and one thus knows that $m_{H^\pm} > 79.3$ GeV.

At the Tevatron, the bounds discussed earlier also assume that the charged Higgs decays into either $c\bar{s}$ or $\tau^+ \nu_\tau$. Logan and MacLennan [53] have analysed these experiments in the context of the flipped 2HDM. Their conclusions are somewhat similar to Fig. 17, with a lower bound in the low-$\tan \beta$ region and an upper bound in the high-$\tan \beta$ region, *cf.* [53] for the
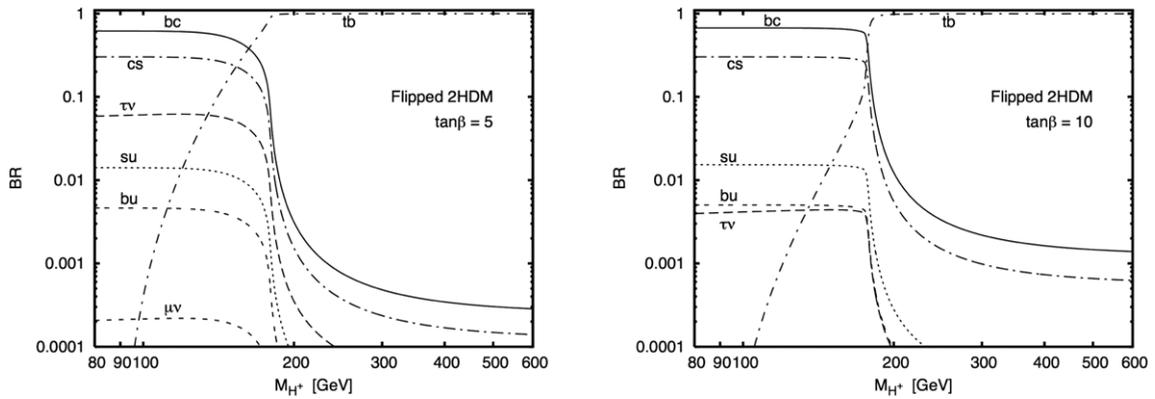
Fig. 19. Branching ratios of the charged Higgs in the flipped 2HDM, from Ref. [53]. Decays into $W^{\pm}$ and a neutral Higgs boson are not considered. The ratio of the dominant channels depends quadratically on the strange quark mass, and thus has a sizeable uncertainty.

detailed plots. Just as in the type II 2HDM, there is a large window at intermediate $\tan \beta$ where the Tevatron is unable to produce bounds on $m_{H^{\pm}}$.

The prospects of detection of the $H^{\pm}$ at the LHC are less promising than in the type II 2HDM. Firstly, suppose that $m_{H^{\pm}}$ is above the top threshold. Then one can study associated production of a charged Higgs with a top. This is similar to what happens in the type II model. There, some hope was obtained by looking at the relatively rare decay $H^+ \rightarrow \tau^+ \nu_{\tau}$. In the flipped 2HDM, however, that decay mode is negligible for moderate or large $\tan \beta$. Thus, there is no advantage to the flipped 2HDM over the type II 2HDM. Secondly, suppose that $m_{H^{\pm}}$ is below the top threshold. The ATLAS and CMS studies looking for top decays into a charged Higgs assumed that the latter primarily decays into $\tau^+ \nu_{\tau}$. This is generally not the case in the flipped 2HDM and it is unlikely that one can get better bounds than the Tevatron. However, an ATLAS study [353,354] of $t \rightarrow H^+ b, H^+ \rightarrow c\bar{s}$ shows that the sensitivity (at 1 fb^{-1} at 7 TeV) is better than the Tevatron. As shown in Refs. [53] and [355], b-tagging could be used to enhance the detection prospects of $t \rightarrow H^+ b, H^+ \rightarrow c\bar{b}$ and distinguish between $H^+ \rightarrow c\bar{b}$ and $H^+ \rightarrow c\bar{s}$. Such a b-tag would provide sensitivity in the region around the W mass which cannot be probed at the Tevatron due to backgrounds.

What about charged-Higgs pair production? For $m_{H^{\pm}} < m_t$ the dominant decay mode [31,32] is (for moderate $\tan \beta$) into $c\bar{b}$, leading to $c\bar{b}\bar{c}b$ final states which will be impossible to detect due to large QCD backgrounds; for a heavier charged Higgs the $t\bar{b}\bar{t}b$ signal will also be impossible to pick out. Thus, Logan and MacLennan [53] concluded that the prospects for detection of the charged Higgs at LHC are not bright in the flipped 2HDM (at the ILC, of course, signatures would be easy to pick out).

4.1.5. Other models

An interesting model that has a potentially exciting charged-Higgs phenomenology is the neutrino-specific model [91]. In it, the second Higgs doublet only couples to right-handed neutrinos and has an extremely small vacuum expectation value, of order eV. Then, the only decay modes of the charged Higgs are into a charged lepton and a right-handed neutrino, the coupling constant being proportional to the neutrino mass times the appropriate lepton-mixing-matrix element. With the normal neutrino mass hierarchy, the decay of H^+ would be primarily into $\mu^+ \nu$ and $\tau^+ \nu$, with a branching ratio of roughly 50% into each; for an inverted mass hierarchy, the decays are primarily into $e^+ \nu$ (roughly 50%), with the remainder into $\mu^+ \nu$ and $\tau^+ \nu$. Thus, the most dramatic signatures would be $H^+ \rightarrow e^+ \nu$ and $H^+ \rightarrow \mu^+ \nu$.

Davidson and Logan [91] have considered the process $pp \rightarrow H^+ H^- \rightarrow \ell^+ \nu \ell'^- \bar{\nu}$. Backgrounds to this process come from the production of $W^+ W^-, ZZ, Z\gamma$, or $t\bar{t}$ (in which the b quarks in the decay of the tops are missed). For $m_{H^{\pm}} = 100$ (300) GeV, the production cross section is 300 (5) fb. They have looked at various cuts in order to increase the signal-to-background ratio. In the $m_{H^{\pm}} = 100$ GeV case, the luminosity needed for a 5σ discovery is 10 to 70 fb^{-1} for normal neutrino mass hierarchy and 8 to 15 fb^{-1} for inverted hierarchy (the range depends on the allowed range of neutrino masses and mixings). In the $m_{H^{\pm}} = 300$ GeV case, the luminosity in the normal-hierarchy case is 55 to 450 fb^{-1} and, in the inverted-hierarchy case, 25 to 55 fb^{-1} (these calculations assumed a 14 TeV LHC). Thus, the signatures in this model are dramatic and within reach of the upgraded LHC.

The primary interest in the Inert Doublet (ID) model is dark matter and not the charged Higgs, since a candidate for that matter naturally arises in the model; most phenomenological studies of the ID model have focused on its neutral sector, and especially on the possibility that the Standard-Model Higgs particle decays into the dark-matter particle. Since the charged Higgs in the ID model does not couple to fermions, it will only decay into $HW^{\pm}$ or $AW^{\pm}$, where the $W^{\pm}$ is either real or virtual, depending on the masses. Cao et al. [175] have considered the associated production of the charged Higgs and a pseudoscalar at LHC and have found cross sections (which are very sensitive to the masses of both the charged Higgs and the pseudoscalar) of several hundred fb for masses below 200 GeV. They have also considered charged-Higgs pair production and concluded that, even for relatively low values of the charged-Higgs mass, the backgrounds overwhelm the signals. Later,

Miao et al. [356,357] found results for the production cross section similar to those of Cao et al. and performed a detailed analysis of possible cuts. They considered eight different benchmark points and tailored the cuts to each of those points. For two of the points a 5σ discovery of the charged Higgs is possible with a luminosity of 300 fb^{-1}; in both of these points $m_{H^\pm}$ is only 110 GeV—a heavier charged Higgs probably cannot be discovered. Huitu et al. [358] have shown that, if one extends the ID model by including a scalar singlet, then the charged Higgs could be long-lived, leading to other detection possibilities. Finally, as noted by Ginzburg [262], the ILC will easily be able to detect and study the charged Higgs up to its kinematic limit.

As pointed out in Section 3, the Lee–Wick Standard Model is a very unusual 2HDM, in which the second doublet states have wrong sign kinetic terms and masses and negative norms. The neutral sector of the theory is similar to an inert model, but the charged sector is identical [198] to a type II 2HDM with $\tan\beta = 1$ and with the sign differences for the three point Higgs-Higgs-γ, Z and the Yukawa couplings, as well as a factor of -1 for each charged Higgs propagator. Carone and Primulando [198] analysed $B_q - \bar{B}_q$ mixing, $B \to X_s\gamma$, R_b from Z-decay, Generally, the effects are of opposite sign to the conventional 2HDM, but the constraints turn out to be similar. They found that the lower bound on the charged Higgs mass from $B_d - \bar{B}_d$, $B_s - \bar{B}_s$ and $b \to X_s\gamma$ are given by 303, 354 and 463 GeV, respectively, and that no further bounds can be found from R_b.

4.2. Models with tree-level flavour-changing neutral currents

4.2.1. The type III model

The most general Yukawa couplings of two Higgs doublets are given by [359,360]

$$\bar{Q}_L \eta_1^U U_R \tilde{\Phi}_1 + \bar{Q}_L \eta_1^D D_R \Phi_1 + \bar{Q}_L \eta_2^U U_R \tilde{\Phi}_2 + \bar{Q}_L \eta_2^D D_R \Phi_2 + \text{H.c.}, \tag{92}$$

where the η_k ($k = 1, 2$) are 3×3 matrices; we have not included leptonic Yukawa couplings for simplicity. With our standard definitions and assuming real vevs, the fermion mass matrices are

$$M^F = \frac{v}{\sqrt{2}} \left(\eta_1^F \cos\beta + \eta_2^F \sin\beta \right), \tag{93}$$

with $F = U, D$. Defining [46]

$$\kappa^F \equiv \eta_1^F \cos\beta + \eta_2^F \sin\beta \tag{94}$$

and the orthogonal combination

$$\rho^F \equiv -\eta_1^F \sin\beta + \eta_2^F \cos\beta, \tag{95}$$

one can move to the ("Higgs") basis where only one doublet, called H_1, has a vev, thus generating fermion mass matrices and coupling with κ^F. The other doublet, H_2, has zero vev and couples with ρ^F. The coupling of the charged Higgs boson then is

$$H^+ \bar{U} \left(V \rho^D P_R - \rho^U V P_L \right) D + \text{H.c.}, \tag{96}$$

where V is the CKM matrix, and the $\rho^{U,D}$ matrices have been rotated by the same unitary matrices that diagonalize $\kappa^{U,D}$. Of course, all the charged-Higgs couplings are flavour-changing, but the neutral sector will have tree-level FCNC couplings unless ρ^U and ρ^D are diagonal. This occurs if either $\eta_1^U = \eta_1^D = 0$ (the type I model) or $\eta_1^D = \eta_2^U = 0$ (the type II model); these relations translate into $\rho^F = \kappa^F \cot\beta$ and $\rho^F = -\kappa^F \tan\beta$, respectively, and, since the κ^F are diagonal, the ρ^F then are diagonal too.

In the type III 2HDM, neither of these assumptions is made. The Cheng–Sher [206] *Ansatz*, discussed in Section 4, is

$$\rho_{ij}^F = \lambda_{ij}^F \frac{\sqrt{2m_i m_j}}{v}, \tag{97}$$

with the λ^F of O(1). In Section 3 we have shown how this *Ansatz* is being challenged by current experiments.

Other versions of the 2HDM include the aligned 2HDM [37], in which it is assumed that the Yukawa-coupling matrices $\eta_{1,2}^F$ are proportional. That model has no tree-level FCNC and was discussed in the last section (see [246] for a detailed analysis of the charged-Higgs phenomenology of the aligned 2HDM). In another 2HDM, Mahmoudi and Stal [46] used the Cheng–Sher parametrization and noted that, if one assumes that the λ_{ij} are diagonal, then both the aligned 2HDM and the various Z_2 models arise as special cases; we shall include that model in this section. One can also find interesting phenomenological possibilities by considering specific Yukawa textures [361,362].

We first discuss direct searches. The LEP bounds on charged-Higgs masses will still apply, but the relative branching ratios to $\tau^+\nu_\tau$ vs. $c\bar{s}$ are parameter-dependent. Still, a bound of around 75 to 80 GeV is expected based on the energy scale. Hadronic colliders can extend the reach. In 1999, He and Yuan [316] discussed a new method of detecting charged Higgs in models with tree-level FCNC. The idea is that, if a model has a large $\bar{t}cH$ coupling (H is a neutral Higgs), then, from isospin

symmetry, there will be a large $\bar{b}cH^+$ coupling and one can produce a charged Higgs through the s channel, since the b and c content of the proton is not negligible. Although this process can exist in standard 2HDMs, it is suppressed by the small V_{cb} CKM-matrix element.

In the type III 2HDM, $(\rho^U V)_{cb} \approx \rho_{ct}^U V_{tb} + \rho_{cc}^U V_{cb} \approx \rho_{ct}^U$, whereas $(V\rho^D)_{cb} \approx \rho_{sb}^D$. The latter is, in the Cheng–Sher *Ansatz*, much smaller, therefore the dominant vertex involves $\bar{c}_R b_L$ and not – as in the standard Type I or Type II 2HDM – $\bar{c}_L b_R$. With the Cheng–Sher *Ansatz* this vertex will be quite large, $O(\sqrt{m_c m_b/v^2}) \sim 1\%$. He and Yuan [316] find that, for a 14 TeV LHC with 100 fb^{-1} luminosity, one would get, for $m_{H+} = 300\,(800)$ GeV, over 130,000 (380) single-top events from the charged-Higgs decay and (if kinematically accessible) about 180,000 (4000) events $H^+ \to W^+ h \to W^+ (b\bar{b}, \tau^+ \tau^-)$. With appropriate cuts, this can by seen over the Standard-Model rate (from W^*) up to charged Higgs masses below 350 GeV, whereas for heavier masses, the signature is not observable.

A study of this process $H^+ \to t\bar{b}$ at the Tevatron, where the rates are of course smaller, was carried out by the D0 Collaboration [363], which was however unable to set bounds on $m_{H\pm}$ unless the coupling of the charged Higgs to $q_i \bar{q}_j$ was substantially higher than expected in the type III 2HDM. Further analysis of the D0 results was carried out by Cardenas and Rodríguez [364]. They looked at a charged Higgs above the $t\bar{b}$ threshold and found that, if $\lambda_{tt} = \lambda_{tc} = 2.8\,(5.0)$, then $m_{H\pm}$ must be above 230 (264) GeV. Below the $t\bar{b}$ threshold, they found from top decays that the lower bound on $m_{H\pm}$ is 145 (160) GeV for $\lambda_{tt} = \lambda_{tc} = 2.8\,(5.0)$ (the value 2.8 is the perturbative validity upper bound [365]). For $\lambda_{tt} = \lambda_{tc} = 1$, no substantive bound could be found on $m_{H\pm}$. Single-top production at the LHC will provide a good probe of the model.

As in the type I and type II models, one can set bounds on the parameters of the type III 2HDM from indirect processes. The first discussion of bounds on $m_{H\pm}$ from $B \to X_s \gamma$ in the type III 2HDM was Ref. [216], but this was long before NLO corrections, which are important, had been calculated. A comprehensive study is difficult because of the large number of parameters in the type III 2HDM. A simplified analysis, in which only λ_{tt} and λ_{bb} are non-zero, was performed by Bowser-Chao et al. [232]. This has the advantage of a smaller parameter space and of the fact that only charged-Higgs loops are relevant—unlike the general case, where neutral-Higgs loops are important too. They allowed for the possibility that those two couplings are complex—the relative phase determines whether the charged-Higgs loops interfere destructively or constructively, and can also give rise to a non-zero neutron electric dipole moment (EDM). Their results are shown in Fig. 20 for three different values of $|\lambda_{tt}\lambda_{bb}|$. On the x axis is the phase angle, which is 180° in the type I and type II models. The type II model corresponds to the middle figure at $\theta = 180°$; and one can see the bound of roughly 300 GeV of that model. The shaded region is excluded by the experimental upper bound on the neutron EDM. Note that, for a reduction by a factor of two of $|\lambda_{tt}\lambda_{bb}|$, the lower bound on $m_{H\pm}$ drops almost to the observed lower bound from LEP (this work was done in 1999 and the experimental constraints are much tighter now, but NLO effects have been calculated and lower the curves only a little, resulting in similar results). They also gave bounds from B–$\bar{B}$ mixing and R_b, but the experimental results have changed a bit since then and the results are rather sensitive to those changes. The important aspect of the paper by Bowser-Chao et al. is that, through slightly lower values of $|\lambda_{tt}\lambda_{bb}|$, one can substantially lower the bound on $m_{H\pm}$ in the type III 2HDM.

The most comprehensive analysis of $B \to X_s \gamma$ and of B–$\bar{B}$ mixing in the type III 2HDM was the one by Xiao and Guo [233] and also, with a specific focus on $B \to (K^*, \rho)\gamma$, the study of Xiao and Zhuang [366]. There are numerous plots in these two papers. From the B–$\bar{B}$ mass difference, and assuming that only λ_{bb} and λ_{tt} are non-zero, they [233,366] find that $\lambda_{tt} < 1.7$ for any $m_{H\pm} < 300$ GeV; moreover, even for such small values of λ_{tt}, there is an excluded region: from 0.75 to 1.6 for $m_{H\pm} = 200$ GeV, from 0.95 to 1.4 for $m_{H\pm} = 300$ GeV, and from 1.15 to 1.25 for $m_{H\pm} = 400$ GeV. They also present results similar to those of Bowser-Chao et al. for $B \to X_s \gamma$, although they use a different set of parameters and also improved, updated data. They provide a comprehensive set of formulae which can easily be employed for other parameter choices.

Of course, if additional λ_{ij} are taken to be non-zero the parameter space becomes quite large. Some recent analyses that discuss other λ_{ij} include the work of Diaz et al. [367], in which the effects of non-zero λ_{sb} and λ_{tc} were discussed (λ_{sb} was found to be unbounded, λ_{tc} was found to be smaller than 1), and of Idarraga et al. [368], who also considered bounds from leptonic B decays but chose very large $\lambda_{\mu\mu}$ and $\lambda_{\tau\tau}$. Also note that, if neutral fields are allowed in the loop, one could get an effect proportional to λ_{bs}^2, which is, however, negligible.

The possibility of s-channel charged-Higgs production at the LHC may provide the best method of detection, and detailed simulations would be welcomed.

As discussed in Section 3, the BGL and MFV couplings can be obtained from the type III model with the Cheng–Sher ansatz replaced by another ansatz (which may depend on parameters and which will vary for different models). Thus, everything in this section would apply to these models, but with λ_{ij} not being of O(1).

5. The scalar sector of the 2HDM

The scalar sector of the 2HDM has many interesting features. In its most general form, the potential apparently has 14 independent parameters. However, the fact that the Higgs doublets Φ_1 and Φ_2 are *not* physical observables – only the scalar mass eigenstates are physical particles – means that we have the freedom to redefine those doublets, provided we preserve the form of their kinetic terms. These basis changes of the Higgs doublets allow one to absorb some of the parameters in the potential and are essential to understand the number of physical parameters really present in it.
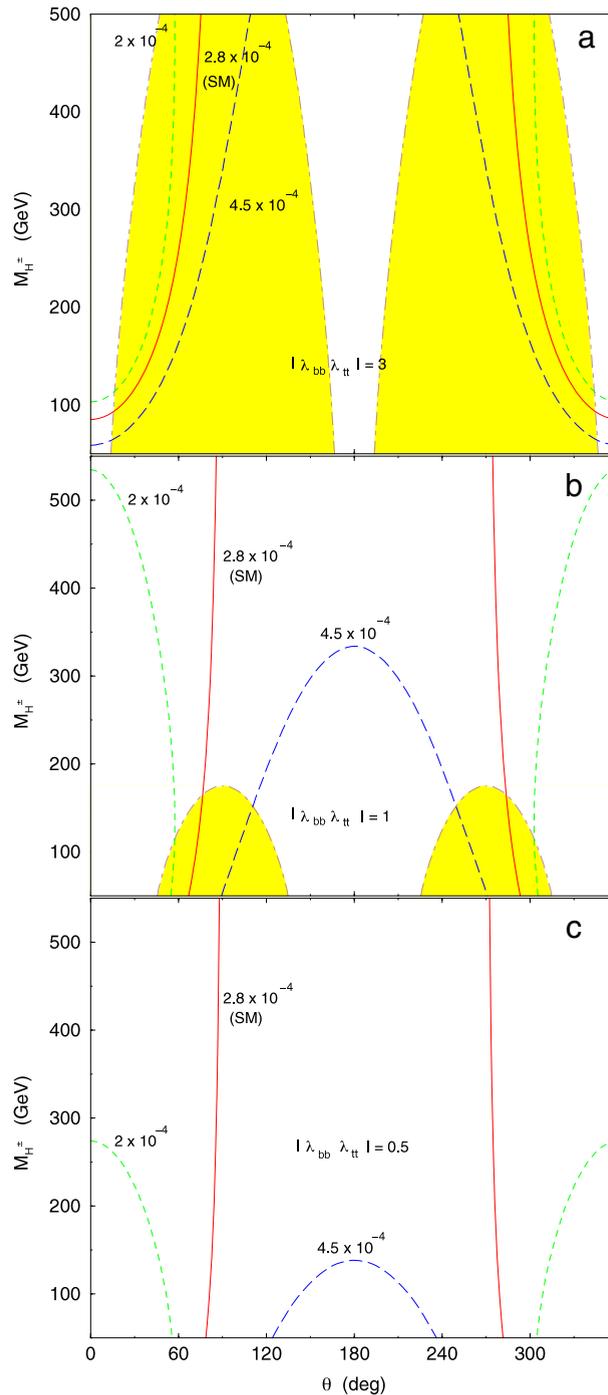
Fig. 20. Contours of the branching ratio of $b \to s\gamma$ as a function of the charged-Higgs mass and of the phase θ between λ_{tt} and λ_{bb}, for three different values of $|\lambda_{tt}\lambda_{bb}|$. The shaded areas are excluded by the experimental bound on the neutron electric dipole moment.
Source: This figure is from Ref. [232].

For several good reasons – the most usual of which is preventing the occurrence of FCNC in the 2HDM – it is common to impose a variety of global symmetries on the 2HDM, thus reducing the number of free parameters. In a highly non-intuitive result, it has been proven that there are only *six* such symmetries which have distinct effects on the scalar potential. The six models resulting from each of those symmetries have different physical implications: different spectra of scalars, different interactions with gauge bosons, and, in some cases, predictions of massless axions or potential dark matter candidates.

It is the scalar potential that determines the vacuum of the 2HDM, and that vacuum, unlike what happens in the SM, is not unique: with two doublets the possibility arises that the vacuum of the model spontaneously breaks the CP symmetry—which in fact is precisely the reason why T. D. Lee first proposed the 2HDM in 1973 [9]. However, for certain values of the parameters in the potential, vacua which violate the electromagnetic symmetry, giving mass to the photon, are also possible. Those have, of course, to be avoided. Even if one considers only vacua which preserve both CP and the usual gauge symmetries of the SM, the 2HDM has a rich vacuum structure: some of the possible 2HDM potentials can display so-called "inert vacua", in which one of the neutral scalars does not couple to gauge bosons at all (and can easily be made to decouple from fermions as well). Some other potentials may have two different electromagnetism-preserving minima, with different predictions for the masses of the gauge bosons, for instance. The 2HDM has however a feature which distinguishes it from other multi-Higgs models, such as SUSY, the Zee model, or the 3HDM: its vacua are stable and no tunnelling from a neutral, CP-conserving vacuum to a deeper, CP- or charge-breaking vacuum, is possible. And vice-versa: any CP- or charge-breaking minimum that one finds is guaranteed to be the global minimum of the model.

However, not all values of the parameters of the 2HDM potential ensure that there is a stable minimum, unless one can be sure that the potential is bounded from below. This basic requirement allows one to impose constraints on the quartic scalar couplings. A renormalization-group improvement of those constraints translates in possibly severe bounds on the masses of the physical scalar particles. In the following we shall analyse these questions. We shall discuss the vacuum structure of the potential, derive general formulae for the scalar masses, present bounds on quartic couplings obtained from the requirement that the potential is bounded from below, and discuss the symmetries of the scalar potential and their extension to the Yukawa sector.

5.1. The scalar potential, notation 1

The most general renormalizable, *i.e.* quartic, scalar potential may be written [369]

$$
\begin{aligned}
V_H ={}& m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - \left(m_{12}^2 \Phi_1^\dagger \Phi_2 + \text{H.c.} \right) \\
&+ \frac{1}{2}\lambda_1 \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{1}{2}\lambda_2 \left(\Phi_2^\dagger \Phi_2 \right)^2 + \lambda_3 \left(\Phi_1^\dagger \Phi_1 \right) \left(\Phi_2^\dagger \Phi_2 \right) + \lambda_4 \left(\Phi_1^\dagger \Phi_2 \right) \left(\Phi_2^\dagger \Phi_1 \right) \\
&+ \left[\frac{1}{2}\lambda_5 \left(\Phi_1^\dagger \Phi_2 \right)^2 + \lambda_6 \left(\Phi_1^\dagger \Phi_1 \right) \left(\Phi_1^\dagger \Phi_2 \right) + \lambda_7 \left(\Phi_2^\dagger \Phi_2 \right) \left(\Phi_1^\dagger \Phi_2 \right) + \text{H.c.} \right],
\end{aligned}
\tag{98}
$$

where "H.c." stands for the Hermitian conjugate. The parameters m_{11}^2, m_{22}^2, and $\lambda_{1,2,3,4}$ are real. In general, m_{12}^2 and $\lambda_{5,6,7}$ are complex. Thus, the Higgs potential in Eq. (98) depends on six real and four complex parameters, *i.e.* a total of fourteen degrees of freedom. However, as we shall see below, the freedom to redefine the basis means that in reality only eleven degrees of freedom are physical.

In Eq. (98) we are following the definitions of Davidson and Haber [359]; often other definitions are used, in which the same symbol may be employed for quantities which differ from ours in sign, a factor of two, or complex conjugation.

5.2. The scalar potential, notation 2

An alternative notation for the scalar potential, which has been championed by Botella and Silva [370], is

$$
V_H = \sum_{a,b=1}^{2} \mu_{ab}\, \Phi_a^\dagger \Phi_b + \frac{1}{2} \sum_{a,b,c,d=1}^{2} \lambda_{ab,cd} \left(\Phi_a^\dagger \Phi_b \right) \left(\Phi_c^\dagger \Phi_d \right),
\tag{99}
$$

where, by definition,

$$
\lambda_{ab,cd} = \lambda_{cd,ab}.
\tag{100}
$$

In Eq. (99) Hermiticity implies

$$
\mu_{ab} = \mu_{ba}^* \quad \text{and} \quad \lambda_{ab,cd} = \lambda_{ba,dc}^*.
\tag{101}
$$

The notation of Eq. (99) is useful for the study of invariants, basis transformations, and symmetries. The correspondence between notations 1 and 2 is given by

$$
\begin{aligned}
&\mu_{11} = m_{11}^2, \qquad \mu_{22} = m_{22}^2, \\
&\mu_{12} = -m_{12}^2, \qquad \mu_{21} = -m_{12}^{2\,*} \\
&\lambda_{11,11} = \lambda_1, \qquad \lambda_{22,22} = \lambda_2, \\
&\lambda_{11,22} = \lambda_{22,11} = \lambda_3, \qquad \lambda_{12,21} = \lambda_{21,12} = \lambda_4, \\
&\lambda_{12,12} = \lambda_5, \qquad \lambda_{21,21} = \lambda_5^*, \\
&\lambda_{11,12} = \lambda_{12,11} = \lambda_6, \qquad \lambda_{11,21} = \lambda_{21,11} = \lambda_6^*, \\
&\lambda_{22,12} = \lambda_{12,22} = \lambda_7, \qquad \lambda_{22,21} = \lambda_{21,22} = \lambda_7^*.
\end{aligned}
\tag{102}
$$

Once again, one must be careful when confronting Eq. (99) to similar equations written in other papers, since the same symbol may be used in different papers for quantities which differ in sign, a factor of two, or complex conjugation.

5.3. The scalar potential, notation 3

The previous two notations consider the scalar doublets Φ_a $(a = 1, 2)$ individually. A third notation emphasizes the presence of field bilinears $\Phi_a^\dagger \Phi_b$ in the scalar potential. An early use of bilinears is due to Velhinho et al. [371]; the notation has later been much employed by Nagel [372], by Maniatis et al. [373–375], by Nishi [376–378], and by Ivanov [379–381]. Following Nishi [376] we write:

$$V_H = \sum_{\mu=0}^{3} M_\mu r_\mu + \sum_{\mu,\nu=0}^{3} \Lambda_{\mu\nu} r_\mu r_\nu, \tag{103}$$

where

$$\Lambda_{\mu\nu} = \Lambda_{\nu\mu} \tag{104}$$

and

$$
\begin{aligned}
r_0 &= \frac{1}{2}\left(\Phi_1^\dagger \Phi_1 + \Phi_2^\dagger \Phi_2\right), \\
r_1 &= \frac{1}{2}\left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1\right) = \mathrm{Re}\left(\Phi_1^\dagger \Phi_2\right), \\
r_2 &= -\frac{i}{2}\left(\Phi_1^\dagger \Phi_2 - \Phi_2^\dagger \Phi_1\right) = \mathrm{Im}\left(\Phi_1^\dagger \Phi_2\right), \\
r_3 &= \frac{1}{2}\left(\Phi_1^\dagger \Phi_1 - \Phi_2^\dagger \Phi_2\right).
\end{aligned}
\tag{105}
$$

In Eq. (103) we have adopted an Euclidean metric. It differs from the notation of Ivanov [380], who pointed out that r_μ parametrizes the gauge orbits of the Higgs fields in a space equipped with a Minkowski metric.

Notation 3 is convenient for studies of features such as the existence and number of minima of the scalar potential. Since the Yukawa couplings involve the Higgs doublets individually rather than bilinears, notation 3 cannot be applied for studies of the full theory with both scalars and fermions.

The correspondence between notations 1 and 3 is given by

$$M_\mu = \left(m_{11}^2 + m_{22}^2, \ -2\mathrm{Re}\left(m_{12}^2\right), \ 2\mathrm{Im}\left(m_{12}^2\right), \ m_{11}^2 - m_{22}^2\right), \tag{106}$$

$$\Lambda_{\mu\nu} = \begin{pmatrix}
(\lambda_1 + \lambda_2)/2 + \lambda_3 & \mathrm{Re}(\lambda_6 + \lambda_7) & -\mathrm{Im}(\lambda_6 + \lambda_7) & (\lambda_1 - \lambda_2)/2 \\
\mathrm{Re}(\lambda_6 + \lambda_7) & \lambda_4 + \mathrm{Re}(\lambda_5) & -\mathrm{Im}(\lambda_5) & \mathrm{Re}(\lambda_6 - \lambda_7) \\
-\mathrm{Im}(\lambda_6 + \lambda_7) & -\mathrm{Im}(\lambda_5) & \lambda_4 - \mathrm{Re}(\lambda_5) & -\mathrm{Im}(\lambda_6 - \lambda_7) \\
(\lambda_1 - \lambda_2)/2 & \mathrm{Re}(\lambda_6 - \lambda_7) & -\mathrm{Im}(\lambda_6 - \lambda_7) & (\lambda_1 + \lambda_2)/2 - \lambda_3
\end{pmatrix}. \tag{107}$$

The correspondence between notations 2 and 3 is given by

$$M_\mu = \sum_{a,b=1}^{2} (\sigma_\mu)_{ab}\, \mu_{ba}, \tag{108}$$

$$\Lambda_{\mu\nu} = \frac{1}{2} \sum_{a,b,c,d=1}^{2} (\sigma_\mu)_{ab} (\sigma_\nu)_{cd}\, \lambda_{ba,dc}, \tag{109}$$

where σ_0 is the 2×2 identity matrix and $\sigma_{1,2,3}$ are the three Pauli matrices.

5.4. Basis transformations

The doublets Φ_a are not physical—only the scalar mass eigenstates, corresponding to particles, are physical. Thus, any combination of the doublets which respects the symmetries of the theory will produce the same physical predictions. We call any combination of (Φ_1, Φ_2) a *basis* for the doublets. We may rewrite the potential in terms of new doublets Φ_a', obtained from the original ones by a (global) basis transformation

$$\Phi_a' = \sum_{b=1}^{2} U_{ab}\Phi_b, \tag{110}$$

where U is a 2×2 unitary matrix. Under this unitary basis transformation the gauge-kinetic terms are unchanged, but the coefficients of the potential in notation 2 transform as

$$\mu'_{ab} = \sum_{c,d=1}^{2} U_{ac}\, \mu_{cd}\, U^*_{bd} = \left(U\mu U^{\dagger}\right)_{ab}, \tag{111}$$

$$\lambda'_{ab,cd} = \sum_{e,f,g,h=1}^{2} U_{ae}\, U_{cg}\, \lambda_{ef,gh}\, U^*_{bf}\, U^*_{dh}. \tag{112}$$

The basis transformation of Eq. (110) induces a transformation of the 4-vector in Eq. (105) given by

$$r'_0 = r_0, \qquad r'_i = \sum_{j=1}^{3} R_{ij}(U)\, r_j \tag{113}$$

where [376]

$$R_{ij}(U) = \frac{1}{2}\, \mathrm{tr}\left(U^{\dagger}\sigma_i U \sigma_j\right). \tag{114}$$

The matrix R belongs to SO(3) and the transformation of U into $R(U)$ is the $SU(2) \to$ SO(3) two-to-one[13] mapping. Under this rotation of the 3-vector r_i, the parameters of the scalar potential in notation 3 transform as

- scalars,

$$M'_0 = M_0, \qquad \Lambda'_{00} = \Lambda_{00}, \tag{115}$$

- vectors,

$$M'_i = \sum_{j=1}^{3} R_{ij}M_j, \qquad \Lambda'_{0i} = \sum_{j=1}^{3} R_{ij}\Lambda_{0j}, \tag{116}$$

- and a symmetric tensor

$$\Lambda'_{ij} = \sum_{k,l=1}^{3} R_{ik}R_{jl}\Lambda_{kl}. \tag{117}$$

We see in Eqs. (111) and (112) that the overall phase of U does not impact the change of the parameters of the potential. As a result, one may consider $U \in SU(2)$. We then parametrize

$$U = \begin{pmatrix} e^{i\chi}c_{\psi} & e^{i(\chi-\xi)}s_{\psi} \\ -e^{i(\xi-\chi)}s_{\psi} & e^{-i\chi}c_{\psi} \end{pmatrix}, \tag{118}$$

where $c_{\psi} = \cos\psi$ and $s_{\psi} = \sin\psi$; a similar notation will be used below for multiples of the angle ψ. Then, in notation 1, the parameters m_{ij}^2 and λ_i of Eq. (98) transform as

$$m_{11}^{2\,'} = m_{11}^2 c_{\psi}^2 + m_{22}^2 s_{\psi}^2 - \mathrm{Re}\left(m_{12}^2 e^{i\xi}\right) s_{2\psi}, \tag{119}$$

$$m_{22}^{2\,'} = m_{11}^2 s_{\psi}^2 + m_{22}^2 c_{\psi}^2 + \mathrm{Re}\left(m_{12}^2 e^{i\xi}\right) s_{2\psi}, \tag{120}$$

$$m_{12}^{2\,'} = e^{i(2\chi-\xi)}\left[\frac{1}{2}\left(m_{11}^2 - m_{22}^2\right)s_{2\psi} + \mathrm{Re}\left(m_{12}^2 e^{i\xi}\right)c_{2\psi} + i\,\mathrm{Im}\left(m_{12}^2 e^{i\xi}\right)\right], \tag{121}$$

$$\lambda'_1 = \lambda_1 c_{\psi}^4 + \lambda_2 s_{\psi}^4 + \frac{1}{2}\lambda_{345}s_{2\psi}^2 + 2s_{2\psi}\left[c_{\psi}^2\,\mathrm{Re}\left(\lambda_6 e^{i\xi}\right) + s_{\psi}^2\,\mathrm{Re}\left(\lambda_7 e^{i\xi}\right)\right], \tag{122}$$

$$\lambda'_2 = \lambda_1 s_{\psi}^4 + \lambda_2 c_{\psi}^4 + \frac{1}{2}\lambda_{345}s_{2\psi}^2 - 2s_{2\psi}\left[s_{\psi}^2\,\mathrm{Re}\left(\lambda_6 e^{i\xi}\right) + c_{\psi}^2\,\mathrm{Re}\left(\lambda_7 e^{i\xi}\right)\right], \tag{123}$$

$$\lambda'_3 = \lambda_3 + \frac{1}{4}s_{2\psi}^2\left(\lambda_1 + \lambda_2 - 2\lambda_{345}\right) - s_{2\psi}c_{2\psi}\,\mathrm{Re}\left[(\lambda_6 - \lambda_7)\, e^{i\xi}\right], \tag{124}$$

$$\lambda'_4 = \lambda_4 + \frac{1}{4}s_{2\psi}^2\left(\lambda_1 + \lambda_2 - 2\lambda_{345}\right) - s_{2\psi}c_{2\psi}\,\mathrm{Re}\left[(\lambda_6 - \lambda_7)\, e^{i\xi}\right], \tag{125}$$

---

[13] Note that $R(U) = R(-U)$.

$$\lambda_5' = e^{2i(2\chi-\xi)} \left\{ \frac{1}{4} s_{2\psi}^2 (\lambda_1 + \lambda_2 - 2\lambda_{345}) + \text{Re}\left(\lambda_5 e^{2i\xi}\right) + i c_{2\psi} \text{Im}\left(\lambda_5 e^{2i\xi}\right) \right.$$
$$\left. - s_{2\psi} c_{2\psi} \text{Re}\left[(\lambda_6 - \lambda_7) e^{i\xi}\right] - i s_{2\psi} \text{Im}\left[(\lambda_6 - \lambda_7) e^{i\xi}\right] \right\}, \tag{126}$$

$$\lambda_6' = e^{i(2\chi-\xi)} \left\{ -\frac{1}{2} s_{2\psi} \left[\lambda_1 c_\psi^2 - \lambda_2 s_\psi^2 - \lambda_{345} c_{2\psi} - i \text{Im}\left(\lambda_5 e^{2i\xi}\right)\right] \right.$$
$$\left. + c_\psi c_{3\psi} \text{Re}\left(\lambda_6 e^{i\xi}\right) + s_\psi s_{3\psi} \text{Re}\left(\lambda_7 e^{i\xi}\right) + i c_\psi^2 \text{Im}\left(\lambda_6 e^{i\xi}\right) + i s_\psi^2 \text{Im}\left(\lambda_7 e^{i\xi}\right) \right\}, \tag{127}$$

$$\lambda_7' = e^{i(2\chi-\xi)} \left\{ -\frac{1}{2} s_{2\psi} \left[\lambda_1 s_\psi^2 - \lambda_2 c_\psi^2 + \lambda_{345} c_{2\psi} + i \text{Im}\left(\lambda_5 e^{2i\xi}\right)\right] \right.$$
$$\left. + s_\psi s_{3\psi} \text{Re}\left(\lambda_6 e^{i\xi}\right) + c_\psi c_{3\psi} \text{Re}\left(\lambda_7 e^{i\xi}\right) + i s_\psi^2 \text{Im}\left(\lambda_6 e^{i\xi}\right) + i c_\psi^2 \text{Im}\left(\lambda_7 e^{i\xi}\right) \right\}, \tag{128}$$

where

$$\lambda_{345} := \lambda_3 + \lambda_4 + \text{Re}\left(\lambda_5 e^{2i\xi}\right). \tag{129}$$

A basis transformation may be utilized to eliminate some of the degrees of freedom in the scalar potential. This implies that not all the parameters in that potential have physical significance. Thus, the three parameters in Eq. (118) may be used to absorb three out of the 14 parameters in the scalar potential. As a result, there are only 11 physical degrees of freedom in the potential and, thus, only eleven independent observables. Note, though, that we are still discussing the most general potential; when one imposes a global symmetry (see Section 5.6 ahead) on the 2HDM, the number of parameters which may be eliminated through basis transformations may be less than three.

5.5. GCP transformations

The standard CP transformation for a Higgs doublet reads

$$\Phi\left(t, \vec{x}\right) \rightarrow \Phi^{\text{CP}}\left(t, \vec{x}\right) = \Phi^*\left(t, -\vec{x}\right). \tag{130}$$

The reference to the time (t) and space $(\vec{x})$ coordinates will henceforth be suppressed.

However, in the presence of several identical doublets, the possibility of arbitrary basis transformations should be included in the definition of the CP transformation. Let us illustrate this problem with a simple example. We start from the basis (Φ_1, Φ_2) and the usual definition of CP:

$$\Phi_1^{\text{CP}} = \Phi_1^*, \qquad \Phi_2^{\text{CP}} = \Phi_2^*. \tag{131}$$

Now we perform a basis transformation

$$\Phi_1 = \frac{1}{\sqrt{2}} \left(\Phi_1' + e^{i\pi/4} \Phi_2'\right),$$
$$\Phi_2 = \frac{1}{\sqrt{2}} \left(-e^{-i\pi/4} \Phi_1' + \Phi_2'\right). \tag{132}$$

Substituting Eqs. (132) into Eqs. (131), we obtain

$$\frac{1}{\sqrt{2}} \left[(\Phi_1')^{\text{CP}} + e^{i\pi/4} (\Phi_2')^{\text{CP}}\right] = \frac{1}{\sqrt{2}} \left[(\Phi_1')^* + e^{-i\pi/4} (\Phi_2')^*\right]$$
$$\frac{1}{\sqrt{2}} \left[-e^{-i\pi/4} (\Phi_1')^{\text{CP}} + (\Phi_2')^{\text{CP}}\right] = \frac{1}{\sqrt{2}} \left[-e^{i\pi/4} (\Phi_1')^* + (\Phi_2')^*\right],$$

which leads to

$$(\Phi_1')^{\text{CP}} = \frac{1+i}{2} (\Phi_1')^* - \frac{i}{\sqrt{2}} (\Phi_2')^*,$$
$$(\Phi_2')^{\text{CP}} = -\frac{i}{\sqrt{2}} (\Phi_1')^* + \frac{1-i}{2} (\Phi_2')^*. \tag{133}$$

This is clearly *not* the usual CP transformation, which means that Eq. (131) is too restrictive as a definition of CP. Even the usual CP transformation may look different from Eq. (131) in a different basis.

As a result, we must consider a more general version of the CP transformation, which we denote with the superscript 'GCP'[14]:

$$
\begin{aligned}
\Phi_a \rightarrow \Phi_a^{\text{GCP}} &= \sum_{b=1}^{2} X_{ab} \Phi_b^*, \\
\Phi_a^\dagger \rightarrow \Phi_a^{\text{GCP}\,\dagger} &= \sum_{b=1}^{2} X_{ab}^* \Phi_b^T,
\end{aligned}
\tag{134}
$$

where X is an arbitrary unitary matrix. GCP transformations were first discussed by Lee and Wick [399]. Their explicit use for quarks is due to Bernabéu et al. [400]. GCP transformations in the scalar sector were developed by the Vienna group [382,383,401,402]. Note that, unlike the standard CP transformation of Eq. (130), GCP transformations are such that the square of the transformation is *not*, in general, equal to unity (see Section 5.6.3)—as we already see in the practical example of Eq. (133).

Under the GCP transformation in Eq. (134), the gauge-kinetic terms stay invariant but the coefficients μ_{ab} and $\lambda_{ab,cd}$ transform as

$$
\mu_{ab}^{\text{GCP}} = \sum_{c,d=1}^{2} X_{ac} \mu_{cd}^* X_{bd}^* = \left(X \mu^* X^\dagger \right)_{ab},
\tag{135}
$$

$$
\lambda_{ab,cd}^{\text{GCP}} = \sum_{e,f,g,h=1}^{2} X_{ae} X_{cf} \lambda_{eg,fh}^* X_{bg}^* X_{dh}^*.
\tag{136}
$$

We next turn to the interplay between GCP transformations and basis transformations. If the GCP transformation is given by Eq. (134), then $\Phi_a' = \sum_b U_{ab} \Phi_b$ has a GCP transformation given by

$$
\Phi_a' \rightarrow \left(\Phi_a' \right)^{\text{GCP}} = \sum_{b=1}^{2} X_{ab}' \Phi_b'^*,
\tag{137}
$$

where

$$
X' = U X U^T.
\tag{138}
$$

The fact that U^T, rather than $U^\dagger$, appears in Eq. (138) is crucial. If one had $U^\dagger$, then one would be able to find a basis such that X was diagonal. Because one has U^T in Eq. (138) *it is not possible to reduce, through a basis transformation, all GCP transformations to the standard CP transformation* of Eq. (130). However, Ecker et al. [383] have proved that for every matrix X there exists a unitary matrix U such that

$$
X' = U X U^T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},
\tag{139}
$$

with $0 \leq \theta \leq \pi/2$. The value of θ may be determined through either one of two ways: (i) the (twice degenerate) eigenvalue of $\left(X + X^T \right) \left(X^* + X^\dagger \right)$ is $4\cos^2\theta$; (ii) the eigenvalues of XX^* are $e^{\pm 2i\theta}$.

The GCP transformation in Eq. (134) induces the following transformation of the four-vector in Eq. (105):

$$
r_0^{\text{GCP}} = r_0, \qquad r_i^{\text{GCP}} = \sum_{j=1}^{3} \bar{R}_{ij}(X)\, r_j
\tag{140}
$$

where [375]

$$
\begin{aligned}
\bar{R}(X) &= R(X)\,\bar{R}_2, \\
R_{ij}(X) &= \frac{1}{2}\,\text{tr}\left(X^\dagger \sigma_i X \sigma_j \right), \\
\bar{R}_2 &= \text{diag}\,(1, -1, 1).
\end{aligned}
\tag{141}
$$

Both $\bar{R}_2$ and $\bar{R}(X)$ are improper rotations, *i.e.* O(3) matrices with determinant -1. For the simplified form of X in Eq. (139) one has

$$
\bar{R}(X) = \begin{pmatrix} \cos(2\theta) & 0 & -\sin(2\theta) \\ 0 & -1 & 0 \\ \sin(2\theta) & 0 & \cos(2\theta) \end{pmatrix}.
\tag{142}
$$

---

[14] These are known in the literature by many authors as Generalized CP transformations [382–398].

5.6. The six classes of symmetry-constrained scalar potentials

The large number of free parameters in the scalar potential of the 2HDM reduces the theory's predictive power. Any symmetry that we may impose on the 2HDM to constrain its scalar potential is therefore welcome. Also, as we discussed in previous sections, the 2HDM is in general plagued by flavour-changing neutral currents; these, however, may be eliminated – or strongly suppressed – by imposing an internal symmetry on the 2HDM.

Symmetries leaving the kinetic terms unchanged may be of either one of two types:

1. One may relate Φ_a to some unitary transformation of Φ_b,

$$\Phi_a \to \Phi_a^S = \sum_{b=1}^2 S_{ab} \Phi_b, \tag{143}$$

where S is a unitary matrix. We then require the potential to be invariant under this transformation. As a result of this invariance,

$$\mu_{ab} = \sum_{c,d=1}^2 S_{ac} \mu_{cd} S_{bd}^*, \tag{144}$$

$$\lambda_{ab,cd} = \sum_{e,f,g,h=1}^2 S_{ae} S_{cf} \lambda_{eg,fh} S_{bg}^* S_{dh}^*. \tag{145}$$

These are known as Higgs Family (HF) symmetries.[15]

2. One may relate Φ_a with some unitary transformation of Φ_b^*:

$$\Phi_a \to \Phi_a^{\text{GCP}} = \sum_{b=1}^2 X_{ab} \Phi_b^*, \tag{146}$$

where X is an arbitrary unitary matrix. We then require that the potential be invariant under this symmetry:

$$\mu_{ab} = \sum_{c,d=1}^2 X_{ac} \mu_{cd}^* X_{bd}^*, $$

$$\lambda_{ab,cd} = \sum_{e,f,g,h=1}^2 X_{ae} X_{cf} \lambda_{eg,fh}^* X_{bg}^* X_{dh}^*. \tag{147}$$

These are known as GCP symmetries.

Under the basis transformation $\Phi_a \to \Phi_a' = U_{ab} \Phi_b$ of Eq. (110) the specific forms of the HF and GCP symmetries get altered, respectively, into:

$$S' = USU^\dagger, \tag{148}$$

$$X' = UXU^T. \tag{149}$$

Therefore, a symmetry relation among the coefficients of the scalar potential will, in general, appear as a different relation if the coefficients of the potential are written using a different basis for the Higgs doublets.

One may, of course, impose on a theory several HF symmetries and/or GCP symmetries simultaneously. Ivanov [380] has proved that, no matter what combination of HF and/or GCP symmetries one imposes on the scalar potential of the 2HDM, one always ends up with one of six distinct classes of potentials. This issue was studied further by Ferreira et al. [386]. In Table 5 we present an example of a symmetry in each of the six classes of symmetries found by Ivanov, and the constraints on the parameters of the potential following from that specific symmetry. The number of physical parameters in the potential may in general, within each one of Ivanov's classes, be further reduced by choosing a specific basis for the scalar doublets, much in the same way as the general 2HDM potential has 14 parameters which may, however, be reduced to 11 through a suitable basis choice; the number of physical parameters for each class is given in the last column of Table 5.

The six specific symmetries given as examples in Table 5 are the following:

- $U(2)$ is the strongest (most general) HF symmetry,

$$S = \begin{pmatrix} e^{-i\xi} \cos\theta & e^{-i\psi} \sin\theta \\ -e^{i\psi} \sin\theta & e^{i\xi} \cos\theta \end{pmatrix}, \tag{150}$$

where ξ, θ, and ψ are arbitrary.

---

[15] Notice that this is *not* the situation considered in Eqs. (110)–(112). There, the coefficients of the Lagrangian *do change* under the transformation. In contrast, Eqs. (143)–(145) imply the existence of a HF symmetry of the scalar potential because the coefficients of V_H are *unchanged*.

Table 5
The six classes of symmetries (I–VI) of the scalar potential and a practical example of a symmetry in each of those classes. The number in the last column is the minimal number of parameters (n) in the scalar potential, obtainable in a specific basis.

Class	Symmetry	m_{11}^2	m_{22}^2	m_{12}^2	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	n
I	$U(2)$		m_{11}^2	0		λ_1		$\lambda_1 - \lambda_3$	0	0	0	3
II	CP3		m_{11}^2	0		λ_1			$\lambda_1 - \lambda_3 - \lambda_4$	0	0	4
III	CP2		m_{11}^2	0		λ_1					$-\lambda_6$	5
IV	$U(1)$			0					0	0	0	6
V	Z_2			0						0	0	7
VI	CP1			real					real	real	real	8

- CP3 is a GCP symmetry with

$$X = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{151}$$

  the first-quadrant angle θ being generic, *i.e.* different from the two specific values 0 and $\pi/2$.
- CP2 is the GCP symmetry of Eq. (151) but with $\theta = \pi/2$, *i.e.*

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{152}$$

- $U(1)$ is a restricted version of the HF symmetry of Eq. (150) with

$$S = \begin{pmatrix} e^{-i\xi} & 0 \\ 0 & e^{i\xi} \end{pmatrix}, \tag{153}$$

  where ξ is arbitrary.[16]
- Z_2 is the symmetry under $\phi_2 \to -\phi_2$,

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{154}$$

- CP1 is the standard CP symmetry, with

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{155}$$

Some of these models may be further simplified by choosing an appropriate basis for the Higgs doublets:

- A specific basis may be chosen in the CP2 model [359] such that λ_5 is real and $\lambda_6 = \lambda_7 = 0$, hence that model only has five parameters.
- In the $U(1)$ model one may render λ_5 real through a rephasing of the doublets, so that the model only has seven parameters.
- One may perform a real rotation of the doublets in the CP1 model such that m_{12}^2 becomes zero [9], hence that model only has nine parameters.

It should be noted that the scalar potential of the minimal supersymmetric Standard Model (MSSM) does *not* fall into any of Ivanov's symmetry classes. In fact, in the tree-level scalar potential of the MSSM one has [47,373,403]

$$\lambda_1 = \lambda_2 = \frac{g^2 + g'^2}{4}, \qquad \lambda_3 = \frac{g^2 - g'^2}{4}, \qquad \lambda_4 = -\frac{g^2}{2}, \qquad \lambda_5 = \lambda_6 = \lambda_7 = 0, \tag{156}$$

where g and g' are the gauge coupling constants of $SU(2)$ and $U(1)$, respectively. This is much similar to the $U(2)$-symmetric 2HDM, with the crucial difference that the latter has $\lambda_3 + \lambda_4 = \lambda_1$, while in the MSSM $\lambda_3 + \lambda_4 = -\lambda_1$. It is often stated that the scalar sector of the MSSM is a particular case of a 2HDM, but that statement is potentially misleading: the relations among the quartic couplings in Eq. (156) are renormalization-group (RG) invariant in the MSSM, but only due to the presence of extra particles—namely the gauginos; analogous relations among the couplings of a 2HDM are *not* RG-protected if the 2HDM is not supersymmetrized [392,403]. An RG analysis of the relations between couplings shown in Table 5 was undertaken in [393,394]. The list of possible potential symmetries increases when one considers field transformations which do *not* leave the gauge-kinetic terms invariant, as was shown recently by Battye et al. [395] and by Pilaftsis [396]; of course, those extra symmetries lead to relations among the parameters of the potential which are not RG-invariant. See also the discussion on custodial symmetry in Appendix F.

The remainder of this section is devoted to a careful explanation of Table 5, following the presentation in [386].

---

[16] See, though, the discussion after Eq. (162).

5.6.1. Higgs family symmetries

Higgs Family symmetries have a long history in the 2HDM. Glashow and Weinberg [27] and, separately, Paschos [28] have introduced the discrete Z_2 symmetry

$$Z_2: \quad \Phi_1 \rightarrow \Phi_1, \quad \Phi_2 \rightarrow -\Phi_2, \tag{157}$$

and extended it to the quark sector in order to avoid flavour-changing neutral currents. This symmetry enforces $m_{12}^2 = 0$ and $\lambda_6 = \lambda_7 = 0$.

We may consider the Z_2 symmetry in a different scalar basis,

$$\begin{aligned} \Phi_1' &= 2^{-1/2} \left(\Phi_1 + \Phi_2 \right), \\ \Phi_2' &= 2^{-1/2} \left(\Phi_1 - \Phi_2 \right) \end{aligned} \tag{158}$$

obtaining the interchange symmetry

$$\Pi_2: \quad \Phi_1' \leftrightarrow \Phi_2'. \tag{159}$$

This is equivalent to applying Eq. (148) in the form

$$\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{160}$$

The Π_2 symmetry enforces $m_{22}^2 = m_{11}^2$, $\text{Im}\left(m_{12}^2\right) = 0$, $\lambda_2 = \lambda_1$, $\lambda_7 = \lambda_6^*$, and $\text{Im}\left(\lambda_5\right) = 0$. Thus,

- the constraints obtained by applying Z_2 are apparently different from those obtained by applying Π_2;
- however, the two symmetries are equivalent, since applying Z_2 in a given basis is the same as applying Π_2 in a basis obtained from the first one through the transformation (158);
- the Z_2-symmetric and Π_2-symmetric potentials must lead to exactly the same physical predictions – we say that they are in the same class – because physical observables cannot depend on the basis in which we choose to write the Higgs doublets.

Eq. (148) constitutes a conjugacy relation within the group $U(2)$. Thus, HF symmetries associated with matrices S and S' which are in the same conjugacy class of $U(2)$ correspond to the same model. Moreover, symmetries S and S' related by an overall phase transformation ($S' = e^{i\xi}S$) also lead to the same physics, since that overall phase transformation does not affect the bilinears $\Phi_a^\dagger \Phi_b$.

Ferreira and Silva [404] have shown that there are only two classes of HF symmetries generated by one single generator in the scalar potential of the 2HDM: the discrete Z_2 symmetry and a continuous $U(1)$ symmetry

$$\Phi_1 \rightarrow e^{-i\theta} \Phi_1, \qquad \Phi_2 \rightarrow e^{i\theta} \Phi_2, \tag{161}$$

for an arbitrary θ. This $U(1)$ symmetry (suitably extended to the quark sector) was first introduced by Peccei and Quinn [12] in connection with the strong-CP problem. The Higgs potential invariant under $U(1)$ has $m_{12}^2 = 0$ and $\lambda_5 = \lambda_6 = \lambda_7 = 0$ and is therefore also invariant under Z_2.

It is important to note that, for instance, a potential invariant under

$$S_{2/3} = \begin{pmatrix} e^{-i2\pi/3} & 0 \\ 0 & e^{i2\pi/3} \end{pmatrix} \tag{162}$$

is automatically invariant under the full Peccei–Quinn $U(1)$ group. Even though we only want to enforce a symmetry group $Z_3 = \{S_{2/3}, S_{2/3}^2, S_{2/3}^3 = 1\}$, we automatically obtain a potential with full $U(1)$ symmetry. In fact, invariance under any Z_n group, with $n > 2$, will lead us to a $U(1)$-invariant potential. Another possibility of obtaining the same result is to choose an irrational multiple of π for the angle θ in Eq. (161). This is an important point because continuous symmetries, when broken, may lead to massless scalars (Goldstone bosons). An innocent-looking discrete symmetry may have the same effect on the scalar potential as a continuous symmetry and therefrom arises the possibility of undesired massless scalars.

We must however point out two caveats to the discussion in the preceding paragraph. The first caveat is that we are assuming a renormalizable theory, from which we exclude all terms in the potential with dimension larger than four. If, however, we take the reasonable view that the 2HDM is just the low-energy limit of a larger theory, and decide to include effective operators of dimensions five, six, or above, then the equivalence between different symmetries (such as the Z_n with $n > 2$, all of them leading to the same $U(1)$-invariant scalar potential) might no longer be verified. The second caveat pertains to the fermionic sector: given a specific symmetry of the scalar sector, there are in general many ways of extending that symmetry to the fermion sector, often with completely different effects on the Yukawa terms. We shall return to this issue in more detail in Section 5.12.

One may also impose a symmetry with multiple generators on the scalar potential. For example, the scalar potential invariant under both Z_2 and Π_2 *in the same basis* has $m_{11}^2 = m_{22}^2$, $m_{12}^2 = 0$, $\lambda_1 = \lambda_2$, and $\lambda_6 = \lambda_7 = \text{Im}\left(\lambda_5\right) = 0$. Thus,

the potential invariant under $Z_2 \times \Pi_2$ only has five parameters (m_{11}^2, λ_1, λ_3, λ_4, and Re (λ_5)) and, indeed, one may show [359,386] that it is equivalent, in a different basis, to a potential with CP2 symmetry:

$$Z_2 \times \Pi_2 \Leftrightarrow \text{CP2} \Leftrightarrow \text{Class III}. \tag{163}$$

The $U(2)$-invariant potential may similarly be obtained [386] through the imposition of the CP3 and $U(1)$ symmetries in the same basis, as can easily be seen in Table 5.

One can also prove [397] that the existence of either the Z_2 (or, equivalently, Π_2), U(1), or $U(2)$ symmetries is sufficient to guarantee the existence of a basis choice in which all the parameters of the scalar potential are real. That is, the corresponding scalar Higgs sectors are explicitly CP-conserving. Therefore, all models belonging to the classes in Table 5 have CP-conserving scalar potentials.

5.6.2. CP symmetries

We now want to discuss the potentials obtained by imposing one single GCP symmetry. Writing X as in Eq. (139), Ferreira et al. [386] have shown that there are only three classes of potentials obtainable by imposing a single GCP symmetry: CP1 (class VI), CP2 (class III), and CP3 (class II). The potential CP1 results from applying the GCP symmetry with the matrix X in Eq. (139) with $\theta = 0$:

$$\Phi_1 \to \Phi_1^*, \qquad \Phi_2 \to \Phi_2^*; \tag{164}$$

this is the standard CP symmetry, which forces all coefficients in the potential to be real. The potential CP2 results from applying the GCP symmetry with the matrix X in Eq. (139) with $\theta = \pi/2$:

$$\Phi_1 \to \Phi_2^*, \qquad \Phi_2 \to -\Phi_1^*. \tag{165}$$

The potential CP3 results from applying the GCP symmetry with any other (arbitrary) angle $\theta \neq 0, \pi/2$. The theories with symmetry CP2 and CP3 are (of course) CP conserving, but they have potentials more restrictive than CP1.

In the CP3 symmetry, any single angle θ different from 0 or $\pi/2$ in Eq. (139) leads to the same potential. However, if one wants to extend the CP symmetry to the Yukawa sector, different values of θ will have different consequences for the quark masses−only $\theta = \pi/3$ allows for six massive quarks [391].

As mentioned regarding Eq. (163), we may reach class III of 2HDM scalar potentials either by requiring symmetry under the GCP transformation of Eq. (152) or, alternatively, by requiring joint symmetry under Z_2 and Π_2. There are, indeed, many other ways to obtain the class III scalar potential. Similarly [386]

$$U(1) \times \Pi_2 \equiv \text{CP3} \equiv \text{Class IV}. \tag{166}$$

In general, there are many possible symmetries leading into any of the six classes of symmetry-constrained 2HDM potentials. The different symmetries are equivalent with respect to the scalar potential, but they may differ when one tries to extend them to the Yukawa sector. Ferreira et al. [386] have proved that (i) except for the class VI potential, all other five classes of potentials can be obtained through multiple applications of HF symmetries; (ii) one can obtain all six classes of scalar potentials through multiple applications of the standard CP symmetry in different bases. An interesting geometric interpretation of these properties is presented in [398].

5.6.3. The square of the GCP transformation

Applying the GCP transformation twice to the scalar fields,

$$\left(\Phi_a^{\text{GCP}}\right)^{\text{GCP}} = \sum_{b=1}^{2} X_{ab} \left(\Phi_b^{\text{GCP}}\right)^* = \sum_{b,c=1}^{2} X_{ab} X_{bc}^* \Phi_c, \tag{167}$$

one obtains a HF symmetry with $S = XX^*$. Thus, (GCP)^2 provides a distinction among the three GCP symmetries:

$$\begin{aligned} (CP1)^2 &= \mathbf{1}, \\ (CP2)^2 &= -\mathbf{1}, \\ (CP3)^2 &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}. \end{aligned} \tag{168}$$

While (CP2)^2 is reduced to the identity transformation through a global hypercharge transformation [375], (CP3)^2 is a non-trivial HF symmetry of the class II scalar potential [386].

5.6.4. Symmetries and bilinears

Ivanov's description of the possible classes of scalar potentials in the 2HDM is most conveniently summarized by looking at the corresponding vectors and tensor of Eq. (116) and (117). We start by looking at the implications of the symmetries we have studied so far on the vector $\vec{r} = \{r_1, r_2, r_3\}$, whose components were introduced in Eq. (105). Notice that a unitary transformation U on the fields Φ_a induces an orthogonal transformation R on the vector of bilinears $\vec{r}$, given by Eq. (114). For every pair of unitary transformations $\pm U$ of $SU(2)$, one can find some corresponding transformation R of $SO(3)$, in a two-to-one correspondence. We then see what these symmetries imply for the coefficients of Eq. (103) (recall the $\Lambda_{\mu\nu}$ is a symmetric matrix). Below, we list the transformation of $\vec{r}$ under which the scalar potential is invariant, followed by the corresponding constraints on the quadratic and quartic scalar potential parameters, M_μ and $\Lambda_{\mu\nu}$.

- The class I symmetry implies

$$\vec{r} \to R\vec{r}, \qquad \begin{bmatrix} M_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & 0 & 0 & 0 \\ 0 & \Lambda_{11} & 0 & 0 \\ 0 & 0 & \Lambda_{11} & 0 \\ 0 & 0 & 0 & \Lambda_{11} \end{bmatrix}, \tag{169}$$

where R is an arbitrary 3×3 orthogonal matrix of unit determinant. Thus the vectors M_i and Λ_{0i} vanish while the tensor Λ_{ij} is proportional to the unit matrix. It is evident that class I has only three parameters in the scalar potential.

- For class II, the vectors M_i and Λ_{0i} once again vanish while the tensor Λ_{ij} has two – instead of three, as in class I – degenerate eigenvalues. For instance, if one applies the GCP transformation of Eq. (151)

$$\vec{r} \to \begin{bmatrix} c_2 & 0 & s_2 \\ 0 & -1 & 0 \\ -s_2 & 0 & c_2 \end{bmatrix} \vec{r}, \qquad \begin{bmatrix} M_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & 0 & 0 & 0 \\ 0 & \Lambda_{11} & 0 & 0 \\ 0 & 0 & \Lambda_{22} & 0 \\ 0 & 0 & 0 & \Lambda_{11} \end{bmatrix}. \tag{170}$$

It is evident that for this symmetry class there are four parameters in the scalar potential.

- In class III, the vectors M_i and Λ_{0i} vanish and the eigenvalues of the tensor Λ_{ij} are all different. For instance, if one applies the GCP transformation of Eq. (152)

$$\vec{r} \to \begin{bmatrix} -r_1 \\ -r_2 \\ -r_3 \end{bmatrix}, \qquad \begin{bmatrix} M_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & 0 & 0 & 0 \\ 0 & \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ 0 & \Lambda_{12} & \Lambda_{22} & \Lambda_{23} \\ 0 & \Lambda_{13} & \Lambda_{23} & \Lambda_{33} \end{bmatrix}, \tag{171}$$

but the 3×3 real symmetric matrix

$$\begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{12} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{13} & \Lambda_{23} & \Lambda_{33} \end{pmatrix}$$

may be diagonalized and has in general three distinct eigenvalues. Therefore, in this symmetry class the scalar potential has five parameters.

- For class IV, the vectors M_i and Λ_{0i} are parallel and, moreover, in the subspace orthogonal to those vectors the tensor Λ_{ij} has degenerate eigenvalues. For instance, with the $U(1)$ symmetry of Eq. (153), we have

$$\vec{r} \to \begin{bmatrix} c_2 & -s_2 & 0 \\ s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{r}, \qquad \begin{bmatrix} M_0 \\ 0 \\ 0 \\ M_3 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & 0 & 0 & \Lambda_{03} \\ 0 & \Lambda_{11} & 0 & 0 \\ 0 & 0 & \Lambda_{11} & 0 \\ \Lambda_{03} & 0 & 0 & \Lambda_{33} \end{bmatrix}. \tag{172}$$

It is evident that there are six independent parameters in this scalar potential.

- For class V, the vectors M_i and Λ_{0i} are parallel and the tensor Λ_{ij} has three non-degenerate eigenvalues. For instance, with the Z_2 symmetry of Eq. (154),

$$\vec{r} \to \begin{bmatrix} -r_1 \\ -r_2 \\ r_3 \end{bmatrix}, \qquad \begin{bmatrix} M_0 \\ 0 \\ 0 \\ M_3 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & 0 & 0 & \Lambda_{03} \\ 0 & \Lambda_{11} & \Lambda_{12} & 0 \\ 0 & \Lambda_{12} & \Lambda_{22} & 0 \\ \Lambda_{03} & 0 & 0 & \Lambda_{33} \end{bmatrix}, \tag{173}$$

but the 2×2 real symmetric matrix

$$\begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12} & \Lambda_{22} \end{pmatrix}$$

may be diagonalized and has in general distinct eigenvalues. Therefore, the class V scalar potential has seven parameters.

- For class VI, the vectors M_i and Λ_{0i} lie on the same plane and the tensor Λ has one eigenvector orthogonal to that plane. For instance, with the CP1 symmetry of Eq. (155),

$$\vec{r} \rightarrow \begin{bmatrix} r_1 \\ -r_2 \\ r_3 \end{bmatrix}, \qquad \begin{bmatrix} M_0 \\ M_1 \\ 0 \\ M_3 \end{bmatrix}, \qquad \begin{bmatrix} \Lambda_{00} & \Lambda_{01} & 0 & \Lambda_{03} \\ \Lambda_{01} & \Lambda_{11} & 0 & \Lambda_{13} \\ 0 & 0 & \Lambda_{22} & 0 \\ \Lambda_{03} & \Lambda_{13} & 0 & \Lambda_{33} \end{bmatrix}, \qquad (174)$$

but the 2×2 real symmetric matrix

$$\begin{pmatrix} \Lambda_{11} & \Lambda_{13} \\ \Lambda_{13} & \Lambda_{33} \end{pmatrix}$$

may be diagonalized and has in general distinct eigenvalues. Therefore, in this symmetry class the scalar potential has nine parameters.

Since each unitary transformation of the fields Φ_a induces an $SO(3)$ transformation on the vector of bilinears $\vec{r}$, and since the standard CP transformation corresponds to an inversion of r_2, *i.e.* to a Z_2 transformation on the vector $\vec{r}$, Ivanov [380] actually considered all possible $O(3)$ transformations acting on $\vec{r}$. He has identified the following six classes of transformations: (i) Z_2; (ii) $(Z_2)^2$; (iii) $(Z_2)^3$; (iv) $O(2)$; (v) $O(2) \otimes Z_2$; and (vi) full $O(3)$. No other independent symmetry transformations are possible.

5.7. Bounded from below limits

Stability of the 2HDM potential requires that it be bounded from below, *i.e.* that there is no direction in field space along which the potential tends to minus infinity. This is a basic requirement for any physical theory – the existence of a stable minimum, around which one can perform perturbative calculations – which is satisfied by the scalar potential of the SM through the trivial condition $\lambda > 0$, where λ is the quartic coupling of the SM scalar potential.

The 2HDM scalar potential of Eq. (98) is much more complicated than the SM's, and ensuring its stability requires that one studies all possible directions along which the fields Φ_1 and Φ_2 (or rather, their respective eight component fields) tend to arbitrarily large values. In general, the existence of a non-trivial minimum – by which we mean the fields Φ_i acquiring non-zero vacuum expectation values – implies two conditions on the potential's parameters. They have to be such that: the quartic part of the scalar potential, V_4, is positive for arbitrarily large values of the component fields, but the quadratic part of the scalar potential, V_2, can take negative values for at least some values of the fields.

The restrictions on V_4 need to be handled carefully:

- Demanding that $V_4 > 0$ for all $\Phi_i \rightarrow \infty$ is a *strong stability* requirement. This may, however, be too strong, since several interesting models are excluded by it. For instance, in tree-level SUSY potentials there is a direction ($\langle \Phi_1 \rangle = \langle \Phi_2 \rangle$) for which $V_4 = 0$.
- We can also demand stability in a *marginal* sense, by requiring $V_4 \geq 0$, for any direction in field space tending to infinity.
- The equality in the marginal stability bound comes at a price: if there is a given direction in field space such that $V_4 \rightarrow 0$, it is necessary to demand that, along *that* specific direction, one has $V_2 \geq 0$.

A simple way to obtain *necessary* conditions on the quartic parameters of the potential is to study its behaviour along specific field directions. Considering for instance the direction $|\Phi_1| \rightarrow \infty$ and $|\Phi_2| = 0$, the expression (98) for the potential renders it obvious that one can have positive values for V_4 if and only if $\lambda_1 \geq 0$. Likewise, the direction $|\Phi_1| = 0$ and $|\Phi_2| \rightarrow \infty$ gives us the condition $\lambda_2 \geq 0$. And if one takes $|\Phi_1|^2 = r \cos \theta$, $|\Phi_2|^2 = r \sin \theta$ (with $0 < \theta < \pi/2$ and $r \rightarrow +\infty$) but such that $\Phi_1^\dagger \Phi_2 = 0$ (for instance considering only non-zero upper components for Φ_1 and lower ones for Φ_2), the bounded-from-below condition becomes

$$\lim_{r \rightarrow +\infty} r^2 \left(\frac{\lambda_1}{2} \cos^2 \theta + \frac{\lambda_2}{2} \sin^2 \theta + \lambda_3 \sin \theta \cos \theta \right) = r^2 f(\theta) \geq 0 , \ \forall_\theta. \qquad (175)$$

Minimizing $f(\theta)$ with respect to θ to obtain its smallest value and demanding that it be larger or equal to zero, one finds that the coefficients λ_i need to obey $\lambda_3 \geq -\sqrt{\lambda_1 \lambda_2}$. By studying several such directions, it is possible to reach other conditions on the couplings, and we can gather all as

$$\lambda_1 \geq 0, \qquad \lambda_2 \geq 0,$$
$$\lambda_3 \geq -\sqrt{\lambda_1 \lambda_2}, \qquad \lambda_3 + \lambda_4 - |\lambda_5| \geq -\sqrt{\lambda_1 \lambda_2} , \qquad (176)$$

where we have taken λ_5 to be real. In [373,406] it was proven that, in potentials where one has $\lambda_6 = \lambda_7 = 0$, these are actually necessary and sufficient conditions to ensure the positivity of the quartic potential along all directions. As we can see from the discussion in Section 5.6, most of the possible symmetry-constrained 2HDM scalar potentials fall unto this category—there is at least a basis where $\lambda_6 = \lambda_7 = 0$ holds, for the Z_2, $U(1)$, CP2, CP3 and $U(2)$ models. For the remaining possibilities – a model with CP1 symmetry, or a model with no symmetry at all (other than the gauge ones), one can find

necessary and sufficient conditions for boundedness, involving λ_6 and λ_7 in the work of [373] – unfortunately, they do not have a simple analytical expression, like those of Eq. (176), but they can be handled numerically. In [373], all possible cases – strong stability, marginal stability, analysis of the quadratic terms – were considered. Equivalent conditions, for the strong stability requirements, were found in [380]. Again, they are not easily translated into analytical bounds. It is not difficult, though, to find *necessary* conditions involving λ_6 and λ_7. For the case they are real [405,407], one finds

$$2\,|\lambda_6 + \lambda_7| \; < \; \frac{\lambda_1 + \lambda_2}{2} + \lambda_3 + \lambda_4 + \lambda_5. \tag{177}$$

As mentioned earlier, the requirement of strong stability is too constraining, in the sense that it actually excludes potentially interesting models. The obvious example is the SUSY potential, for which the quartic couplings are related to the gauge coupling constants, and they are given by Eq. (156). With these couplings, the last bound in Eq. (176) is saturated. Likewise, in the CP3 model, for which $\lambda_5 = \lambda_3 + \lambda_4 - \lambda_1$ and $\lambda_1 = \lambda_2$, the last bound in Eq. (176) gives us $\lambda_1 + \lambda_3 + \lambda_4 \geq 0$ and $\lambda_1 \geq \lambda_1$, the latter clearly demanding the equality, lest the model be excluded.

For most cases, the conditions (176) are all we need. They become specially important if one wishes to analyse the stability of the potential including higher order corrections. Clearly they were obtained through tree-level analysis, requiring that the tree-level potential always be convex. But one may wonder whether those conditions ensure that the one-loop corrected effective potential is also bounded from below. The one-loop corrections to the tree-level potential are of the form

$$V(\phi_i) = V_{tree}(\phi_i) + \frac{1}{64\pi^2} \sum_\alpha m_\alpha^4(\phi_i) \left[\log\left(\frac{m_\alpha^2(\phi_i)}{\mu^2} \right) - \frac{3}{2} \right], \tag{178}$$

where the sum runs over all helicity states of the particles of masses m_α present in the theory, and μ is the renormalization scale. We have kept the dependence on the fields ϕ_i explicit. Unbounded from below limits are found analysing the behaviour of the potential for very large values of the fields ϕ_i, at which point one should worry about the appearance of potentially large contributions from the logarithms in the expression above. A renormalization group (RG) improvement of the bounds then amounts to considering only the tree-level expressions we have already discussed, but considering the values of the couplings which appear in those expressions at different renormalization scales. In other words, we take the bounds from Eqs. (176) and run the couplings therein, using the β-functions of the model (see Appendix E), along a range of scales μ— from the weak scale M_Z to an upper scale Λ; at all scales in the interval chosen, the bounds must hold, and in this way combinations of parameters which at one scale might be acceptable would violate the bounds at another scale.

This type of analysis was performed in the SM [408–413]. There, the Higgs potential quartic coupling λ has a β-function with a sizeable negative contribution from the top quark Yukawa. The top being so heavy, this term tends to decrease the value of λ at higher renormalization scales. Thus, if the starting value (at the weak scale, say) of λ is small, the coupling may well become negative at some higher scale, and the potential would suddenly be unbounded from below. In this manner we can thus put a *lower* bound on λ and thus on the Higgs mass. On the other hand, if the starting value of λ is too large, the RG evolution of the coupling will increase its value immensely and eventually the theory becomes non-perturbative. We will return to this when we consider another class of theoretical bounds, in which one requires unitarity of all 2HDM processes involving scalars, which we will look into in Appendix A. Thus, the RG analysis allows us to impose both higher and lower bound on the mass of the Higgs particles.

In the 2HDM the same type of phenomena can occur, and was treated in, for instance, [97,407,409,414–417]. If for instance the Φ_1 is made to couple to the up quarks, the β-function for the λ_1 quartic coupling will have a large negative top Yukawa contribution, and a similar analysis to the SM case will hold. Now, however, many other quartic couplings are present and more bounds need to be obeyed. Nonetheless, the main conclusions hold: smaller values for some of the λ_i at the weak scale are disfavoured as they lead to unbounded from below potentials at higher scales; and large values of those couplings lead to Landau poles at high scales, thus a breakdown of perturbation theory. These translate into bounds on the several Higgs masses. Several observations are in order:

- Clearly the bounds obtained will depend on what the upper renormalization scale Λ is. Usually this is taken to be the gauge unification scale, $\sim 10^{16}$ GeV. Varying this scale will change (mostly) the upper bounds on the masses—the upper bound on the lightest CP-even scalar can change from about 300 to about 100 GeV, if one varies Λ from 10^3 to 10^{16} GeV [97].
- The precise values for the bounds have a noticeable dependence on the value of the top pole mass. This is to expected since the top quark Yukawa is what drives most of the quartic coupling RG evolution. Typically, at most some of the bounds can change by roughly ~ 10 GeV for a 5 GeV change in the top pole mass [97,407].
- Most of the analysis performed considered a 2HDM with an intact Z_2 symmetry. However, as shown in [97], the soft breaking term m_{12} has a crucial importance in the bounds obtained, which are increasingly relaxed the more the magnitude of m_{12} increases. Roughly speaking, with $\Lambda = 10^{16}$ GeV, the exact Z_2 model gives us an upper bound on the lightest CP-even Higgs mass of roughly 100 GeV, which is easily raised to about 185 GeV for the soft broken model [97,407]. For some cases, like the charged Higgs mass, large values of the soft breaking term eliminate the upper bound obtained for the exact symmetry (roughly 160 GeV).

- In [407] other 2HDM theories were considered, namely the inert vacua (see Section 5.8) of the Z_2 potential, the Peccei–Quinn $U(1)$ model and the model without Z_2 symmetry (dubbed the *CP*1 model in Section 5.6. The Peccei–Quinn and CP1 model bounds obtained do not differ significantly from the Z_2 case. The inert model has quite restrictive bounds on the two CP-even scalars.
- In [407] these bounds were also applied to theories in which the vacuum of the theory spontaneously breaks CP (again, see Section 5.8). Such theories always have the m_{12} term, which can be quite large. However, it was found that the lightest Higgs scalar has a very low upper bound, roughly 85 GeV.
- These analyses only take into account the top Yukawa coupling, coupled to only one of doublets. Even for a theory with flavour changing neutral interactions, this can be considered a good approximation.

The general conclusion is that requiring that the potential be bounded from below in the range of scales from M_Z to Λ can severely limit the parameters of the theory, namely the masses of the scalar eigenstates. However, such bounds are heavily model-dependent—they depend on whether the 2HDM considered has a symmetry, or if that symmetry is softly broken, and they depend immensely on what the upper scale Λ is taken to be. Notice, for instance, that we simply do not know what the value of Λ should be; if one thinks of the 2HDM as an effective theory Λ should be taken as the scale above which new fields have to be considered, and it could be as low as 1 TeV. Hence, great care must be exercised when applying such bounds, lest one exclude regions of parameter space which may well be important.

5.8. Spontaneous symmetry breaking

If the scalar potential of the 2HDM is bounded from below, being a quartic polynomial function it will certainly have a global minimum somewhere. This same argument applies to the SM, but there we can only have two types of minima: the "trivial" one, for which the Higgs acquires zero vevs, and the usual one, where electroweak symmetry breaking occurs, away from the origin, for $\langle \Phi \rangle = v/\sqrt{2}$. In particular, vacua which break electric charge or CP conservation are impossible in the SM. In what follows, we consider a vacuum any stationary point of the potential, regardless of whether it is a minimum or not.

In the 2HDM, the vacuum structure is much richer. We can have three types of vacua (other than the trivial case, $\langle \Phi_1 \rangle = \langle \Phi_2 \rangle = 0$):

- "Normal" (N) vacua, with vevs which do not have any complex relative phase and can thus be trivially rendered real:

$$\langle \Phi_1 \rangle_N = \begin{pmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{pmatrix}, \qquad \langle \Phi_2 \rangle_N = \begin{pmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{pmatrix}, \tag{179}$$

where $v = \sqrt{v_1^2 + v_2^2} = 246$ GeV[17] and one defines $\tan \beta = v_2/v_1$. This solution, of course, is the 2HDM equivalent of the SM vacuum. We can distinguish a special case here, in which the minimization conditions allow for one of the vevs v_1, v_2 to be zero. These are called "inert models", already discussed in Section 2.3. Notice that, unlike the passage to the Higgs basis (in which only one of the doublets has a vev, as well), the inert vacua are found in the basis where a Z_2 (or $U(1)$, or other) symmetry is manifest.
- CP breaking vacua, where the vevs do have a relative complex phase, that is

$$\langle \Phi_1 \rangle_{CP} = \begin{pmatrix} 0 \\ \frac{\bar{v}_1}{\sqrt{2}} e^{i\theta} \end{pmatrix}, \qquad \langle \Phi_2 \rangle_{CP} = \begin{pmatrix} 0 \\ \frac{\bar{v}_2}{\sqrt{2}} \end{pmatrix}, \tag{180}$$

with real values for $\bar{v}_1$ and $\bar{v}_2$. The moniker "CP breaking" is not the most appropriate, since such vacua are possible even in potentials where the CP symmetry is not defined (due to it being explicitly broken)—see for instance [418]. Also, the presence of a phase in (180) is not a guarantee of spontaneous CP breaking (see Section 6).
- Charge breaking (CB) vacua, in which one of the vevs carries electric charge,

$$\langle \Phi_1 \rangle_{CB} = \begin{pmatrix} \frac{\alpha}{\sqrt{2}} \\ \frac{v_1'}{\sqrt{2}} \end{pmatrix}, \qquad \langle \Phi_2 \rangle_{CB} = \begin{pmatrix} 0 \\ \frac{v_2'}{\sqrt{2}} \end{pmatrix}, \tag{181}$$

with real numbers v_1', v_2', α. Due to the presence of a non-zero vev in an upper component (charged) of the fields, this vacuum breaks electrical charge conservation, causing the photon to acquire a mass. Thus, they are to be avoided at all costs.

Given our definition for the charge (such that the lower components of fields are neutral), the vacuum will break the charge and lead to a massive photon if and only if

$$\langle \varphi_1^+ \rangle \langle \varphi_2^0 \rangle = \alpha v_2' \neq 0. \tag{182}$$

---

[17] Notice, however, that certain 2HDM potentials can have more than one solution of this type, with different values for v. See the discussion for Eq. (202).

One can give a definition for a charge breaking vacuum which does not depend on our definition of charge or on our basis choice [419],

$$\left| \langle \varphi_1^+ \rangle \langle \varphi_2^0 \rangle \right|^2 + \left| \langle \varphi_1^0 \rangle \langle \varphi_2^+ \rangle \right|^2 - 2 \operatorname{Re} \left(\langle \varphi_1^+ \rangle \langle \varphi_2^0 \rangle, \langle \varphi_1^{0*} \rangle \langle \varphi_2^{+*} \rangle \right) \neq 0. \tag{183}$$

Charge preserving vacua (many times called "aligned vacua") for the 2HDM and for models with additional singlets or triplets were studied in Ref. [420].

That all possible vacua in the 2HDM reduce to one of the three forms of Eqs. (179)–(181) can be seen using the freedom to choose a particular gauge in $SU(2)_L \times U(1)_Y$. Let us write the two doublets in the following simplified manner:

$$\Phi_1 = \begin{pmatrix} |\varphi_1^+| \, e^{i\theta_1^+} \\ |\varphi_1^0| \, e^{i\theta_1^0} \end{pmatrix}, \qquad \Phi_2 = \begin{pmatrix} |\varphi_2^+| \, e^{i\theta_2^+} \\ |\varphi_2^0| \, e^{i\theta_2^0} \end{pmatrix}. \tag{184}$$

The complex phases above will, in general, be functions of the space–time coordinates. Then, the local gauge transformation U_1, given by the $SU(2)_L$ matrix

$$U_1 = \begin{pmatrix} u_{11} & u_{12} \\ -u_{12}^* & u_{11}^* \end{pmatrix} \tag{185}$$

with

$$u_{11} = \frac{|\varphi_2^0|}{\sqrt{|\varphi_2^0|^2 + |\varphi_2^+|^2}}$$

$$u_{12} = -\frac{|\varphi_2^+|}{\sqrt{|\varphi_2^0|^2 + |\varphi_2^+|^2}} \, e^{i(\theta_2^+ - \theta_2^0)}, \tag{186}$$

eliminates the upper components of Φ_2. A combined hypercharge and $SU(2)_L$ gauge transformation can then be used to eliminate the phases of the upper and lower components of Φ_1, through the matrix

$$U_2 = \begin{pmatrix} e^{-i\theta_1^+} & 0 \\ 0 & e^{-i\theta_1^0} \end{pmatrix}. \tag{187}$$

Of course, these phases θ_1^+ and θ_1^0 are not the same that appear in Eq. (184) (they have been changed by the gauge transformation U_1), but that is irrelevant for the argument. The final form of the doublets is thus

$$\Phi_1 = \begin{pmatrix} \operatorname{Re}(\varphi_1^+) \\ \operatorname{Re}(\varphi_1^0) \end{pmatrix}, \qquad \Phi_2 = \begin{pmatrix} 0 \\ \varphi_2^0 \end{pmatrix}. \tag{188}$$

This is almost of the form of Eq. (181), but it has a complex lower component in Φ_2 [409]. However, that phase is physically irrelevant, since it can always be absorbed through a trivial basis transformation—a rephasing of Φ_2, $\Phi_2 \to e^{-i\theta_2^0} \Phi_2$. As such, the form of the vevs in Eq. (181) is indeed the most general one we need: there exists always a basis for which the most generic vacuum will have that form. This conclusion could also have been reached through a series of basis changes, plus a gauge transformation [419]. That method has the advantage of being easily generalized for an N-Higgs doublet model.

Let us now look in more detail at the solutions of the minimization conditions. Writing the potential in terms of the vevs $\tilde{v}_i$ (for any of the three sets (179)–(181)), a stationary point of the potential is found if the set of equations $\partial V / \partial \tilde{v}_i = 0$ has solutions. In terms of the notation introduced in Section 5.2, and for completely general vacuum expectation values such that $\langle \Phi_a \rangle = \tilde{v}_a / \sqrt{2}$,[18] the extremum conditions may be written as

$$\sum_{b=1}^{2} \left[\mu_{ab} + \frac{1}{4} \sum_{c,d=1}^{2} \lambda_{ab,cd} \, \tilde{v}_d^* \tilde{v}_c \right] \tilde{v}_b = 0 \quad \text{(for } a = 1, 2\text{)}. \tag{189}$$

Multiplying by $\tilde{v}_a^*$ leads to

$$\sum_{a,b=1}^{2} \mu_{ab} (\tilde{v}_a^* \tilde{v}_b) = -\frac{1}{4} \sum_{a,b,c,d=1}^{2} \lambda_{ab,cd} \, (\tilde{v}_a^* \tilde{v}_b) \, (\tilde{v}_d^* \tilde{v}_c). \tag{190}$$

---

[18] Notice that we are using $\tilde{v}_a / \sqrt{2}$ for the complex vev $\langle \Phi_a \rangle$, while $v_a = |\tilde{v}_a|$ is real.

If one performs a basis transformation such as the one presented in (110), the vevs are transformed as

$$\tilde{v}_a \to \tilde{v}'_a = \sum_{b=1}^{2} U_{ab} \tilde{v}_b. \tag{191}$$

And, for a GCP transformation like in (134), the vevs transform as

$$\tilde{v}_a \to \tilde{v}_a^{\text{GCP}} = X_{a\alpha} \sum_{\alpha=1}^{2} \tilde{v}_\alpha^*. \tag{192}$$

The different CB and CP stationary points are determined by a set of three equations, a normal one by only two. In fact, since the 2HDM potential depends on eight real component fields, any stationary point ought to be the solution of a set of eight equations on eight unknowns which arise from (189). However, given that one can always choose the simplified forms of the vevs in Eqs. (179)–(181), most of those equations are trivially satisfied.

As was shown in [405] (and later in [373,377,380]), the CB vevs can always be obtained analytically, and are given by

$$\begin{pmatrix} v'^2_1 + \alpha^2 \\ v'^2_2 \\ v'_1 v'_2 \end{pmatrix} = 2 \begin{pmatrix} \lambda_1 & \lambda_3 & 2\text{Re}(\lambda_6) \\ \lambda_3 & \lambda_2 & 2\text{Re}(\lambda_7) \\ 2\text{Re}(\lambda_6) & 2\text{Re}(\lambda_7) & 2(\lambda_4 + \text{Re}(\lambda_5)) \end{pmatrix}^{-1} \begin{pmatrix} m_{11}^2 \\ m_{22}^2 \\ -2\text{Re}(m_{12}^2) \end{pmatrix}. \tag{193}$$

This expression has an important consequence: if Eq. (193) admits a solution, it is unique, up to trivial sign changes ($\alpha \to -\alpha$, $v'_1 \to -v'_1$ and $v'_2 \to -v'_2$) with no physical impact. Charge breaking is in fact impossible in several symmetry-constrained 2HDM.

Likewise, the CP vacua vevs can always be obtained analytically in terms of the potential's parameters. Restricting ourselves to potentials where the CP symmetry is defined – *i.e.*, where it is not explicitly broken, see Section 6 – we obtain

$$\begin{pmatrix} \bar{v}_1^2 \\ \bar{v}_2^2 \\ \bar{v}_1 \bar{v}_2 \cos\theta \end{pmatrix} = 2 \begin{pmatrix} \lambda_1 & \lambda_3 + \lambda_4 - \text{Re}(\lambda_5) & 2\text{Re}(\lambda_6) \\ \lambda_3 + \lambda_4 - \text{Re}(\lambda_5) & \lambda_2 & 2\text{Re}(\lambda_7) \\ 2\text{Re}(\lambda_6) & 2\text{Re}(\lambda_7) & 4\text{Re}(\lambda_5) \end{pmatrix}^{-1} \begin{pmatrix} m_{11}^2 \\ m_{22}^2 \\ -2\text{Re}(m_{12}^2) \end{pmatrix}. \tag{194}$$

Again, up to physically irrelevant sign changes, the CP vacuum is *unique*.

The normal vacuum turns out to be the most difficult to solve. In fact, for many potentials, the minimization conditions cannot be solved analytically. The equations $\partial V/\partial v_1 = 0$ and $\partial V/\partial v_2 = 0$ give, for the most general 2HDM potential,

$$m_{11}^2 v_1 - \text{Re}(m_{12}^2) v_2 + \frac{\lambda_1}{2} v_1^3 + \frac{\lambda_{345}}{2} v_1 v_2^2 + \frac{1}{2} \left[3\text{Re}(\lambda_6) v_1^2 v_2 + \text{Re}(\lambda_7) v_2^3 \right] = 0$$

$$m_{22}^2 v_2 - \text{Re}(m_{12}^2) v_1 + \frac{\lambda_2}{2} v_2^3 + \frac{\lambda_{345}}{2} v_2 v_1^2 + \frac{1}{2} \left[\text{Re}(\lambda_6) v_1^3 + 3\text{Re}(\lambda_7) v_2 v_1^2 \right] = 0, \tag{195}$$

where $\lambda_{345} = \lambda_3 + \lambda_4 + \text{Re}(\lambda_5)$. A few important observations about these equations:

- For some models (unbroken Z_2, $U(1)$, CP2, CP3 and $U(2)$) these equations can be solved analytically. However, the presence of soft breaking terms may prevent that.
- For any models in which $m_{12}^2 = \lambda_6 = \lambda_7 = 0$ (for instance, models with a Z_2 symmetry or higher), Eqs. (195) may admit solutions of the form $v_1 \neq 0$, $v_2 = 0$ and $v_1 = 0$, $v_2 \neq 0$. In the first case we have

$$v_1^2 = -\frac{2m_{11}^2}{\lambda_1}, \tag{196}$$

in the second

$$v_2^2 = -\frac{2m_{22}^2}{\lambda_2}, \tag{197}$$

as long that, of course, $m_{11}^2 < 0$ or $m_{22}^2 < 0$. These lead to the so-called Inert models [160,162]. The name derives from the fact that, unlike solutions of Eq. (195) which have $v_1 \neq 0$, $v_2 \neq 0$, these vacua lead to scalar particles which do not couple to gauge bosons, and can easily be made to decouple from fermions. As such, these models provide excellent candidates for dark matter.

- The Eqs. (195) do not, in general, have a *unique* solution. Even when they do not admit inert vacua they can lead, depending on the values of the parameters, to several sets of vevs $\{v_1, v_2\}$ which are not related by trivial sign changes [418]. In [421] it was however proven that there can be no more than *two* such solutions which are simultaneously minima of the potential.

5.9. Vacuum stability

In the SM there is only *one* possible type of vacuum, other than the trivial one. Indeed, in that theory the scalar potential is such that one can only have one minimum. In theories with more than one scalar, however, there is the possibility that minima of different natures occur, and thus that the theory may allow for tunnelling from one minimum to another. An example of this behaviour occurs in SUSY models, where the existence of many charged and/or coloured scalar fields gives rise to possible charge and/or colour breaking minima [422]. These minima would imply massive photons and/or gluons, and as such one wishes to avoid them. Hence, it is desirable to impose bounds on the theory's parameters to ensure that the *global* minimum preserves the SM's symmetries.

Colour breaking is impossible in the 2HDM, but we have already seen, in Section 5.8, that charged vevs are possible. Also, one may have CP breaking vevs. Thus, the question arises: can these vacua of different natures coexist with one another? Could one tunnel, for instance, from a normal minimum to a deeper charge-breaking one? In other words, given a minimum in the 2HDM, is it stable? The limited number of scalars, and the inexistence of cubic terms in the potential, in the 2HDM allows us to treat this question in a fully analytical way. It has been possible to show that [405,418,423]:

- For a potential where a normal stationary point and a charge breaking one exist, with vevs as given by Eqs. (195) and (193), the difference in the values of the scalar potential at both those vacua (respectively V_N and V_{CB}) is given by

$$V_{CB} - V_N = \left(\frac{M_{H^\pm}^2}{4v^2}\right)_N \left[(v_1' v_2 - v_2' v_1)^2 + \alpha^2 v_2^2\right], \tag{198}$$

where $(M_{H^\pm}^2/4v^2)_N$ is the ratio of the squared mass of the charged scalar to the sum of the square of vevs, $v^2 = v_1^2 + v_2^2$, as computed in the normal stationary point.

The significance of (198) is plain: *if* the normal stationary point is a *minimum* (which implies that $M_{H^\pm}^2 > 0$) then one will necessarily have $V_{CB} - V_N > 0$. That is, if there is a normal minimum, any CB stationary point will lie *above* it—the normal minimum is stable against charge breaking. In [405] it was also proven that in that case the CB stationary point is necessarily a saddle point. Thus, normal and CB minima cannot coexist in the 2HDM. Of course, it is possible to choose sets of parameters of the potential such that the global minimum of the potential breaks charge—but in that case no normal minima will exist.

- For a potential where a normal stationary point and a CP breaking one exist, with vevs as given by Eqs. (195) and (194), the difference in the values of the scalar potential at both those vacua (respectively V_N and V_{CP}) is given by

$$V_{CP} - V_N = \left(\frac{M_A^2}{4v^2}\right)_N \left[(\bar{v}_1 v_2 \cos\theta - \bar{v}_2 v_1)^2 + \bar{v}_1^2 v_2^2 \sin^2\theta\right], \tag{199}$$

where $(M_A^2/4v^2)_N$ is the ratio of the squared mass of the pseudoscalar to the sum of the square of vevs, $v^2 = v_1^2 + v_2^2$, as computed in the normal stationary point.

The significance of (199) is plain: *if* the normal stationary point is a *minimum* (which implies that $M_A^2 > 0$) then one will necessarily have $V_{CP} - V_N > 0$. That is, if there is a normal minimum, any CP stationary point will lie *above* it—the normal minimum is stable against CP breaking. In [421] it was also proven that in that case the CP stationary point is necessarily a saddle point. Thus, normal and CP minima cannot coexist in the 2HDM. Of course, it is possible to choose sets of parameters of the potential such that the global minimum of the potential breaks CP—but in that case no normal minima will exist.

- No CB and CP minima can coexist either. This derives from the fact that for the CP vacuum the square of the charged Higgs mass is given by

$$(M_{H^\pm}^2)_{CP} = -\frac{1}{2}[\lambda_4 - \text{Re}(\lambda_5)](\bar{v}_1^2 + \bar{v}_2^2), \tag{200}$$

whereas in a CB vacuum one of the squared mass matrix eigenvalues is

$$M_{CB}^2 = \frac{1}{2}[\lambda_4 - \text{Re}(\lambda_5)](v_1'^2 + v_2'^2 + \alpha^2). \tag{201}$$

As we see, the sign of $\lambda_4 - \text{Re}(\lambda_5)$ determines that both these vacua cannot be simultaneously minima. Thus, if a CP minimum exists the (unique) CB stationary point, if it exists, *cannot* be a minimum as well, and vice-versa.

- Unlike the CB and CP cases, the normal minimization conditions allow for multiple solutions, so that one can have an N_1 vacuum with vevs $\{v_{1,1}, v_{2,1}\}$ and an N_2 vacuum with different vevs $\{v_{1,2}, v_{2,2}\}$. In that case, the difference in the values of the potential in those two vacua (respectively, V_{N_1} and V_{N_2}) is given by

$$V_{N_2} - V_{N_1} = \frac{1}{4}\left[\left(\frac{M_{H^\pm}^2}{v^2}\right)_{N_1} - \left(\frac{M_{H^\pm}^2}{v^2}\right)_{N_2}\right](v_{1,1} v_{2,2} - v_{2,1} v_{1,2})^2, \tag{202}$$

where $(M_{H^\pm}^2/v^2)_{N_1}$ is the ratio of the squared mass of the charged scalar to the sum of the square of vevs, $(v^2)_{N_1} = v_{1,1}{}^2 + v_{2,1}{}^2$, as computed in the N_1 stationary point, and analogously for $(M_{H^\pm}^2/v^2)_{N_2}$.

Eq. (202) shows us that there is nothing favouring N_1 over N_2, the deepest stationary point will be determined by the values of the parameters—as it should be, since both vacua have the same symmetries.

In Ref. [421] it was proven that it is possible to have two coexisting normal minima. Numerical examples of this were found in [418] for the particular case of a softly broken $U(1)$ model, where it was shown that one may have a curious situation: a minimum N_1 with $\sqrt{v_{1,1}^2 + v_{2,1}^2} = 246$ GeV, and all particles having their known masses; and a *deeper* N_2 minimum, for which $\sqrt{v_{1,2}^2 + v_{2,2}^2} \neq 246$ GeV (possibly much larger or smaller). The N_2 minimum would have the same unbroken symmetries as N_1 *but* with a completely different mass spectrum of scalars, fermions and gauge bosons. This was seen to happen only for a very small portion of the parameter space. Of course, it is very easy to pick sets of parameters for which N_1 would be the global minimum, with N_2 above it or not even existing.

The main consequences of this vacuum analysis are:

- Minima of different natures cannot coexist in the 2HDM.
- Whenever a normal minimum exists in the 2HDM, the global minimum of the potential is normal. No tunnelling to a deeper CB or CP minimum is possible.
- If a CP (CB) violating minimum exists, it is the global minimum of the theory, and thoroughly stable. No tunnelling to a deeper normal or CB (CP) minimum can occur.

5.10. Mass matrices for neutral minima

In order to determine whether a given stationary point is a minimum, one needs to analyse the second derivatives of the potential, meaning the scalar mass matrices. For the sake of completeness, we include here the expressions for these matrices for neutral minima. The mass matrices for CB stationary points can be found in [405]. A discussion in the MSSM with explicit CP violation, which approaches the 2HDM in a given limit, can be found in Ref. [424].

- **Normal minima**

With vevs given by (195) and determined by (179), the squared mass for the charged scalar is given by the eigenvalues of a 2×2 matrix whose entries are

$$[M_{H^\pm}^2]_{ij} = \frac{\partial^2 V_H}{\partial \varphi_i^+ \partial \varphi_j^-}. \tag{203}$$

This matrix has a zero eigenvalue (corresponding to the Goldstone boson which gives mass to the W) so that the charged scalar squared mass is

$$M_{H^\pm}^2 = -\frac{v^2}{2v_1 v_2} \bar{V}, \tag{204}$$

where we have defined the quantity

$$\bar{V} = -2\mathrm{Re}(m_{12}^2) + [\lambda_4 + \mathrm{Re}(\lambda_5)]v_1 v_2 + \mathrm{Re}(\lambda_6)v_1^2 + \mathrm{Re}(\lambda_7)v_2^2. \tag{205}$$

Let us now assume for a moment the potential is explicitly CP conserving (*i.e.* we will work in a basis without imaginary couplings). The pseudoscalar mass matrix is the 2×2 matrix of the second derivatives of the imaginary parts of the neutral components,

$$[M_A^2]_{ij} = \frac{1}{2}\frac{\partial^2 V_H}{\partial \mathrm{Im}(\varphi_i^0)\partial \mathrm{Im}(\varphi_j^0)} = \frac{v_1 v_2}{v^2} M_A^2 \begin{pmatrix} \dfrac{v_2}{v_1} & -1 \\ -1 & \dfrac{v_1}{v_2} \end{pmatrix}, \tag{206}$$

where the entries of this matrix have been simplified through the minimization conditions. It has one zero eigenvalue (corresponding to the Goldstone boson which gives mass to the Z), so that the pseudoscalar squared mass is found to be

$$M_A^2 = M_{H^\pm}^2 + \frac{1}{2}[\lambda_4 - \mathrm{Re}(\lambda_5)]\,v^2. \tag{207}$$

As for the CP-even scalars, they are the eigenvalues of the symmetric 2×2 matrix, given by

$$[M_h^2]_{ij} = \frac{1}{2}\frac{\partial^2 V_H}{\partial \mathrm{Re}(\varphi_i^0)\partial \mathrm{Re}(\varphi_j^0)} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}, \tag{208}$$

where the matrix's entries are given by

$$A = m_{11}^2 + \frac{3\lambda_1}{2} v_1^2 + \frac{\lambda_{345}}{2} v_2^2 + 3\text{Re}(\lambda_6) v_1 v_2,$$

$$B = m_{22}^2 + \frac{3\lambda_2}{2} v_2^2 + \frac{\lambda_{345}}{2} v_1^2 + 3\text{Re}(\lambda_7) v_1 v_2,$$

$$C = -\text{Re}(m_{12}^2) + \frac{3}{2} \left[\text{Re}(\lambda_6) v_1^2 + \text{Re}(\lambda_7) v_2^2 \right] + \lambda_{345} v_1 v_2, \tag{209}$$

where we have defined $\lambda_{345} = \lambda_3 + \lambda_4 + \text{Re}(\lambda_5)$. The mass eigenstates of this matrix are traditionally represented as h and H, respectively the lightest and heaviest state. The diagonalization angle α of the matrix (208) is defined[19] as

$$H = -\cos\alpha \, \text{Re}(\varphi_1^0) - \sin\alpha \, \text{Re}(\varphi_2^0)$$

$$h = \sin\alpha \, \text{Re}(\varphi_1^0) - \cos\alpha \, \text{Re}(\varphi_2^0) \tag{210}$$

so that one gets, after trivial calculations,

$$\tan 2\alpha = \frac{2C}{A - B}. \tag{211}$$

If we are dealing with the most general 2HDM potential with complex couplings, then there will be mixing between the CP even and odd scalar particles. The neutral scalars will be the eigenvalues of a 4×4 matrix,

$$[M_N^2]_{ij} = \begin{pmatrix} [M_h^2] & [M_I^2] \\ [M_I^2]^T & [M_A^2] \end{pmatrix}, \tag{212}$$

composed of three 2×2 blocks. $[M_h^2]$ and $[M_A^2]$ are as given in Eqs. (208) and (206), respectively, whereas $[M_I^2]$ is given by the matrix

$$[M_I^2] = \begin{pmatrix} A_I & B_I \\ C_I & D_I \end{pmatrix}, \tag{213}$$

whose entries are

$$A_I = \frac{1}{2} v_2 \left[\text{Im}(\lambda_5) v_2 + 2\,\text{Im}(\lambda_6) v_1 \right],$$

$$B_I = \text{Im}(m_{12}^2) - \text{Im}(\lambda_5) v_1 v_2 - \frac{3}{2} \text{Im}(\lambda_6) v_1^2 - \frac{1}{2} \text{Im}(\lambda_7) v_2^2,$$

$$C_I = -\text{Im}(m_{12}^2) + \text{Im}(\lambda_5) v_1 v_2 + \frac{1}{2} \text{Im}(\lambda_6) v_1^2 - \frac{3}{2} \text{Im}(\lambda_7) v_2^2,$$

$$D_I = -\frac{1}{2} v_1 \left[\text{Im}(\lambda_5) v_1 + 2\,\text{Im}(\lambda_7) v_2 \right]. \tag{214}$$

Obviously, one of the eigenvalues of $[M_N^2]$ will be zero.

- **Inert minima**

Recall that one needs $m_{12}^2 = \lambda_6 = \lambda_7 = 0$ to obtain inert vacua, and they obey the minimization conditions (196) or (197). Considering, for example, the case $v_2 = 0$ and $v_1 = v/\sqrt{2}$, the scalar mass spectrum is greatly simplified. The CP-even mass matrix is diagonal (so that in this basis one may consider $\alpha = \beta = 0$), and the expressions for the masses are:

$$M_{H^\pm}^2 = m_{22}^2 + \frac{1}{2} \lambda_3 v^2$$

$$M_A^2 = M_{H^\pm}^2 + \frac{1}{2} [\lambda_4 - \text{Re}(\lambda_5)] v^2$$

$$M_H^2 = M_A^2 + \text{Re}(\lambda_5) v^2$$

$$M_h^2 = \lambda_1 v^2. \tag{215}$$

In inert models where we further have $\lambda_5 = 0$ (those arising from a $U(1)$ symmetry, for instance) A and H will be degenerate. Also, notice that though we maintained the notation h and H, it is now not guaranteed that they correspond to the lightest and heaviest CP-even states—that will depend on the specific values for the parameters. In fact, notice that the state 'h' can

---

[19] Clearly this is a basis-dependent definition, as is that of $\tan\beta$. None of these angles are basis invariant quantities, though their difference is [359].

be made much heavier or much lighter than the remaining three—it is the only one depending on the coupling λ_1. It can be shown that the H and A states do not couple to the Z or to the fermions, hence the name "inert".

- **CP breaking minima**

It only makes sense to speak of minima with spontaneous CP violation if that symmetry is defined—*i.e.* if it is not explicitly broken by the potential. As explained in Section 6, that corresponds to the existence of a basis where all parameters are real. We now write all masses in such a basis, for a vacuum with vevs such as (180). The charged scalar mass is now, as was mentioned before,

$$M^2_{H^\pm} = -\frac{1}{2}[\lambda_4 - \lambda_5](\bar{v}_1^2 + \bar{v}_2^2). \tag{216}$$

Given that the vevs now have an imaginary component, there will be a mixing between the real and imaginary components of φ_i^0. The neutral Higgs squared masses are thus the eigenvalues of a 4×4 symmetric matrix M^2_{CP}. Using the minimization conditions (194) to simplify, the entries of this matrix are given by:

$$M^2_{CP}(1, 1) = \lambda_1 \bar{v}_1^2 \cos^2 \theta + 2\lambda_6 \bar{v}_1 \bar{v}_2 \cos \theta + \lambda_5 \bar{v}_2^2$$
$$M^2_{CP}(1, 2) = \lambda_6 \bar{v}_1^2 \cos^2 \theta + (\lambda_3 + \lambda_4) \bar{v}_1 \bar{v}_2 \cos \theta + \lambda_7 \bar{v}_2^2$$
$$M^2_{CP}(1, 3) = (\lambda_6 \bar{v}_2 + \lambda_1 \bar{v}_1 \cos \theta) \bar{v}_1 \sin \theta$$
$$M^2_{CP}(1, 4) = (\lambda_6 \bar{v}_1 \cos \theta + \lambda_5 \bar{v}_2) \bar{v}_1 \sin \theta$$
$$M^2_{CP}(2, 2) = \lambda_5 \bar{v}_1^2 \cos^2 \theta + 2\lambda_7 \bar{v}_1 \bar{v}_2 \cos \theta + \lambda_2 \bar{v}_2^2$$
$$M^2_{CP}(2, 3) = [\lambda_7 \bar{v}_1 \cos \theta + (\lambda_3 + \lambda_4 + \lambda_5) \bar{v}_2] \bar{v}_1 \sin \theta$$
$$M^2_{CP}(2, 4) = (\lambda_7 \bar{v}_2 + \lambda_5 \bar{v}_1 \cos \theta) \bar{v}_1 \sin \theta$$
$$M^2_{CP}(3, 3) = \lambda_1 \bar{v}_1^2 \sin^2 \theta$$
$$M^2_{CP}(3, 4) = \lambda_6 \bar{v}_1^2 \sin^2 \theta$$
$$M^2_{CP}(3, 4) = \lambda_5 \bar{v}_1^2 \sin^2 \theta. \tag{217}$$

This matrix has a zero eigenvalue, corresponding to the Z Goldstone boson.

5.11. The Higgs basis

After spontaneous electroweak symmetry breaking with neutral vacua, the fields acquire the vacuum expectation values $v_1/\sqrt{2}$ and $v_2 e^{i\delta}/\sqrt{2}$; where v_1 and v_2 are real, without loss of generality. It is convenient to rotate into a new basis of scalar fields such that the vev is all in the first field, while the second field has no vev. This is known as the Higgs basis $\{H_1, H_2\}$, obtained through $H_a = \sum_{b=1}^2 U_{ab} \Phi_b$, where [425]

$$U = \frac{1}{v} \begin{bmatrix} v_1 & v_2 \, e^{-i\delta} \\ -v_2 & v_1 \, e^{-i\delta} \end{bmatrix} \tag{218}$$

$$= \frac{e^{-i\delta/2}}{v} \begin{bmatrix} v_1 \, e^{i\delta/2} & v_2 \, e^{-i\delta/2} \\ -v_2 \, e^{i\delta/2} & v_1 \, e^{-i\delta/2} \end{bmatrix}, \tag{219}$$

is unitary, and $v = \sqrt{v_1^2 + v_2^2} = (\sqrt{2}G_F)^{-1/2} = 246$ GeV. This rotates the vev into H_1, allowing us to parametrize

$$H_1 = \begin{bmatrix} G^+ \\ (v + H + iG^0)/\sqrt{2} \end{bmatrix}, \qquad H_2 = \begin{bmatrix} H^+ \\ (R + iI)/\sqrt{2} \end{bmatrix}, \tag{220}$$

where G^+ and G^0 are the Goldstone bosons, which, in the unitary gauge, become the longitudinal components of the W^+ and of the Z^0, and H, R and I are real neutral fields.

Notice that there are infinitely many Higgs basis. Indeed, we may change the phase of H_2,

$$H_2 \rightarrow e^{i\xi} H_2 \tag{221}$$

while keeping the vev in H_1. Under this phase transformation, the fields R and I are rotated by

$$\begin{pmatrix} R \\ I \end{pmatrix} \rightarrow \begin{pmatrix} \cos \xi & \sin \xi \\ -\sin \xi & \cos \xi \end{pmatrix} \begin{pmatrix} R \\ I \end{pmatrix}. \tag{222}$$

In going from a generic basis to the Higgs basis, the couplings in the scalar potential get rotated. Comparing Eq. (118) with Eq. (219) without the overall phase $e^{-i\delta/2}$, we find

$$\tan \beta = \frac{v_2}{v_1},$$
$$\chi = \delta/2,$$
$$\xi = \delta. \tag{223}$$

Thus, the potential coefficients in the Higgs basis are obtained from those in the generic basis, through Eqs. (119)–(128), with β, χ, and ξ as defined in Eq. (223). The quadratic (quartic) coefficients of the scalar potential in the Higgs basis are denoted by $\bar{m}_{ij}^2$ ($\bar{\lambda}_i$).

In the Higgs basis, the stationarity conditions are simply given by

$$\bar{m}_{11}^2 = -\frac{1}{2}\bar{\lambda}_1 v^2,$$
$$\bar{m}_{12}^2 = \frac{1}{2}\bar{\lambda}_6 v^2. \tag{224}$$

Let us count the parameters in the Higgs basis: the complex parameter $\bar{m}_{12}^2$ is determined by $\bar{m}_{11}^2$, $\bar{\lambda}_1$, and $\bar{\lambda}_6$. Thus, we would seem to have 12 parameters for the most general 2HDM potential. However, the fact that one may rephase H_2 implies that only the relative phases of the complex parameters $\bar{\lambda}_5$, $\bar{\lambda}_6$, and $\bar{\lambda}_7$ have physical significance. We are thus left with 11 physical parameters, as expected (we can view m_{11}^2 as determining v^2). However, as observed earlier, potentials to which symmetries have been imposed will display, in the Higgs basis, a smaller number of parameters. As first pointed out by Lavoura [426] and subsequently greatly expanded by Davidson and Haber [359], the coefficients of the potential in the Higgs basis are observable, up to the overall phase of the complex parameters. The only physically meaningful phases are $\text{Im}(\bar{\lambda}_5^* \bar{\lambda}_6^2)$, $\text{Im}(\bar{\lambda}_5^* \bar{\lambda}_7^2)$, and $\text{Im}(\bar{\lambda}_6 \bar{\lambda}_7^*)$. These are proportional to the quantities J_1, J_2, and J_3 (respectively), introduced by Lavoura and Silva [425] as basis-invariant signals of the CP violation in the 2HDM present after spontaneous electroweak symmetry breaking. Notice that only two are independent; for example, J_1 and J_3. However, since in a given theory $\bar{\lambda}_6$ may vanish while $\text{Im}(\bar{\lambda}_5^* \bar{\lambda}_7^2)$ does not, all three must be considered when searching for CP violation.

5.12. Yukawa couplings in the Higgs basis

The most generic Yukawa interactions that one can write with two doublets and the fermionic content of the SM are given by

$$\mathcal{L}_Y = -\sum_{j=1}^{2} \left[\bar{Q}_L \left(\Phi_j Y_j^d n_R + \tilde{\Phi}_j Y_j^u p_R \right) + \bar{L}_L \Phi_j Y_j^e \ell_R \right] + \text{H.c.} \tag{225}$$

In this equation, $\tilde{\Phi}_j = i\tau_2 \Phi_j^*$; Q_L, L_L, n_R, p_R and ℓ_R are 3-vectors in flavour space—the n_R correspond to the negative-charged quarks, and p_R to the positive-charged ones[20]; L_L and ℓ_R are the leptonic fields; and Y_j^d, Y_j^u and Y_j^e are generic 3×3 complex matrices containing the Yukawa couplings for, respectively, the down, up and leptonic sector. In the initial basis (Φ_1, Φ_2), the doublet Φ_1 has vacuum expectation value (vev) $\tilde{v}_1/\sqrt{2}$ and the doublet Φ_2 has vev $\tilde{v}_2/\sqrt{2}$, where $\tilde{v}_1$ and $\tilde{v}_2$ are allowed to be *complex*. We define

$$v = \sqrt{|\tilde{v}_1|^2 + |\tilde{v}_2|^2}. \tag{226}$$

Experimentally, $v \approx 246$ GeV. Notice that v is, by definition, real and positive.

Let us define the Higgs basis (H_1, H_2) by H_1 having vev $v/\sqrt{2}$. while H_2 has vanishing vev.[21] The transformation from one basis to the other is

$$\Phi_1 = \frac{1}{v} \left(\tilde{v}_1 H_1 + \tilde{v}_2^* H_2 \right), \tag{227}$$

$$\Phi_2 = \frac{1}{v} \left(\tilde{v}_2 H_1 - \tilde{v}_1^* H_2 \right). \tag{228}$$

---

[20] Which, after diagonalization, will yield the down and up type quarks.

[21] The definition used in Eq. (218) was useful for comparison with the basis-change formulae of Eqs. (119)–(128). This definition differs from the first by an irrelevant phase such as in Eq. (221).

Let us define the matrices

$$M_n = \frac{1}{\sqrt{2}} \left(\tilde{v}_1 Y_1^d + \tilde{v}_2 Y_2^d \right), \tag{229}$$

$$N_n = \frac{1}{\sqrt{2}} \left(\tilde{v}_2^* Y_1^d - \tilde{v}_1^* Y_2^d \right). \tag{230}$$

Then

$$\sum_{j=1}^{2} \Phi_j Y_j^d = \frac{\sqrt{2}}{v} \left(M_n H_1 + N_n H_2 \right). \tag{231}$$

The matrix M_n in Eq. (229) is the mass matrix of the down-type quarks. Notice that N_n in Eq. (230) may be written as

$$N_n = \frac{\tilde{v}_2^*}{\tilde{v}_1} M_n - \frac{v^2}{\sqrt{2}\tilde{v}_1} Y_2^d \tag{232}$$

$$= -\frac{\tilde{v}_1^*}{\tilde{v}_2} M_n + \frac{v^2}{\sqrt{2}\tilde{v}_2} Y_1^d. \tag{233}$$

It is convenient to pass to the mass basis of the quarks, in which the mass matrices are diagonal. To do so, we bi-diagonalize M_n, via a simultaneous rotation on the left-handed and right-handed quark fields:

$$U_L^\dagger M_n U_R^n = M_d, \tag{234}$$

$$U_L^\dagger N_n U_R^n = N_d, \tag{235}$$

where $M_d = \mathrm{diag}(m_d, m_s, m_b)$ is diagonal with real and positive diagonal elements. In the sector of the up-type quarks,

$$M_p = \frac{1}{\sqrt{2}} \left(\tilde{v}_1^* Y_1^u + \tilde{v}_2^* Y_2^u \right), \tag{236}$$

$$N_p = \frac{1}{\sqrt{2}} \left(\tilde{v}_2 Y_1^u - \tilde{v}_1 Y_2^u \right), \tag{237}$$

and

$$\sum_{j=1}^{2} \tilde{\Phi}_j Y_j^u = \frac{\sqrt{2}}{v} \left[M_p \left(i\tau_2 H_1^* \right) + N_p \left(i\tau_2 H_2^* \right) \right]. \tag{238}$$

The bi-diagonalization proceeds as

$$U_L^\dagger M_p U_R^p = M_u, \tag{239}$$

$$U_L^\dagger N_p U_R^p = N_u, \tag{240}$$

where $M_u = \mathrm{diag}(m_u, m_c, m_t)$ is diagonal with real and positive diagonal elements. Notice that the U_L matrix has to be the same in Eqs. (235) and (240).

If, after the bi-diagonalization, the matrix N_d (N_u) is *not* diagonal, then there are scalar tree-level flavour-changing neutral interactions in the down (up) sector, and the FCNC couplings for those interactions are obtained from the entries of N_d (N_u). In the generic basis of Eq. (225), the condition for non-existence of FCNC is also quite simple: if the matrices Y_1^d and Y_2^d (Y_1^u and Y_2^u) commute, there is no tree-level FCNC in the down-quark (up-quark) sector [383]. This, of course, is trivial if, for instance, Y_2^x is zero, as is obtained in models with Natural Flavour Conservation, which were discussed in Section 2.

5.13. Basis transformations and Yukawa couplings

We start from the Lagrangian

$$\mathcal{L} = \mathcal{L}_\mathrm{H} + \mathcal{L}_\mathrm{Y}, \tag{241}$$

where

$$-\mathcal{L}_\mathrm{H} == \sum_{a,b=1}^{2} \mu_{ab} (\Phi_a^\dagger \Phi_b) + \frac{1}{2} \sum_{a,b,c,d=1}^{2} \lambda_{ab,cd} (\Phi_a^\dagger \Phi_b)(\Phi_c^\dagger \Phi_d), \tag{242}$$

$$-\mathcal{L}_\mathrm{Y} = \bar{q}_L \left[\left(Y_1^d \Phi_1 + Y_2^d \Phi_2 \right) n_R + \left(Y_1^u \tilde{\Phi}_1 + Y_2^u \tilde{\Phi}_2 \right) p_R \right] + \mathrm{H.c.} \tag{243}$$

The Lagrangian can be rewritten in terms of new fields obtained from the original ones by simple basis transformations

$$\Phi_a \to \Phi'_a = \sum_{\alpha=1}^{2} U_{a\alpha} \, \Phi_\alpha,$$

$$q_L \to q'_L = U_L \, q_L,$$

$$n_R \to n'_R = U_{nR} \, n_R,$$

$$p_R \to p'_R = U_{pR} \, p_R, \tag{244}$$

where $U \in U(2)$ is a 2×2 unitary matrix, while $\{U_L, U_{nR}, U_{pR}\} \in U(3)$ are 3×3 unitary matrices. Under these unitary basis transformations, the gauge-kinetic terms are unchanged, but the coefficients μ_{ab} and $\lambda_{ab,cd}$ are transformed as

$$\mu_{ab} \to \mu'_{ab} = \sum_{\alpha,\beta=1}^{2} U_{a\alpha} \, \mu_{\alpha\beta} \, U^*_{b\beta}, \tag{245}$$

$$\lambda_{ab,cd} \to \lambda'_{ab,cd} = \sum_{\alpha,\beta,\gamma,\delta=1}^{2} U_{a\alpha} \, U_{c\gamma} \, \lambda_{\alpha\beta,\gamma\delta} \, U^*_{b\beta} \, U^*_{d\delta}, \tag{246}$$

while the Yukawa matrices change as

$$Y_a^d \to Y_a^{d\,\prime} = \sum_{\alpha=1}^{2} U_L \, Y_\alpha^d \, U^\dagger_{nR} \, \left(U^\dagger\right)_{\alpha a}$$

$$Y_a^u \to Y_a^{u\,\prime} = \sum_{\alpha=1}^{2} U_L \, Y_\alpha^u \, U^\dagger_{pR} \, \left(U^\top\right)_{\alpha a}. \tag{247}$$

Notice that we have kept the notation of showing explicitly the indices in scalar-space, while using matrix formulation for the quark flavour spaces. The basis transformations may be utilized in order to absorb some of the degrees of freedom of μ, λ, Y^d, and/or Y^u, which implies that not all parameters in the Lagrangian have physical significance.

5.14. Symmetries and Yukawa couplings

The symmetries we have discussed in Section 5.6 were imposed on the scalar sector, but they have to be extended to the full Lagrangian. By their definition, they leave the gauge kinetic terms invariant, but will affect the Yukawa terms, where the scalars are coupled to the fermions. As such, one needs to consider how the fermion fields transform under such symmetries, and whether or not their impact on the Yukawa sector leads to viable models.

5.14.1. Family symmetries

We will now assume that the Lagrangian in Eq. (241) is invariant under the symmetry

$$\Phi_a \to \Phi_a^S = \sum_{\alpha=1}^{2} S_{a\alpha} \, \Phi_\alpha,$$

$$q_L \to q_L^S = S_L \, q_L,$$

$$n_R \to n_R^S = S_{nR} \, n_R,$$

$$p_R \to p_R^S = S_{pR} \, p_R, \tag{248}$$

where $S \in U(2)$, while $\{S_L, S_{nR}, S_{pR}\} \in U(3)$. As a result of this symmetry, the parameters in the Lagrangian have to obey the following equations:

$$\mu_{ab} = \sum_{\alpha,\beta=1}^{2} S_{a\alpha} \, \mu_{\alpha\beta} \, S^*_{b\beta}, \tag{249}$$

$$\lambda_{ab,cd} = \sum_{\alpha,\beta,\gamma,\delta=1}^{2} S_{a\alpha} \, S_{c\gamma} \, \lambda_{\alpha\beta,\gamma\delta} \, S^*_{b\beta} \, S^*_{d\delta}, \tag{250}$$

$$Y_a^d = \sum_{\alpha=1}^{2} S_L \, Y_\alpha^d \, S^\dagger_{nR} \, \left(S^\dagger\right)_{\alpha a}, \tag{251}$$

$$Y_a^u = \sum_{\alpha=1}^{2} S_L \, Y_\alpha^u \, S^\dagger_{pR} \, \left(S^\top\right)_{\alpha a}. \tag{252}$$

Under the basis transformation of Eq. (244), the specific form of the symmetry in Eq. (248) is altered as

$$S' = U S U^\dagger,$$ (253)

$$S'_L = U_L S_L U_L^\dagger,$$ (254)

$$S'_{nR} = U_{nR} S_{nR} U_{nR}^\dagger,$$ (255)

$$S'_{pR} = U_{pR} S_{pR} U_{pR}^\dagger.$$ (256)

5.14.2. CP symmetries

We will now assume that the Lagrangian in Eq. (241) is invariant under the CP symmetry

$$\Phi_a \to \sum_{\alpha=1}^2 X_{a\alpha} \Phi_\alpha^*,$$

$$q_L \to X_L \gamma^0 C q_L^*,$$

$$n_R \to X_{nR} \gamma^0 C n_R^*,$$

$$p_R \to X_{pR} \gamma^0 C p_R^*.$$ (257)

As a result of this symmetry,[22] the parameters of the Lagrangian need to obey

$$\mu_{ab}^* = \sum_{\alpha,\beta=1}^2 X_{\alpha a}^* \mu_{\alpha\beta} X_{\beta b},$$ (258)

$$\lambda_{ab,cd}^* = \sum_{\alpha,\beta,\gamma,\delta=1}^2 X_{\alpha a}^* X_{\gamma c}^* \lambda_{\alpha\beta,\gamma\delta} X_{\beta b} X_{\delta d},$$ (259)

$$Y_a^{d*} = \sum_{\alpha=1}^2 X_{a\alpha} X_L Y_\alpha^d X_{nR}^\dagger,$$ (260)

$$Y_a^{u*} = \sum_{\alpha=1}^2 X_{a\alpha}^* X_L Y_\alpha^u X_{pR}^\dagger.$$ (261)

Under the basis transformation of Eq. (244), the specific form of the symmetry in Eq. (257) is altered as

$$X' = U X U^\top,$$ (262)

$$X'_L = U_L X_L U_L^\top,$$ (263)

$$X'_{nR} = U_{nR} X_{nR} U_{nR}^\top,$$ (264)

$$X'_{pR} = U_{pR} X_{pR} U_{pR}^\top.$$ (265)

5.14.3. Symmetries of the scalar–scalar and scalar–fermion interactions

We may now ask whether the six symmetry classes of the Higgs potential shown in Table 5 can be extended to the fermion sector in a way consistent with experiment. This issue is complicated by the fact that the fermion fields can transform, for a given scalar symmetry, under infinitely many ways, as detailed in Eqs. (251) and (252) for Higgs-family symmetries, and Eqs. (260) and (261) for GCP symmetries. Let us deal with those two types of symmetries separately.

- Higgs family symmetries extended to the Yukawa sector

As an example of how complex this issue can become, consider that even for a simple Z_2 symmetry in the scalars, for which the Φ_2 doublet flips its sign, we can choose an extremely elaborate transformation law for the quark fields, with arbitrary unitary 3×3 matrices S_L, S_{nR} and S_{pR}. To further complicate matters, the field transformation laws can correspond to an Abelian symmetry (with a single generator, or a set of generators which commute amongst themselves), or a more general and complex non-Abelian one (with several non-commuting generators). In the scalar sector, all but the class I models (with a full $U(2)$ symmetry) can be obtained via an Abelian symmetry.

---

[22] Eq. (258) can be written in Higgs-family matrix form as $\mu^* = X^\dagger \mu X$, in an obvious notation. This is equivalent to $\mu^* = X^* \mu X^\top$. Similar rewritings are also possible for Eqs. (259)–(261), sometimes complicating comparisons.

As such, extensions of scalar symmetries to the Yukawa sector are usually specific examples, where one chooses a particular form for the fermion transformation matrices. There are many such examples. For instance, the various implementations of the Z_2 symmetry (or the $U(1)$ one) which preclude the occurrence of tree-level FCNC [27,28]—the so-called type I, II, X (lepton specific) and Y (flipped) 2HDMs. Such models are said to have "Natural Flavour Conservation", and their phenomenology was already discussed in Section 2. Another example would be the BGL model [237], which does contain FCNC which are naturally small due to a flavour-dependent fermionic symmetry. A third example would be a recent application of a Z_3 symmetry to the 2HDM Lagrangian, with interesting consequences regarding the origin of CP violation [427].

The extension of generic Abelian symmetries into the quark sector was only fully mapped recently by Ferreira and Silva [428]. That calculation was greatly simplified by the usage of the full freedom of choosing a basis of both scalar and fermionic fields, as explained in the previous sections: it turns out that it is *always* possible, for a transformation involving a single transformation matrix S_X, to go to a basis of fields for which *all* such matrices are diagonal, containing, in general, only complex phases. Meaning, it is always possible to choose a basis of fields such that the fermionic transformation matrices of Eqs. (251) and (252) are reduced, through transformations like (254)–(256), to the form

$$S_L = \text{diag}(e^{i\alpha_1}, e^{i\alpha_2}, e^{i\alpha_3}),$$
$$S_{nR} = \text{diag}(e^{i\beta_1}, e^{i\beta_2}, e^{i\beta_3}),$$
$$S_{pR} = \text{diag}(e^{i\gamma_1}, e^{i\gamma_2}, e^{i\gamma_3}). \tag{266}$$

Thus, all the freedom in choosing the fermions' transformation laws is reduced to the choice of the arbitrary (real) phases α_i, β_i and γ_i. Even this immense simplification, though, yields millions of possible specific symmetry implementations—that is, of possible different models.

However, as explained in detail in [428], it turns out that the simultaneous requirements of six massive quarks[23] and an acceptable CKM matrix are extremely powerful and curtail immensely the number of allowed models. The reason for that is that the effect of the phases α_i on the CKM matrix, or in the quark squared-mass matrices, is tantamount to setting many of their entries to zero. Very easily, for arbitrary choices of the α_i, one obtains a line or column of zeros in the mass matrices, or a diagonal block on the CKM matrix. Surprisingly, then, the transformation laws of the quark fields are extremely constrained. The authors of [428] have shown that there are only 246 possible forms for the Yukawa matrices for both the up and down quarks. Up to physically unimportant permutations, these involve only 34 forms of Yukawa matrices. As such, there is effectively a maximum of only 34 possible ways of extending the Abelian Higgs-family symmetries (that is, the models like classes IV and V, with symmetries like Z_2 or the Peccei–Quinn $U(1)$) to the quark sector. Most of these symmetry-constrained Yukawa matrices lead to tree-level scalar FCNC. However some of them have already been shown to have naturally small FCNC, which are CKM suppressed, the aforementioned BGL models; or to be such that one can easily find values of parameters which satisfy all experimental constraints, such as the mass differences of neutral kaons and B-mesons, which typically are difficult to accommodate when tree-level FCNC are present.

In addition, the general analysis of these Abelian symmetries leads to some wide-reaching conclusions:

- Imposing Z_2 on the scalars does not imply a continuous symmetry in the Higgs sector, but it may or not imply a continuous symmetry in the Yukawa sector, depending on how the symmetry is extended into the fermions.
- Imposing Z_3 on the scalars does imply a continuous symmetry in the Higgs sector, but it may or not imply a continuous symmetry in the Yukawa sector, depending on how the symmetry is extended into the fermions.
- Imposing Z_n, with $n \geq 4$, on the scalars implies always a continuous symmetry, both in the Higgs sector and in the Yukawa sector.

This analysis also permits us to qualify the statement we made in Section 5.6.1, following Eq. (162): the imposition on the scalar potential of any discrete symmetry Z_n, with $n > 2$, always leads to the same potential, the Peccei–Quinn one of class IV. However, as we mentioned there, the extension of those symmetries to the fermion sector might be able to "lift the degeneracy" of these symmetries—and in fact it does. The results of [428] show very clearly that: (a) imposing Z_2 on the Yukawa terms is different from imposing Z_n, with $n > 2$; (b) imposing Z_3 on the Yukawa terms is different from imposing any other Z_n, $n > 3$; and (c), that imposing any Z_n, with $n \geq 4$, *always* leads to the same form of Yukawa matrices. As such, in terms of Z_n symmetries, the 2HDM *Lagrangian* falls under three classes: Lagrangians with a Z_2 symmetry; Lagrangians with a Z_3 symmetry; and Lagrangians with any other Z_n, $n \geq 4$, which always lead to the same symmetry constraints, regardless of the value of n.

Also, notice that, since any finite discrete group has an Abelian subgroup, the classification of [428] is important even when considering non-Abelian family symmetries. The discussion above concerns extensions of the Z_2 and $U(1)$ scalar symmetries to the fermion sector. That leaves out, of the three possible Higgs-family symmetries, the models invariant under the full $U(2)$ group—dubbed class I in Section 5.6. However, up until now it has been impossible to extend the $U(2)$ symmetry in a satisfactory way to the Yukawa sector—all attempts to do so have lead to zero mass quarks, for instance. A proof of impossibility has not yet been obtained, and the question remains open.

---

[23] Or three massive charged leptons, for the argument is trivially extended to the leptonic sector.

- GCP symmetries extended to the Yukawa sector

By comparison with Higgs-family symmetries, it is much simpler to extend the three possible scalar GCP symmetries to the Yukawa sector. In fact, as we discussed in Section 5.6.2, any GCP transformation on the doublets can be reduced to a simple rotation matrix of the form of Eq. (139). As it turns out, the Vienna group has also shown [383] that a similar result is attained for the generic 3×3 transformation matrices of the quark fields in Eqs. (260) and (261). That is, it is always possible, through a judicious choice of basis of quark fields, to reduce the transformation matrix of the left doublets to the form

$$X_L \implies \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{267}$$

with some angle $0 \leq \alpha \leq \pi/2$. A similar form is obtained for the matrices X_{nR} and X_{pR} in Eqs. (260) and (261), with independent angles β and γ in the same range as α. The extremely simple form of the fermion transformation matrices imposes severe constraints on the Yukawa couplings. In fact, the constraints are so serious that no ambiguity occurs in the fermionic sector when one extends the scalar GCP symmetries to it: each of the three GCP models have only one possible implementation on the fermion sector. Recalling that the three GCP scalar models can be parameterized in terms of the angle θ in the simplified GCP transformation of Eq. (139), it was concluded that [391]:

- For the CP1 symmetry (class VI), with $\theta = 0$, there is only one way to extend the scalar symmetry to the fermion sector which does not entail massless quarks or charged leptons: by forcing all Yukawa couplings to be real. We are thus left with a Lagrangian with generic real matrices Y_1^d, Y_2^d, Y_1^u and Y_2^u—as such the model has tree-level scalar FCNC, which are not in any way "naturally suppressed". In this model, CP violation must needs arise spontaneously, through a relative phase between the two vevs.
- For the CP2 symmetry (class III), with $\theta = \pi/2$, there is no way to extend the symmetry to the Yukawa sector without obtaining at least one massless charged fermion. As such, the CP2 model may be considered ruled out by experiment. However, one might also take the point of view that the CP2 symmetry is an *approximate* one, broken by some manner of mechanism, and as such the massless fermions it predicts will gain a (small) mass somehow, corresponding to the first generations of particles. The phenomenology of such models was explored in great detail in [387–389,429].
- For the CP3 model (class II), with *any* $0 < \theta < \pi/2$, a remarkable thing happens: all values of $\theta \neq \pi/3$ lead to a massless quark or charged lepton. Only $\theta = \pi/3$ leads to an acceptable fermion mass spectrum. The Yukawa matrices which result from such a symmetry are extremely constrained—the quark sector ends up depending only on ten independent parameters (seven moduli and three phases). Nonetheless, this model is capable of fitting all quark masses and the elements of the CKM matrix with relative ease. The model does possess tree-level FCNC, but they end up being quite suppressed, in a "natural" way. The model also possesses a unique feature, in the sense that CP violation arises in a completely novel way—we will return to this point in Section 6.7. However, the value of the Jarlskog invariant predicted by this model is several orders of magnitude below its SM value, which leads to values of the unitarity triangle angles α and β practically equal—a prediction of the model in contradiction with the most recent experimental data [293].

In conclusion, when one extends the three GCP scalar symmetries to the Yukawa sector, one obtains: arbitrary FCNC for the CP1 case; massless quarks and charged leptons for the CP2 case; a single CP3 symmetry leading to three massive generations of fermions, with naturally small FCNC but predictions for heavy meson phenomenology which do not agree with experiment.

6. CP violation

6.1. CP invariance and CP violation at the Lagrangian level: scalar potential

6.1.1. Two Higgs doublets

The scalar potential for $n_d SU(2)$ doublets is the most general renormalizable polynomial consistent with the gauge invariance and can be written

$$V = \sum_{a,b=1}^{n_d} \mu_{ab} \Phi_a^\dagger \Phi_b + \frac{1}{2} \sum_{a,b,c,d=1}^{n_d} \lambda_{ab,cd} \left(\Phi_a^\dagger \Phi_b \right) \left(\Phi_c^\dagger \Phi_d \right). \tag{268}$$

Hermiticity of V implies:

$$\mu_{ab}^* = \mu_{ba}, \qquad \lambda_{ab,cd}^* = \lambda_{ba,dc}. \tag{269}$$

One may redefine the n_d doublets through unitary transformations without changing the physics. Those transformations are called Higgs-basis transformations (HBT), defined by:

$$\Phi_a \xrightarrow{\text{HBT}} \Phi_a' = \sum_{b=1}^{n_d} V_{ab} \Phi_b, \qquad \Phi_a^\dagger \xrightarrow{\text{HBT}} \Phi_a'^\dagger = \sum_{b=1}^{n_d} V_{ab}^* \Phi_b^\dagger, \tag{270}$$

where V is an $n_d \times n_d$ unitary matrix acting in the space of the Higgs doublets. Under a HBT the couplings μ and λ transform as:

$$\mu_{ab} \xrightarrow{\text{HBT}} \mu'_{ab} = \sum_{m,n=1}^{n_d} V_{am}\mu_{mn}V^\dagger_{nb},$$

$$\lambda_{ab,cd} \xrightarrow{\text{HBT}} \lambda'_{ab,cd} = \sum_{m,n,p,q=1}^{n_d} V_{am}V_{cp}\lambda_{mn,pq}V^\dagger_{nb}V^\dagger_{qd}. \tag{271}$$

The most general CP transformation that leaves the kinetic energy invariant is:

$$\Phi_a \xrightarrow{\text{CP}} \sum_{b=1}^{n_d} U_{ab}\Phi^*_b, \qquad \Phi^\dagger_a \xrightarrow{\text{CP}} \sum_{b=1}^{n_d} U^*_{ab}\Phi^T_b, \tag{272}$$

where U is an $n_d \times n_d$ unitary matrix operating in the space of the Higgs doublets. This is the definition of a CP transformation for the scalar doublets in models with several such doublets. It combines what would be the CP transformation for a single Higgs doublet with a Higgs basis transformation, due to the existence of several doublets with the same quantum numbers.

For two Higgs doublets the most general Higgs potential is explicitly written in Eq. (98). Hermiticity only allows for four of the coefficients – m^2_{12}, λ_5, λ_6, and λ_7 – to be complex. However, that potential contains an excess of parameters. There is no loss of generality in redefining the two Higgs doublets in such a way that the quadratic terms are diagonal, thus eliminating m^2_{12}. Furthermore, one of the remaining three phases can still be eliminated through a rephasing of one of the doublets. Thus, there are (at most) only two independent CP-violating phases in the potential of a 2HDM.

In this section we address the question of what are the necessary and sufficient conditions for the potential V to be CP-invariant. At this stage we analyse the potential prior to gauge-symmetry breaking. We want to derive HBT-invariant conditions following the general method proposed in [400] and described in detail in [430]. We investigate what restrictions for the couplings μ and λ are implied by CP invariance [431].

From the potential of Eq. (268) and the definition of the CP transformation given by Eq. (272), it is clear that the necessary and sufficient condition for V to conserve CP is the existence of an $n_d \times n_d$ unitary matrix U satisfying the following relations:

$$\mu^*_{ab} = \sum_{m,n=1}^{n_d} U^\dagger_{am}\mu_{mn}U_{nb}, \qquad \lambda^*_{ab,cd} = \sum_{m,n,p,q=1}^{n_d} U^\dagger_{am}U^\dagger_{cp}\lambda_{mn,pq}U_{nb}U_{qd}. \tag{273}$$

From Eq. (273), using the property of invariance of the trace under similarity transformations, one can derive necessary conditions for CP invariance fully expressed in terms of the couplings μ and λ. Examples of such relations are:

$$I_1 \equiv \text{Tr}\left(\mu Z_Y \widehat{Z} - \widehat{Z}Z_Y\mu\right) = 0, \tag{274}$$

$$I_2 \equiv \text{Tr}\left(\mu Z_2 \widetilde{Z} - \widetilde{Z}Z_2\mu\right) = 0, \tag{275}$$

where we have introduced the following $n_d \times n_d$ Hermitian matrices:

$$\widehat{Z}_{ab} \equiv \sum_{m=1}^{n_d} \lambda_{ab,\,mm}, \tag{276}$$

$$\widetilde{Z}_{ab} \equiv \sum_{m=1}^{n_d} \lambda_{am,mb}, \tag{277}$$

$$(Z_Y)_{ab} \equiv \sum_{m,n=1}^{n_d} \lambda_{ab,nm}\mu_{mn}, \tag{278}$$

$$(Z_2)_{ab} \equiv \sum_{m,n=1}^{n_d} \lambda_{ap,nm}\lambda_{mn,pb}. \tag{279}$$

It is clear that the Eqs. (274) and (275) are HBT invariant. These two conditions have the remarkable property of being necessary and sufficient conditions for V to conserve CP in the case of two Higgs doublets, barring the consideration of special isolated points such as $m^2_{11} = m^2_{22}$ in the basis where $m^2_{12} = 0$, or the special isolated point where $\lambda_6 = -\lambda_7$ with generic λ_1 and λ_2 [431]. It should be emphasized that these isolated points are of measure zero and are unstable under renormalization, since they do not correspond to any symmetry. In order to check that these two conditions are sufficient let us express I_1 and I_2 in terms of the parameters of the potential after diagonalization of the quadratic terms, i.e. when $m^2_{12} = 0$:

$$I_1 = \frac{i}{2}\left(m^2_{11} - m^2_{22}\right)^2 \text{Im}\left(\lambda_6\lambda^*_7\right), \tag{280}$$

$$I_2 = i\left(m^2_{11} - m^2_{22}\right) \text{Im}\left[\lambda_5\lambda^{*2}_6 + \lambda_5\lambda^{*2}_7 + 2\lambda_5\lambda^*_6\lambda^*_7 + \lambda^*_6\lambda_7(\lambda_2 - \lambda_1)\right]. \tag{281}$$

Choosing λ_5 to be real, which can be done without loss of generality, from $I_1 = 0$ we obtain that λ_6 and λ_7 have either equal phases of phases differing by π. Then the condition $I_2 = 0$ implies

$$\text{Im}\left[\lambda_5\left(\lambda_6^* + \lambda_7^*\right)^2\right] = 0. \tag{282}$$

In the case where λ_6 and λ_7 have equal phases, we find from Eq. (282) that the phase is either 0 or $\pi/2$. The case $\arg \lambda_6 = \arg \lambda_7 = 0$ obviously corresponds to a CP-invariant V, with the matrix U in Eq. (272) being the 2×2 identity matrix. It can be easily checked that the case $\arg \lambda_6 = \arg \lambda_7 = \pi/2$ also corresponds to a CP-invariant V with

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{283}$$

Notice that at this stage we are assuming $m_{11}^2 \neq m_{22}^2$ and $|\lambda_6| \neq |\lambda_7|$.

Let us now consider the singular points which are unstable under renormalization and, therefore, of limited interest. It is clear from the explicit form of I_1 and I_2 that in the special cases $m_{11}^2 = m_{22}^2$ or $\lambda_6 = -\lambda_7$ both conditions $I_1 = I_2 = 0$ are trivially verified irrespective of the values of the phases. Yet, it is still possible to have CP violation in this region of parameters. In these cases, barring again cases of special isolated points, the following basis-invariant necessary condition for CP conservation is useful:

$$I_3 \equiv \text{Tr}\left(Z_2 Z_3 \widehat{Z} - \widehat{Z} Z_3 Z_2\right) = 0, \tag{284}$$

where Z_3 is one further $n_d \times n_d$ Hermitian matrix given by

$$(Z_3)_{ab} \equiv \sum_{m,n,p,r,s=1}^{n_d} \lambda_{am,rp}\lambda_{mr,ns}\lambda_{pn,sb}. \tag{285}$$

An important characteristic of I_3 is the fact that, unlike I_1 and I_2, it is built exclusively from the quartic couplings λ. Therefore, non-vanishing of this invariant necessarily signals hard CP violation.

Let us consider the case where $m_{12}^2 = 0$, $\lambda_1 = \lambda_2$, and $\lambda_6 = \lambda_7$. In that case the quartic couplings by themselves conserve CP, provided one chooses

$$U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{286}$$

This CP symmetry is only broken by $m_{11}^2 \neq m_{22}^2$, i.e. by the quadratic terms, and therefore one may say that this potential corresponds to "hidden" soft CP breaking. In this example, both I_1 and I_3 vanish identically, whilst $I_2 = 4i\left(m_{11}^2 - m_{22}^2\right)\text{Im}\left(\lambda_5\lambda_6^{*2}\right)$.

Another interesting example of soft symmetry breaking with two Higgs doublets is obtained by taking $\lambda_1 = \lambda_2$ and $\lambda_6 = -\lambda_7$. In this case the quartic couplings by themselves conserve CP, provided one now chooses U

$$U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{287}$$

Once again we have an example of hidden soft CP breaking for $m_{11}^2 \neq m_{22}^2$. In this example all three CP-odd invariants defined above vanish and an additional CP-odd invariant is necessary. In Ref. [432] the following condition was provided:

$$I_4 \equiv \text{Im} \sum_{a,b,\ldots,i=1}^{2} \left(\lambda_{ac,bd}\lambda_{ce,dg}\lambda_{eh,fi}\mu_{ga}\mu_{hb}\mu_{if}\right). \tag{288}$$

It was pointed out in [359] that, under Higgs basis transformations, the given relations $\lambda_1 = \lambda_2$ and $\lambda_6 = -\lambda_7$ remain invariant, this case being a special isolated point in the 2HDM scalar-potential parameter space. In the language of Ref. [379] this corresponds to the absence of the triplet in the decomposition of the quartic Higgs potential into irreducible representations of the $SU(2)$ Higgs basis transformation for two Higgs doublets.

More examples of hidden soft symmetry breaking can be written, based on different U matrices, such as

$$U = \begin{pmatrix} 0 & i \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad U = \begin{pmatrix} 0 & -i \\ 1 & 0 \end{pmatrix}. \tag{289}$$

In both these cases I_1 does not automatically vanish.

Under a HBT the specific form of U for a given CP transformation changes in the following way:

$$U' = VUV^T. \tag{290}$$

Invariance under CP of the 2HDM potential for U given by Eq. (286) will therefore look different in a different Higgs basis. Taking V as

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{291}$$

one obtains:

$$U' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{292}$$

This two Higgs doublet model with the reflection symmetry $\Phi_2 \to -\Phi_2$ softly broken by the quadratic terms proportional to m_{12}^2 in Eq. (98) was considered in Refs. [433–435].

6.1.2. Three Higgs doublets

In this subsection we briefly mention the case of three Higgs doublets. The total number of independent CP-violating phases in the scalar potential is given by [431]:

$$N_{\text{phases}} = \frac{1}{4} \left[n_d^2 (n_d^2 - 1) \right] - (n_d - 1), \tag{293}$$

the second term corresponds to the number of phases one can eliminate by rephasing the Higgs fields. In general there are sixteen independent phases for $n_d = 3$. Models with three Higgs doublets are much more involved than those with only two. Three Higgs doublets were considered [436] in an attempt to introduce CP violation in an extension of the SM with NFC [27,28]. Natural flavour conservation means that the Higgs neutral currents conserve all quark flavours for all values of the parameters of the theory, i.e., this conservation is a consequence of the group structure and the representation content of the theory and not of a special choice of parameters. NFC with several Higgs doublets was implemented in Ref. [436] by requiring invariance of the Lagrangian under separate reflections under which any one of the doublets (and perhaps some fermions) change sign. Discrete symmetries of this kind were used to insure that only one Higgs couples to the right-handed up quarks and another to the right-handed down quarks. It was pointed out in Ref. [436] that for three or more Higgs doublets the scalar potential that is invariant under these reflections need not conserve CP. On the other hand it was shown in [437–439] that with three Higgs doublets it is possible to violate CP spontaneously while having NFC. In the three Higgs doublets model proposed by Weinberg [436], a $Z_2 \times Z_2 \times Z_2$ symmetry, under separate reflections of the Higgs doublets of the form $\phi_i \to -\phi_i$, together with an appropriately chosen transformation for the quark fields, ensures NFC and leads to a strong reduction in the number of parameters. The Higgs potential is given by:

$$V = \sum_{i=1}^{3} \left[m_i \, \Phi_i^\dagger \Phi_i + a_{ii} \left(\Phi_i^\dagger \Phi_i \right)^2 \right] + \sum_{i<j} \left\{ 2 b_{ij} \left(\Phi_i^\dagger \Phi_i \right) \left(\Phi_j^\dagger \Phi_j \right) \right.$$
$$\left. + 2 c_{ij} \left(\Phi_i^\dagger \Phi_j \right) \left(\Phi_j^\dagger \Phi_i \right) + \left[d_{ij} \, e^{i\theta_{ij}} \left(\Phi_i^\dagger \Phi_j \right)^2 + h.c. \right] \right\}. \tag{294}$$

There are three different $d_{ij} \, e^{i\theta_{ij}}$ terms, and only these can be complex. It was pointed out by Weinberg that in general one cannot rotate away simultaneously the three phases θ_{ij}. A relevant CP-odd invariant relevant to this model is:

$$I_2^W = \text{Im} \sum_{a,\ldots,i=1}^{n_d} [Z_{abcd} \, Y_{be} \, Z_{efgh} \, \widehat{Z}_{fi} \, Z_{icha}]. \tag{295}$$

Its explicit form is:

$$I_2^W = d_{12} d_{13} d_{23} \left[(m_3 - m_2)(a_{11} - b_{23}) - (m_3 - m_1)(a_{22} - b_{13}) \right.$$
$$\left. + (m_2 - m_1)(a_{33} - b_{12}) \right] \sin(\theta_{12} - \theta_{13} + \theta_{23}). \tag{296}$$

Assuming non-degenerate values for the m_i, a non-vanishing I_2^W indicates a non-vanishing $(\theta_{12} - \theta_{13} + \theta_{23})$.

In the general case of three Higgs doublets it is possible to build relevant simpler CP-odd invariants [431], which are irrelevant for the case of two Higgs, since they trivially vanish in that case.

The softly broken three Higgs doublets model of Weinberg and relevant CP-odd invariant in this case is also discussed in Ref. [431].

6.2. CP violation after spontaneous symmetry breaking

6.2.1. Spontaneous CP violation

The idea of spontaneous CP breaking was suggested by Lee [9] in the early stages of unified gauge theories. In order to have spontaneous CP violation, one must have a Lagrangian which is CP invariant but after spontaneous gauge symmetry breaking the vacuum is not CP invariant. One has to be careful in correctly identifying a vacuum as CP violating. This has to do with the fact that often a CP invariant Lagrangian allows not only for a single CP transformation, but for a whole class of CP transformations. In order to have a genuine spontaneous CP violation, the following two conditions have to be satisfied:

 (i) The Lagrangian is invariant under a CP transformation which may be physically interpreted as CP.
(ii) There is no transformation which can be physically interpreted as CP which leaves both the vacuum and the Lagrangian invariant.

In the Standard Model there is only one Higgs doublet. Hermiticity requires that the parameters of the scalar potential be real and, as a result the scalar potential of the Standard Model cannot violate CP. Furthermore, spontaneous CP violation is also ruled out, in this case, due to the possibility of using a $U(1)$ gauge transformation to make the vacuum expectation value of the neutral Higgs real and positive.

Since we are working in the framework of relativistic quantum field theory, the CPT theorem applies [440–443], so spontaneous CP breaking also implies spontaneous T breaking and vice-versa.

For definiteness, let us consider an extension of the SM where n $SU(2) \times U(1)$ scalar doublets are introduced. In order to include the possible existence of symmetries of the Lagrangian under which the scalar doublets transform non-trivially, one has to consider the most general CP transformation which leaves invariant the kinetic energy terms of the scalar potential. Thus, we consider the following CP transformation for the scalar doublets:

$$CP\Phi_i(CP)^\dagger = \sum_{j=1}^n U_{ij}\Phi_j^* \tag{297}$$

corresponding to Eq. (272). Assuming that the vacuum is CP invariant, meaning that:

$$CP|0\rangle = |0\rangle \tag{298}$$

one can readily derive from Eqs. (297) and (298) the following relation [444]:

$$\sum_{j=1}^n U_{ij}\langle 0|\Phi_j|0\rangle^* = \langle 0|\Phi_i|0\rangle. \tag{299}$$

If the vacuum is such that none of the CP symmetries allowed by the Lagrangian satisfy Eq. (299), then this means that the vacuum is not CP invariant and we say that CP is spontaneously broken.

From the previous discussion, one concludes that in the presence of extra symmetries imposed on the Lagrangian, one has to be specially careful in analysing whether a particular vacuum violates CP or not. The point is that the presence of an extra symmetry in the Lagrangian may allow for non-trivial possibilities for the matrix U in Eq. (297), which may satisfy Eq. (299), even in the case of complex minima. This can be best understood through a simple example:

6.2.2. An example

Let us consider an extension of the SM with two Higgs doublets, where a Z_2 symmetry is introduced, with the scalars transforming as:

$$\Phi_1 \longrightarrow -\Phi_1, \qquad \Phi_2 \longrightarrow \Phi_2. \tag{300}$$

A possible motivation for the introduction of a Z_2 symmetry, is the requirement of natural flavour conservation (NFC) in the Higgs sector [27,28]. Indeed, if the right-handed down quarks are odd under Z_2,

$$d_R \longrightarrow -d_R \tag{301}$$

while all other fields are even, then down quarks receive mass only from Φ_1, while up quarks receive mass only from Φ_2, thus satisfying the NFC principle.

The most general gauge symmetry invariant Higgs potential, consistent with the Z_2 symmetry, can be written as

$$V = V_0 + \left[\lambda_5 \left(\Phi_1^\dagger \Phi_2\right)\left(\Phi_1^\dagger \Phi_2\right) + \text{h.c.}\right] \tag{302}$$

where V_0 denotes the part of the potential which does not depend on the relative phase of the Φ_i. It is clear that for $\lambda_5 > 0$, the minimum of the potential is at:

$$\langle \phi_1^0 \rangle = v_1 e^{i\frac{\pi}{2}}; \qquad \langle \phi_2^0 \rangle = v_2. \tag{303}$$

One could be tempted to think that this vacuum violates CP "maximally". This is *not the case*, as it can be seen from Eq. (299). Indeed, the presence of the Z_2 symmetry allows for various choices of matrix U which defines the CP properties of Φ_1, Φ_2. Apart from the trivial one, one may choose:

$$U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{304}$$

It is clear that with this choice of U, Eq. (299) is satisfied,

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} v_1 e^{i\frac{\pi}{2}} \\ v_2 \end{bmatrix}^* = \begin{bmatrix} v_1 e^{i\frac{\pi}{2}} \\ v_2 \end{bmatrix} \tag{305}$$

thus proving that the vacuum of Eq. (303) is CP invariant. At this stage, it is worth noting that one encounters an entirely analogous situation in the Minimal Supersymmetric Standard Model (MSSM) where one of the possible minima also has a relative phase of $\pi/2$, which does not lead to spontaneous CP violation. It has been shown [445,446] that one may achieve spontaneous CP violation in an extension of the MSSM where one introduces a singlet scalar.

Coming back to the previous discussion, one may wonder whether this is a generic feature of vacua with "calculable" vacuum phases. This question was addressed in detail by Branco et al. in Ref. [444]. It was shown that indeed this is the case for most vacua with calculable phases arising from the presence of extra symmetries in the Higgs potential. Another interesting example is the case of S_3 symmetry introduced in a Higgs potential with three Higgs doublets, transforming as a three dimensional reducible representation of S_3. The most general renormalizable Higgs potential can be written [447]:

$$V = V_0 + \lambda_1 \left(\Phi_i^\dagger \Phi_j \right) + \lambda_2 \left[\left(\Phi_i^\dagger \Phi_i \right) \left(\Phi_i^\dagger \Phi_j + \text{h.c.} \right) \right] + \lambda_3 \left[\left(\Phi_i^\dagger \Phi_i \right) \left(\Phi_j^\dagger \Phi_k + \text{h.c.} \right) \right]$$
$$+ \lambda_4 \left[\left(\Phi_i^\dagger \Phi_j \right) \left(\Phi_i^\dagger \Phi_k \right) + \text{h.c.} \right] + \lambda_5 \left[\left(\Phi_i^\dagger \Phi_j \right) \left(\Phi_k^\dagger \Phi_i \right) + \text{h.c.} \right] + \lambda_6 \left[\left(\Phi_i^\dagger \Phi_j \right)^2 + \text{h.c.} \right] \tag{306}$$

where V_0 denotes the part of the potential with no phase dependence. In each square bracket a sum is understood over all independent permutation of i, j, k with $i \neq j \neq k$. It can be readily verified that there is a region of parameter space where the vacuum has the following phase structure:

$$\langle \phi_i^0 \rangle = v \exp\left[i\frac{2\pi}{3}(k-1) \right]; \quad k = 1, 2, 3 \qquad \langle \phi_2^0 \rangle = v_2. \tag{307}$$

It can be easily shown using an argument analogous to the one used in the Z_2 example that, contrary to naïve intuition, the vacuum of Eq. (307) is CP and T invariant. One may wonder whether this is a universal feature. Namely, one may ask whether calculability of vacuum phases necessarily implies CP invariant vacua. In Ref. [444] it was shown that this is not the case. A counterexample was found [444] based on the group $\Delta(27)$ which is a dihedral-like subgroup of $SU(3)$ with 27 elements. In this example, a CP-violating vacuum with calculable phases was found.

6.2.3. A survey of models with spontaneous CP violation

At this stage, the following question is in order: What is the minimal extension of the SM where one may achieve spontaneous CP violation while at the same time not entering in conflict with experiment? In order for a given model to be a candidate for a realistic example of spontaneous CP violation, it should satisfy the following conditions:

(i) The CP violating phase arising from the vacuum should be able to create a complex CKM matrix, leading to CP violation through W-mediated weak currents.

(ii) The model should be able to avoid in a plausible way, i.e. without unreasonable fine tuning, the stringent experimental constraints arising from FCNC processes, as well as from the knowledge of the location of the upper vertex of the Unitarity Triangle [7].

First let us explain why the model should satisfy the condition (i). At present, there is strong evidence for a complex CKM matrix even if one allows for the presence of New Physics beyond the SM [448]. The best evidence arises from the non-vanishing of the rephasing invariant angle γ [7] which does not receive important contributions from New Physics. It is non-trivial for a model with spontaneous CP violation to satisfy at the same time both constraints (i) and (ii). The reason has to do with the fact that in order to have spontaneous CP violation, the Lagrangian has to be CP invariant, which requires real Yukawa couplings. Then the phase arising from the vacuum expectation value must give rise to complex quark mass matrices M_u, M_d, in such a way that the weak basis invariant I_{CP} defined by

$$I_{CP} \equiv \text{Tr}\,[H_u, H_d]^3 \tag{308}$$

does not vanish [400], with $H_u = M_u M_u^\dagger$ and $H_d = M_d M_d^\dagger$. The non-vanishing of I_{CP} has to be achieved without generating too large FCNC, so that condition (ii) is satisfied. This is the source of the difficulty in constructing realistic models of spontaneous CP violation. In the sequel we shall present a simple model which satisfies both conditions in an elegant way. We consider next some of the minimal models with spontaneous CP violation which have been considered in the literature.

A. Lee's model [9]

As previously mentioned, this is the first model of spontaneous CP violation proposed in the literature. Lee introduced two doublets but no extra symmetry in the Higgs potential. In this case it was shown by Lee that there is a non-singular region of the Higgs parameter space where the vacuum conserves electric charge but violates CP. At the time Lee suggested his model, there were only two incomplete generations, since charm had not been discovered. Therefore in Lee's model with only two generations CP arises exclusively from Higgs exchange. If we implement Lee's model in the framework of three generations, it turns out that although there is only one phase (θ) arising from the vacuum, one can generate a non-trivial CP violating phase in V_{CKM}. This can be seen by noting that the two quark mass matrices can be written

$$M_d = \frac{1}{\sqrt{2}}(v_1 Y_1^d + v_2 e^{i\alpha} Y_2^d), \qquad M_u = \frac{1}{\sqrt{2}}(v_1 Y_1^u + v_2 e^{-i\alpha} Y_2^u), \tag{309}$$

where Y_i^d, Y_i^u are real matrices. From Eq. (309) it follows that one generates a complex H_d, with its imaginary part given by:

$$\mathrm{Im}\, H_d = v_1 v_2 \left[Y_2^d Y_1^{d\dagger} - Y_1^d Y_2^{d\dagger} \right] \sin\theta \tag{310}$$

with an analogous expression for H_u. It is clear from Eq. (310) that in spite of having only one phase θ arising from the vacuum, the structure of H_d, H_u is such that the invariant I_{CP}, given in Eq. (308), does not vanish, which is sufficient to have a non-trivially complex CKM matrix. One concludes that in Lee's model with three generations, there are two sources of CP violation, Higgs exchange plus CKM mechanism. Therefore, the Lee model satisfies condition (i). However, the model has difficulty in satisfying condition (ii), since it leads, in general, to too large FCNC, unless one assumes very large Higgs masses, in the range 1–10 TeV, or invokes some suppression mechanism [433].

B. Models with natural flavour conservation in the Higgs sector

We have seen in the previous section that in the case of two Higgs doublets, the imposition of NFC in the Higgs sector eliminates the possibility of generating spontaneous CP violation. However, this is no longer true for three Higgs doublets, where it was shown [436] that one can generate genuine spontaneous CP violation. However, it was also shown [437] that in this model V_{CKM} is real, which is in disagreement with present experiment [7,448].

C. A minimal realistic model with spontaneous CP violation

Now we present what we consider to be the minimal extension of the SM where one can generate spontaneous CP violation, leading to a complex CKM matrix, having no conflict with the experimentally suppressed FCNC processes.

Let us consider an extension of the SM which consists of the addition of a vector-like singlet quark D and a complex scalar singlet S. The vector-like quark may be the down-type or the up-type. For definiteness, we consider that it is of the down-type with electrical charge $Q = -1/3$. For simplicity, we introduce a Z_2 symmetry, under which the new fields are odd,

$$Z_2: \quad D_L \longrightarrow -D_L, \quad D_R \longrightarrow -D_R, \quad S \longrightarrow -S \tag{311}$$

while all the SM fields are even. Strictly speaking, the introduction of the Z_2 symmetry is not necessary. However, its presence in the model provides a simple solution [449] to the strong CP problem [450–455]. As a result of the Z_2 symmetry, the couplings $\overline{d_{Li}} D_R \Phi$ are forbidden but the mass term and couplings:

$$M \overline{D_L} D_R + (f_j S + f_j' S^*) \overline{D_L} d_R^j + \mathrm{h.c.} \tag{312}$$

are allowed by gauge and Z_2 invariance. As a result the 4×4 quark mass matrix has the form:

$$\mathcal{M}_d = \begin{pmatrix} m_d & 0 \\ M_D & M \end{pmatrix}, \tag{313}$$

where m_d stands for the 3×3 mass matrix connecting standard quarks, the zero reflects the presence of the Z_2 symmetry and M_D is a 1×3 matrix given by:

$$(M_D)_j = \left(f_j V \exp(i\alpha) + f_j' V \exp(-i\alpha) \right). \tag{314}$$

We have assumed that CP is broken by the vacuum with:

$$\langle S \rangle = V \exp(i\alpha). \tag{315}$$

It can be shown [456] that the presence in the scalar potential of terms like $(S^2 + S^{*2})$, $(S^4 + S^{*4})$, implies that there is a region of parameters where the minimum is at the $\langle S \rangle$ of Eq. (315), with α a non-trivial phase. This vacuum breaks CP spontaneously. Note that although one has only one phase α arising from the vacuum, due to the arbitrariness of the real couplings f_j, f_j', M_d is an arbitrary 1×3 complex matrix. The matrix $\mathcal{M}_d$ is diagonalized by the usual bi-unitary transformation:

$$U_L^\dagger \mathcal{M}_d U_R = \begin{pmatrix} d & 0 \\ 0 & D \end{pmatrix} \tag{316}$$

where $d = \mathrm{diag.}(m_d, m_s, m_b)$ and D denotes the mass of the physical $Q = -1/3$ vector-like quark. The unitary matrix U_L can be written in block form:

$$U_L = \begin{pmatrix} K & R \\ S & T \end{pmatrix}. \tag{317}$$

The matrix K stands for the usual 3×3 CKM matrix, which is determined by the relation:

$$K^{-1}(m_{eff} m_{meff}^{\dagger}) K = d^2 \tag{318}$$

where

$$m_{eff} m_{meff}^{\dagger} = m_d m_d^{\dagger} - \frac{m_d M_D^{\dagger} M_D m_d^{\dagger}}{\tilde{M}^2} \tag{319}$$

where $\tilde{M}^2 = M_D M_D^{\dagger} + M^2$. The crucial point is that the two terms contributing to $m_{eff} m_{meff}^{\dagger}$ are of the same order of magnitude, since both M_D and M are $SU(2) \times U(1)$ invariant mass terms which are expected to be of the same large mass scale. As a result, a non-trivial complex CKM matrix is generated and the CP violating phase is not suppressed by the large scale of M_D and M. Of course, this can be explicitly checked by evaluating $\mathrm{Tr}\left[H_u, m_{eff} m_{meff}^{\dagger}\right]^3$. It is worth summarizing the main features of this class of models, where spontaneous CP violation is achieved through the introduction of at least one vector-like isosinglet quark and a complex singlet scalar:

(i) They provide a simple framework for having spontaneous CP violation, while at the same time generating a complex CKM matrix.

(ii) The 3×3 quark mixing matrix connecting standard quarks is not unitary. However, deviations from unitarity are naturally suppressed by the ratio m^2/M^2 where m and M denote the mass of the standard quarks and the mass of the heavy isosinglet quark(s), respectively. These deviations from unitarity lead to Z-mediated flavour changing neutral currents, (FCNC) which are again naturally suppressed by the ratio m^2/M^2. This is a general feature of models with vector-like quarks [457–466]. If the mass of the new quark is of order 1 TeV, the FCNC are sufficiently suppressed so that they do not enter in conflict with the stringent limits on $\Delta S = 2$ tree level transitions. Yet, one may have significant contributions to B_d–$\overline{B_d}$ mixing [467–473] and B_s–$\overline{B_s}$ mixing which could be detected at LHCb and in super-B factories.

6.3. CP-violating quantities from the scalar potential

As shown in Section 5.11, in the 2HDM one may define the 'Higgs basis' as a basis (H_1, H_2) for the scalar $SU(2)$ doublets such that the neutral component of H_1 has real and positive vacuum expectation value $v/\sqrt{2}$ while H_2 has vanishing vev. In the Higgs basis the doublets are given by Eq. (220), where $G^{\pm}$ and G^0 are the three 'would-be Goldstone bosons' while $H^{\pm}$ are the two physical charged scalars. The three physical neutral scalars $S_{1,2,3}$ are the linear combinations of H, R, and I given by Eq. (373), where T is a matrix of SO(3).

However, as was also stressed in Section 5.11, the Higgs basis is *not* uniquely defined, because, when one rotates H_2 as in Eq. (221), the conditions for the Higgs basis, *viz.* the vev of H_2 being zero and the vev of H_1 being real and positive, remain satisfied. The rotation (221) implies

$$\begin{pmatrix} R \\ I \end{pmatrix} \to O \begin{pmatrix} R \\ I \end{pmatrix}, \tag{320}$$

where $O \in$ SO(2), *cf.* Eq. (222).

Under a CP transformation, besides the change in the sign of the space coordinates, H_1 transforms to its complex conjugate[24] while H_2 transforms to its complex conjugate apart from an arbitrary phase:

$$H_1\left(t, \vec{r}\right) \overset{CP}{\to} H_1^*\left(t, -\vec{r}\right), \qquad H_2\left(t, \vec{r}\right) \overset{CP}{\to} e^{i\zeta} H_2^*\left(t, -\vec{r}\right). \tag{321}$$

Thus, in a CP transformation,

$$\begin{pmatrix} R \\ I \end{pmatrix} \overset{CP}{\to} O' \begin{pmatrix} R \\ I \end{pmatrix}, \tag{322}$$

where $O' \in$ O(2) but $\det O' = -1$.

Our task in this section consists in finding quantities which depend solely on the scalar potential and *are invariant under basis transformations* but *change sign under CP*. The solution to this problem [425] hinges on the matrix

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{323}$$

---

[24] In second-quantized field theory one must employ, instead of the complex conjugate, the transpose of the Hermitian conjugate.

which has the property

$$O \, \epsilon \, O^T = \epsilon \det O \tag{324}$$

for any $O \in O(2)$.

The scalar potential of a general 2HDM in the Higgs basis is in Eq. (376). The quantities $\bar{m}_{11}^2$ and $\bar{m}_{12}^2$ are functions of the vev v and of the quartic couplings $\bar{\lambda}_1$ and $\bar{\lambda}_6$, respectively, through Eq. (224). When one expands the scalar potential as a function of the component fields of H_1 and H_2, one finds terms of the following forms[25]:

$a^S S$, where a^S is a real coefficient and S is a neutral combination of fields which is invariant under both the basis transformation of Eq. (320) and the CP transformation in Eq. (322). For instance, S may be H^2, H^3, G^{02}, $H^- H^+$, $G^- G^+$, $R^2 + I^2$, and so on.

$\left(b_1^S R + b_2^S I\right) S$, where b_1^S and b_2^S are real coefficients.

$\left[c_{11}^S R^2 + c_{22}^S I^2 + \left(c_{12}^S + c_{21}^S\right) RI\right] S$, where the c_{ab}^S ($a, b = 1, 2$) are real coefficients and, without loss of generality, $c_{21}^S = c_{12}^S$.

Considering for instance V_2 in Eqs. (378)–(380), we see that $b_1^H = v^2 \operatorname{Re} \bar{\lambda}_6$, $b_2^H = -v^2 \operatorname{Im} \bar{\lambda}_6$, and that $c_{11} = -c_{22} = (v^2/4)\operatorname{Re} \bar{\lambda}_5$, $c_{12} = c_{21} = -(v^2/4)\operatorname{Im} \bar{\lambda}_5$ (in V_2 there are also terms $m_+^2 H^- H^+$, $(v^2/2)\bar{\lambda}_1 H^2$, and $(m_+^2/2 + v^2\bar{\lambda}_4/4)(R^2 + I^2)$, which are of the form $a^S S$). If we also consider V_3 in Eq. (383f), we find for instance $b_1^{H^- H^+} = v \operatorname{Re} \bar{\lambda}_7$ and $b_2^{H^- H^+} = -v \operatorname{Im} \bar{\lambda}_7$.

It is clear that, under the basis transformation of Eq. (320),

$$\begin{pmatrix} b_1^S \\ b_2^S \end{pmatrix} \to O \begin{pmatrix} b_1^S \\ b_2^S \end{pmatrix}, \qquad \begin{pmatrix} c_{11}^S & c_{12}^S \\ c_{21}^S & c_{22}^S \end{pmatrix} \to O \begin{pmatrix} c_{11}^S & c_{12}^S \\ c_{21}^S & c_{22}^S \end{pmatrix} O^T, \tag{325}$$

and similarly, under the CP transformation of Eq. (322),

$$\begin{pmatrix} b_1^S \\ b_2^S \end{pmatrix} \overset{CP}{\to} O' \begin{pmatrix} b_1^S \\ b_2^S \end{pmatrix}, \qquad \begin{pmatrix} c_{11}^S & c_{12}^S \\ c_{21}^S & c_{22}^S \end{pmatrix} \overset{CP}{\to} O' \begin{pmatrix} c_{11}^S & c_{12}^S \\ c_{21}^S & c_{22}^S \end{pmatrix} O'^T. \tag{326}$$

Using the matrix ϵ of Eq. (323), it is then easy to construct basis-invariant, CP-violating quantities like

$$\left(b_1^S, \quad b_2^S\right) \epsilon \begin{pmatrix} b_1^{S'} \\ b_2^{S'} \end{pmatrix} \quad \text{or} \quad \left(b_1^S, \quad b_2^S\right) \begin{pmatrix} c_{11}^{S'} & c_{12}^{S'} \\ c_{21}^{S'} & c_{22}^{S'} \end{pmatrix} \epsilon \begin{pmatrix} b_1^S \\ b_2^S \end{pmatrix}.$$

For instance, from the tensors explicitly given above one obtains

$$\left(b_1^H, \quad b_2^H\right) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \epsilon \begin{pmatrix} b_1^H \\ b_2^H \end{pmatrix} = -\frac{v^6}{4} \operatorname{Im}\left(\bar{\lambda}_6^2 \bar{\lambda}_5^*\right), \tag{327}$$

$$\left(b_1^{H^- H^+}, \quad b_2^{H^- H^+}\right) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \epsilon \begin{pmatrix} b_1^{H^- H^+} \\ b_2^{H^- H^+} \end{pmatrix} = -\frac{v^4}{4} \operatorname{Im}\left(\bar{\lambda}_7^2 \bar{\lambda}_5^*\right), \tag{328}$$

$$\left(b_1^H, \quad b_2^H\right) \epsilon \begin{pmatrix} b_1^{H^- H^+} \\ b_2^{H^- H^+} \end{pmatrix} = v^3 \operatorname{Im}\left(\bar{\lambda}_6 \bar{\lambda}_7^*\right). \tag{329}$$

One easily sees that, indeed,

$$J_1 \propto \operatorname{Im}\left(\bar{\lambda}_6^2 \bar{\lambda}_5^*\right), \qquad J_2 \propto \operatorname{Im}\left(\bar{\lambda}_7^2 \bar{\lambda}_5^*\right), \qquad J_3 \propto \operatorname{Im}\left(\bar{\lambda}_6 \bar{\lambda}_7^*\right), \tag{330}$$

are the only basis-invariant CP-violating quantities in the potential of Eq. (376).

By taking into account Eq. (381) and $T \in SO(3)$, the quantity in Eq. (327) may be written [474]

$$-\frac{v^6}{2} \operatorname{Im}\left(\bar{\lambda}_6^2 \bar{\lambda}_5^*\right) = M_{23}\left[(M_{13})^2 - (M_{12})^2\right] + (M_{22} - M_{33}) M_{12} M_{13}$$

$$= \left(m_1^2 - m_2^2\right)\left(m_1^2 - m_3^2\right)\left(m_2^2 - m_3^2\right) T_{11} T_{12} T_{13}. \tag{331}$$

This shows that there is CP violation in the 2HDM if all the matrix elements in the first row of the mixing matrix T are non-zero and, moreover, the three physical neutral scalars are non-degenerate.[26] This can be confirmed by considering, for

---

[25] We do not consider terms containing $G^{\pm} H^{\mp}$, which are trickier to handle.

[26] If two of the physical neutral spin-0 fields, say S_j and S_k, are degenerate, then the matrix T may be redefined in such a way that either T_{1j} or T_{1k} becomes zero, which is a CP-conserving situation.

instance, the interactions in Eq. (375h). They tell us that, for $j = 1, 2, 3$, if $T_{1j} \neq 0$, then S_j is a scalar; therefore, $T_{11}T_{12}T_{13} \neq 0$ implies that all three physical neutral spin-0 fields are scalars, contrary to our knowledge that in the CP-conserving 2HDM one of those fields must be a pseudoscalar.

If we then consider the interactions in line (383e) and take into account Eq. (384), we find that

$$v^2 \text{Im} \left(\bar{\lambda}_6 \bar{\lambda}_7^* \right) = \sum_{j,k,l=1}^{3} \epsilon_{jkl} m_k^2 T_{1k} c_j T_{1l}. \tag{332}$$

Indeed, the interactions in line (383e) indicate that, if $c_j \neq 0$, then S_j is a scalar. CP conservation would then require either S_k or S_l ($j \neq k \neq l \neq j$) to be a pseudoscalar; correspondingly, either T_{1j} or T_{1k}, respectively, ought to vanish.

One should note that in the generic 2HDM J_1, J_2, and J_3 are not all independent. One might in principle choose only J_1 and J_2 as the two independent basis-invariant signals of CP violation. However, $J_1 = J_2 = 0$ in a particular 2HDM with $\bar{\lambda}_5 = 0$, yet there might still be CP violation through $J_3 \neq 0$. Different particular cases of the 2HDM may require different choices for a minimum set of independent J-invariant and, to cover all the particular cases, we need J_1, J_2, and J_3, even though they are not all independent.

An important point made in Refs. [431,432,475] concerns spontaneous symmetry breaking (SSB). Quantities like J_1, J_2, and J_3 involve the vev v and therefore refer to CP violation *after* SSB; the Lagrangian before SSB does not involve v. Of course, all laboratory CP-violating observables concern CP violation after SSB, but early Universe phenomena, such as leptogenesis, involve CP violation *before* SSB. The quantities J_1, J_2, and J_3 are the only ones needed to study CP violation in the scalar potential *after* SSB. In an earlier section, we found that one needs the four invariants $I_{1,2,3,4}$ in order to study CP violation *before* SSB. Interestingly, four quantities are needed in order to study CP violation in the generic 2HDM before SSB, but only three quantities are required after SSB.

The comparison between the I's and the J's is also of theoretical interest. If some I is non-vanishing, then there is CP violation at the Lagrangian level. If all the I's vanish but some J do not, then there is spontaneous CP violation. The theory is CP conserving only when all the I's and J's vanish. For discussions see, for instance, [375,376,397,476].

We end this section with an open problem. Through the minimization conditions, one can determine (at least implicitly, or numerically) the vevs in terms of the couplings of the scalar potential. Therefore, one should be able to write $J_{1,2,3}$ in terms of $I_{1,2,3,4}$ together with, possibly, some CP-conserving quantities. As far as we know, this has not yet been achieved.

6.4. CP-violating quantities with scalars and fermions

6.4.1. The general method

In this section we discuss a systematic method for the construction of basis-invariant quantities which was developed by Botella and Silva [370]. The main focus of their work was on basis-invariant signals of CP violation, which we designate by J-invariants. Yet, they pointed out that their strategy applies to any other property; for example, their method has later been applied [477] to R-parity in supersymmetric models. The method applies with any gauge group G and can also be used in an effective field theory including nonrenormalizable interactions.

To illustrate the main idea, we start with a generic Lagrangian of the form

$$\mathcal{L}_I = \left(\sum_{i,j} g_{ij} \alpha_i \beta_j + \sum_{k,l} h_{kl} \alpha_k \gamma_l \right) \Phi + \text{H.c.}, \tag{333}$$

where the $g_{i,j}$ and $h_{k,l}$ are coupling constants and the α, β, and γ are field operators with their respective $U(\alpha), U(\beta)$, and $U(\gamma)$ flavour spaces, and also transforming like some multiplet of the gauge group G. As an example, in the SM we have, after SSB,

$$\mathcal{L}_I = \sum_{i,j} \left(\bar{u}_L, \quad \bar{d}_L \right)_i \left[(M_u)_{ij} \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_{Rj} + (M_d)_{ij} \begin{pmatrix} 0 \\ 1 \end{pmatrix} d_{Rj} \right] + \text{H.c.}, \tag{334}$$

and the flavour spaces are $U(3)_L$, $U(3)_{uR}$, and $U(3)_{dR}$, respectively. In perturbation theory, one can generate interactions mediated by any power of $\mathcal{L}_I$. For example, to second order in perturbation theory, we will find interactions mediated by

$$(g_{ij} \alpha_i \beta_j \Phi) (h_{kl} \alpha_k \gamma_l \Phi). \tag{335}$$

Hence, a given property of the theory (say CP violation) may show up at some order of perturbation theory as a suitable product of couplings.

Under a basis transformation the couplings transform as,

$$g_{ij} \to \sum_{kl} U(\alpha)_{ki} g_{kl} U(\beta)_{lj},$$

$$h_{ij} \to \sum_{kl} U(\alpha)_{ki} h_{kl} U(\gamma)_{lj}. \tag{336}$$

The strategy in looking for basis invariant quantities consists in taking products of couplings (as in the perturbative expansion), contracting over the internal flavour spaces and taking a trace at the end. For example, the quantities

$$H_u = M_u M_u^\dagger, \qquad H_d = M_d M_d^\dagger, \qquad H_u H_d \tag{337}$$

are tensors in the $U(3)_L$ space and their traces are weak-basis invariants. The same is true for the trace of the $U(3)_{uR}$ tensor $M_u^\dagger M_u$.

In so doing, we have already traced over the basis transformations that could lead to the spurious phases brought about by basis transformations. Therefore, the imaginary parts of such traces are unequivocal signs of CP violation [478]. For example, the J-invariant of the three-family SM is

$$J \propto \mathrm{Im}\left[\mathrm{tr}\left(H_u H_d H_u^2 H_d^2\right)\right]. \tag{338}$$

A detail concerns spontaneous symmetry breaking (SSB). After SSB, the physical degrees of freedom of the neutral scalars are shifted fields η_i, related to the original ones (ϕ_i) by the vevs (v_i) as

$$\phi_i = v_i + \eta_i. \tag{339}$$

This reparameterizes a Lagrangian term; for instance,

$$\mu_{ij}\phi_i^\dagger \phi_j = \mu_{ij} v_i^* v_j + \mu_{ij} v_i^* \eta_j + \mu_{ij} \eta_i^\dagger v_j + \mu_{ij} \eta_i^\dagger \eta_j. \tag{340}$$

In this way, v_i becomes an integral part of new couplings like $\left(\mu_{ij} v_i^*\right) \eta_j$. Thus, in the construction of basis invariants involving the scalar sector, one must consider combinations of couplings both with and without vevs. This greatly simplifies the study of the scalar sector over the analysis in the previous section. In addition, the minimization conditions provide relations between the couplings in the scalar potential which must be used in identifying the correct number of independent CP-violating invariants.

This discussion motivates the following prescription for the construction of J-invariants:

- identify all the scalar and fermion flavour spaces in the theory;
- make a list of all the couplings according to their transformation properties under weak basis transformations, including the vacuum expectation values (which transform as vectors under the scalar basis change); make use of the stationarity conditions of the scalar potential to reduce the number of parameters;
- construct invariants by contracting over internal flavour spaces in all possible ways, taking traces at the end (in order to be systematic it is advisable to do this firstly in the fermion sector, using this to define new scalar tensors, and then perform the analysis of the scalar sector);
- take the imaginary part to obtain a basis-invariant signal of CP violation.

In general, a minimal set of CP-violating quantities is not easy to find, since one could in principle go to arbitrary order in perturbation theory. The identification of the number of independent J-invariants may, at best, be done in a case-by-case way through a careful study of the sources of CP violation in the model. Moreover, different particular cases of a model may require different choices for the minimum set of fundamental J-invariants.

6.4.2. Invariants with scalars

We can use the technique of the previous section to reproduce the CP-violating invariants of the scalar sector after SSB. One finds [370,432],

$$J_1 \propto \mathrm{Im}\left(\sum_{a,\ldots,f=1}^{2} v_a^* Y_{ae} v_b^* Y_{bf} Z_{ecfd} v_c v_d\right), \tag{341}$$

$$J_2 \propto \mathrm{Im}\left(\sum_{a,\ldots,h=1}^{2} v_b^* v_c^* Z_{bgge} Z_{chhf} Z_{eafd} v_a v_d\right), \tag{342}$$

$$J_3 \propto \mathrm{Im}\left(\sum_{a,b,c,d=1}^{2} v_a^* Y_{ab} Z_{bd,dc} v_c\right). \tag{343}$$

These, it may be shown, are equivalent to the J invariants defined in Eqs. (330).

6.4.3. Invariants with scalars and fermions

In looking for invariants probing CP violation and involving both scalars and fermions we follow Botella and Silva [370]. We start from the Yukawa Lagrangian in Eq. (225). The Yukawa-coupling matrices Y_a^d ($a = 1, 2$) are 3×3 complex matrices. Their rows (columns) are acted upon by unitary $U(3)_L$ ($U(3)_{dR}$) transformations on the space of left-handed quark doublets (right-handed down-type quark singlets). Similarly, the Yukawa matrices Y_a^u are 3×3 complex matrices. Its rows (columns) are acted upon by unitary $U(3)_L$ ($U(3)_{uR}$) transformations on the space of left-handed quark doublets (right-handed, up-type singlets). Combinations such as $Y_a^d Y_b^{d\dagger}$ and $Y_a^u Y_b^{u\dagger}$ have both indices in the left-handed-doublet space. As a result, in

$$T_{ab}^d = \mathrm{tr}\left(Y_a^d Y_b^{d\dagger}\right), \tag{344a}$$

$$T_{ab}^u = \mathrm{tr}\left(Y_a^u Y_b^{u\dagger}\right), \tag{344b}$$

all quark spaces have been traced over. These quantities depend only on the scalar indices a and b. As a result, they can be combined with the v_a, Y_{ab}, and Z_{abcd} to construct weak-basis invariants depending on both scalars and fermions. The lowest order basis-invariant measures of CP violation involving both scalars and fermions are [370]

$$J^d = \mathrm{Im}\left(\sum_{a,b,c=1}^{2} v_a v_b^* Y_{bc} T_{ca}^d\right) \tag{345a}$$

$$J^u = \mathrm{Im}\left(\sum_{a,b,c=1}^{2} v_a v_b^* Y_{bc} T_{ca}^u\right). \tag{345b}$$

Expressing these invariants in the quark mass basis, we find

$$J^d \propto \mathrm{Im}\left[m_{12}^2 \sum_{i=1}^{n_G} m_{di} (N_d)_{ii}^*\right], \tag{346a}$$

$$J^u \propto \mathrm{Im}\left[m_{12}^2 \sum_{i=1}^{n_G} m_{ui} (N_u)_{ii}^*\right], \tag{346b}$$

where n_G is the number of generations and m_{di} (m_{ui}) is the mass of the i-th down-type (up-type) quark. In the simplest case of $n_G = 1$, there is no CP violation in the CKM matrix, there are only two independent CP-violating invariants – which we may take to be J_1 and J_3 – in the scalar sector, and there are only two independent invariants – J^d and J^u – involving both the scalars and the fermions. For $n_G = 2$, there is still no CP violation in the CKM matrix, J_1 and J_3 apply to the scalar sector, and there are a total of eight invariants in the scalar–fermion interactions, which were explicitly constructed in Ref. [370]. For $n_G = 3$, J_1 and J_3 apply to the scalar sector, there are 18 invariants in the scalar–fermion interactions, and there is now also one CP-violating invariant in the CKM matrix, given by Eq. (338).

6.5. CP basis invariants and the bilinear formalism

The basis invariant quantities of Eqs. (274), (275), (284) and (288) determine whether or not a given 2HDM scalar potential is explicitly CP-conserving. They have extremely simplified expressions in terms of the bilinear formalism introduced in Section 5.3, as was shown in Refs. [375,376]. We follow the notation of the Heidelberg group [372,373,375] and introduce the vectors ξ and η and the matrix E, given, in terms of the parameters of the 2HDM potential defined in Eq. (98), as

$$\boldsymbol{\xi} = \frac{1}{2}\begin{pmatrix} -2\,\mathrm{Re}(m_{12}^2) \\ 2\,\mathrm{Im}(m_{12}^2) \\ m_{11}^2 - m_{22}^2 \end{pmatrix}, \qquad \boldsymbol{\eta} = \frac{1}{4}\begin{pmatrix} \mathrm{Re}(\lambda_6 + \lambda_7) \\ -\mathrm{Im}(\lambda_6 + \lambda_7) \\ \frac{1}{2}(\lambda_1 - \lambda_2) \end{pmatrix},$$

$$E = \frac{1}{4}\begin{pmatrix} \lambda_4 + \mathrm{Re}(\lambda_5) & -\mathrm{Im}(\lambda_5) & \mathrm{Re}(\lambda_6 - \lambda_7) \\ -\mathrm{Im}(\lambda_5) & \lambda_4 - \mathrm{Re}(\lambda_5) & -\mathrm{Im}(\lambda_6 - \lambda_7) \\ \mathrm{Re}(\lambda_6 - \lambda_7) & -\mathrm{Im}(\lambda_6 - \lambda_7) & \frac{1}{2}(\lambda_1 + \lambda_2) - \lambda_3 \end{pmatrix}. \tag{347}$$

Then, the I invariants of Section 6.1.1 may be written as[27]

$$I_1 = (\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}} \cdot E\boldsymbol{\xi}, \tag{348}$$

---

[27] In fact, these four invariants are linear combinations of those in Section 6.1.1, but their usage is equivalent.

$$I_2 = (\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}} . E\boldsymbol{\eta}, \tag{349}$$

$$I_3 = (\boldsymbol{\xi} \times (E\boldsymbol{\xi}))^{\mathrm{T}} . E^2\boldsymbol{\xi}, \tag{350}$$

$$I_4 = (\boldsymbol{\eta} \times (E\boldsymbol{\eta}))^{\mathrm{T}} . E^2\boldsymbol{\eta}. \tag{351}$$

The extremely simple form of these equations, and the appearance of external products between vectors in them, leads to meaningful geometrical interpretations for CP-violating/preserving potentials [375,376]. See also [380,421].

As for the basis invariant quantities J_i of Eqs. (341)–(343), which determine whether a given vacuum preserves or breaks CP, they too can be written, in an extremely simplified manner, in terms of the bilinear formalism. Let us introduce the vector $\mathbf{r} = (r_1, r_2, r_3)^T$, with the r_i defined in Eq. (105). Then [375],

$$J_1 = (\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}} . \langle \mathbf{r} \rangle, \tag{352}$$

$$J_2 = (\boldsymbol{\xi} \times (E\boldsymbol{\xi}))^{\mathrm{T}} . \langle \mathbf{r} \rangle, \tag{353}$$

$$J_3 = (\boldsymbol{\eta} \times (E\boldsymbol{\eta}))^{\mathrm{T}} . \langle \mathbf{r} \rangle, \tag{354}$$

where $\langle \mathbf{r} \rangle$ corresponds to the vector $\mathbf{r}$ evaluated at some stationary point of the theory.[28]

6.6. CP violation and symmetries

As discussed in Section 5.6, there are six classes of 2HDM symmetry-constrained potentials. Those symmetries, and their impact on the parameters of the potential, were shown in Table 5. Leaving aside the extension of these symmetries to the Yukawa sector – where, as we have seen in Section 5.14.3, each symmetry has very different consequences – each of those models corresponds to a very specific and different type of scalar physics. For instance, a model with a Peccei–Quinn [12] $U(1)$ symmetry (class IV) can have a vacuum with a massless scalar, an axion. That is not possible, whatever the vacuum, for models with the Z_2 or CP1 symmetries (classes V and VI).

Another aspect for which the six classes behave very differently concerns the possibility of CP breaking – explicitly or spontaneously – for each potential. This question is better handled using the bilinear formalism formulae presented in Section 6.5. A systematic analysis of all possible potentials was carried out in [397]. Briefly, this consists of:

- One may wish to consider models with an *exact* symmetry, out of the six considered, or to softly-break that symmetry via the inclusion of generic dimension-two terms (real or, in the case of m_{12}^2, even complex).
- Prior to spontaneous symmetry breaking, it is necessary to determine whether CP is a valid symmetry of the potential, or if it is explicitly broken. This is best achieved computing the four I invariants of Eqs. (348)–(351).
- For any given model, the extremum conditions need to be solved, to prove that a certain vacuum, which might break CP spontaneously, is possible.
- Having proved that a given sets of vevs is a possible solution of the extremum conditions, one must verify whether that vacuum effectively breaks CP. This is best achieved through the calculation of the three invariant quantities of Eqs. (352)–(354).

If any of the I_i invariants is different from zero, the potential is not CP-conserving, and CP is not a defined symmetry of the potential (for *all* possible CP definitions of the form of Eq. (134)). If all $I_i = 0$ then the potential is CP-conserving; a given vacuum is CP-conserving if and only if all invariants $J_i = 0$. A CP-conserving scalar sector has some very distinct physical properties: there is a well-defined pseudoscalar state A, and there are two well-defined CP-even states h and H; as a result, though triple vertices of the form ZZh and ZZH are possible, no vertex like ZZA is allowed.

In [397] the bilinear formalism of Refs. [373–375] was used to compute all invariant quantities. As shown in the previous section, the formulae for the I and J CP basis invariants are extremely simple in the bilinear formalism, and that allowed a general analysis of all possible models, without even an explicit calculation of the vevs (a major simplification, since solving the extremum conditions can be analytically impossible in some models). The conclusions of the study of [397] are summarized in Table 6. Some obvious observations are drawn from this table:

- Any scalar potential with an exact symmetry is CP-conserving.
- If the symmetry of the potential is *continuous* (i.e., classes I, II and IV) no CP violation is possible, be it explicitly or spontaneously, even with generic soft-breaking terms.
- Discrete symmetries allow for the possibility of spontaneous symmetry breaking; however, with the exception of the class VI model, that is only possible via the inclusion of a soft-breaking term.[29]

Though a general analysis of CP breaking in the scalar sector is achievable, the study of the CP properties of the theory requires that one takes into account the Yukawa terms. And some of the models of Table 6 may well end up having CP conserved in the scalar sector but violated by the fermion–scalar interactions, much like the SM.

---

[28] These, too, are linear combinations of the invariants of Eqs. (341)–(343).

[29] For the potential with a Z_2 symmetry this was well-known for a long time [433].

Table 6
CP properties of the six symmetry-constrained classes of 2HDM scalar potentials. We consider both the case of exact symmetries and their soft breaking via dimension-two terms. In this table, "Yes" means that it is possible to choose the parameters of the potential such as to enable that particular form of CP violation.

Symmetry class	Exact		Softly-broken	
	Explicit CPV	Spontaneous CPV	Explicit CPV	Spontaneous CPV
I - $U(2)$	–	–	–	–
II - CP3	–	–	–	–
III - CP2	–	–	Yes	Yes
IV - $U(1)$	–	–	–	–
V - Z_2	–	–	Yes	Yes
VI - CP1	–	Yes	Yes	Yes

6.7. Two models with an original source of CP violation

In the SM, the source of CP violation is *explicit breaking*, via Yukawa matrices which are complex. This leads to a complex CKM matrix, and in turn to a non-zero Jarlskog invariant. In such models, then, CP is not defined *a priori*, since it is not a symmetry of the Lagrangian—it is broken by hard, dimension four, Yukawa terms. The MSSM can be another example: in [479] it was shown that radiative Higgs-sector CP violation can be quite large in the MSSM. In effect, after quantum effects are included, one obtains, as a limit, the potential of a general CP-violating 2HDM.

In the 2HDM, as has been explained in previous sections, there is the possibility of *spontaneous breaking* of CP. A complex phase can appear in the vacuum of the theory and lead to CP breaking. The presence of such a phase, however, is not sufficient to guarantee CP violation—one needs to calculate the J invariants of Eqs. (341)–(343) and verify whether at least one of them is non-zero. An example of a model where this happens is the classic paper by Branco and Rebelo [433]—they considered a scalar potential with a Z_2 symmetry, extended to the fermion sector in a particular way, and also required CP conservation at the Lagrangian level. Hence, all of the model's parameters are real. Since the exact symmetry forbids any CP-breaking vacuum, they added a *real* soft breaking term to the scalar potential, and thus generated a complex CKM matrix. In such a model some (or all) of the J_i invariants are non-zero, and there is CP violation in the scalar–scalar interactions, as well as in the fermion sector.

In this section we will briefly describe two versions of the 2HDM in which CP violation arises in ways which are different from the two usual ones described above.

6.7.1. The CP3 model

In [391] a 2HDM was built with the CP3 symmetry extended to the fermion sector. We recall (see Eq. (151)) that this corresponds to a transformation on the scalar fields of the form

$$\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \Phi_1^* \\ \Phi_2^* \end{pmatrix}. \tag{355}$$

In [391] it was shown that the only value of the angle θ which leads to six massive quarks (and three massive charged leptons) is $\theta = \pi/3$. Also, the transformation laws of the quark fields under this symmetry are uniquely determined (again, by the requirement of six massive quarks). One finds that the Yukawa coupling matrices for the down quarks have a very simple form,

$$Y_1^d = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & -a_{11} & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix}, \qquad Y_2^d = \begin{bmatrix} a_{12} & -a_{11} & -a_{23} \\ -a_{11} & -a_{12} & a_{13} \\ -a_{32} & a_{31} & 0 \end{bmatrix}, \tag{356}$$

in a special basis where all the a_{ij} coefficients are real. Indeed, in that basis *all* the parameters of the potential are real. An analogous form is found for the Y^u matrices for the up quarks, with different coefficients b_{ij}.

The scalar potential of the CP3 model (class II of Table 5) has a continuous symmetry which will be broken if both scalar fields acquire a vev—which then implies the appearance of a massless axion. As such, one needs to add *real* soft breaking terms to the potential, which give mass to the would-be axion. One then finds that, with such a soft breaking term, a vacuum with a complex relative phase between the vevs is possible—of the form $\langle\Phi_1\rangle = v_1$, $\langle\Phi_2\rangle = v_1 e^{i\delta}$. However, as shown in Table 6, even such a vacuum does not provoke CP breaking in the scalar sector—all J_i invariants are equal to zero, even if $\delta \neq 0$. Still, it was shown in [391] that the Jarlskog invariant is directly proportional to $\sin\delta$. As such, in this model a vacuum with a complex relative phase is possible and does lead to CP violation, even if its scalar sector preserves CP. This then, is a new type of CP violation:

- The Lagrangian does preserve CP because there is a basis for which all of its parameters are real. Thus, no explicit CP breaking occurs, as in the SM.

- The scalar sector preserves CP because, even for a vacuum with a complex phase δ, all of the basis-invariant quantities J_i which measure CP violation in the scalar–scalar interactions are zero.
- However, CP breaking *does* occur, and it is spontaneous, since $\delta \neq 0$ implies a non-zero Jarlskog invariant.

It is interesting that the scalar sector does the deed (spontaneously break the symmetry) but it is the fermion sector which pays the consequences (providing CP violation). To the best of our knowledge, this type of CP violation is unheard of in the literature—it arises spontaneously, but the scalar sector remains CP-conserving.

The model's interest is increased by the limited number of free parameters it contains—12 independent Yukawa couplings, and 6 scalar potential parameters (counting the soft breaking terms). It is easy to fit the six quark masses and a quartet of moduli of CKM matrix elements (in theory all one needs to fit the entire CKM matrix [430]). This model has FCNC (the matrices Y_1^d and Y_2^d do not in general commute with one another), so care must be taken to ensure that it is in agreement with the stringent FCNC bounds which arise from meson physics. Remarkably, the model does manage to fit the mass differences of the K, B_s and B_d mesons, as well as the ϵ_K parameters and the unitarity triangle angle β, as obtained from B meson decays. However, the fits performed lead to a value of the Jarlskog invariant at least three orders of magnitude smaller than the SM value; this would seem to indicate a relation between the unitarity triangle angles, $\alpha \simeq \beta$, which goes against observational data [293].

6.7.2. A specific Z_3 model

Recently [427] a model was proposed wherein one extended a Z_3 symmetry in the scalar sector to the Yukawa terms, in a particular way. The scalar sector therein resulting is identical to that of the Peccei–Quinn model; the Yukawa matrices one obtains are extremely simple, having the form

$$Y_1^d, Y_1^u \sim \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \times \\ \times & \times & 0 \end{pmatrix}, \qquad Y_2^d \sim \begin{pmatrix} \times & \times & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \times \end{pmatrix}, \qquad Y_2^u \sim \begin{pmatrix} 0 & 0 & \times \\ \times & \times & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where the symbol $\times$ denotes a non-zero matrix entry, in general complex. These Yukawa matrices were first found in [428].

If, on top of the Z_3 symmetry, we apply the standard CP transformation, all terms in the Lagrangian are forced to be real. Again, since the matrices Y_1^d and Y_2^d (and Y_1^u and Y_2^u) do not in general commute, this model has FCNC.

Once more, the Peccei–Quinn model has an axion if both scalar fields acquire a non-zero vev. Thus, it is necessary to add a soft-breaking m_{12}^2 term to the potential. However, unlike what happened for the CP3 model of the previous section, it is impossible to obtain a vacuum of the form $\langle \Phi_1 \rangle = v_1$, $\langle \Phi_2 \rangle = v_1 e^{i\delta}$, with $\delta \neq 0$, *unless* the soft breaking term m_{12}^2 is itself complex (and has phase $-\delta$).[30] However, even despite the introduction of a complex soft breaking term, the scalar potential remains CP-conserving, *before of after* spontaneous symmetry breaking—check Table 6. As such, this too is an unusual source of CP violation:

- The model has a scalar sector which is CP-conserving.
- CP violation occurs due to an *explicit* breaking of the CP symmetry, via a complex coefficient in the scalar potential.
- However, unlike the case of the SM, this CP breaking is *soft*, not *hard*, since the complex coefficient is a dimension-two term of the Lagrangian.

The FCNC which arise in this model have an extra surprise: it is easy to show that all FCNC couplings are real, so that no CP violation occurs in FCNC interactions. Indeed, this model reproduces perfectly the type of CP violation one obtains in the SM—all quark masses, CKM matrix elements, meson mass differences and CP-violating quantities can be fitted with the model's eleven parameters (see [427]). Remarkably, one is capable of fitting all observables with some of the scalar masses as low as ∼150 GeV, even in the presence of FCNC. Further, the model's scalar sector also satisfies constraints on New Physics arising from the oblique parameters of Appendix D, as well as the LEP2 constraints on the lightest Higgs mass [7].

As such, this model suggests that even though the CKM mechanism for CP violation is well established experimentally, the origin for a complex CKM matrix need not necessarily be that of the SM—a hard breaking of CP through dimension-four terms. In the 2HDM with a Z_3 symmetry, the possibility exists that the origin of all CP violation is a soft, dimension-two, term.

7. Recent results from the LHC

The 35 inverse picobarns of the 2010 Large Hadron Collider (LHC) run were not useful in constraining the Higgs sector. Over the summer of 2011, the LHC collected more than one inverse femtobarn of data, and at the Hadron Collider Workshop in November 2011 the ATLAS and CMS Collaborations presented [480] the combined results for that summer run. They presented upper bounds on the cross section for Standard Model Higgs boson production as a function of the Higgs boson mass. The main plot, the so-called "Brazil bands" plot, is shown in Fig. 21. The solid line in this plot gives the upper bound, at 95% confidence level, on the measured cross section for Higgs boson production relative to the Standard Model cross section, *assuming Standard Model decay signatures*. One cannot quickly draw conclusions about 2HDMs from this plot,

---

[30] The reason for this drastic difference in behaviour is the fact that in the CP3 model the λ_5 quartic coupling is non-zero, unlike what happens for the Peccei–Quinn potential.
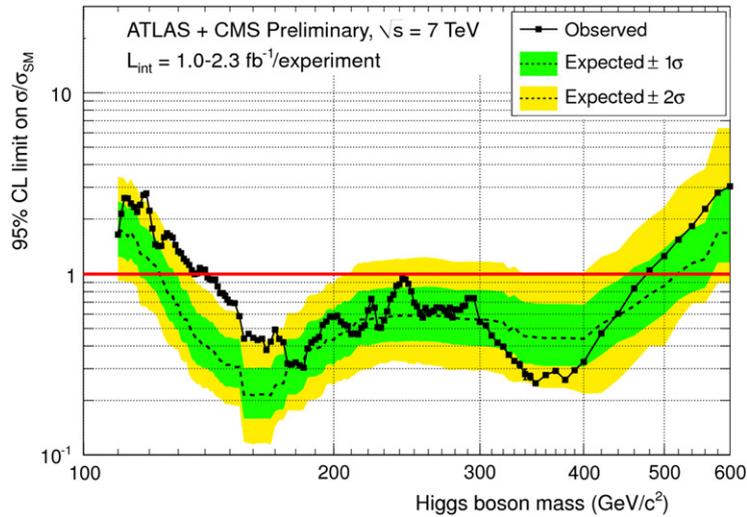
Fig. 21. Combined results of the ATLAS and CMS Collaborations presented at the Hadron Collider Workshop in November 2011. The solid line is the experimental upper bound, at 95% confidence level, on the cross section for Higgs boson production divided by the cross section of the Standard Model. The dashed line is the expected 95% exclusion bound, and the green and yellow bands are the one- and two-standard deviation bands around that expected bound.

which is meaningful only for Standard Model Higgs boson decays. As seen throughout this report, in many 2HDMs, and for a substantial region of the parameter space, one expects very different branching ratios from those of the Standard Model.

Nonetheless, there is one circumstance in which the plot can directly lead to information about 2HDMs. That is when the primary decay modes of the h and H of the (CP-conserving) 2HDM are into W^+W^- and ZZ. In the Standard Model, these decays dominate for Higgs boson masses above 130 GeV, and thus the branching ratios into W^+W^- and ZZ would be the same, since the total branching ratio is 100% and the only difference between the W^+W^- and ZZ branching ratios would be the $SU(2)$ couplings and phase–space.

If the primary decays are into W^+W^- and ZZ, then the only differences in the event rate between the Standard Model Higgs boson and the h and H scalars of the 2HDM would occur in the production cross section, which proceeds through top-quark loops. For the h (H), this is equal to the Standard Model cross section times $\cos^2\alpha/\sin^2\beta$ ($\sin^2\alpha/\sin^2\beta$). Clearly, one can adjust α such as to make either one or the other of the neutral scalars h or H invisible, but the sum of the cross sections of h and H is independent of α. Note that the one exception would be in the type II model with large $\tan\beta$, since then bottom-quark loops can affect the production cross section of h and H. Including this possibility would add an additional parameter and, for illustrative purposes, we shall not include those loops here. One should keep in mind that, as noted in Section 2, large $\tan\beta$ has difficulties with perturbativity and unitarity, and requires fine-tuning.

Consider the case $\tan\beta = 1$. Then, the sum of the cross sections for h and H production, where both the h and H decay into W^+W^- and ZZ, is twice the cross section for the Higgs boson of the Standard Model. That means that at least one of those cross sections must be larger than the one of the Standard Model. From the plot in Fig. 21, one can see that if *both* the h and H have masses in between 140 and 480 GeV, then the production cross section for each of them is below the Standard Model; therefore, this case is experimentally excluded. Suppose instead that $\tan\beta$ is very large. Then, $\sin^2\beta \sim 1$ and the sum of the cross sections for h and H production is similar to the one of the Standard Model.

We can thus exclude the regions shown in Fig. 22. For $\tan\beta = 1$, the entire parameter space in between 140 and 480 GeV is excluded, as well as some regions beyond. For $\tan\beta = 10$ a much smaller region is excluded. It is a rather complicated shape due to the complex structure of the lines in Fig. 21. As more data is collected, the islands are expected to gradually merge and move outwards, eventually excluding most of the region—unless, of course, a signal is found, in which case a region will remain. It must be once again emphasized that *these exclusion regions are only valid if the primary decay of both the h and H are into W^+W^- and ZZ*—and, as can be seen from Section 2, this is not the case for a substantial part of the parameter space in various 2HDMs.

In general, for each of the four models described in Section 2 it is possible to find bounds analogous to those depicted in Fig. 22, but for various values of α. However, these bounds will often be quite weak, since a value of α can always be found that makes either h or H gauge-phobic, thereby eliminating all LHC bounds.

All of this may have changed due to LHC experimenters presenting, in December 2011 [481,482], evidence for a peak in the $\gamma\gamma$ channel at a mass of approximately 125 GeV. If this evidence is confirmed, then Fig. 22 will become irrelevant. What would such a discovery mean for 2HDMs?

The cross section for $pp \to h \to \gamma\gamma$ is consistent with the Standard Model, but substantial discrepancies are still allowed. In the Standard Model, a Higgs boson with mass 125 GeV has a branching ratio of approximately 0.002 to $\gamma\gamma$. In the 2HDM,
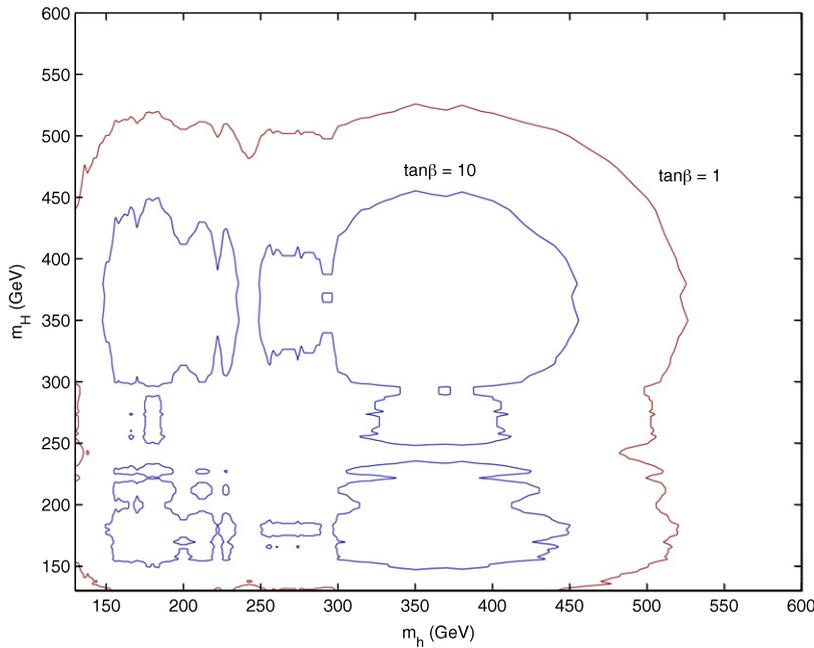
Fig. 22. For $\tan\beta = 1$ and $\tan\beta = 10$, and assuming that *both h and H decay primarily into W^+W^- and ZZ*. The excluded regions are those inside the contours.

the branching ratios can differ, as seen in many of the plots in Section 2, although there is usually a region of parameter space in which the branching ratio is similar. The production cross sections will also be different. It is clear that one will usually be able to find a set of parameters in which the recent results can be accommodated. If the branching ratios into W^+W^-, ZZ, $b\bar{b}$, and $\tau^+\tau^-$ can also be observed, then these will correspond to other regions of parameter space, which might be mutually exclusive. Thus it might be possible to rule out some of the 2HDMs in Section 2—or else, if the Standard Model predictions are not verified, these 2HDMs might constitute viable alternatives. This possibility has recently been discussed [483]. Note that failure to find a heavy Higgs boson at the LHC is not necessarily a problem; the more Standard-Model-like is the h field, the more gauge-phobic will the H be, and thus it could easily evade detection.

Clearly, one will need information about the other decay modes of this 125 GeV state, and such information should be forthcoming shortly.

8. Conclusions

The Large Hadron Collider has begun taking data, and a flood of information is imminent. Within a short time, the deepest question in particle physics in a generation, the nature of electroweak symmetry breaking, will definitively be addressed. In the Standard Model, one assumes the simplest possible scalar sector, and the properties of the Higgs boson in that model have been studied extensively.

The simplest extension of the electroweak Higgs sector is the addition of another scalar doublet. The purpose of this review article is to discuss this extension in detail. The phenomenology of Two-Higgs-Doublet models is extremely rich, since it contains a charged Higgs, a pseudoscalar and two neutral scalars, flavour-changing neutral currents, and more possibilities for CP violation and baryogenesis.

The most general 2HDM has tree level flavour changing neutral currents, which can be phenomenologically problematic. The most studied versions of the 2HDM use a discrete symmetry to avoid tree level FCNCs, and the phenomenology of the neutral sector of these versions was studied in Section 2, focusing on the bounds from the Tevatron and expectations from the LHC. In Section 3, we studied models that do contain FCNC at tree level. In Section 4, the phenomenology of the charged Higgs in these models was analysed.

The theoretical structure of 2HDMs is quite complex. One can have CP-violating, CP-conserving and charge breaking minima, there are several bases that one can choose, as well as a number of invariants. The theoretical structure is discussed in Section 5, including renormalization group analyses, vacuum stability bounds, symmetry-constrained Lagrangians, etc. In Section 6, CP violation in the Higgs sector is studied in detail.

Many important topics have not been included. The most important is supersymmetry, which automatically requires at least two Higgs doublets. There are extremely comprehensive reviews of the Higgs sector of supersymmetric models, and we referred the reader to those reviews. Other than a brief discussion in Section 6, we have not included models with three or more Higgs doublets, and we have, for simplicity, not considered models with singlets. In addition, the phenomenological

focus has been on the Tevatron and the LHC. While a linear collider would be enormously helpful in a detailed study of the scalar sector, it is sufficiently far in the future that we did not discuss it here; the questions that a linear collider would answer will likely be very different within a couple of years.

While the LHC will help elucidate the nature of the Higgs sector, it will be several years before the various couplings can be measured precisely enough to test many of the models in this review, although some will be tested sooner. Of course, paraphrasing Sidney Coleman, it is possible that the entire structure will be swept into the dustbin of history by a thunderbolt from Geneva. We may know very shortly.

Note added in proof

Following the announcement, in December 2011, of evidence for a neutral state with mass 125 GeV, several papers have considered the implications of such a state in the context of 2HDMs. Burdman et al. [484] posed the question whether this could be the pseudoscalar of the 2HDM. If this were the case, then the decays of the 125 GeV state into ZZ would be absent; a definitive answer as to whether these decays proceed will be forthcoming within the next few months. Ferreira et al. [485] considered the possibility that the 125 GeV state could be the *heavier neutral* scalar of a 2HDM; then, the lighter Higgs would have to be gauge-phobic in order to have escaped detection at LEP. Chen and He [486] studied the 125 GeV state in the context of a 2HDM with a fourth generation of fermions. Then, the additional fermions in the loop would substantially increase the production cross section, but the possible decay of the scalar into fourth generation neutrinos would suppress the branching ratios of visible decays; Chen and He looked at this problem considering the 125 GeV state to be either the pseudoscalar or the light scalar of a 2HDM. Blum and D'Agnolo [487] used an effective field theory approach, integrating out the heavy fields and parameterizing their effects as modifications to the couplings of the 125 GeV state. Blankenburg et al. [488] considered neutral flavour-changing decays of the 125 GeV state, somewhat in the spirit of the type III model but in a more general context.

Acknowledgements

Appendix A. Unitarity limits

The theoretical bounds discussed in Section 5.7 arise from the condition that the potential must have a minimum, *viz.* they prevent the existence of directions in field space along which the potential is unbounded from below. There are other theoretical bounds that one must impose on the potential, namely, all the (tree-level) scalar–scalar scattering amplitudes must respect unitarity. This is equivalent to requiring that the $J = 0$ partial waves (usually denoted a_0) for scalar–scalar scattering satisfy $|a_0| < 1/2$ in the high-energy limit.[31] In the SM, this requirement is equivalent to ensuring that the quartic coupling in the scalar potential is not too large; in the original work by Lee et al. [489,490] the bound $m_H < \sqrt{8\pi\sqrt{2}/(3G_F)} \simeq 1\,\text{TeV}$ was thus obtained. Extending this bound to the 2HDM is complicated, due to the richer scalar spectrum and, consequently, to the need to take into account many scattering amplitudes; the existence of many quartic couplings also complicates matters. This leads to an analysis of the eigenvalues of the S matrix in the scalar sector.

Early work on the unitarity bounds in the 2HDM was undertaken in [491–493]. A comprehensive study of all the scattering amplitudes and of their relevance for a CP-conserving scalar potential with a Z_2 symmetry was presented in [494]. That work was later generalized in [98,495] by allowing the presence the following quantities:

$$L_1 = \frac{1}{2v^2}\left[m_H^2\cos^2\alpha + m_h^2\sin^2\alpha + \frac{\sin 2\alpha}{2\tan\beta}\left(m_h^2 - m_H^2\right)\right] + \frac{2m_{12}^2}{v^2\sin 2\beta}\left(1 - \tan^2\beta\right),$$

$$L_2 = \frac{1}{2v^2}\left[m_H^2\sin^2\alpha + m_h^2\cos^2\alpha + \frac{\sin 2\alpha\tan\beta}{2}\left(m_h^2 - m_H^2\right)\right] + \frac{2m_{12}^2}{v^2\sin 2\beta\tan^2\beta}\left(\tan^2\beta - 1\right),$$

$$L_3 = \frac{1}{v^2\sin 2\beta}\left[\frac{\sin 2\alpha}{2}\left(m_H^2 - m_h^2\right) - 2m_{12}^2\right],$$

$$L_4 = \frac{2m_+^2}{v^2}, \qquad L_5 = \frac{2m_{12}^2}{v^2\sin 2\beta}, \qquad L_6 = \frac{2M_A^2}{v^2}, \tag{357}$$

---

[31] The same condition must be satisfied by the amplitudes for gauge boson–scalar scattering, but we need not worry about those, since they yield no new bounds.

which are functions of the masses of the physical scalars, of the angles α and β, and the soft-breaking parameter m_{12}^2 (notice that we are assuming a CP-conserving vacuum). The unitarity bounds can be, and usually are, expressed in terms of these quantities. We shall however, for consistency, express them in the notation adopted in this review. The L_i are written, in terms of the couplings of the scalar potential in Eq. (98), as

$$\lambda_1 = 2(L_1 + L_3), \tag{358}$$
$$\lambda_2 = 2(L_1 + L_3), \tag{359}$$
$$\lambda_3 = 2(L_3 + L_4), \tag{360}$$
$$\lambda_4 = \frac{1}{2}(L_5 + L_6 - 2L_4), \tag{361}$$
$$\lambda_5 = \frac{1}{2}(L_5 - L_6). \tag{362}$$

The computation of the S matrix for scalar–scalar scattering amplitudes allows the determination of its eigenvalues; the relevant ones are given by[32]

$$a_\pm = \frac{3}{2}(\lambda_1 + \lambda_2) \pm \sqrt{\frac{9}{4}(\lambda_1 - \lambda_2)^2 + (2\lambda_3 + \lambda_4)^2}, \tag{363}$$
$$b_\pm = \frac{1}{2}(\lambda_1 + \lambda_2) \pm \frac{1}{2}\sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_4^2}, \tag{364}$$
$$c_\pm = \frac{1}{2}(\lambda_1 + \lambda_2) \pm \frac{1}{2}\sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_5^2}, \tag{365}$$
$$e_1 = \lambda_3 + 2\lambda_4 - 3\lambda_5 \tag{366}$$
$$e_2 = \lambda_3 - \lambda_5, \tag{367}$$
$$f_+ = \lambda_3 + 2\lambda_4 + 3\lambda_5, \tag{368}$$
$$f_- = \lambda_3 + \lambda_5, \tag{369}$$
$$f_1 = \lambda_3 + \lambda_4, \tag{370}$$
$$p_1 = \lambda_3 - \lambda_4. \tag{371}$$

The requirement of tree-level perturbative unitarity translates as

$$|a_\pm|, \ |b_\pm|, \ |c_\pm|, \ |f_\pm|, \ \left|e_{1,2}\right|, \ |f_1|, \ |p_1| < 8\pi. \tag{372}$$

From all these conditions, the one on $|a_\pm|$ is the most restrictive one, but all others also contribute to place severe upper bounds on the scalar masses. Under the assumption that the Z_2 symmetry is unbroken and for small values of $\tan\beta \simeq 0.5$, Akeroyd et al. have thus found that $m_+ < 691\,\text{GeV}$, $m_A < 695\,\text{GeV}$, $m_h < 435\,\text{GeV}$, and $m_H < 638\,\text{GeV}$; for larger values of $\tan\beta$ the bound on m_h becomes quite stronger, dropping below 100 GeV for $\tan\beta \simeq 6$. Despite the constraints from LEP data [7], such low values for m_h are not forbidden, since the production cross section of a Zh pair at LEP is suppressed by a factor of $\sin^2(\alpha - \beta)$.

However, the presence of the soft-breaking term m_{12}^2 greatly relaxes these bounds—for large enough values of m_{12}^2 the bound on m_h becomes independent of β and approximately equal to 670 GeV. The unitarity bounds of the 2HDM thus are quite dependent on the values of some parameters, but may in some cases be quite constraining, even ruling out entire sections of parameter space due to conflicts with experimental findings.

Appendix B. Gauge interactions

In the Higgs basis of Eq. (220), the charged field H^+ is physical and has mass m_+, the neutral fields H, R, I are linear combinations of the physical neutral fields S_1, S_2, S_3:

$$\begin{pmatrix} H \\ R \\ I \end{pmatrix} = T \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}, \tag{373}$$

where the 3×3 matrix T is orthogonal. Without loss of generality we shall assume T to have determinant $+1$. The field S_j ($j = 1, 2, 3$) has mass m_j.

---

[32] For a CP-violating potential, in which the λ_5 coupling is complex, just replace λ_5 by $|\lambda_5|$ [496].

The gauge-kinetic Lagrangian

$$\mathcal{L}_g = \sum_{k=1}^{2} (D^\mu H_k)^\dagger (D_\mu H_k) \tag{374}$$

may be developed as [497]

$$\mathcal{L}_g = (\partial^\mu G^-)(\partial_\mu G^+) + (\partial^\mu H^-)(\partial_\mu H^+) + \frac{(\partial^\mu G^0)(\partial_\mu G^0)}{2} + \sum_{j=1}^{3} \frac{(\partial^\mu S_j)(\partial_\mu S_j)}{2} \tag{375a}$$

$$+ m_W^2 W^{\mu-} W_\mu^+ + m_Z^2 \frac{Z^\mu Z_\mu}{2} \tag{375b}$$

$$+ m_Z Z_\mu \partial^\mu G^0 + i m_W (W_\mu^- \partial^\mu G^+ - W_\mu^+ \partial^\mu G^-) \tag{375c}$$

$$- \left(e m_W A^\mu + g s_W^2 m_Z Z^\mu \right) \left(W_\mu^- G^+ + W_\mu^+ G^- \right) \tag{375d}$$

$$+ i \left[e A_\mu + \frac{g \left(s_W^2 - c_W^2 \right)}{2 c_W} Z_\mu \right] \left[(G^+ \partial^\mu G^- - G^- \partial^\mu G^+) + (H^+ \partial^\mu H^- - H^- \partial^\mu H^+) \right] \tag{375e}$$

$$+ \frac{g}{2 c_W} Z_\mu \left[\sum_{j=1}^{3} T_{1j} \left(S_j \partial^\mu G^0 - G^0 \partial^\mu S_j \right) + T_{13} \left(S_1 \partial^\mu S_2 - S_2 \partial^\mu S_1 \right) + T_{12} \left(S_3 \partial^\mu S_1 - S_1 \partial^\mu S_3 \right) \right.$$

$$\left. + T_{11} \left(S_2 \partial^\mu S_3 - S_3 \partial^\mu S_2 \right) \right] \tag{375f}$$

$$+ i \frac{g}{2} \left\{ W_\mu^+ \sum_{j=1}^{3} T_{1j} \left(G^- \partial^\mu S_j - S_j \partial^\mu G^- \right) + W_\mu^- \sum_{j=1}^{3} T_{1j} \left(S_j \partial^\mu G^+ - G^+ \partial^\mu S_j \right) \right.$$

$$+ i W_\mu^+ \left(G^- \partial^\mu G^0 - G^0 \partial^\mu G^- \right) + i W_\mu^- \left(G^0 \partial^\mu G^+ - G^+ \partial^\mu G^0 \right)$$

$$+ W_\mu^+ \sum_{j=1}^{3} \left(T_{2j} + i T_{3j} \right) \left(H^- \partial^\mu S_j - S_j \partial^\mu H^- \right)$$

$$\left. + W_\mu^- \sum_{j=1}^{3} \left(T_{2j} - i T_{3j} \right) \left(S_j \partial^\mu H^+ - H^+ \partial^\mu S_j \right) \right\} \tag{375g}$$

$$+ g \left(m_W W_\mu^+ W^{\mu-} + \frac{m_Z}{c_W} \frac{Z_\mu Z^\mu}{2} \right) \sum_{j=1}^{3} T_{1j} S_j \tag{375h}$$

$$- \left(\frac{eg}{2} A^\mu + \frac{g^2 s_W^2}{2 c_W} Z^\mu \right) \left[\left(W_\mu^+ G^- + W_\mu^- G^+ \right) \sum_{j=1}^{3} T_{1j} S_j + i \left(W_\mu^+ G^- - W_\mu^- G^+ \right) G^0 \right.$$

$$\left. + W_\mu^+ H^- \sum_{j=1}^{3} (T_{2j} + i T_{3j}) S_j + W_\mu^- H^+ \sum_{j=1}^{3} (T_{2j} - i T_{3j}) S_j \right] \tag{375i}$$

$$+ \left(\frac{g^2}{2} W^{\mu-} W_\mu^+ + \frac{g^2}{2 c_W^2} \frac{Z^\mu Z_\mu}{2} \right) \left[\frac{(G^0)^2}{2} + \sum_{j=1}^{3} \frac{S_j^2}{2} \right] \tag{375j}$$

$$+ \left[\frac{g^2}{2} W^{\mu-} W_\mu^+ + 2 e^2 \frac{A^\mu A_\mu}{2} + \frac{g^2 \left(s_W^2 - c_W^2 \right)^2}{2 c_W^2} \frac{Z^\mu Z_\mu}{2} + \frac{eg}{c_W \left(s_W^2 - c_W^2 \right)} A^\mu Z_\mu \right] \left(G^- G^+ + H^- H^+ \right). \tag{375k}$$

Appendix C. Scalar–scalar interactions

The scalar potential in the Higgs basis is given by

$$V = \bar{m}_{11}^2 H_1^\dagger H_1 + \bar{m}_{22}^2 H_2^\dagger H_2 - \left(\bar{m}_{12}^2 H_1^\dagger H_2 + \text{H.c.} \right)$$

$$+ \frac{\bar{\lambda}_1}{2} \left(H_1^\dagger H_1 \right)^2 + \frac{\bar{\lambda}_2}{2} \left(H_2^\dagger H_2 \right)^2 + \bar{\lambda}_3 \, H_1^\dagger H_1 \, H_2^\dagger H_2 + \bar{\lambda}_4 \, H_1^\dagger H_2 \, H_2^\dagger H_1$$

$$+ \left[\frac{\bar{\lambda}_5}{2} \left(H_1^\dagger H_2 \right)^2 + \left(\bar{\lambda}_6 \, H_1^\dagger H_1 + \bar{\lambda}_7 \, H_2^\dagger H_2 \right) H_1^\dagger H_2 + \text{H.c.} \right], \tag{376}$$

where, as before, the parameters in the Higgs basis are denoted by $\bar{m}_{ij}$ and $\bar{\lambda}_i$. The vacuum expectation value of the potential is

$$V_0 \equiv \langle 0 \, | V | \, 0 \rangle = \bar{m}_{11} \frac{v^2}{4} = -\bar{\lambda}_1 \frac{v^4}{8}. \tag{377}$$

We now present general expressions for masses and couplings written in the Higgs basis, for any neutral vacuum. These are, of course, written in a different basis from the results of Section 5.10 and as such cannot be trivially compared. The mass terms for the scalars are given by the part of the potential which is bilinear in the fields:

$$V_2 = m_+^2 H^- H^+ + \frac{1}{2} \begin{pmatrix} H & R & I \end{pmatrix} M \begin{pmatrix} H \\ R \\ I \end{pmatrix}. \tag{378}$$

The mass of the charged Higgs is given by

$$m_+^2 = \bar{m}_{22}^2 + \bar{\lambda}_3 \frac{v^2}{2}, \tag{379}$$

while

$$M = \begin{pmatrix} \bar{\lambda}_1 \, v^2 & \text{Re} \, \bar{\lambda}_6 \, v^2 & -\text{Im} \, \bar{\lambda}_6 \, v^2 \\ \text{Re} \, \bar{\lambda}_6 \, v^2 & m_+^2 + \left(\bar{\lambda}_4 + \text{Re} \, \bar{\lambda}_5 \right) v^2 / 2 & -\text{Im} \, \bar{\lambda}_5 \, v^2 / 2 \\ -\text{Im} \, \bar{\lambda}_6 \, v^2 & -\text{Im} \, \bar{\lambda}_5 \, v^2 / 2 & m_+^2 + \left(\bar{\lambda}_4 - \text{Re} \, \bar{\lambda}_5 \right) v^2 / 2 \end{pmatrix}. \tag{380}$$

The symmetric matrix M is diagonalized by the orthogonal matrix T:

$$T^T M T = \text{diag} \left(m_1^2, m_2^2, m_3^2 \right), \tag{381}$$

so that

$$V_2 = m_+^2 H^- H^+ + \sum_{j=1}^{3} m_j^2 \frac{S_j^2}{2}. \tag{382}$$

Like we said before, we assume $\det T = +1$ without loss of generality.

The part of V which is trilinear in the fields may be written

$$V_3 = \left[G^- G^+ + \frac{(G^0)^2}{2} \right] \sum_{j=1}^{3} S_j T_{1j} \frac{m_j^2}{v} \tag{383a}$$

$$+ G^+ H^- \sum_{j=1}^{3} S_j \left(T_{2j} + i T_{3j} \right) \frac{m_j^2 - m_+^2}{v} \tag{383b}$$

$$+ G^- H^+ \sum_{j=1}^{3} S_j \left(T_{2j} - i T_{3j} \right) \frac{m_j^2 - m_+^2}{v} \tag{383c}$$

$$+ G^0 \left(S_1 S_2 T_{13} \frac{m_1^2 - m_2^2}{v} + S_3 S_1 T_{12} \frac{m_3^2 - m_1^2}{v} + S_2 S_3 T_{11} \frac{m_2^2 - m_3^2}{v} \right) \tag{383d}$$

$$+ v H^- H^+ \sum_{j=1}^{3} c_j S_j \tag{383e}$$

$$+ \sum_{j,k,l=1}^{3} S_j S_k S_l \left[\frac{v}{2} c_j \left(\delta_{kl} - T_{1k} T_{1l} \right) + \frac{1}{v} \left(m_+^2 - \frac{m_j^2}{2} \right) T_{1j} T_{1k} T_{1l} + \frac{m_k^2 - m_+^2}{v} T_{1j} \delta_{kl} \right] \tag{383f}$$

where [425]

$$c_j \equiv T_{1j} \bar{\lambda}_3 + T_{2j} \, \text{Re} \, \bar{\lambda}_7 - T_{3j} \, \text{Im} \, \bar{\lambda}_7. \tag{384}$$

A thorough analysis of the trilinear Higgs couplings for the most general 2HDM potential, considering one-loop corrections, was undertaken in [498].

Appendix D. The oblique parameters

D.1. Definition

Let the vacuum polarization tensors be written

$$\Pi_{VV'}^{\mu\nu}(q) = g^{\mu\nu} A_{VV'}\left(q^2\right) + q^\mu q^\nu B_{VV'}\left(q^2\right), \tag{385}$$

where VV' may be either $\gamma\gamma$, γZ^0, $Z^0 Z^0$, or $W^+ W^-$, and $q = (q^\alpha)$ is the four-momentum of the gauge bosons. Let us moreover define

$$\bar{A}_{VV'}\left(q^2\right) = A_{VV'}\left(q^2\right)\big|_{2HDM} - A_{VV'}\left(q^2\right)\big|_{SM}, \tag{386}$$

where SM denotes the Standard Model with a Higgs particle of mass m_H. Then, the oblique parameters of the 2HDM are defined [499]

$$S = \frac{16\pi c_W^2}{g^2}\left[\frac{\bar{A}_{Z^0 Z^0}\left(m_Z^2\right) - \bar{A}_{Z^0 Z^0}(0)}{m_Z^2} - \frac{\partial \bar{A}_{\gamma\gamma}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=0} + \frac{c_W^2 - s_W^2}{c_W s_W}\frac{\partial \bar{A}_{\gamma Z^0}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=0}\right], \tag{387a}$$

$$T = \frac{4\pi}{g^2 s_W^2}\left[\frac{\bar{A}_{W^+ W^-}(0)}{m_W^2} - \frac{\bar{A}_{Z^0 Z^0}(0)}{m_Z^2}\right], \tag{387b}$$

$$U = \frac{16\pi}{g^2}\left[\frac{\bar{A}_{W^+ W^-}\left(m_W^2\right) - \bar{A}_{W^+ W^-}(0)}{m_W^2} - c_W^2\frac{\bar{A}_{Z^0 Z^0}\left(m_Z^2\right) - \bar{A}_{Z^0 Z^0}(0)}{m_Z^2} - s_W^2\frac{\partial \bar{A}_{\gamma\gamma}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=0}\right.$$
$$\left. + 2c_W s_W\frac{\partial \bar{A}_{\gamma Z^0}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=0}\right], \tag{387c}$$

$$V = \frac{4\pi}{g^2 s_W^2}\left[\frac{\partial \bar{A}_{Z^0 Z^0}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=m_Z^2} - \frac{\bar{A}_{Z^0 Z^0}\left(m_Z^2\right) - \bar{A}_{Z^0 Z^0}(0)}{m_Z^2}\right], \tag{387d}$$

$$W = \frac{4\pi}{g^2 s_W^2}\left[\frac{\partial \bar{A}_{W^+ W^-}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=m_W^2} - \frac{\bar{A}_{W^+ W^-}\left(m_W^2\right) - \bar{A}_{W^+ W^-}(0)}{m_W^2}\right], \tag{387e}$$

$$X = \frac{4\pi c_W}{g^2 s_W}\left[\frac{\partial \bar{A}_{\gamma Z^0}\left(q^2\right)}{\partial q^2}\bigg|_{q^2=0} - \frac{\bar{A}_{\gamma Z^0}\left(m_Z^2\right)}{m_Z^2}\right]. \tag{387f}$$

These parameters are finite and, in principle, observable. In practice, they will be measurable in practical electroweak experiments on the 2HDM provided

the measurements are performed at one of the energy scales $q^2 \approx 0$, $q^2 = m_W^2$, or $q^2 = m_Z^2$, and
the measurements are performed with light fermions which couple mainly to the gauge bosons γ, Z^0, and $W^\pm$, but couple only very weakly to the scalar particles.

D.2. Formulae

The expressions for the oblique parameters in multi-Higgs-doublet models have been derived in [497,500]. We give here those expressions in the particular case of the 2HDM. For T one has [497]

$$T = \frac{1}{16\pi s_W^2 m_W^2}\left\{\sum_{j=1}^3\left(1 - T_{1j}^2\right) F\left(m_+^2, m_j^2\right)\right. \tag{388a}$$

$$-T_{11}^2 F\left(m_2^2, m_3^2\right) - T_{12}^2 F\left(m_3^2, m_1^2\right) - T_{13}^2 F\left(m_1^2, m_2^2\right) \tag{388b}$$

$$+3\sum_{j=1}^3 T_{1j}^2\left[F\left(m_Z^2, m_j^2\right) - F\left(m_W^2, m_j^2\right)\right] \tag{388c}$$

$$\left. -3\left[F\left(m_Z^2, m_H^2\right) - F\left(m_W^2, m_H^2\right)\right]\right\}, \tag{388d}$$

where

$$F(x, y) = \begin{cases} \dfrac{x + y}{2} - \dfrac{xy}{x - y} \ln \dfrac{x}{y} & \Leftarrow x \neq y, \\ 0 & \Leftarrow x = y. \end{cases} \tag{389}$$

Line (388a) is a positive definite contribution to T coming from the neutral scalars not having the same mass as the charged scalar, while line (388b) is a negative definite contribution to T resulting from the neutral scalars not all having the same mass. Line (388d) is the subtraction of the SM equivalent of line (388c).

The expressions for S, U, and X involve the following two functions:

$$\begin{aligned} G(x, y) = {}& -\frac{16}{3} + 5(x + y) - 2(x - y)^2 + 3\left[\frac{x^2 + y^2}{x - y} - x^2 + y^2 + \frac{(x - y)^3}{3}\right] \ln \frac{x}{y} \\ & + \left[1 - 2(x + y) + (x - y)^2\right] f\left(x + y - 1, 1 - 2(x + y) + (x - y)^2\right), \end{aligned} \tag{390}$$

$$\hat{G}(x) = -\frac{79}{3} + 9x - 2x^2 + \left(-10 + 18x - 6x^2 + x^3 - 9\frac{x + 1}{x - 1}\right) \ln x + \left(12 - 4x + x^2\right) f\left(x, x^2 - 4x\right), \tag{391}$$

where

$$f(z, w) = \begin{cases} \sqrt{w} \ln \left|\dfrac{z - \sqrt{w}}{z + \sqrt{w}}\right| & \Leftarrow w > 0, \\ 0 & \Leftarrow w = 0, \\ 2\sqrt{-w} \arctan \dfrac{\sqrt{-w}}{z} & \Leftarrow w < 0. \end{cases} \tag{392}$$

One has

$$\begin{aligned} S = {}& \frac{1}{24\pi} \left\{ \left(s_W^2 - c_W^2\right)^2 G(z_+, z_+) + T_{11}^2 G(z_2, z_3) + T_{12}^2 G(z_3, z_1) + T_{13}^2 G(z_1, z_2) \right. \\ & \left. + \sum_{j=1}^{3} \left[T_{1j}^2 \hat{G}(z_j) + \ln \frac{m_j^2}{m_+^2}\right] - \hat{G}(z_H) - \ln \frac{m_H^2}{m_+^2} \right\}, \end{aligned} \tag{393}$$

$$\begin{aligned} U = {}& \frac{1}{24\pi} \left\{ \sum_{j=1}^{3} \left(1 - T_{1j}^2\right) G(w_+, w_j) - \left(s_W^2 - c_W^2\right)^2 G(z_+, z_+) - T_{11}^2 G(z_2, z_3) - T_{12}^2 G(z_3, z_1) - T_{13}^2 G(z_1, z_2) \right. \\ & \left. + \sum_{j=1}^{3} T_{1j}^2 \left[\hat{G}(w_j) - \hat{G}(z_j)\right] - \hat{G}(w_H) + \hat{G}(z_H) \right\}, \end{aligned} \tag{394}$$

$$X = \frac{c_W^2 - s_W^2}{48\pi} G(z_+, z_+), \tag{395}$$

where

$$z_a \equiv \frac{m_a^2}{m_Z^2} \quad \text{and} \quad w_a \equiv \frac{m_a^2}{m_W^2} \tag{396}$$

for $a = +, 1, 2, 3, H$.

The expressions for the oblique parameters V and W involve the following two functions:

$$\begin{aligned} H(x, y) = {}& 2 - 9(x + y) + 6(x - y)^2 + 3\left[-\frac{x^2 + y^2}{x - y} + 2(x^2 - y^2) - (x - y)^3\right] \ln \frac{x}{y} \\ & + 3\left[x + y - (x - y)^2\right] f\left(x + y - 1, 1 - 2(x + y) + (x - y)^2\right), \end{aligned} \tag{397}$$

$$\hat{H}(x) = 47 - 21x + 6x^2 + 3\left(7 - 12x + 5x^2 - x^3 + 3\frac{x + 1}{x - 1}\right) \ln x + 3\left(28 - 20x + 7x^2 - x^3\right) \frac{f\left(x, x^2 - 4x\right)}{x - 4}. \tag{398}$$

One has

$$\begin{aligned} V = {}& \frac{1}{96\pi c_W^2 s_W^2} \left[\left(s_W^2 - c_W^2\right)^2 H(z_+, z_+) + T_{11}^2 H(z_2, z_3) + T_{12}^2 H(z_3, z_1) + T_{13}^2 H(z_1, z_2) \right. \\ & \left. + \sum_{j=1}^{3} T_{1j}^2 \hat{H}(z_j) - \hat{H}(z_H) \right], \end{aligned} \tag{399}$$

$$W = \frac{1}{96\pi s_W^2} \left[\sum_{j=1}^{3} \left(1 - T_{1j}^2\right) H\left(w_+, w_j\right) + \sum_{j=1}^{3} T_{1j}^2 \hat{H}\left(w_j\right) - \hat{H}\left(w_H\right) \right]. \tag{400}$$

Appendix E. Renormalization-group equations

The one-loop renormalization-group (RG) equations for a general gauge theory were presented in [501,502]. For the specific case of the multi-Higgs-doublet $SU(2) \times U(1)$ gauge theory, they were given in [39,403,503].

Let μ be the mass parameter used in the regularization of ultraviolet divergences in loop integrals. Let $\mathcal{D}$ denote the dimensionless differential operator $16\pi^2 \, (\mathrm{d}/\mathrm{d} \ln \mu) = 16\pi^2 \mu \, (\mathrm{d}/\mathrm{d}\mu)$. Let g_s, g, and g' denote the gauge coupling constants of $SU(3)_{\mathrm{colour}}$, $SU(2)$, and $U(1)$, respectively. The normalization of g' is such that the $U(1)$ charge of the Higgs doublets is $+1/2$. The one-loop RG equations for the gauge coupling constants do not depend of the Yukawa and scalar couplings and are

$$\mathcal{D}g_s = \left(-11 + \frac{4}{3} n_F\right) g_s^3, \tag{401}$$

$$\mathcal{D}g = \left(-\frac{22}{3} + \frac{4}{3} n_F + \frac{1}{6} n_H\right) g^3, \tag{402}$$

$$\mathcal{D}g' = \left(\frac{20}{9} n_F + \frac{1}{6} n_H\right) g'^3, \tag{403}$$

where n_F is the number of fermion generations and n_H is the number of Higgs doublets. In the 2HDM with three families of fermions this is therefore

$$\mathcal{D}g_s = -7g_s^3, \tag{404}$$

$$\mathcal{D}g = -3g^3, \tag{405}$$

$$\mathcal{D}g' = 7g'^3. \tag{406}$$

The one-loop RG equations for the Yukawa-coupling matrices defined in Eq. (225) do not depend on the scalar couplings and are

$$\mathcal{D}Y_j^d = a_d Y_j^d + \sum_{k=1}^{n_H} T_{jk} Y_k^d + \sum_{k=1}^{n_H} \left(-2 Y_k^u Y_k^{u\dagger} Y_k^d + \frac{1}{2} Y_k^u Y_k^{u\dagger} Y_j^d + Y_j^d Y_k^{d\dagger} Y_k^d + \frac{1}{2} Y_k^d Y_k^{d\dagger} Y_j^d\right), \tag{407}$$

$$\mathcal{D}Y_j^u = a_u Y_j^u + \sum_{k=1}^{n_H} T_{jk}^* Y_k^u + \sum_{k=1}^{n_H} \left(-2 Y_k^d Y_k^{d\dagger} Y_k^u + \frac{1}{2} Y_k^d Y_k^{d\dagger} Y_j^u + Y_j^u Y_k^{u\dagger} Y_k^u + \frac{1}{2} Y_k^u Y_k^{u\dagger} Y_j^u\right), \tag{408}$$

$$\mathcal{D}Y_j^e = a_e Y_j^e + \sum_{k=1}^{n_H} T_{jk} Y_k^e + \sum_{k=1}^{n_H} \left(Y_j^e Y_k^{e\dagger} Y_k^e + \frac{1}{2} Y_k^e Y_k^{e\dagger} Y_j^e\right), \tag{409}$$

where

$$a_d = -8g_s^2 - \frac{9}{4} g^2 - \frac{5}{12} g'^2, \tag{410}$$

$$a_u = -8g_s^2 - \frac{9}{4} g^2 - \frac{17}{12} g'^2, \tag{411}$$

$$a_e = -\frac{9}{4} g^2 - \frac{15}{4} g'^2, \tag{412}$$

and

$$T_{jk} = 3 \, \mathrm{tr}\left(Y_j^d Y_k^{d\dagger} + Y_j^{u\dagger} Y_k^u\right) + \mathrm{tr}\left(Y_j^e Y_k^{e\dagger}\right). \tag{413}$$

Of course, in the context of the 2HDM we should set $n_H = 2$ in Eq. (407)–(409).

Let the scalar potential be

$$V = \sum_{j,k} \mu_k^j \Phi_j^\dagger \Phi_k + \frac{1}{2} \sum_{j,k,l,m} \Lambda_{km}^{jl} \Phi_j^\dagger \Phi_k \Phi_l^\dagger \Phi_m, \tag{414}$$

with $\Lambda_{km}^{jl} = \Lambda_{mk}^{lj}$.[33] The one-loop RG equations for the quartic couplings are

$$\mathcal{D}\Lambda_{km}^{jl} = 2 \sum_{p,q=1}^{n_H} \left(2\Lambda_{kq}^{jp}\Lambda_{pm}^{ql} + \Lambda_{kq}^{jp}\Lambda_{pm}^{lq} + \Lambda_{qk}^{jp}\Lambda_{pm}^{ql} + \Lambda_{pq}^{jl}\Lambda_{km}^{pq} + \Lambda_{pm}^{jq}\Lambda_{kq}^{pl} \right)$$

$$- \left(9g^2 + 3g'^2 \right) \Lambda_{km}^{jl} + \frac{9g^4 - 6g^2 g'^2 + 3g'^4}{4} \delta_k^j \delta_m^l + 3g^2 g'^2 \delta_m^j \delta_k^l$$

$$+ \sum_{p=1}^{n_H} \left(T_{kp}\Lambda_{pm}^{jl} + T_{mp}\Lambda_{kp}^{jl} + T_{jp}^*\Lambda_{km}^{pl} + T_{lp}^*\Lambda_{km}^{jp} \right) - 4\,\mathrm{tr} \left(Y_j^{e\dagger} Y_k^e Y_l^{e\dagger} Y_m^e \right)$$

$$- 12\,\mathrm{tr} \left(Y_j^{d\dagger} Y_k^d Y_l^{d\dagger} Y_m^d + Y_k^{u\dagger} Y_j^u Y_m^{u\dagger} Y_l^u + Y_j^{d\dagger} Y_l^u Y_m^{u\dagger} Y_k^d \right.$$

$$\left. + Y_k^{u\dagger} Y_m^d Y_l^{d\dagger} Y_j^u - Y_m^{u\dagger} Y_k^d Y_l^{d\dagger} Y_j^u - Y_j^{d\dagger} Y_l^u Y_k^{u\dagger} Y_m^d \right). \tag{415}$$

The one-loop RG equations for the quadratic couplings are

$$\mathcal{D}\mu_k^j = 2 \sum_{p,q=1}^{n_H} \mu_q^p \left(2\Lambda_{kp}^{jq} + \Lambda_{pk}^{jq} \right). \tag{416}$$

In the specific case of the 2HDM, with the notation of Eq. (98), one has

$$\mu_1^1 = m_{11}^2, \qquad \mu_2^2 = m_{22}^2, \qquad \mu_2^1 = -m_{12}^2, \qquad \mu_1^2 = -m_{12}^2{}^*, \tag{417}$$

$$\Lambda_{11}^{11} = \lambda_1, \qquad \Lambda_{22}^{22} = \lambda_2, \qquad \Lambda_{12}^{12} = \Lambda_{21}^{21} = \lambda_3, \qquad \Lambda_{21}^{12} = \Lambda_{12}^{21} = \lambda_4, \tag{418}$$

$$\Lambda_{22}^{11} = \lambda_5, \qquad \Lambda_{11}^{22} = \lambda_5^*, \tag{419}$$

$$\Lambda_{21}^{11} = \Lambda_{12}^{11} = \lambda_6, \qquad \Lambda_{11}^{21} = \Lambda_{11}^{12} = \lambda_6^*, \qquad \Lambda_{22}^{12} = \Lambda_{22}^{21} = \lambda_7, \qquad \Lambda_{12}^{22} = \Lambda_{21}^{22} = \lambda_7^*. \tag{420}$$

Therefore, from Eq. (416),

$$\mathcal{D}m_{11}^2 = 6\lambda_1 m_{11}^2 + (4\lambda_3 + 2\lambda_4) m_{22}^2 - 12\,\mathrm{Re}\left(m_{12}^2 \lambda_6^* \right), \tag{421}$$

$$\mathcal{D}m_{22}^2 = (4\lambda_3 + 2\lambda_4) m_{11}^2 + 6\lambda_2 m_{22}^2 - 12\,\mathrm{Re}\left(m_{12}^2 \lambda_7^* \right), \tag{422}$$

$$\mathcal{D}m_{12}^2 = -6\left(\lambda_6 m_{11}^2 + \lambda_7 m_{22}^2 \right) + (2\lambda_3 + 4\lambda_4) m_{12}^2 + 6\lambda_5 m_{12}^2{}^*. \tag{423}$$

For the quartic couplings, in a 2HDM where the up-type quarks couple only to the doublet Φ_1, and if we only consider the contribution from the top-quark Yukawa coupling λ_t (with the normalization $m_t = \lambda_t v/\sqrt{2}$, with $v = 246$ GeV), the one-loop RG equations are

$$\mathcal{D}\lambda_1 = 12\lambda_1^2 + 4\lambda_3^2 + 4\lambda_3\lambda_4 + 2\lambda_4^2 + 2|\lambda_5|^2 + 24|\lambda_6|^2$$
$$+ \frac{3}{4}(3g^4 + g'^4 + 2g^2 g'^2) - 3\lambda_1(3g^2 + g'^2 - 4\lambda_t^2) - 12\lambda_t^4, \tag{424a}$$

$$\mathcal{D}\lambda_2 = 12\lambda_2^2 + 4\lambda_3^2 + 4\lambda_3\lambda_4 + 2\lambda_4^2 + 2|\lambda_5|^2 + 24|\lambda_7|^2 + \frac{3}{4}(3g^4 + g'^4 + 2g^2 g'^2) - 3\lambda_2(3g^2 + g'^2), \tag{424b}$$

$$\mathcal{D}\lambda_3 = (\lambda_1 + \lambda_2)(6\lambda_3 + 2\lambda_4) + 4\lambda_3^2 + 2\lambda_4^2 + 2|\lambda_5|^2 + 4\left(|\lambda_6|^2 + |\lambda_7|^2\right) + 16\,\mathrm{Re}\left(\lambda_6\lambda_7^*\right)$$
$$+ \frac{3}{4}(3g^4 + g'^4 - 2g^2 g'^2) - 3\lambda_3(3g^2 + g'^2 - 2\lambda_t^2), \tag{424c}$$

$$\mathcal{D}\lambda_4 = 2(\lambda_1 + \lambda_2)\lambda_4 + 8\lambda_3\lambda_4 + 4\lambda_4^2 + 8|\lambda_5|^2 + 10\left(|\lambda_6|^2 + |\lambda_7|^2\right) + 4\,\mathrm{Re}\left(\lambda_6\lambda_7^*\right)$$
$$+ 3g^2 g'^2 - 3\lambda_4(3g^2 + g'^2 - 2\lambda_t^2), \tag{424d}$$

$$\mathcal{D}\lambda_5 = (2\lambda_1 + 2\lambda_2 + 8\lambda_3 + 12\lambda_4)\lambda_5 + 10\left(\lambda_6^2 + \lambda_7^2\right) + 4\lambda_6\lambda_7 - 3\lambda_5(3g^2 + g'^2 - 2\lambda_t^2), \tag{424e}$$

$$\mathcal{D}\lambda_6 = (12\lambda_1 + 6\lambda_3 + 8\lambda_4)\lambda_6 + (6\lambda_3 + 4\lambda_4)\lambda_7 + 10\lambda_5\lambda_6^* + 2\lambda_5\lambda_7^* - 3\lambda_6(3g^2 + g'^2 - 3\lambda_t^2), \tag{424f}$$

$$\mathcal{D}\lambda_7 = (12\lambda_2 + 6\lambda_3 + 8\lambda_4)\lambda_7 + (6\lambda_3 + 4\lambda_4)\lambda_6 + 10\lambda_5\lambda_7^* + 2\lambda_5\lambda_6^* - 3\lambda_7(3g^2 + g'^2 - \lambda_t^2). \tag{424g}$$

---

[33] The correspondence with the notation of Eq. (99) is $\Lambda_{km}^{jl} = \lambda_{jk,lm}$; using the tensor Λ allows for a more compact writing of the formulae below.

Appendix F. Custodial symmetry

The experimentally measured value for the observable $\rho = m_W^2 / \left(m_Z^2 \cos^2 \theta_W \right)$ is extremely close to one [7]. In the SM, the tree-level prediction for that observable is exactly one. The fundamental reason for that prediction is an approximate symmetry that the SM Lagrangian possesses. Indeed, if one writes the SM Higgs doublet as $\Phi = \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix}$, then the SM scalar potential only depends on $\Phi^\dagger \Phi = \varphi_1^2 + \varphi_2^2 + \varphi_3^2 + \varphi_4^2$. Therefore, the potential automatically has $SO(4)$ symmetry. The group $SO(4)$ is isomorphic to $SU(2) \times SU(2)$, which is larger than the SM gauge group $SU(2)_L \times U(1)_Y$. On the other hand, this symmetry $SO(4)$ is respected neither by the scalar gauge-kinetic terms – specifically, those involving the weak-hypercharge coupling g' – nor by the Yukawa terms, linear in Φ, since the up-type and down-type quarks have different masses. Thus, $SO(4)$ is not a symmetry of the full SM Lagrangian, rather a symmetry only of the scalar potential; it is usually regarded as an *approximate* symmetry, since in the scalar sector it is only broken by (small) g' terms in the kinetic energy, and dubbed 'custodial symmetry' [8,504].

In the 2HDM, there is no $SO(4)$ symmetry in the most general scalar potential, thus the possibility of large contributions to ρ from that sector alone arises. If one wants to avoid them, one may impose custodial symmetry on the 2HDM potential. Following the work of Pomarol and Vega [505], we first define the 2×2 matrices

$$M_{ij} = \left(\tilde{\Phi}_i \mid \Phi_j \right) = \begin{pmatrix} \varphi_i^{0*} & \varphi_j^+ \\ -\varphi_i^- & \varphi_j^0 \end{pmatrix}. \tag{425}$$

We further define an $SU(2)_L \times SU(2)_R$ group under which these matrices transform as

$$M_{ij} \rightarrow L M_{ij} R^\dagger, \tag{426}$$

with $L, R \in SU(2)$. The quantities $\mathrm{tr}(M_{ij}^\dagger M_{kl})$ are invariant under this $SO(4) = SU(2)_L \times SU(2)_R$. This corresponds to the same $SO(4)$ custodial symmetry of the SM scalar potential.

Pomarol and Vega have considered two separate situations. In case 1 they have used only M_{11} and M_{22}. The most general scalar potential invariant under $SO(4)$ is then

$$\begin{aligned}
V_1 = {} & \frac{1}{2} m_{11}^2 \mathrm{tr}(M_{11}^\dagger M_{11}) + \frac{1}{2} m_{22}^2 \mathrm{tr}(M_{22}^\dagger M_{22}) - m_{12}^2 \mathrm{tr}(M_{11}^\dagger M_{22}) \\
& + \frac{1}{8} \lambda_1 \left[\mathrm{tr}(M_{11}^\dagger M_{11}) \right]^2 + \frac{1}{8} \lambda_2 \left[\mathrm{tr}(M_{22}^\dagger M_{22}) \right]^2 + \frac{1}{4} \lambda_3 \mathrm{tr}(M_{11}^\dagger M_{11}) \mathrm{tr}(M_{22}^\dagger M_{22}) \\
& + \frac{1}{2} \lambda_4 \left[\mathrm{tr}(M_{11}^\dagger M_{22}) \right]^2 + \frac{1}{2} \left[\lambda_6 \mathrm{tr}(M_{11}^\dagger M_{11}) + \lambda_7 \mathrm{tr}(M_{22}^\dagger M_{22}) \right] \mathrm{tr}(M_{11}^\dagger M_{22}).
\end{aligned} \tag{427}$$

Notice that $\mathrm{tr}(M_{ii}^\dagger M_{jj}) = \Phi_i^\dagger \Phi_j + \Phi_j^\dagger \Phi_i$. This allows us to identify, in Eq. (427), the relations that one obtains for the usual parameters of the 2HDM potential in Eq. (98). We find

- Case 1: all the parameters are real, and $\lambda_4 = \lambda_5$.

At a neutral vacuum $\langle \varphi_i^0 \rangle = v_i$, one has

$$\langle M_{ij} \rangle = \begin{pmatrix} v_i^* & 0 \\ 0 & v_j \end{pmatrix}. \tag{428}$$

This vacuum is *not* invariant under the full group $SU(2)_L \times SU(2)_R$. However, if $v_i^* = v_j$, then $\langle M_{ij} \rangle$ is proportional to the 2×2 identity matrix and the vacuum preserves a group $SU(2)_V$ (the "V" stands for "vectorial"), corresponding to identical matrices, *i.e.* $L = R$, in Eq. (426). This remaining group preserved by the vacuum is the custodial-symmetry group. However, most authors refer to the potential invariant under $SU(2)_L \times SU(2)_R$ as displaying a custodial symmetry, and we shall also employ that terminology.

In case 2 of Pomarol and Vega the potential is built only with M_{12}. It reads

$$\begin{aligned}
V_2 = {} & m_{11}^2 \mathrm{tr}(M_{12}^\dagger M_{12}) - m_{12}^2 \left(\det M_{12} + \text{h.c.} \right) + \frac{1}{2} \lambda_1 \left[\mathrm{tr}(M_{12}^\dagger M_{12}) \right]^2 + \lambda_4 \det(M_{12}^\dagger M_{12}) + \frac{1}{2} \left[\lambda_5 \det(M_{12})^2 + \text{h.c.} \right] \\
& + \left[\lambda_6 \det M_{12} \, \mathrm{tr}(M_{12}^\dagger M_{12}) + \text{h.c.} \right].
\end{aligned} \tag{429}$$

The following constraints on the parameters of the 2HDM scalar potential ensue:

- Case 2: $m_{11}^2 = m_{22}^2$, $\lambda_1 = \lambda_2 = \lambda_3$, and $\lambda_6 = \lambda_7$.

Notice that m_{12}^2, λ_5, λ_6, and λ_7 remain complex in this case. The vacuum preserves $SU(2)_V$ if and only if $v_1 = v_2^*$.

In both cases, there is a dramatic prediction for the scalar masses: the charged Higgs $H^\pm$ is degenerate with the pseudoscalar A. This is easy to see from Eq. (207) in the Case 1 model. In both cases of Pomarol and Vega, the potential conserves CP, even with the complex couplings of case 2. There is thus a well-defined pseudoscalar particle.

Gérard and Herquet [506] have proposed a *twisted custodial symmetry* which generalizes the formalism presented above. They observed that the transformation matrix R need not be the same for M_{11} and M_{22}, namely

$$M_{11} \rightarrow LM_{11}R^\dagger, \quad M_{22} \rightarrow LM_{22}R'^\dagger. \tag{430}$$

This extra freedom has a limitation, though: since the hypercharge is proportional to the diagonal generator of $SU(2)_R$, the matrices R and R' must be related through $R' = X^\dagger R X$, with $X = \text{diag}\left(e^{i\gamma/2}, e^{-i\gamma/2}\right)$. A specific choice for the phase γ yields the custodial mass relation $m_\pm^2 = m_A^2$; but a different choice imposes degeneracy between the charged Higgs and one of the CP-even scalars, $m_\pm^2 = m_h^2$.

As we know, in the general 2HDM the masses of the $H^\pm$, H, A, and h are arbitrary, since they depend on arbitrary quartic terms in the potential. In the MSSM, the quartic terms are constrained to be gauge couplings, and one finds the relationship $m_{H^\pm}^2 = m_A^2 + m_W^2$. This would rule out decays $H^\pm \rightarrow W^\pm A$. The twisted symmetry allows for a scenario where the pseudoscalar is light (leading to possible decays $h \rightarrow AA$, as discussed in Section 2). A mass spectrum with $m_A < m_{H^\pm}$, $m_H < m_h$ leads to the possibility of the Standard Model-like Higgs decaying into charged Higgs. Any of the four flavour conserving models can be accommodated in this scenario. A detailed discussion of the phenomenology of the model can be found in Ref. [507].

One might wonder about the apparent ambiguity that one has in constructing the 2HDM custodial-symmetric potential in the manner displayed above. For instance, $\text{tr}(M_{11}^\dagger M_{12})$ is also invariant under $SU(2)_L \times SU(2)_R$, and we might insert one such term in the potential, together with many others. So there would seem to be many more models to be considered. Grzadkowski et al. [508], Haber and O'Neil [390], and Nishi [509] have shown that the models Case 1 and Case 2 are indeed related by a basis change—they are not different models, rather they have the same physical predictions. Furthermore, basis-invariant methods were developed to identify whether a given potential has custodial symmetry or not. Grzadkowski et al. and Nishi have used the bilinear formalism to find such conditions for the scalar sector alone. Haber and O'Neil used a different formulation and analysed the Yukawa sector as well. The basis-invariant conditions of Grzadkowski et al. are particularly simple: a necessary and sufficient condition for a potential to possess custodial symmetry is the existence of a three-vector $\mathbf{v}$ such that

$$E\mathbf{v} = 0, \quad \boldsymbol{\xi} \cdot \mathbf{v} = 0, \quad \boldsymbol{\eta} \cdot \mathbf{v} = 0, \tag{431}$$

with E, $\boldsymbol{\xi}$, and $\boldsymbol{\eta}$ defined in Section 6.5. Still regarding the twisted symmetry, Haber and O'Neil [390] argued that the choice of the phase γ discussed above is tantamount to a basis choice. They also showed that in the region of parameter space where the twisted scenario arises CP is not perfectly well defined, and thus the distinction between scalar and pseudoscalar particles is not clear.

Extending the custodial symmetry to the Yukawa sector [390,510] requires the equality of the up-type and down-type-quark mass matrices. Thus, the custodial symmetry is broken by the hypercharge interactions and also by the mass differences amongst quarks.

Notice that, even though we refer to the custodial symmetry as being a symmetry, it does not correspond to any of the six proper symmetries of the 2HDM potential discussed in Section 5.6. For instance, the renormalization-group (RG) running preserves the relations among parameters of the potential following from any of the six symmetries of Table 5, whereas it does *not* preserve the relations following from custodial symmetry. It is easy to see, by using the β-functions of Eqs. (424d)–(424e), that the Case 1 relation $\lambda_4 = \lambda_5$ is unstable under the RG, *i.e.* $\beta_{\lambda_4} \neq \beta_{\lambda_5}$.[34] The fundamental reason why custodial symmetry cannot appear in Table 5 is that those symmetries were obtained by requiring invariance of the scalar gauge-kinetic terms, while the custodial transformations do not leave them invariant. In fact, the largest symmetry group of the 2HDM scalar gauge-kinetic terms was identified in [511] as $U(2) \times SU(2)$, promoted to $Sp(2) \times SO(4)$ in the limit $g' \rightarrow 0$. If one relaxes the requirement of invariance of the gauge-kinetic terms, then the list of possible symmetries of the potential increases, as shown in [395,396], and includes custodial symmetry.

References

[1]   P.W. Higgs, Phys. Rev. Lett. 13 (1964) 508.
[2]   P.W. Higgs, Phys. Rev. 145 (1966) 1156.
[3]   F. Englert, R. Brout, Phys. Rev. Lett. 13 (1964) 321.
[4]   G.S. Guralnik, C.R. Hagen, T.W.B. Kibble, Phys. Rev. Lett. 13 (1964) 585.
[5]   T.W.B. Kibble, Phys. Rev. 155 (1967) 1554.
[6]   P. Langacker, Phys. Rep. 72 (1981) 185.
[7]   K. Nakamura, et al., [Particle Data Group], J. Phys. G 37 (2010) 075021.
[8]   M.S. Chanowitz, M. Golden, Phys. Lett. B 165 (1985) 105.
[9]   T.D. Lee, Phys. Rev. D 8 (1973) 1226.
[10] H.E. Haber, G.L. Kane, Phys. Rep. 117 (1985) 75.
[11] J.E. Kim, Phys. Rep. 150 (1987) 1.
[12] R.D. Peccei, H.R. Quinn, Phys. Rev. Lett. 38 (1977) 1440.
[13] M. Trodden, hep-ph/9805252.

---

[34] This occurs even when one sets the Yukawa couplings to zero, due to the $g^2 g'^2$ term in β_{λ_4}.

[14] N. Turok, J. Zadrozny, Nuclear Phys. B 358 (1991) 471.
[15] M. Joyce, T. Prokopec, N. Turok, Phys. Rev. D 53 (1996) 2958. hep-ph/9410282.
[16] K. Funakubo, A. Kakuto, K. Takenaga, Prog. Theor. Phys. 91 (1994) 341. hep-ph/9310267.
[17] A.T. Davies, C.D. Froggatt, G. Jenkins, R.G. Moorhouse, Phys. Lett. B 336 (1994) 464.
[18] J.M. Cline, K. Kainulainen, A.P. Vischer, Phys. Rev. D 54 (1996) 2451. hep-ph/9506284.
[19] J.M. Cline, P.-A. Lemieux, Phys. Rev. D 55 (1997) 3873. hep-ph/9609240.
[20] M. Laine, K. Rummukainen, Nuclear Phys. B 597 (2001) 23. hep-lat/0009025.
[21] L. Fromme, S.J. Huber, M. Seniuch, J. High Energy Phys. 0611 (2006) 038. hep-ph/0605242.
[22] A. Kozhushko, V. Skalozub, Ukr. J. Phys. 56 (2011) 431. arXiv:1106.0790 [hep-ph].
[23] J.M. Cline, K. Kainulainen, M. Trott, J. High Energy Phys. 1111 (2011) 089. arXiv:1107.3559 [hep-ph].
[24] A. Djouadi, Phys. Rep. 459 (2008) 1. hep-ph/0503173.
[25] B. Mcwilliams, L.F. Li, Nuclear Phys. B 179 (1981) 62.
[26] O.U. Shanker, Nuclear Phys. B 206 (1982) 253.
[27] S.L. Glashow, S. Weinberg, Phys. Rev. D 15 (1977) 1958.
[28] E.A. Paschos, Phys. Rev. D 15 (1977) 1966.
[29] M.S. Carena, H.E. Haber, Prog. Part. Nucl. Phys. 50 (2003) 63. hep-ph/0208209.
[30] V.D. Barger, J.L. Hewett, R.J.N. Phillips, Phys. Rev. D 41 (1990) 3421.
[31] Y. Grossman, Nuclear Phys. B 426 (1994) 355. hep-ph/9401311.
[32] A.G. Akeroyd, W.J. Stirling, Nuclear Phys. B 447 (1995) 3.
[33] A.G. Akeroyd, Phys. Lett. B 377 (1996) 95. hep-ph/9603445.
[34] R.M. Barnett, G. Senjanović, L. Wolfenstein, D. Wyler, Phys. Lett. B 136 (1984) 191.
[35] R.M. Barnett, G. Senjanović, D. Wyler, Phys. Rev. D 30 (1984) 1529.
[36] M. Aoki, S. Kanemura, K. Tsumura, K. Yagyu, Phys. Rev. D 80 (2009) 015017. arXiv:0902.4665 [hep-ph].
[37] A. Pich, P. Tuzón, Phys. Rev. D 80 (2009) 091702. arXiv:0908.1554 [hep-ph].
[38] P. Tuzón, A. Pich, Acta Phys. Polon. Supp. 3 (2010) 215. arXiv:1001.0293 [hep-ph].
[39] P.M. Ferreira, L. Lavoura, J.P. Silva, Phys. Lett. B 688 (2010) 341. arXiv:1001.2561 [hep-ph].
[40] J. Bijnens, J. Lu, J. Rathsman, arXiv:1111.5760 [hep-ph].
[41] H. Serôdio, Phys. Lett. B 700 (2011) 133. arXiv:1104.2545 [hep-ph].
[42] I. de M. Varzielas, Phys. Lett. B 701 (2011) 597. arXiv:1104.2601 [hep-ph].
[43] J.L. Díaz-Cruz, A. Díaz-Furlong, J.H. Montes de Oca, arXiv:1010.0950 [hep-ph].
[44] G. Cree, H.E. Logan, Phys. Rev. D 84 (2011) 055021. arXiv:1106.4039 [hep-ph].
[45] C.B. Braeuninger, A. Ibarra, C. Simonetto, Phys. Lett. B 692 (2010) 189. arXiv:1005.5706 [hep-ph].
[46] F. Mahmoudi, O. Stal, Phys. Rev. D 81 (2010) 035016. arXiv:0907.1791 [hep-ph].
[47] J.F. Gunion, H.E. Haber, G.L. Kane, S. Dawson, Front. Phys. 80 (2000) 1.
[48] A.G. Akeroyd, J. Phys. G 24 (1998) 1983. hep-ph/9803324.
[49] A.G. Akeroyd, M.A. Díaz, J.W.F. Valle, Phys. Lett. B 441 (1998) 224. hep-ph/9806382.
[50] V. Barger, H.E. Logan, G. Shaughnessy, Phys. Rev. D 79 (2009) 115018. arXiv:0902.0170 [hep-ph].
[51] A. Arhrib, R. Benbrik, C.H. Chen, R. Guedes, R. Santos, J. High Energy Phys. 0908 (2009) 035. arXiv:0906.0387 [hep-ph].
[52] D. Eriksson, J. Rathsman, O. Stal, Comput. Phys. Commun. 181 (2010) 189. arXiv:0902.0851 [hep-ph].
[53] H.E. Logan, D. Maclennan, Phys. Rev. D 81 (2010) 075016. arXiv:1002.4916 [hep-ph].
[54] S. Su, B. Thomas, Phys. Rev. D 79 (2009) 095014. arXiv:0903.0667 [hep-ph].
[55] H.E. Logan, D. Maclennan, Phys. Rev. D 79 (2009) 115022. arXiv:0903.2246 [hep-ph].
[56] H.S. Goh, L.J. Hall, P. Kumar, J. High Energy Phys. 0905 (2009) 097. arXiv:0902.0814 [hep-ph].
[57] M.S. Boucenna, S. Profumo, Phys. Rev. D 84 (2011) 055011. arXiv:1106.3368 [hep-ph].
[58] J. Cao, P. Wan, L. Wu, J.M. Yang, Phys. Rev. D 80 (2009) 071701. arXiv:0909.5148 [hep-ph].
[59] F. Jegerlehner, A. Nyffeler, Phys. Rep. 477 (2009) 1. arXiv:0902.3360 [hep-ph].
[60] M. Aoki, S. Kanemura, O. Seto, Phys. Rev. Lett. 102 (2009) 051805. arXiv:0807.0361 [hep-ph].
[61] H.E. Haber, G.L. Kane, T. Sterling, Nuclear Phys. B 161 (1979) 493.
[62] M. Ambroso, V. Braun, B.A. Ovrut, J. High Energy Phys. 0810 (2008) 046. arXiv:0807.3319 [hep-th].
[63] A. Stange, W.J. Marciano, S. Willenbrock, Phys. Rev. D 49 (1994) 1354. hep-ph/9309294.
[64] M.A. Díaz, T.J. Weiler, hep-ph/9401259.
[65] V.D. Barger, N.G. Deshpande, J.L. Hewett, T.G. Rizzo, in: Argonne 1993, Physics at current accelerators and supercolliders. hep-ph/9211234.
[66] A.G. Akeroyd, Phys. Lett. B 368 (1996) 89. hep-ph/9511347.
[67] L. Brücher, R. Santos, Eur. Phys. J. C 12 (2000) 87. hep-ph/9907434.
[68] L. Brücher, R. Santos, hep-ph/0002027.
[69] A. Barroso, L. Brücher, R. Santos, Phys. Rev. D 60 (1999) 035005. hep-ph/9901293.
[70] H. Pois, T.J. Weiler, T.C. Yuan, Phys. Rev. D 47 (1993) 3886. hep-ph/9303277.
[71] A.G. Akeroyd, Nuclear Phys. B 544 (1999) 557. hep-ph/9806337.
[72] A.G. Akeroyd, A. Arhrib, E.M. Naimi, Eur. Phys. J. C 12 (2000) 451. hep-ph/9811431;
     A.G. Akeroyd, A. Arhrib, E.M. Naimi, Eur. Phys. J. C 14 (2000) 371 (erratum).
[73] A.G. Akeroyd, M.A. Díaz, F.J. Pacheco, Phys. Rev. D 70 (2004) 075002. hep-ph/0312231.
[74] A.G. Akeroyd, M.A. Díaz, Phys. Rev. D 67 (2003) 095007. hep-ph/0301203.
[75] A. Dedes, H.E. Haber, J. High Energy Phys. 0105 (2001) 006. hep-ph/0102297.
[76] M. Krawczyk, J. Zochowski, Phys. Rev. D 55 (1997) 6968. hep-ph/9608321.
[77] A.G. Akeroyd, S. Baek, Phys. Lett. B 525 (2002) 315. hep-ph/0105228.
[78] S. Litsey, M. Sher, Phys. Rev. D 80 (2009) 057701. arXiv:0908.0502 [hep-ph].
[79] M. Ciuchini, G. Degrassi, P. Gambino, G.F. Giudice, Nuclear Phys. B 527 (1998) 21. hep-ph/9710335.
[80] F. Borzumati, C. Greub, Phys. Rev. D 58 (1998) 074004. hep-ph/9802391.
[81] M. Misiak, et al., Phys. Rev. Lett. 98 (2007) 022002. hep-ph/0609232.
[82] M. Krawczyk, D. Sokolowska, in: Proceedings of 2007 International Linear Collider Workshop (LCWS07 and ILC07), Hamburg, Germany,
     30 May–3 June 2007. arXiv:0711.4900 [hep-ph].
[83] E.L. Berger, T. Han, J. Jiang, T. Plehn, Phys. Rev. D 71 (2005) 115012. hep-ph/0312286.
[84] J. Alwall, J. Rathsman, J. High Energy Phys. 0412 (2004) 050. hep-ph/0409094.
[85] T. Plehn, Phys. Rev. D 67 (2003) 014018. hep-ph/0206121.
[86] E. Asakawa, O. Brein, S. Kanemura, Phys. Rev. D 72 (2005) 055017. hep-ph/0506249.
[87] M. Krawczyk, D. Temes, Eur. Phys. J. C 44 (2005) 435. hep-ph/0410248.
[88] A.W. El Kaffas, P. Osland, O.M. Ogreid, Phys. Rev. D 76 (2007) 095001. arXiv:0706.2997 [hep-ph].
[89] E. Ma, Phys. Rev. Lett. 86 (2001) 2502. hep-ph/0011121.
[90] S. Gabriel, S. Nandi, Phys. Lett. B 655 (2007) 141. hep-ph/0610253.
[91] S.M. Davidson, H.E. Logan, Phys. Rev. D 80 (2009) 095008. arXiv:0906.3335 [hep-ph].
[92] T. Morozumi, H. Takata, K. Tamai, Phys. Rev. D 85 (2012) 055002. arXiv:1107.1026 [hep-ph].

 [93] A. Djouadi, Phys. Rep. 457 (2008) 1. hep-ph/0503172.
 [94] V.D. Barger, M.S. Berger, P. Ohmann, Phys. Rev. D 47 (1993) 1093. hep-ph/9209232.
 [95] M.S. Carena, M. Olechowski, S. Pokorski, C.E.M. Wagner, Nuclear Phys. B 426 (1994) 269. hep-ph/9402253.
 [96] S. Kanemura, Eur. Phys. J. C 17 (2000) 473. hep-ph/9911541.
 [97] S. Kanemura, T. Kasai, Y. Okada, Phys. Lett. B 471 (1999) 182. hep-ph/9903289.
 [98] A.G. Akeroyd, A. Arhrib, E.M. Naimi, Phys. Lett. B 490 (2000) 119. hep-ph/0006035.
 [99] A. Arhrib, hep-ph/0012353.
[100] J.F. Gunion, H.E. Haber, Phys. Rev. D 67 (2003) 075019. hep-ph/0207010.
[101] S. Mantry, M. Trott, M.B. Wise, Phys. Rev. D 77 (2008) 013006. arXiv:0709.1505 [hep-ph].
[102] L. Randall, J. High Energy Phys. 0802 (2008) 084. arXiv:0711.4360 [hep-ph].
[103] C. Amsler, et al., [Particle Data Group], Phys. Lett. B 667 (2008) 1.
[104] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, K. Moenig, M. Schott, J. Stelzer, arXiv:1107.0975 [hep-ph].
[105] S. Schael, et al., [ALEPH, DELPHI, L3, and OPAL Collaborations], Eur. Phys. J. C 47 (2006) 547. hep-ex/0602042.
[106] J. Kalinowski, M. Krawczyk, Phys. Lett. B 361 (1995) 66. hep-ph/9506291.
[107] R.S. Gupta, J.D. Wells, Phys. Rev. D 81 (2010) 055012. arXiv:0912.0267 [hep-ph].
[108] P. Posch, Phys. Lett. B 696 (2011) 447. arXiv:1001.1759 [hep-ph].
[109] R. Barate, et al., [LEP Working Group for Higgs Boson Searches], Phys. Lett. B 565 (2003) 61. hep-ex/0306033.
[110] T. Aaltonen, et al., [CDF Collaboration], Phys. Rev. Lett. 103 (2009) 061803. arXiv:0905.0413 [hep-ex].
[111] V.M. Abazov, et al., [D0 Collaboration], Phys. Rev. Lett. 101 (2008) 051801. arXiv:0803.1514 [hep-ex].
[112] G.L. Landsberg, [CDF and D0 Collaborations], arXiv:0705.2855 [hep-ex].
[113] A.G. Akeroyd, M.A. Díaz, M.A. Rivera, Phys. Rev. D 76 (2007) 115012. arXiv:0708.1939 [hep-ph].
[114] J. Abdallah, et al., [DELPHI Collaboration], Eur. Phys. J. C 38 (2004) 1. hep-ex/0410017.
[115] Ahmes, The Rhind Papyrus 1900 B.C.
[116] J. Conway, Private communication.
[117] R. Dermisek, J.F. Gunion, Phys. Rev. Lett. 95 (2005) 041801. hep-ph/0502105.
[118] R. Dermisek, J.F. Gunion, Phys. Rev. D 73 (2006) 111701. hep-ph/0510322.
[119] R. Dermisek, J.F. Gunion, Phys. Rev. D 75 (2007) 075019. hep-ph/0611142.
[120] R. Dermisek, J.F. Gunion, B. Mcelrath, Phys. Rev. D 76 (2007) 051105. hep-ph/0612031.
[121] S. Chang, R. Dermisek, J.F. Gunion, N. Weiner, Ann. Rev. Nucl. Part. Sci. 58 (2008) 75. arXiv:0801.4554 [hep-ph].
[122] G. Abbiendi, et al., [OPAL Collaboration], Eur. Phys. J. C 27 (2003) 311. hep-ex/0206022.
[123] S. Schael, et al., [ALEPH Collaboration], J. High Energy Phys. 1005 (2010) 049. arXiv:1003.0705 [hep-ex].
[124] R. Dermisek, J.F. Gunion, Phys. Rev. D 79 (2009) 055014. arXiv:0811.3537 [hep-ph].
[125] C. Kao, G. Lovelace, L.H. Orr, Phys. Lett. B 567 (2003) 259. hep-ph/0305028.
[126] D. Kominis, Nuclear Phys. B 427 (1994) 575. hep-ph/9402339.
[127] H. Baer, C. Kao, X. Tata, Phys. Lett. B 303 (1993) 284.
[128] S. Abdullin, H. Baer, C. Kao, N. Stepanov, X. Tata, Phys. Rev. D 54 (1996) 6728. hep-ph/9603433.
[129] S. Bar-Shalom, S. Nandi, A. Soni, Phys. Rev. D 84 (2011) 053009. arXiv:1105.6095 [hep-ph].
[130] A.W. El Kaffas, P. Osland, O.M. Ogreid, Nonlin. Phenom. Complex Syst. 10 (2007) 347. hep-ph/0702097.
[131] M.A. Luty, Phys. Rev. D 41 (1990) 2893.
[132] G. Burdman, C.E.F. Haluch, J. High Energy Phys. 1112 (2011) 038. arXiv:1109.3914 [hep-ph].
[133] A. Belyaev, R. Guedes, S. Moretti, R. Santos, J. High Energy Phys. 1007 (2010) 051. arXiv:0912.2620 [hep-ph].
[134] A. Belyaev, R. Guedes, S. Moretti, R. Santos, PoS D IS2010 (2010) 207. arXiv:1006.0144 [hep-ph].
[135] S. Kanemura, K. Tsumura, H. Yokoya, arXiv:1111.6089 [hep-ph].
[136] T. Plehn, D.L. Rainwater, Phys. Lett. B 520 (2001) 108. hep-ph/0107180.
[137] T. Han, B. Mcelrath, Phys. Lett. B 528 (2002) 81. hep-ph/0201023.
[138] S. Su, B. Thomas, Phys. Lett. B 677 (2009) 296. arXiv:0812.1798 [hep-ph].
[139] F. Wang, W. Wang, J.M. Yang, Europhys. Lett. 76 (2006) 388. hep-ph/0601018.
[140] M. Sher, C. Triola, Phys. Rev. D 83 (2011) 117702. arXiv:1105.4844 [hep-ph].
[141] S. Zhou, Phys. Rev. D 84 (2011) 038701. arXiv:1106.3880 [hep-ph].
[142] E. Ma, M. Raidal, Phys. Rev. Lett. 87 (2001) 011802. hep-ph/0102255;
       E. Ma, M. Raidal, Phys. Rev. Lett. 87 (2001) 159901 (erratum).
[143] N. Haba, M. Hirotsu, Eur. Phys. J. C 69 (2010) 481. arXiv:1005.1372 [hep-ph].
[144] N. Haba, T. Horita, Phys. Lett. B 705 (2011) 98. arXiv:1107.3203 [hep-ph].
[145] N. Haba, K. Tsumura, J. High Energy Phys. 1106 (2011) 068. arXiv:1105.1409 [hep-ph].
[146] N. Haba, O. Seto, Prog. Theor. Phys. 125 (2011) 1155. arXiv:1102.2889 [hep-ph].
[147] N. Haba, O. Seto, Phys. Rev. D 84 (2011) 103524. arXiv:1106.5354 [hep-ph].
[148] J. Schaarschmidt, Ph.D. Thesis, University of Dresden, 2007.
[149] G. Aad, et al., [ATLAS Collaboration], JINST 3 (2008) S08003.
[150] J. Schaarschmidt, [ATLAS Collaboration], AIP Conf. Proc. 1078 (2009) 220.
[151] T. Plehn, M. Spira, P.M. Zerwas, Nuclear Phys. B 479 (1996) 46. hep-ph/9603205;
       T. Plehn, M. Spira, P.M. Zerwas, Nuclear Phys. B 531 (1998) 655 (erratum).
[152] U. Baur, T. Plehn, D.L. Rainwater, Phys. Rev. D 67 (2003) 033003. hep-ph/0211224.
[153] U. Baur, T. Plehn, D.L. Rainwater, Phys. Rev. D 69 (2004) 053004. hep-ph/0310056.
[154] M. Moretti, S. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, hep-ph/0411039.
[155] A. Djouadi, W. Kilian, M. Mühlleitner, P.M. Zerwas, Eur. Phys. J. C 10 (1999) 45. hep-ph/9904287.
[156] S. Dawson, S. Dittmaier, M. Spira, Phys. Rev. D 58 (1998) 115012. hep-ph/9805244.
[157] M. Moretti, S. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, J. High Energy Phys. 0502 (2005) 024. hep-ph/0410334.
[158] M. Moretti, S. Moretti, F. Piccinini, R. Pittau, J. Rathsman, J. High Energy Phys. 0712 (2007) 075. arXiv:0706.4117 [hep-ph].
[159] M. Moretti, S. Moretti, F. Piccinini, R. Pittau, J. Rathsman, J. High Energy Phys. 1011 (2010) 097. arXiv:1008.0820 [hep-ph].
[160] N.G. Deshpande, E. Ma, Phys. Rev. D 18 (1978) 2574.
[161] E. Ma, Phys. Rev. D 73 (2006) 077301. hep-ph/0601225.
[162] R. Barbieri, L.J. Hall, V.S. Rychkov, Phys. Rev. D 74 (2006) 015007. hep-ph/0603188.
[163] R. Barbieri, A. Strumia, hep-ph/0007265.
[164] H. Martínez, A. Melfo, F. Nesti, G. Senjanović, Phys. Rev. Lett. 106 (2011) 191802. arXiv:1101.3796 [hep-ph].
[165] E. Ma, Modern Phys. Lett. A 21 (2006) 1777. hep-ph/0605180.
[166] D. Majumdar, A. Ghosal, Modern Phys. Lett. A 23 (2008) 2011. hep-ph/0607067.
[167] L.L. Honorez, E. Nezri, J.F. Oliver, M.H.G. Tytgat, JCAP 0702 (2007) 028. hep-ph/0612275.
[168] N. Sahu, U. Sarkar, Phys. Rev. D 76 (2007) 045014. hep-ph/0701062.
[169] M. Gustafsson, E. Lundstrom, L. Bergstrom, J. Edsjö, Phys. Rev. Lett. 99 (2007) 041301. arXiv:astro-ph/0703512.
[170] M. Lisanti, J.G. Wacker, arXiv:0704.2816 [hep-ph].

[171] T. Hambye, F.S. Ling, L.L. Honorez, J. Rocher, J. High Energy Phys. 0907 (2009) 090. arXiv:0903.4010 [hep-ph];
      T. Hambye, F.S. Ling, L.L. Honorez, J. Rocher, J. High Energy Phys. 1005 (2010) 066 (erratum).
[172] A. Melfo, M. Nemevsek, F. Nesti, G. Senjanović, Y. Zhang, Phys. Rev. D 84 (2011) 034009. arXiv:1105.4611 [hep-ph].
[173] I.F. Ginzburg, K.A. Kanishev, M. Krawczyk, D. Sokolowska, Phys. Rev. D 82 (2010) 123533. arXiv:1009.4593 [hep-ph].
[174] D.S. Akerib, et al., [CDMS Collaboration], Phys. Rev. Lett. 96 (2006) 011302. arXiv:astro-ph/0509259.
[175] Q.H. Cao, E. Ma, G. Rajasekaran, Phys. Rev. D 76 (2007) 095011. arXiv:0708.2939 [hep-ph].
[176] E. Lundstrom, M. Gustafsson, J. Edsjö, Phys. Rev. D 79 (2009) 035013. arXiv:0810.3924 [hep-ph].
[177] O.J.P. Eboli, D. Zeppenfeld, Phys. Lett. B 495 (2000) 147. hep-ph/0009158.
[178] B. Grzadkowski, O.M. Ogreid, P. Osland, Phys. Rev. D 80 (2009) 055013. arXiv:0904.2173 [hep-ph].
[179] B. Grzadkowski, O.M. Ogreid, P. Osland, A. Pukhov, M. Purmohammadi, J. High Energy Phys. 1106 (2011) 003. arXiv:1012.4680 [hep-ph].
[180] S.M. Barr, T.W. Kephart, Phys. Rev. D 85 (2012) 035010. arXiv:1111.0963 [hep-ph].
[181] T.D. Lee, G.C. Wick, Nuclear Phys. B 9 (1969) 209.
[182] T.D. Lee, G.C. Wick, Phys. Rev. D 2 (1970) 1033.
[183] B. Grinstein, D. O'Connell, M.B. Wise, Phys. Rev. D 77 (2008) 025012. arXiv:0704.1845 [hep-ph].
[184] B. Grinstein, D. O'Connell, M.B. Wise, Phys. Rev. D 79 (2009) 105019. arXiv:0805.2156 [hep-th].
[185] T.G. Rizzo, J. High Energy Phys. 0706 (2007) 070. arXiv:0704.3458 [hep-ph].
[186] T.R. Dulaney, M.B. Wise, Phys. Lett. B 658 (2008) 230. arXiv:0708.0567 [hep-ph].
[187] E. Alvarez, L. Da Rold, C. Schat, A. Szynkman, J. High Energy Phys. 0804 (2008) 026. arXiv:0802.1061 [hep-ph].
[188] T.E.J. Underwood, R. Zwicky, Phys. Rev. D 79 (2009) 035016. arXiv:0805.3296 [hep-ph].
[189] C.D. Carone, R.F. Lebed, Phys. Lett. B 668 (2008) 221. arXiv:0806.4555 [hep-ph].
[190] C.D. Carone, R.F. Lebed, J. High Energy Phys. 0901 (2009) 043. arXiv:0811.4150 [hep-ph].
[191] B. Grinstein, D. O'Connell, Phys. Rev. D 78 (2008) 105005. arXiv:0801.4034 [hep-ph].
[192] C.D. Carone, Phys. Lett. B 677 (2009) 306. arXiv:0904.2359 [hep-ph].
[193] B. Fornal, B. Grinstein, M.B. Wise, Phys. Lett. B 674 (2009) 330. arXiv:0902.1585 [hep-th].
[194] A.M. Shalaby, Phys. Rev. D 80 (2009) 025006. arXiv:0812.3419 [hep-th].
[195] F. Wu, M. Zhong, Phys. Rev. D 78 (2008) 085010. arXiv:0807.0132 [hep-ph].
[196] J.R. Espinosa, B. Grinstein, D. O'Connell, M.B. Wise, Phys. Rev. D 77 (2008) 085002. arXiv:0705.1188 [hep-ph].
[197] F. Krauss, T.E.J. Underwood, R. Zwicky, Phys. Rev. D 77 (2008) 015012. arXiv:0709.4054 [hep-ph];
      F. Krauss, T.E.J. Underwood, R. Zwicky, Phys. Rev. D 83 (2011) 019902 (erratum).
[198] C.D. Carone, R. Primulando, Phys. Rev. D 80 (2009) 055020. arXiv:0908.0342 [hep-ph].
[199] R.S. Chivukula, A. Farzinnia, R. Foadi, E.H. Simmons, Phys. Rev. D 81 (2010) 095015. arXiv:1002.0343 [hep-ph].
[200] E. Alvarez, E.C. Leskow, J. Zurita, Phys. Rev. D 83 (2011) 115024. arXiv:1104.3496 [hep-ph].
[201] Q.H. Cao, M. Carena, S. Gori, A. Menon, P. Schwaller, C.E.M. Wagner, L.T.M. Wang, J. High Energy Phys. 1108 (2011) 002. arXiv:1104.4776 [hep-ph].
[202] T. Aaltonen, et al., [CDF Collaboration], Phys. Rev. Lett. 106 (2011) 171801. arXiv:1104.0699 [hep-ex].
[203] C.H. Chen, C.W. Chiang, T. Nomura, Y. Fusheng, arXiv:1105.2870 [hep-ph].
[204] B. Dutta, S. Khalil, Y. Mimura, Q. Shafi, arXiv:1104.5209 [hep-ph].
[205] J.D. Bjorken, S. Weinberg, Phys. Rev. Lett. 38 (1977) 622.
[206] T.P. Cheng, M. Sher, Phys. Rev. D 35 (1987) 3484.
[207] J.L. Díaz-Cruz, R. Noriega-Papaqui, A. Rosado, Phys. Rev. D 69 (2004) 095002. hep-ph/0401194.
[208] A. Antaramian, L.J. Hall, A. Rašin, Phys. Rev. Lett. 69 (1992) 1871. hep-ph/9206205.
[209] W.S. Hou, Phys. Lett. B 296 (1992) 179.
[210] M. Gronau, R. Johnson, J. Schechter, Phys. Rev. D 39 (1989) 1913.
[211] M. Sher, Y. Yuan, Phys. Rev. D 44 (1991) 1461.
[212] D. Chang, W.S. Hou, W.Y. Keung, Phys. Rev. D 48 (1993) 217. hep-ph/9302267.
[213] M.E. Luke, M.J. Savage, Phys. Lett. B 307 (1993) 387. hep-ph/9303249.
[214] T.S. Kosmas, G.K. Leontaris, J.D. Vergados, Prog. Part. Nucl. Phys. 33 (1994) 397. hep-ph/9312217.
[215] L. Wolfenstein, Y.L. Wu, Phys. Rev. Lett. 73 (1994) 2809. hep-ph/9410253.
[216] D. Atwood, L. Reina, A. Soni, Phys. Rev. D 55 (1997) 3156. hep-ph/9609279.
[217] E. Golowich, J. Hewett, S. Pakvasa, A.A. Petrov, Phys. Rev. D 76 (2007) 095009. arXiv:0705.3650 [hep-ph].
[218] S. Davidson, G.J. Grenier, Phys. Rev. D 81 (2010) 095016. arXiv:1001.0434 [hep-ph].
[219] A. Cordero-Cid, M.A. Pérez, G. Tavares-Velasco, J.J. Toscano, Phys. Rev. D 70 (2004) 074003. hep-ph/0407127.
[220] F. Larios, R. Martínez, M.A. Pérez, Internat. J. Modern Phys. A 21 (2006) 3473. hep-ph/0605003.
[221] J.I. Aranda, A. Cordero-Cid, F. Ramírez-Zavaleta, J.J. Toscano, E.S. Tututi, Phys. Rev. D 81 (2010) 077701. arXiv:0911.2304 [hep-ph].
[222] A.S. Joshipura, B.P. Kodrani, Phys. Rev. D 82 (2010) 115013. arXiv:1004.3637 [hep-ph].
[223] B. Aubert, et al., [BaBar Collaboration], Phys. Rev. Lett. 99 (2007) 201801. arXiv:0708.1303 [hep-ex].
[224] S. Nie, M. Sher, Phys. Rev. D 58 (1998) 097701. hep-ph/9805376.
[225] E. de Rafael, PoS EFT09 (2009) 050.
[226] J. Prades, Acta Phys. Polon. Supp. 3 (2010) 75. arXiv:0909.2546 [hep-ph].
[227] R.A. Díaz, R. Martínez, J.A. Rodríguez, Phys. Rev. D 67 (2003) 075011. hep-ph/0208117.
[228] R.A. Díaz, R. Martínez, C.E. Sandoval, Eur. Phys. J. C 41 (2005) 305. hep-ph/0406265.
[229] W.S. Hou, F.F. Lee, C.Y. Ma, Phys. Rev. D 79 (2009) 073002. arXiv:0812.0064 [hep-ph].
[230] W. Li, Y. Ma, G. Liu, W. Guo, arXiv:0812.0727 [hep-ph].
[231] W.J. Li, Y.Y. Fan, G.W. Liu, L.X. Lu, Internat. J. Modern Phys. A 25 (2010) 4827. arXiv:1007.2894 [hep-ph].
[232] D. Bowser-Chao, K.M. Cheung, W.Y. Keung, Phys. Rev. D 59 (1999) 115006. hep-ph/9811235.
[233] Z.J. Xiao, L. Guo, Phys. Rev. D 69 (2004) 014002. hep-ph/0309103.
[234] J.P. Idarraga, R. Martínez, J.A. Rodríguez, N. Poveda, hep-ph/9509072.
[235] C.S. Huang, J.T. Li, Internat. J. Modern Phys. A 20 (2005) 161. hep-ph/0405294.
[236] A.E. Blechman, A.A. Petrov, G. Yeghiyan, J. High Energy Phys. 1011 (2010) 075. arXiv:1009.1612 [hep-ph].
[237] G.C. Branco, W. Grimus, L. Lavoura, Phys. Lett. B 380 (1996) 119. hep-ph/9601383.
[238] G. D'Ambrosio, G.F. Giudice, G. Isidori, A. Strumia, Nuclear Phys. B 645 (2002) 155. hep-ph/0207036.
[239] R.S. Chivukula, H. Georgi, Phys. Lett. B 188 (1987) 99.
[240] A.J. Buras, P. Gambino, M. Gorbahn, S. Jäger, L. Silvestrini, Phys. Lett. B 500 (2001) 161. hep-ph/0007085.
[241] M. Blanke, A.J. Buras, D. Guadagnoli, C. Tarantino, J. High Energy Phys. 0610 (2006) 003. hep-ph/0604057.
[242] R. Alonso, M.B. Gavela, L. Merlo, S. Rigolin, J. High Energy Phys. 1107 (2011) 012. arXiv:1103.2915 [hep-ph].
[243] L.J. Hall, S. Weinberg, Phys. Rev. D 48 (1993) 979. hep-ph/9303241.
[244] A.S. Joshipura, Modern Phys. Lett. A 6 (1991) 1693.
[245] A.S. Joshipura, S.D. Rindani, Phys. Lett. B 260 (1991) 149.
[246] M. Jung, A. Pich, P. Tuzón, J. High Energy Phys. 1011 (2010) 003. arXiv:1006.0470 [hep-ph].
[247] M. Jung, A. Pich, P. Tuzón, Phys. Rev. D 83 (2011) 074011. arXiv:1011.5154 [hep-ph].
[248] L. Lavoura, Internat. J. Modern Phys. A 9 (1994) 1873.
[249] F.J. Botella, G.C. Branco, M.N. Rebelo, Phys. Lett. B 687 (2010) 194. arXiv:0911.1753 [hep-ph].

[250] F.J. Botella, G.C. Branco, M. Nebot, M.N. Rebelo, J. High Energy Phys. 1110 (2011) 037. arXiv:1102.0520 [hep-ph].
[251] A.S. Joshipura, B.P. Kodrani, Phys. Lett. B 670 (2009) 369. arXiv:0706.0953 [hep-ph].
[252] F.J. Botella, M. Nebot, O. Vives, J. High Energy Phys. 0601 (2006) 106. hep-ph/0407349.
[253] A.J. Buras, M.V. Carlucci, S. Gori, G. Isidori, J. High Energy Phys. 1010 (2010) 009. arXiv:1005.5310 [hep-ph].
[254] V. Cirigliano, B. Grinstein, G. Isidori, M.B. Wise, Nuclear Phys. B 728 (2005) 121. hep-ph/0507001.
[255] S. Davidson, F. Palorini, Phys. Lett. B 642 (2006) 72. hep-ph/0607329.
[256] G.C. Branco, A.J. Buras, S. Jäger, S. Uhlig, A. Weiler, J. High Energy Phys. 0709 (2007) 004. hep-ph/0609067.
[257] H. Georgi, Ann. Rev. Nucl. Part. Sci. 43 (1993) 209.
[258] Y. Tomozawa, Phys. Rev. D 18 (1978) 2556.
[259] J.F. Donoghue, L.F. Li, Phys. Rev. D 19 (1979) 945.
[260] http://www.grid.tsl.uu.se/chargedhiggs2006/.
[261] http://www.grid.tsl.uu.se/chargedhiggs2008/.
[262] http://www.grid.tsl.uu.se/chargedhiggs2010/.
[263] R. Guedes, S. Kanemura, S. Moretti, R. Santos, K. Yagyu, PoS CHarged 2010 (2010) 037. arXiv:1102.3791 [hep-ph].
[264] J.A. Coarasa Pérez, J. Guasch, J. Solà, W. Hollik, Phys. Lett. B 442 (1998) 326. hep-ph/9808278.
[265] S. Bertolini, F. Borzumati, A. Masiero, Phys. Rev. Lett. 59 (1987) 180.
[266] N.G. Deshpande, P. Lo, J. Trampetic, G. Eilam, P. Singer, Phys. Rev. Lett. 59 (1987) 183.
[267] B. Grinstein, M.B. Wise, Phys. Lett. B 201 (1988) 274.
[268] B. Grinstein, R.P. Springer, M.B. Wise, Phys. Lett. B 202 (1988) 138.
[269] R. Grigjanis, P.J. O'donnell, M. Sutherland, H. Navelet, Phys. Lett. B 213 (1988) 355;
      R. Grigjanis, P.J. O'donnell, M. Sutherland, H. Navelet, Phys. Lett. B 286 (1992) 413 (erratum).
[270] G. Cella, G. Curci, G. Ricciardi, A. Vicere, Phys. Lett. B 248 (1990) 181.
[271] G. Cella, G. Curci, G. Ricciardi, A. Vicere, Phys. Lett. B 325 (1994) 227. hep-ph/9401254.
[272] G. Cella, G. Curci, G. Ricciardi, A. Vicere, Nuclear Phys. B 431 (1994) 417. hep-ph/9406203.
[273] M. Ciuchini, E. Franco, G. Martinelli, L. Reina, L. Silvestrini, Phys. Lett. B 316 (1993) 127. hep-ph/9307364.
[274] M. Ciuchini, E. Franco, L. Reina, L. Silvestrini, Nuclear Phys. B 421 (1994) 41. hep-ph/9311357.
[275] A.J. Buras, M. Misiak, M. Münz, S. Pokorski, Nuclear Phys. B 424 (1994) 374. hep-ph/9311345.
[276] A. Ali, C. Greub, Z. Phys. C 60 (1993) 433.
[277] M. Misiak, Nuclear Phys. B 393 (1993) 23;
      M. Misiak, Nuclear Phys. B 439 (1995) 461 (erratum).
[278] K. Adel, Y.P. Yao, Phys. Rev. D 49 (1994) 4945. hep-ph/9308349.
[279] C. Greub, T. Hurth, D. Wyler, Phys. Rev. D 54 (1996) 3350. hep-ph/9603404.
[280] A.J. Buras, A. Kwiatkowski, N. Pott, Phys. Lett. B 414 (1997) 157. hep-ph/9707482;
      A.J. Buras, A. Kwiatkowski, N. Pott, Phys. Lett. B 434 (1998) 459 (erratum).
[281] U. Haisch, arXiv:0805.2141 [hep-ph].
[282] T. Hurth, M. Nakao, Ann. Rev. Nucl. Part. Sci. 60 (2010) 645. arXiv:1005.1224 [hep-ph].
[283] T. Hurth, PoS CHarged 2010 (2010) 020. arXiv:1104.5123 [hep-ph].
[284] P. Ciafaloni, A. Romanino, A. Strumia, Nuclear Phys. B 524 (1998) 361. hep-ph/9710312.
[285] F. Borzumati, C. Greub, Y. Yamada, Phys. Rev. D 69 (2004) 055005. hep-ph/0311151.
[286] C. Bobeth, M. Misiak, J. Urban, Nuclear Phys. B 574 (2000) 291. hep-ph/9910220.
[287] A. Limosani, et al., [Belle Collaboration], Phys. Rev. Lett. 103 (2009) 241801. arXiv:0907.1384 [hep-ex].
[288] W.S. Hou, Phys. Rev. D 48 (1993) 2342.
[289] Y. Grossman, Z. Ligeti, Phys. Lett. B 332 (1994) 373. hep-ph/9403376.
[290] K. Hara, et al., [Belle Collaboration], Phys. Rev. D 82 (2010) 071101. arXiv:1006.4201 [hep-ex].
[291] K. Ikado, et al., [Belle Collaboration], Phys. Rev. Lett. 97 (2006) 251802. hep-ex/0604018.
[292] B. Aubert, et al. [BaBar Collaboration], arXiv:0912.2453 [hep-ex].
[293] Heavy Flavor Averaging Group. http://www.slac.stanford.edu/xorg/hfag/.
[294] A. Czarnecki, W.J. Marciano, Phys. Rev. Lett. 81 (1998) 277. hep-ph/9804252.
[295] M. Tanaka, R. Watanabe, Phys. Rev. D 82 (2010) 034027. arXiv:1005.4306 [hep-ph].
[296] Y. Grossman, H.E. Haber, Y. Nir, Phys. Lett. B 357 (1995) 630. hep-ph/9507213.
[297] H.E. Haber, H.E. Logan, Phys. Rev. D 62 (2000) 015011. hep-ph/9909335.
[298] G. Degrassi, P. Slavich, Phys. Rev. D 81 (2010) 075001. arXiv:1002.1071 [hep-ph].
[299] C.Q. Geng, J.N. Ng, Phys. Rev. D 38 (1988) 2857;
      C.Q. Geng, J.N. Ng, Phys. Rev. D 41 (1990) 1715 (erratum).
[300] O. Deschamps, S. Descotes-Genon, S. Monteil, V. Niess, S. T'jampens, V. Tisserand, Phys. Rev. D 82 (2010) 073012. arXiv:0907.5135 [hep-ph].
[301] A. Djouadi, J. Kalinowski, M. Spira, Comput. Phys. Commun. 108 (1998) 56. hep-ph/9704448.
[302] S. Kanemura, S. Moretti, Y. Mukai, R. Santos, K. Yagyu, Phys. Rev. D 79 (2009) 055017. arXiv:0901.0204 [hep-ph].
[303] LEP Working Group for Higgs Boson Searches. hep-ex/0107031.
[304] T. Aaltonen, et al., [CDF Collaboration], Phys. Rev. Lett. 103 (2009) 101803. arXiv:0907.1269 [hep-ex].
[305] V.M. Abazov, et al., [D0 Collaboration], Phys. Lett. B 682 (2009) 278. arXiv:0908.1811 [hep-ex].
[306] V.M. Abazov, et al., [D0 Collaboration], Phys. Rev. D 84 (2011) 092002. arXiv:1107.1268 [hep-ex].
[307] A. Buzatu, arXiv:1110.5349 [hep-ex].
[308] G. Aad, et al., [ATLAS Collaboration], arXiv:0901.0512 [hep-ex].
[309] M. Baarmand, M. Hashemi, A. Nikitenko, J. Phys. G 32 (2006) N21.
[310] A. Ali, F. Barreiro, J. Llorente, Eur. Phys. J. C 71 (2011) 1737. arXiv:1103.1827 [hep-ph].
[311] M. Hashemi, arXiv:1109.5356 [hep-ph].
[312] S. Yang, Q.-S. Yan, J. High Energy Phys. 1202 (2012) 074. arXiv:1111.4530 [hep-ph].
[313] J.L. Díaz-Cruz, O.A. Sampayo, Phys. Rev. D 50 (1994) 6820.
[314] J.F. Gunion, H.E. Haber, F.E. Paige, W.K. Tung, S.S.D. Willenbrock, Nuclear Phys. B 294 (1987) 621.
[315] S. Moretti, K. Odagiri, Phys. Rev. D 55 (1997) 5627. hep-ph/9611374.
[316] H.J. He, C.P. Yuan, Phys. Rev. Lett. 83 (1999) 28. hep-ph/9810367.
[317] S. Dittmaier, G. Hiller, T. Plehn, M. Spannowsky, Phys. Rev. D 77 (2008) 115001. arXiv:0708.0940 [hep-ph].
[318] A.A.B. Bendezu, B.A. Kniehl, Phys. Rev. D 59 (1999) 015009. hep-ph/9807480.
[319] S. Moretti, K. Odagiri, Phys. Rev. D 59 (1999) 055008. hep-ph/9809244.
[320] O. Brein, W. Hollik, S. Kanemura, Phys. Rev. D 63 (2001) 095001. hep-ph/0008308.
[321] D. Eriksson, S. Hesselbach, J. Rathsman, Eur. Phys. J. C 53 (2008) 267. hep-ph/0612198.
[322] M. Hashemi, Phys. Rev. D 83 (2011) 055004. arXiv:1008.3785 [hep-ph].
[323] R. Enberg, R. Pasechnik, Phys. Rev. D 83 (2011) 095020. arXiv:1104.0889 [hep-ph].
[324] S.-S. Bao, X. Gong, H.-L. Li, S.-Y. Li, Z.-G. Si, arXiv:1112.0086 [hep-ph].
[325] S. Kanemura, C.P. Yuan, Phys. Lett. B 530 (2002) 188. hep-ph/0112165.
[326] Q.H. Cao, S. Kanemura, C.P. Yuan, Phys. Rev. D 69 (2004) 075008. hep-ph/0311083.

[327] A. Belyaev, Q.H. Cao, D. Nomura, K. Tobe, C.P. Yuan, Phys. Rev. Lett. 100 (2008) 061801. hep-ph/0609079.
[328] A.G. Akeroyd, Phys. Rev. D 68 (2003) 077701. hep-ph/0306045.
[329] D.K. Ghosh, S. Moretti, Eur. Phys. J. C 42 (2005) 341. hep-ph/0412365.
[330] A. Alves, T. Plehn, Phys. Rev. D 71 (2005) 115014. hep-ph/0503135.
[331] S. Moretti, J. Phys. G 28 (2002) 2567. hep-ph/0102116.
[332] K. Huitu, S.K. Rai, K. Rao, S.D. Rindani, P. Sharma, J. High Energy Phys. 1104 (2011) 026. arXiv:1012.0527 [hep-ph].
[333] J. Baglio, M. Beccaria, A. Djouadi, G. Macorini, E. Mirabella, N. Orlando, F.M. Renard, C. Verzegnassi, Phys. Lett. B 705 (2011) 212. arXiv:1109.2420 [hep-ph].
[334] S.-S. Bao, Y. Tang, Y.-L. Wu, Phys. Rev. D 83 (2011) 075006. arXiv:1011.1409 [hep-ph].
[335] G.L. Bayatian, et al. [CMS collaboration], "CMS physics: technical design report".
[336] F. Borzumati, A. Djouadi, Phys. Lett. B 549 (2002) 170. hep-ph/9806301.
[337] A.G. Akeroyd, A. Arhrib, E. Naimi, Eur. Phys. J. C 20 (2001) 51. hep-ph/0002288.
[338] M.S. Carena, et al., [Higgs Working Group Collaboration]. hep-ph/0010338.
[339] M. Drees, M. Guchait, D.P. Roy, Phys. Lett. B 471 (1999) 39. hep-ph/9909266.
[340] K.A. Assamagan, Y. Coadou, A. Deandrea, Eur. Phys. J. C 4 (2002) 9. hep-ph/0203121.
[341] A.G. Akeroyd, in: [262].
[342] J.H. Field, Modern Phys. Lett. A 13 (1998) 1937. hep-ph/9801355.
[343] J. Urban, F. Krauss, U. Jentschura, G. Soff, Nuclear Phys. B 523 (1998) 40. hep-ph/9710245.
[344] A.J. Buras, D. Guadagnoli, Phys. Rev. D 78 (2008) 033005. arXiv:0805.3887 [hep-ph].
[345] A.J. Buras, D. Guadagnoli, G. Isidori, Phys. Lett. B 688 (2010) 309. arXiv:1002.3612 [hep-ph].
[346] J. Abdallah, et al., [DELPHI Collaboration], Eur. Phys. J. C 34 (2004) 399. hep-ex/0404012.
[347] G. Abbiendi, et al., [OPAL Collaboration]. arXiv:0812.0267 [hep-ex].
[348] M. Aoki, R. Guedes, S. Kanemura, S. Moretti, R. Santos, K. Yagyu, Phys. Rev. D 84 (2011) 055028. arXiv:1104.3178 [hep-ph].
[349] J.H. Park, J. High Energy Phys. 0610 (2006) 077. hep-ph/0607280.
[350] G. Abbiendi, et al., [OPAL Collaboration], Eur. Phys. J. C 32 (2004) 453. hep-ex/0309014.
[351] S.M. Davidson, H.E. Logan, Phys. Rev. D 82 (2010) 115031. arXiv:1009.4413 [hep-ph].
[352] A. Heister, et al., [ALEPH Collaboration], Phys. Lett. B 543 (2002) 1. hep-ex/0207054.
[353] ATLAS note ATLAS-PHYS-PUB-2010-009.
[354] A. Ferrari, [ATLAS Collaboration], PoS CHARGED 2010 (2010) 010.
[355] A.G. Akeroyd, hep-ph/9509203.
[356] X. Miao, S. Su, B. Thomas, Phys. Rev. D 82 (2010) 035009. arXiv:1005.0090 [hep-ph].
[357] E. Dolle, X. Miao, S. Su, B. Thomas, Phys. Rev. D 81 (2010) 035003. arXiv:0909.3094 [hep-ph].
[358] K. Huitu, K. Kannike, A. Racioppi, M. Raidal, J. High Energy Phys. 1101 (2011) 010. arXiv:1005.4409 [hep-ph].
[359] S. Davidson, H.E. Haber, Phys. Rev. D 72 (2005) 035004. hep-ph/0504050;
      S. Davidson, H.E. Haber, Phys. Rev. D 72 (2005) 099902 (erratum).
[360] H.E. Haber, D. O'neil, Phys. Rev. D 74 (2006) 015018. hep-ph/0602242.
[361] A. Cordero-Cid, O. Felix-Beltrán, J. Hernández-Sánchez, R. Noriega-Papaqui, PoS CHARGED 2010 (2010) 042. arXiv:1105.4951 [hep-ph].
[362] J. Hernández-Sánchez, L. López-Lozano, R. Noriega-Papaqui, A. Rosado, arXiv:1106.5035 [hep-ph].
[363] V.M. Abazov, et al., [D0 Collaboration], Phys. Rev. Lett. 102 (2009) 191802. arXiv:0807.0859 [hep-ex].
[364] H. Cardenas, J.A. Rodríguez, Modern Phys. Lett. A 26 (2011) 1869. arXiv:0810.3046 [hep-ph].
[365] R. Martínez, J.A. Rodríguez, M. Rozo, Phys. Rev. D 68 (2003) 035001. hep-ph/0212236.
[366] Z.J. Xiao, C. Zhuang, Eur. Phys. J. C 33 (2004) 349. hep-ph/0310097.
[367] R.A. Díaz, R. Martínez, C.E. Sandoval, Eur. Phys. J. C 46 (2006) 403. hep-ph/0509194.
[368] J.P. Idarraga, R. Martínez, J.A. Rodríguez, N. Poveda, Braz. J. Phys. 38 (2008) 531.
[369] Y.L. Wu, L. Wolfenstein, Phys. Rev. Lett. 73 (1994) 1762. hep-ph/9409421.
[370] F.J. Botella, J.P. Silva, Phys. Rev. D 51 (1995) 3870. hep-ph/9411288.
[371] J. Velhinho, R. Santos, A. Barroso, Phys. Lett. B 322 (1994) 213.
[372] F. Nagel, Ph.D. Thesis, University of Heidelberg, 2004.
[373] M. Maniatis, A. Von Manteuffel, O. Nachtmann, F. Nagel, Eur. Phys. J. C 48 (2006) 805. hep-ph/0605184.
[374] M. Maniatis, A. Von Manteuffel, O. Nachtmann, Eur. Phys. J. C 49 (2007) 1067. hep-ph/0608314.
[375] M. Maniatis, A. Von Manteuffel, O. Nachtmann, Eur. Phys. J. C 57 (2008) 719. arXiv:0707.3344 [hep-ph].
[376] C.C. Nishi, Phys. Rev. D 74 (2006) 036003. hep-ph/0605153;
      C.C. Nishi, Phys. Rev. D 76 (2007) 119901 (erratum).
[377] C.C. Nishi, Phys. Rev. D 76 (2007) 055013. arXiv:0706.2685 [hep-ph].
[378] C.C. Nishi, Phys. Rev. D 77 (2008) 055009. arXiv:0712.4260 [hep-ph].
[379] I.P. Ivanov, Phys. Lett. B 632 (2006) 360. hep-ph/0507132.
[380] I.P. Ivanov, Phys. Rev. D 75 (2007) 035001. hep-ph/0609018;
      I.P. Ivanov, Phys. Rev. D 76 (2007) 039902 (erratum).
[381] A. Degee, I.P. Ivanov, Phys. Rev. D 81 (2010) 015012. arXiv:0910.4492 [hep-ph].
[382] H. Neufeld, W. Grimus, G. Ecker, Internat. J. Modern Phys. A 3 (1988) 603.
[383] G. Ecker, W. Grimus, H. Neufeld, J. Phys. A 20 (1987) L807.
[384] G. Ecker, W. Grimus, H. Neufeld, Phys. Lett. B 228 (1989) 401.
[385] W. Grimus, M.N. Rebelo, Phys. Rep. 281 (1997) 239. hep-ph/9506272.
[386] P.M. Ferreira, H.E. Haber, J.P. Silva, Phys. Rev. D 79 (2009) 116004. arXiv:0902.1537 [hep-ph].
[387] M. Maniatis, O. Nachtmann, J. High Energy Phys. 0905 (2009) 028. arXiv:0901.4341 [hep-ph].
[388] M. Maniatis, O. Nachtmann, J. High Energy Phys. 1004 (2010) 027. arXiv:0912.2727 [hep-ph].
[389] M. Maniatis, O. Nachtmann, A. von Manteuffel, arXiv:1009.1869 [hep-ph].
[390] H.E. Haber, D. O'neil, Phys. Rev. D 83 (2011) 055017. arXiv:1011.6188 [hep-ph].
[391] P.M. Ferreira, J.P. Silva, Eur. Phys. J. C 69 (2010) 45. arXiv:1001.0574 [hep-ph].
[392] P.M. Ferreira, H.E. Haber, J.P. Silva, Phys. Rev. D 82 (2010) 016001. arXiv:1004.3292 [hep-ph].
[393] E. Ma, M. Maniatis, Phys. Lett. B 683 (2010) 33. arXiv:0909.2855 [hep-ph].
[394] M. Maniatis, O. Nachtmann, J. High Energy Phys. 1111 (2011) 151. arXiv:1106.1436 [hep-ph].
[395] R.A. Battye, G.D. Brawn, A. Pilaftsis, J. High Energy Phys. 1108 (2011) 020. arXiv:1106.3482 [hep-ph].
[396] A. Pilaftsis, Phys. Lett. B 706 (2012) 465. arXiv:1109.3787 [hep-ph].
[397] P.M. Ferreira, M. Maniatis, O. Nachtmann, J.P. Silva, J. High Energy Phys. 1008 (2010) 125. arXiv:1004.3207 [hep-ph].
[398] P.M. Ferreira, H.E. Haber, M. Maniatis, O. Nachtmann, J.P. Silva, Internat. J. Modern Phys. A 26 (2011) 769. arXiv:1010.0935 [hep-ph].
[399] T.D. Lee, G.C. Wick, Phys. Rev. 148 (1966) 1385.
[400] J. Bernabéu, G.C. Branco, M. Gronau, Phys. Lett. B 169 (1986) 243.
[401] G. Ecker, W. Grimus, W. Konetschny, Nuclear Phys. B 191 (1981) 465.
[402] G. Ecker, W. Grimus, H. Neufeld, Nuclear Phys. B 247 (1984) 70.
[403] H.E. Haber, R. Hempfling, Phys. Rev. D 48 (1993) 4280. hep-ph/9307201.

[404] P.M. Ferreira, J.P. Silva, Phys. Rev. D 78 (2008) 116007. arXiv:0809.2788 [hep-ph].
[405] P.M. Ferreira, R. Santos, A. Barroso, Phys. Lett. B 603 (2004) 219. hep-ph/0406231;
      P.M. Ferreira, R. Santos, A. Barroso, Phys. Lett. B 629 (2005) 114 (erratum).
[406] K.G. Klimenko, Theor. Math. Phys. 62 (1985) 58.
[407] P.M. Ferreira, D.R.T. Jones, J. High Energy Phys. 0908 (2009) 069. arXiv:0903.2856 [hep-ph].
[408] M. Lindner, Z. Phys. C 31 (1986) 295.
[409] M. Sher, Phys. Rep. 179 (1989) 273.
[410] M. Sher, Phys. Lett. B 317 (1993) 159. hep-ph/9307342;
      M. Sher, Phys. Lett. B 331 (1994) 448. [Addendum].
[411] J.A. Casas, J.R. Espinosa, M. Quirós, Phys. Lett. B 342 (1995) 171. hep-ph/9409458.
[412] J.A. Casas, J.R. Espinosa, M. Quirós, Phys. Lett. B 382 (1996) 374. hep-ph/9603227.
[413] J.R. Espinosa, M. Quirós, Phys. Lett. B 353 (1995) 257. hep-ph/9504241.
[414] G. Kreyerhoff, R. Rodenberg, Phys. Lett. B 226 (1989) 323.
[415] J. Freund, G. Kreyerhoff, R. Rodenberg, Phys. Lett. B 280 (1992) 267.
[416] B.M. Kastening, hep-ph/9307224.
[417] S. Nie, M. Sher, Phys. Lett. B 449 (1999) 89. hep-ph/9811234.
[418] A. Barroso, P.M. Ferreira, R. Santos, Phys. Lett. B 652 (2007) 181. hep-ph/0702098.
[419] A. Barroso, P.M. Ferreira, R. Santos, J.P. Silva, Phys. Rev. D 74 (2006) 085016. hep-ph/0608282.
[420] J.L. Díaz-Cruz, A. Méndez, Nuclear Phys. B 380 (1992) 39.
[421] I.P. Ivanov, Phys. Rev. D 77 (2008) 015017. arXiv:0710.3490 [hep-ph].
[422] J.-M. Frère, D.R.T. Jones, S. Raby, Nuclear Phys. B 222 (1983) 11.
[423] A. Barroso, P.M. Ferreira, R. Santos, Phys. Lett. B 632 (2006) 684. hep-ph/0507224.
[424] A. Pilaftsis, C.E.M. Wagner, Nuclear Phys. B 553 (1999) 3. hep-ph/9902371.
[425] L. Lavoura, J.P. Silva, Phys. Rev. D 50 (1994) 4619. hep-ph/9404276.
[426] L. Lavoura, Phys. Rev. D 50 (1994) 7089. hep-ph/9405307.
[427] P.M. Ferreira, L. Lavoura, J.P. Silva, Phys. Lett. B 704 (2011) 179. arXiv:1102.0784 [hep-ph].
[428] P.M. Ferreira, J.P. Silva, Phys. Rev. D 83 (2011) 065026. arXiv:1012.2874 [hep-ph].
[429] M. Maniatis, A. von Manteuffel, O. Nachtmann, Eur. Phys. J. C 57 (2008) 739. arXiv:0711.3760 [hep-ph].
[430] G.C. Branco, L. Lavoura, J.P. Silva, CP Violation, Oxford University Press, 1999.
[431] G.C. Branco, M.N. Rebelo, J.I. Silva-Marcos, Phys. Lett. B 614 (2005) 187. hep-ph/0502118.
[432] J.F. Gunion, H.E. Haber, Phys. Rev. D 72 (2005) 095002. hep-ph/0506227.
[433] G.C. Branco, M.N. Rebelo, Phys. Lett. B 160 (1985) 117.
[434] S. Weinberg, Phys. Rev. D 42 (1990) 860.
[435] I.F. Ginzburg, M. Krawczyk, Phys. Rev. D 72 (2005) 115013. hep-ph/0408011.
[436] S. Weinberg, Phys. Rev. Lett. 37 (1976) 657.
[437] G.C. Branco, Phys. Rev. Lett. 44 (1980) 504.
[438] G.C. Branco, Phys. Rev. D 22 (1980) 2901.
[439] G.C. Branco, A.J. Buras, J.-M. Gérard, Nuclear Phys. B 259 (1985) 306.
[440] G. Luders, Kong. Dan. Vid. Sel. Mat. Fys. Med. 28N5 (1954) 1.
[441] W. Pauli, Niels Bohr and the Development of Physics, Pergamon Press, London, 1955.
[442] R. Jost, Helv. Phys. Acta 30 (1957) 409.
[443] R.F. Streater, A.S. Wightman, PCT, Spin and Statistics, and all that, Addison-Wesley Publishing Company, Redwood City, 1989.
[444] G.C. Branco, J.-M. Gérard, W. Grimus, Phys. Lett. B 136 (1984) 383.
[445] G.C. Branco, F. Krüger, J.C. Romão, A.M. Teixeira, J. High Energy Phys. 0107 (2001) 027. hep-ph/0012318.
[446] G.C. Branco, D. Emmanuel-Costa, J.C. Romão, Phys. Lett. B 639 (2006) 661. hep-ph/0604110.
[447] E. Derman, Phys. Rev. D 19 (1979) 317.
[448] F.J. Botella, G.C. Branco, M. Nebot, M.N. Rebelo, Nuclear Phys. B 725 (2005) 155. hep-ph/0502133.
[449] L. Bento, G.C. Branco, P.A. Parada, Phys. Lett. B 267 (1991) 95.
[450] G.'t. Hooft, Phys. Rev. Lett. 37 (1976) 8.
[451] G.'t. Hooft, Phys. Rev. D 14 (1976) 3432.
[452] A.E. Nelson, Phys. Lett. B 136 (1984) 387.
[453] A.E. Nelson, Phys. Lett. B 143 (1984) 165.
[454] S.M. Barr, Phys. Rev. Lett. 53 (1984) 329.
[455] For a review see R.D. Peccei, Lecture Notes in Phys. 741 (2008) 3. hep-ph/0607268.
[456] L. Bento, G.C. Branco, Phys. Lett. B 245 (1990) 599.
[457] F. del Aguila, J. Cortés, Phys. Lett. B 156 (1985) 243.
[458] G.C. Branco, L. Lavoura, Nuclear Phys. B 278 (1986) 738.
[459] F. del Aguila, M.K. Chase, J. Cortés, Nuclear Phys. B 271 (1986) 61.
[460] Y. Nir, D.J. Silverman, Phys. Rev. D 42 (1990) 1477.
[461] D. Silverman, Phys. Rev. D 45 (1992) 1800.
[462] W.S. Choong, D. Silverman, Phys. Rev. D 49 (1994) 2322.
[463] V.D. Barger, M.S. Berger, R.J.N. Phillips, Phys. Rev. D 52 (1995) 1663. hep-ph/9503204.
[464] M. Gronau, D. London, Phys. Rev. D 55 (1997) 2845. hep-ph/9608430.
[465] F. del Aguila, J.A. Aguilar-Saavedra, G.C. Branco, Nuclear Phys. B 510 (1998) 39. hep-ph/9703410.
[466] G.C. Branco, T. Morozumi, P.A. Parada, M.N. Rebelo, Phys. Rev. D 48 (1993) 1167.
[467] G. Barenboim, F.J. Botella, G.C. Branco, O. Vives, Phys. Lett. B 422 (1998) 277. hep-ph/9709369.
[468] G. Barenboim, F.J. Botella, O. Vives, Phys. Rev. D 64 (2001) 015007. hep-ph/0012197.
[469] G. Barenboim, F.J. Botella, O. Vives, Nuclear Phys. B 613 (2001) 285. hep-ph/0105306.
[470] J.A. Aguilar-Saavedra, Phys. Rev. D 67 (2003) 035003. hep-ph/0210112;
      J.A. Aguilar-Saavedra, Phys. Rev. D 69 (2004) 099901 (erratum).
[471] J.A. Aguilar-Saavedra, F.J. Botella, G.C. Branco, M. Nebot, Nuclear Phys. B 706 (2005) 204. hep-ph/0406151.
[472] F.J. Botella, G.C. Branco, M. Nebot, Nuclear Phys. B 768 (2007) 1. hep-ph/0608100.
[473] F.J. Botella, G.C. Branco, M. Nebot, Phys. Rev. D 79 (2009) 096009. arXiv:0805.3995 [hep-ph].
[474] A. Méndez, A. Pomarol, Phys. Lett. B 272 (1991) 313.
[475] J.F. Gunion, Talk given at the CPNSH, CERN, Switzerland, December 2004.
[476] D. Sokolowska, K.A. Kanishev, M. Krawczyk, PoS CHARGED2008 (2008) 016. arXiv:0812.0296 [hep-ph].
[477] S. Davidson, J.R. Ellis, Phys. Lett. B 390 (1997) 210. hep-ph/9609451.
[478] J. Roldán, Ph.D. Thesis, University of València, 1991.
[479] A. Pilaftsis, C.E.M. Wagner, Nuclear Phys. B 553 (1999) 3. hep-ph/9902371.
[480] CERN notes ATLAS-CONF-2011-157 and CMS PAS HIG-11-023, Hadron Collider Physics Symposium, November 2011.
[481] ATLAS Collaboration, ATLAS note ATLAS-CONF-2011-161.

G.C. Branco et al. / Physics Reports 516 (2012) 1–102

[482] CMS Collaboration, talk given by G. Tonelli at CERN on Dec 13 2011.
[483] P.M. Ferreira, R. Santos, M. Sher, J.P. Silva, arxiv:1112.3277 [hep-ph].
[484] G. Burdman, C. Haluch, R. Matheus, arXiv:1112.3961 [hep-ph].
[485] P.M. Ferreira, R. Santos, M. Sher, J.P. Silva, Phys. Rev. D 85 (2012) 035020. arXiv:1201.0019 [hep-ph].
[486] N. Chen, H.-J. He, arXiv:1202.3072 [hep-ph].
[487] K. Blum, R.T. D'Agnolo, arXiv:1202.2364 [hep-ph].
[488] G. Blankenburg, J. Ellis, G. Isidori, arXiv:1202.5704 [hep-ph].
[489] B.W. Lee, C. Quigg, H.B. Thacker, Phys. Rev. Lett. 38 (1977) 883.
[490] B.W. Lee, C. Quigg, H.B. Thacker, Phys. Rev. D 16 (1977) 1519.
[491] R. Casalbuoni, D. Dominici, R. Gatto, C. Giunti, Phys. Lett. B 178 (1986) 235.
[492] R. Casalbuoni, D. Dominici, F. Feruglio, R. Gatto, Phys. Lett. B 200 (1988) 495.
[493] J. Maalampi, J. Sirkka, I. Vilja, Phys. Lett. B 265 (1991) 371.
[494] S. Kanemura, T. Kubota, E. Takasugi, Phys. Lett. B 313 (1993) 155. hep-ph/9303263.
[495] J. Hořejši, M. Kladiva, Eur. Phys. J. C 46 (2006) 81. hep-ph/0510154.
[496] I.F. Ginzburg, I.P. Ivanov, Phys. Rev. D 72 (2005) 115010. hep-ph/0508020.
[497] W. Grimus, L. Lavoura, O.M. Ogreid, P. Osland, J. Phys. G 35 (2008) 075001. arXiv:0711.4022 [hep-ph].
[498] P. Osland, P.N. Pandita, L. Selbuz, Phys. Rev. D 78 (2008) 015003. arXiv:0802.0060 [hep-ph].
[499] I. Maksymyk, C.P. Burgess, D. London, Phys. Rev. D 50 (1994) 529. hep-ph/9306267.
[500] W. Grimus, L. Lavoura, O.M. Ogreid, P. Osland, Nuclear Phys. B 801 (2008) 81. arXiv:0802.4353 [hep-ph].
[501] T.P. Cheng, E. Eichten, L.F. Li, Phys. Rev. D 9 (1974) 2259.
[502] M.E. Machacek, M.T. Vaughn, Phys. Lett. B 103 (1981) 427.
[503] W. Grimus, L. Lavoura, Eur. Phys. J. C 39 (2005) 219. hep-ph/0409231.
[504] P. Sikivie, L. Susskind, M.B. Voloshin, V.I. Zakharov, Nuclear Phys. B 173 (1980) 189.
[505] A. Pomarol, R. Vega, Nuclear Phys. B 413 (1994) 3. hep-ph/9305272.
[506] J.-M. Gérard, M. Herquet, Phys. Rev. Lett. 98 (2007) 251802. hep-ph/0703051.
[507] S. de Visscher, J.-M. Gérard, M. Herquet, V. Lemaitre, F. Maltoni, J. High Energy Phys. 0908 (2009) 042. arXiv:0904.0705 [hep-ph].
[508] B. Grzadkowski, M. Maniatis, J. Wudka, J. High Energy Phys. 1111 (2011) 030. arXiv:1011.5228 [hep-ph].
[509] C.C. Nishi, Phys. Rev. D 83 (2011) 095005. arXiv:1103.0252 [hep-ph].
[510] H.E. Haber, A. Pomarol, Phys. Lett. B 302 (1993) 435. hep-ph/9207267.
[511] K. Olaussen, P. Osland, M.A. Solberg, J. High Energy Phys. 1107 (2011) 020. arXiv:1007.1424 [hep-ph].