Baryogenesis in the two doublet and inert singlet extension of the Standard Model

This content has been downloaded from IOPscience. Please scroll down to see the full text.

JCAP08(2016)057

(http://iopscience.iop.org/1475-7516/2016/08/057)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 130.234.75.50
This content was downloaded on 15/09/2016 at 13:28

Please note that terms and conditions apply.

You may also be interested in:

Scalar representations in the light of electroweak phase transition and cold dark matter phenomenology
Shehu S. AbdusSalam and Talal Ahmed Chowdhury

Detectable gravitational waves from very strong phase transitions in the general NMSSM
Stephan J. Huber, Thomas Konstandin, Germano Nardini et al.

A model for dark matter, naturalness and a complete gauge unification
Kimmo Kainulainen, Kimmo Tuominen and Jussi Virkijärvi

Signatures from scalar dark matter with a vector-like quark mediator
Federica Giacchino, Alejandro Ibarra, Laura Lopez Honorez et al.

Asymmetric dark matter models and the LHC diphoton excess
Mads T. Frandsen and Ian M. Shoemaker

Lower bound on the electroweak wall velocity from hydrodynamic instability
Ariel Mégevand, Federico Agustín Membiela and Alejandro D. Sánchez

Fermi Bubbles under Dark Matter Scrutiny Part II: Particle Physics Analysis
Wei-Chih Huang, Alfredo Urbano and Wei Xue
Baryogenesis in the two doublet and inert singlet extension of the Standard Model

Tommi Alanne, Kimmo Kainulainen, Kimmo Tuominen and Ville Vaskonen

CP³-Origins, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark
Department of Physics, University of Jyväskylä, P.O. Box 35 (YFL), FI-40014 Jyväskylä, Finland
Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland
Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland
E-mail: alanne@cp3.sdu.dk, kimo.kainulainen@jyu.fi, kimo.i.tuominen@helsinki.fi, ville.vaskonen@jyu.fi

Received July 13, 2016
Accepted August 21, 2016
Published August 25, 2016

Abstract. We investigate an extension of the Standard Model containing two Higgs doublets and a singlet scalar field (2HDSM). We show that the model can have a strongly first-order phase transition and give rise to the observed baryon asymmetry of the Universe, consistent with all experimental constraints. In particular, the constraints from the electron and neutron electric dipole moments are less constraining here than in pure two-Higgs-doublet model (2HDM). The two-step, first-order transition in 2HDSM, induced by the singlet field, may lead to strong supercooling and low nucleation temperatures in comparison with the critical temperature, $T_n \ll T_c$, which can significantly alter the usual phase-transition pattern in 2HD models with $T_n \approx T_c$. Furthermore, the singlet field can be the dark matter particle. However, in models with a strong first-order transition its abundance is typically but a thousandth of the observed dark matter abundance.

Keywords: baryon asymmetry, cosmological phase transitions

ArXiv ePrint: 1607.03303

doii:10.1088/1475-7516/2016/08/057
1 Introduction

The matter-antimatter asymmetry in the universe presents one of the major quests for particle cosmology. Due to cosmic inflation, such asymmetry cannot be an initial condition for the thermal history of the universe, but calls for a dynamical explanation. The Standard Model (SM) of elementary particle interactions fails in providing a successful mechanism for baryogenesis, and one must look at different extensions of the SM. In this paper we address these issues in the context of a 2HDSM featuring an extended scalar sector with two gauged Higgs doublets and an extra singlet.

Generation of the matter-antimatter asymmetry in connection with the electroweak phase transition, i.e. electroweak baryogenesis, is a particularly appealing scenario due to the possibility of connecting it with the collider experiments. Generic 2HDMs have been studied earlier in connection with the electroweak baryogenesis problem [1–9]. They provide both a new source of CP violation arising from complex parameters in the 2HDM potential and a strong first-order phase transition arising from the one-loop effective potential. However, observational constraints are placing stringent limits also on 2HDMs [8]. Here we show that these constraints are alleviated when the model is further extended by a real scalar singlet field.

A generic feature of 2HDM, also inherited by the 2HDSM, is the danger of generating large flavour changing neutral currents. To avoid these, one has to constrain the Higgs-fermion couplings in one way or the other. Here we choose to work in the context of universal
Yukawa alignment, which may be argued for by a requirement that the whole Lagrangian is invariant under the group GL\((2,\mathbb{C})\) of linear reparametrization transformations in the doublet space. We also use the reparametrization invariance to develop an elegant way explore the vacuum stability and the phase-transition pattern in the model.

In the 2HDM context large CP violation requires that scalar couplings have large complex phases and strong transition requires that couplings are large in magnitude. When combined, these requirements tend to give too large electron and neutron electric dipole moments (EDMs). We will show that the presence of the additional scalar allows for a strong two-step electroweak phase transition, which does not rely on large radiative corrections to the effective potential. This alleviates the burden on the scalar self-couplings and significantly increases the phase space consistent with EDM constraints in the 2HDSM.

The singlet scalar can also be a dark matter (DM) candidate when a discrete \(Z_2\) symmetry is imposed to stabilize it. However, we will find that a strong first-order phase transition is not consistent with a dominant singlet scalar DM particle. The problem is that a strong two-step transition requires a large coupling between the singlet and doublet sectors and this implies so large annihilation rate for the DM that its relic abundance becomes too small to account for the full observed DM density. This conclusion is generic for all models of this type.

We observe that two-step transitions may also give rise to too strong transitions. It is possible that fields get trapped in the metastable minimum so that electroweak symmetry remains unbroken. Also, the latent heat released in the transition may be so large that the transition walls necessarily become supersonic. However, we find also parameters for which walls may be subsonic, consistent with the electroweak baryogenesis scenario. Overall, we are able to find models that satisfy all observational and experimental constraints and can also give rise to a successful electroweak baryogenesis, accompanied by a subleading DM in the 2HDSM context.

The structure of the paper is as follows: in section 2 we introduce the model and discuss the most general GL\((2,\mathbb{C})\)-reparametrization invariant 2HDSM Lagrangian including Yukawa couplings. Here we also develop methods to study the vacuum stability and the phase-transition patterns in the theory. In section 3 we first go through the experimental constraints on the model and evaluate the DM relic abundance and the DM search limits on model parameters. We then evaluate the strength of the transition and compute the baryon asymmetry created in the electroweak phase transition. The section is concluded by a study of bubble nucleation in the 2HDSM and in the singlet extension of the SM. In section 4 we conclude and outline some directions for future research.

2 The model

We start from the most general two-Higgs-doublet and inert-singlet extension of the SM with the scalar field Lagrangian:

\[
\mathcal{L}_{\text{scalar}} = Z_{ij} (D^\mu H_i)^\dagger D_\mu H_j + \frac{1}{2} (\partial_\mu S)^2 - V(H_1, H_2, S),
\]

(2.1)
where $Z^{ij}$ is an arbitrary Hermitian $2 \times 2$ matrix and the most general potential is given by

$$
V(H_1, H_2, S) = -m_1^2 |H_1|^2 - m_2^2 |H_2|^2 - \left( m_{12} \bar{H}_2 H_1 + \text{h.c.} \right) - \frac{1}{2} m_S^2 S^2
+ \lambda_1 |H_1|^4 + \lambda_2 |H_2|^4 + \lambda_3 |H_1|^2 |H_2|^2 + \lambda_4 (H_1^2 H_2)(H_2^2 H_1)
+ \left( \lambda_5 (H_2^4 H_1) + \lambda_6 |H_1|^2 (H_2^2 H_1) + \lambda_7 |H_2|^2 (H_2^2 H_1) + \text{h.c.} \right)
+ \frac{1}{4} \lambda_S S^4 + \frac{1}{2} \lambda_{S1} S^2 |H_1|^2 + \frac{1}{2} \lambda_{S2} S^2 |H_2|^2 + \left( \frac{1}{2} \lambda_{S12} S^2 H_1^2 H_1 + \text{h.c.} \right).
$$

Both doublets $H_i$ are assumed to be gauged under $SU(2)_L \times U(1)_Y$, while the scalar $S$ is a singlet under all SM gauge interactions. The singlet $S$ is a crucial ingredient in the model because it will disentangle the source of a strongly first-order transition from that of sufficiently strong CP violation.

The Lagrangian (2.1) is invariant under a reparametrization transformation $\Phi \rightarrow \Phi' \equiv P \Phi$ (and a simultaneous rescaling of $S$), where $P$ is an element of the general linear group $GL(2, \mathbb{C})$, and $\Phi$ is the Higgs hyperdoublet:

$$
\Phi \equiv (H_1, H_2)^T.
$$

$GL(2, \mathbb{C})$ is the semidirect product of special linear transformations $SL(2, \mathbb{C})$ and multiplicative group of dilatations $\mathbb{C}^\times$. We can always use the dilatation and a hyperbolic $SL(2, \mathbb{C})$ transformation to bring the kinetic term into the canonical form, $Z^{ij} \rightarrow \text{diag}(1, 1)$, i.e.

$$
Z^{ij}(D^\mu H_i)\bar{D}_\mu H_j \rightarrow |D_\mu H_1|^2 + |D_\mu H_2|^2.
$$

The resulting Lagrangian is still invariant under elliptic $SL(2, \mathbb{C})$ transformations, i.e. the usual $SU(2)$ rotations of the doublets.

A generic 2HDM gives rise to unacceptably large flavour-changing neutral currents (FCNCs) and the presence of a singlet does not change the situation. One way to avoid FCNCs is the Yukawa alignment [10], which assumes that both doublets couple to fermions (FCNCs) and the presence of a singlet does not change the situation. One way to avoid FCNCs is the Yukawa alignment [10], which assumes that both doublets couple to fermions (FCNCs) and the presence of a singlet does not change the situation. One way to avoid FCNCs is the Yukawa alignment [10], which assumes that both doublets couple to fermions (FCNCs) and the presence of a singlet does not change the situation. One way to avoid FCNCs is the Yukawa alignment [10], which assumes that both doublets couple to fermions (FCNCs) and the presence of a singlet does not change the situation. One way to avoid FCNCs is the Yukawa alignment [10], which assumes that both doublets couple to fermions (FCNCs) and the presence of a singlet does not change the situation.

$$
\mathcal{L}_{\text{Yukawa}} = y_u C_{U}^1 \bar{Q}_L \tilde{H}_2 u_R + y_d C_{d}^1 \bar{Q}_L \tilde{H}_2 d_R + y_e C_{e}^1 \bar{L}_L \tilde{H}_2 e_R + \text{h.c.},
$$

where $\tilde{H}_2 \equiv i \sigma_2 H_2^\dagger$. Here $y_a$ are flavour matrices independent of the doublet index, and $C_{q}^i$ are doublet-index dependent complex numbers. In general the alignment may be different in different fermion sectors: $C_{q}^a \neq C_{q}^b$. However, for simplicity, we choose to work in the special case of universal Yukawa alignment, where $C_{q}^a \equiv C_{q}$. In this case we can, without a further loss of generality, choose the basis where only the $H_2$ field couples to fermions. This corresponds to setting $C_{q} = 0$ and $C_{q} = 1$, so that:

$$
\mathcal{L}_{\text{Yukawa}} = y_u \bar{Q}_L \tilde{H}_2 u_R + y_d \bar{Q}_L \tilde{H}_2 d_R + y_e \bar{L}_L \tilde{H}_2 e_R + \text{h.c.}.
$$

\footnote{This actually involves a rotation and a redefining of the scale of $y^a$ matrices.}
The choice of basis leading to (2.5) can be effected by an SU(2) rotation of \( \Phi \), and it exhausts our remaining freedom to perform elliptic SL(2,\( \mathbb{C} \))-reparametrization transformations after diagonalizing the kinetic term.\(^2\)

Let us stress that while the Yukawa sector (2.5) appears to be of type-I 2HDM, we did not impose any discrete symmetry to derive it. This is why we have kept the \( \lambda_6 \) and \( \lambda_7 \) terms in the scalar potential. Note that renormalization does not change the form of the theory; while it both re-introduces a kinetic mixing between doublets and a coupling of \( H_1 \) to fermions, these changes can be countered by another GL(2,\( \mathbb{C} \)) transformation. Also, we point out that the universal Yukawa alignment can be argued for based on reparametrization invariance: only in the context of universal alignment is the complete Lagrangian including the 2HDM and Yukawa sectors invariant under GL(2,\( \mathbb{C} \)) transformations.

Interestingly, the universal alignment structure arises as a low-energy effective theory in models of dynamical electroweak symmetry breaking [11–18]. In the bosonic technicolor [11–13], the ultraviolet theory contains a new gauge theory responsible for dynamically breaking the electroweak symmetry and an elementary scalar doublet \( H \), which communicates the symmetry breaking to the SM fields through its renormalizable Yukawa couplings. At low energies the strong technicolor dynamics is described in terms of an effective Lagrangian for a composite Higgs doublet, which couples with the elementary one through a Lagrangian of the form (2.1), including the non-trivial kinetic mixing. Only the elementary scalar couples to SM fermions, which naturally introduces Yukawa alignment. Moreover, when the kinetic mixing is removed by a non-unitary transformation, the Yukawa Lagrangian becomes naturally of the universally Yukawa-aligned form with \( C_{a} = C_{i} \). After a final SU(2) rotation, the model has a diagonal kinetic term, type-I Yukawa sector (2.5), and the most general potential of eq. (2.2).

2.1 Reparametrization invariance and tree-level vacuum stability

The original Lagrangian with the most general potential, kinetic term, and the universally aligned Yukawa sector has 27 real parameters (not counting the parameters entering the Yukawa-flavour-mixing matrices). We removed the four arbitrary parameters from kinetic terms and three from the complex Yukawa coefficients \( C_{i} \) by the use of the GL(2,\( \mathbb{C} \)) invariance of the theory. This still leaves us with 15 real couplings and five real mass parameters in the model potential \( V(H_1, H_2, S) \). Our next task is to find out which sets of these parameters correspond to physically viable models with a stable potential.

We can use the reparametrization invariance to our advantage in constructing the stable potentials. To this end, it is convenient to rephrase the invariance in terms of Lorentz invariance of the potential, written in terms of bilinears formed from hyperdoublets. Following the analysis of ref. [19–21], we define

\[
\rho^\mu \equiv \Phi^\dagger \sigma^\mu \Phi, \quad \text{where} \quad \sigma^\mu = (1, \sigma_i).
\]

The bilinear four-vector \( \rho^\mu \) is positive definite.\(^3\) That is, \( \rho^\mu \) vectors span the future light cone, \( L C^+ \), of a Minkowski space. Thus, in bilinear representation the elliptic and hyperbolic SL(2,\( \mathbb{C} \)) basis transformations of fields \( \Phi \rightarrow \Phi' \equiv P \Phi \) correspond to proper orthochronous

\(^2\)Note that in the case of general Yukawa alignment, where \( C_{a} \neq C_{i} \), one could still use the SU(2) rotation to set \( C_{i} = 0 \), so that up-type quarks couple only to \( H_2 \). Most of our subsequent analysis would hold also for this scenario, because it is mostly sensitive only to the large top-quark coupling. The only exception is the electron EDM, for which our analysis covers only a part of the full phase space available in the context of general alignment.

\(^3\)Clearly \( r_0 = \sum |H_i|^2 \geq 0 \). Also \( r_\mu r^\mu = 4(|H_1|^2|H_2|^2 - (H_1^\dagger H_2)(H_2^\dagger H_1)) \geq 0 \), by Schwartz inequality.
Lorentz transformations $r^\mu \to r'^\mu = (\Lambda r)^\mu_{\nu} r'^\nu$, where $(\Lambda r)^\mu_{\nu} \in \text{SO}(1,3)^+$. In this notation one can rewrite the Higgs potential (2.2) in a very compact form:

$$V = -\frac{1}{2} m_S^2 S^2 - \frac{1}{2} M^2_{\mu\nu} r^\mu r^\nu + \frac{1}{4} \lambda S^2 r^\mu S^\nu + \frac{1}{4} \lambda_S S^4,$$

(2.7)

where we defined mass and coupling four-vectors

$$M^2_{\mu\nu} \equiv (m_1^2 + m_2^2, 2m_{12}^2 \gamma, -2m_{12}^2 \gamma, m_1^2 - m_2^2),$$

$$\lambda_{S\mu} \equiv (\lambda_{S1} + \lambda_{S2}, 2\lambda_{S12}, -2\lambda_{S12}, \lambda_{S1} - \lambda_{S2})$$

(2.8)

and a symmetric coupling tensor

$$\lambda_{\mu\nu} \equiv \begin{pmatrix} \lambda_1 + \lambda_2 + \lambda_3 & \lambda_6 R + \lambda_7 I & -\lambda_6 I + \lambda_7 I & \lambda_1 - \lambda_2 \\ \lambda_6 R + \lambda_7 I & \lambda_4 + 2\lambda_5 R & -2\lambda_5 I & \lambda_6 R - \lambda_7 R \\ -\lambda_6 I + \lambda_7 I & -2\lambda_5 I & \lambda_4 - 2\lambda_5 R & -\lambda_6 I - \lambda_7 I \\ \lambda_1 - \lambda_2 & \lambda_6 R - \lambda_7 R & -\lambda_6 I - \lambda_7 I & \lambda_1 + \lambda_2 - \lambda_3 \end{pmatrix},$$

(2.9)

where the subscripts $R$ and $I$ refer to the real and imaginary parts of the couplings, respectively. The reparametrization invariance is manifest in eq. (2.7), because $V$ depends only on Lorentz-invariant products of vectors and tensors. This form is particularly suitable for a study of the vacuum stability and phase-transition patterns of the model.

First consider the direction $S = 0$ in the potential (2.7). Here the term $r^\mu \lambda_{\mu\nu} r^\nu$ must be bounded from below. In [19] it was shown that this is the case precisely when $\lambda_{\mu\nu}$ is positive definite in the future light cone. That is, all stable potentials can be written as $\lambda_{\mu\nu} \equiv \Lambda_\mu^\alpha \Lambda_\nu^\beta \lambda^{\alpha\beta}$, where $\Lambda_\mu^\alpha$ is an SO$(1,3)^+$ transformation and

$$\lambda^{\alpha\beta} = \text{diag}(\lambda^{00}_0, -\lambda^{01}_0, -\lambda^{02}_0, -\lambda^{03}_0), \quad \text{with} \quad \lambda^{00}_0 > 0 \quad \text{and} \quad \lambda^{00}_0 > \lambda^{ij}_0.$$ 

(2.10)

Note that the four parameters in $\lambda^{\alpha\beta}$ together with the six parameters in $\Lambda^{\beta\nu}$ add up to the ten real degrees of freedom in the most general 2HDM potential. Second, if we set $r^\mu = 0$ ($H_1 = H_2 = 0$), we see that we must have

$$\lambda_S > 0.$$ 

(2.11)

Finally, we have to consider the directions where both $S$ and $H_i$ are nonzero. First, if the vector $\lambda_{S\mu}$ of couplings which mix $S$ and $H_i$ lies in the future light cone, $\lambda_{S\mu} \in \text{LC}^+$, i.e.

$$\lambda_{S\mu} \lambda_{S\nu}^* = 4 (\lambda_{S1} \lambda_{S2} - |\lambda_{S12}|^2) > 0 \quad \text{and} \quad \lambda^S_S = \lambda_{S1} + \lambda_{S2} > 0,$$

(2.12)

then the mixing term $\frac{1}{2} \lambda_{S\mu} r^\mu S^2$ in the potential, eq. (2.7), is always positive, and no new conditions arise. However, if $\lambda_{S\mu} \notin \text{LC}^+$, there are always directions $r^\mu \in \text{LC}^+$ along which the product $\lambda_{S\mu} r^\mu$ is negative. If we in such cases rewrite the quartic part of the potential as

$$V_{\text{quartic}} = \frac{1}{4} r^\mu \left( \lambda_{\mu\nu} - \frac{1}{4\lambda_S} \lambda_{S\mu} \lambda_{S\nu} \right) r^\nu + \frac{1}{4} \lambda_S \left( S^2 + \frac{\lambda_{S\mu} r^\mu}{2\lambda_S} \right)^2,$$

(2.13)

we see that the potential is most negative as a function of $S$ along direction $2\lambda_S S^2 = -\lambda_{S\mu} r^\mu$. In this subspace, the potential reduces to the form $4V_{\text{quartic}} = r^\mu \lambda_{\mu\nu} r^\nu$, with a new coupling matrix

$$\lambda^\mu_{\mu\nu} \equiv \lambda_{\mu\nu} - \frac{1}{4\lambda_S} \lambda_{S\mu} \lambda_{S\nu}.$$

(2.14)
The matrix $\lambda^S_{\mu\nu}$ is not in general diagonalizable by an SO(1, 3)$^+$ rotation, and it may have also complex eigenvalues. Unfortunately we cannot restrict its properties like we did for $\lambda_{\mu\nu}$, because $\lambda^S_{\mu\nu}$ does not need not be positive definite in the entire future light cone, but only in the subset of $\text{LC}^+$ where $\lambda_{\mu\nu} r^\mu$ is negative. Instead of covering the full range of possibilities, we will require a sufficient (but not necessary) condition that $\lambda^S_{\mu\nu}$ is positive definite in the future light cone whenever $\lambda_{\mu\nu} r^\mu < 0$. Having found an acceptable set, we generate a random Lorentz transformation and define

$$\lambda_{\mu\nu} \equiv \Lambda^\alpha_{\mu} \lambda_D^{\alpha\beta} \Lambda^\beta_{\nu} \quad \text{and} \quad \lambda^S_\mu \equiv \Lambda^\nu_\mu (\lambda^S_\nu)^\nu,$$

where $\Lambda^\mu_\nu \in \text{SO}(1, 3)^+$. 

Let us comment on the role of the kinetic and Yukawa terms in the above construction of the potential. We implicitly assumed that kinetic term becomes diagonal, and the Yukawa term becomes of type-I form in the final frame, after the Lorentz transformation. Thus, they necessarily must be nontrivial in the original, diagonal frame. Indeed, the six degrees of freedom “missing” in the diagonal potential are in this frame evenly divided between the $C_i$ coefficients in the Yukawa Lagrangian and the mixing parameters in the kinetic term, which in the bilinear notation can be written as

$$K_{\mu}(D^\alpha \Phi)^\dagger \sigma^\mu(D_\alpha \Phi). \quad (2.16)$$

Here $K_{\mu}$ is some positive definite, but otherwise arbitrary four-vector of unit length (here $\alpha$ refers to the usual space-time indices and $\mu$ to potential indices).\footnote{In the most general case $K_{\mu}$ has four free parameters, but one parameter is here assumed to be removed by a dilatation, used to bring the length of $K_{\mu}$ to unity.} It is amusing to see that exactly a Lorentz boost (a hyperbolic SL(2, $\mathbb{C}$) transformation on fields) is needed to bring an arbitrary $K_{\mu}$ into the canonical form: $K_{\mu} \rightarrow (1; 0, 0, 0)$, after which the kinetic term is manifestly invariant under Lorentz rotations (elliptic SL(2, $\mathbb{C}$) transformations on fields). These boosts and rotations into the canonical frame activate the whole SO(1, 3)$^+$ group discussed above and thus create all physically viable Lagrangians with bounded potentials.

### 2.2 Spontaneous symmetry breaking

Since we are interested in the cases where the singlet scalar $S$ is a DM candidate, we restrict our considerations to the cases where the potential is unbroken in $S$ direction and hence $\mathbb{Z}_2$ symmetric at low temperatures. However, to enhance the strength of the latter phase transition, we also need the $S$ symmetry to be broken at high temperatures before the symmetry breaks in the doublet direction. To this end, we must have negative quadratic term in the $S$ direction in the potential ($m_S^2 > 0$). This implies that there may be other minima away from the $\langle S \rangle = 0$ vacuum, and we must check that none of these minima is the global one at zero temperature. The extremization conditions are:

$$\frac{\partial V}{\partial S} = S \zeta_S = 0, \quad \text{where} \quad \zeta_S \equiv -m_S^2 + \lambda_S S^2 + \frac{1}{2} \lambda_{\mu\nu} r^\mu,$$  

(2.17)
and \[
\frac{\partial V}{\partial H_i} = \sigma_{ij}^\mu H_j \zeta_\mu = 0, \quad \text{where} \quad \zeta_\mu = -\frac{1}{2} \left( M_\mu^2 - \frac{1}{2} \lambda S_\mu S^2 \right) + \frac{1}{2} \lambda_{\mu\nu} r_\nu.
\] (2.18)

Eqs. (2.18) are complex, so we have five equations relating vacuum fields to the parameters of the theory. Our goal is to have an unbroken singlet and a broken neutral doublet vacuum at zero temperature. In [19] it was shown that the neutral and charged vacua cannot coexist in the pure 2HDM case. When \( S = 0 \), our potential reduces to the pure 2HDM case, and so the above statement applies here as well. Therefore, if we find a neutral vacuum, we know it is the global one in the doublet space and the charged extremum may at best be a saddle point. Moreover, even with \( S \neq 0 \), the doublet-symmetry-breaking pattern is formally similar to the pure 2HDM case, only with an effective mass parameter

\[- m_\mu^2 \equiv -M_\mu^2 + \frac{1}{2} \lambda S_\mu S^2. \quad (2.19)\]

This implies that also any neutral minimum with \( S \) breaking is the lowest one in the doublet space.

The most general neutral vacuum in 2HDM field space is given by

\[\langle H_1 \rangle = \frac{1}{\sqrt{2}} \left( \begin{array}{c} v_1 \\ 0 \end{array} \right) e^{i\theta_1}, \quad \text{and} \quad \langle H_2 \rangle = \frac{1}{\sqrt{2}} \left( \begin{array}{c} 0 \\ v_2 e^{i\theta_2} \end{array} \right). \] (2.20)

Of course the local SU(2)-gauge invariance guarantees that only the relative phase \( \theta \equiv \theta_2 - \theta_1 \) is physical, and one could rotate for example \( \theta_1 \to 0 \). Explicitly we find that this corresponds to

\[r_0^\mu = (v_1^2 + v_2^2, 2v_1v_2 \cos \theta, 2v_1v_2 \sin \theta, v_1^2 - v_2^2); \] (2.21)

so that \( r_0^2 = 0 \) as it should for a neutral vacuum [19]. This construction actually only ensures that the special point (2.20) is an extremum of the potential. To check that this extremum is a also a minimum, one needs to compute second derivative matrices of the potential corresponding to scalar field masses and require that there are no negative eigenvalues. For the mass of the physical excitation in the singlet direction, this corresponds to requiring that

\[\left. \frac{d^2V}{dS^2} \right|_{\text{vev}} = -m_S^2 + \frac{1}{2} \lambda S_\mu r_0^\mu \equiv M_S^2 > 0, \quad (2.22)\]

where the left hand side is evaluated at the extremum \( \langle S \rangle = 0 \) and \( \langle H_i \rangle \neq 0 \) as defined in eq. (2.20). These requirements on the spectrum allow one to set the Lagrangian mass parameters in terms of physical masses and vacuum expectation values (vevs) of the fields. However, it still remains to check that the vacuum with \( S = 0 \) is the global minimum for any given set of parameters. It is straightforward to show that the value of the potential at the desired \( \langle S \rangle = 0, \langle H_i \rangle \neq 0 \) vacuum is

\[V(\langle S \rangle = 0, \langle H_i \rangle \neq 0) = -\frac{1}{4} r^\mu_0 \lambda_{\mu\nu} r^\nu. \] (2.23)

In the direction \( H_i = 0 \), the potential has at least a directional minimum at \( \langle S \rangle^2 = m_S^2/2\lambda_S \), and the value of the potential in this directional minimum is

\[V(\langle S \rangle \neq 0, \langle H_i \rangle = 0) = -\frac{1}{4\lambda_S} \left( M_S^2 - \frac{1}{2} \lambda S_\mu r_0^\mu \right)^2. \] (2.24)
We impose the condition that the minimum in eq. (2.23) is below that in eq. (2.24). Finally, there is an extremum where both $\langle S \rangle \neq 0$ and $\langle H_i \rangle \neq 0$, but this is a local maximum.

In addition to the massive scalar $S$, the physical spectrum contains the usual states arising from the two Higgs doublets: the three neutral scalars $h_0$, $H_0$, and $A_0$, and two charged scalars $H^\pm$. The diagonalization of the mass matrices is presented in detail in appendix A. The lightest neutral scalar state $h_0$ is identified with the 125 GeV Higgs particle observed at LHC, while the masses of the heavier neutral and charged scalar states are constrained to lie above the current limits.

### 2.3 Finite-temperature potential

The final ingredient we need for our analysis is the effective potential at finite temperature. Here we only consider the leading corrections to the potential, which bring about the symmetry restoration at high temperatures. In the high-temperature limit, these corrections are accounted for by the thermal masses:

$$m_a^2(T) = -m_a^2 + c_a T^2_{12},$$  \hspace{1cm} (2.25)

where $(a = 1, 2, 12, S)$:

$$c_1 = c_{SM} + 6\lambda_1 + 2\lambda_3 + \lambda_4 + \frac{1}{2}\lambda_{S1},$$

$$c_2 = c_{SM} + 6\lambda_2 + 2\lambda_3 + \lambda_4 + \frac{1}{2}\lambda_{S2},$$

$$c_{12} = c_{SM} + 3\lambda_{6R} + \lambda_{7R} + \frac{1}{2}\lambda_{S12R} - i \left( 3\lambda_{6I} - \lambda_{7I} + \frac{1}{2}\lambda_{S12I} \right),$$

$$c_S = 3\lambda_S + 2(\lambda_{S1} + \lambda_{S2}),$$ \hspace{1cm} (2.26)

and

$$c_{SM} = \frac{9}{4}g_L^2 + \frac{3}{4}g_Y^2 + 3g_t^2$$ \hspace{1cm} (2.27)

is the common SM contribution from the SU(2)$_L$ and U(1)$_Y$ gauge fields and the top quark.

### 3 Results

We now test our model against experimental constraints and for the consistency of its cosmological predictions. We scan the parameter space by solving $m_1^2$, $m_2^2$ and $m_{12}^2$ from the vacuum conditions and create Monte Carlo chains in the coupling constant space by the algorithm described in the previous section. We first subject the models to various theoretical and experimental constraints. For models that pass these constraints, we compute the DM abundance and the strength of the phase transition. We find that the model can provide either the observed DM abundance or a strong transition, but not both simultaneously. For the points providing a strong transition, we compute the prediction for the baryon-to-entropy ratio $\eta_B$. We show that the parameter space where $\eta_B$ is sufficiently large is strongly constrained but not excluded by the current experimental limit on the electric dipole moment of the electron. Finally, we compute the nucleation temperatures $T_n$ and show that typically, and in particular for strong transitions, $T_n$ is much smaller than the critical temperature $T_c$. 

– 8 –
3.1 Theoretical and experimental constraints

The couplings between scalar fields tend to run strongly, which typically leads to a relatively low cut-off scale for the validity of the effective theory. To be specific, we demand that our model is consistent up to $\Lambda = 1.5 \text{ TeV}$, i.e. we check that the vacuum is stable and that all couplings remain perturbative up to $\Lambda$. The 1-loop renormalization-group equations needed for this calculation are summarized in appendix B. Moreover, at zero temperature we impose

$$|\langle H_1 \rangle|^2 + |\langle H_2 \rangle|^2 = v_0^2, \quad \langle S \rangle = 0 \quad (3.1)$$

and check that this choice gives the global minimum of the potential as detailed in previous section.

Second, we address the current experimental constraints. We only accept points that give a light scalar of mass $m_{h_0} = 125 \text{ GeV}$, and for which the heavy scalars satisfy the mass limits from direct searches at LEP on charged particles, $m_{H^\pm} > 500 \text{ GeV}$ [22], and at LHC on heavy neutral scalars, $m_{H^0}, m_{A^0} > 600 \text{ GeV}$ [23]. For consistency, we also constrain the heavy scalar masses to be below the cut-off $\Lambda = 1.5 \text{ TeV}$. We also take into account the electroweak precision data using the $S$ and $T$ parameters [24]. The necessary formulae for computing $S$ and $T$ can be extracted from [25], where the oblique parameters have been calculated for models with extra doublets and extra singlets. The current values of $S$ and $T$ parameters are $S = 0.00 \pm 0.08, T = 0.05 \pm 0.07$ with a correlation factor $\rho = 0.90$ [26]. We accept points that lie within the $2\sigma$ region in the $(S,T)$ plane.

Moreover we take into account constraints on the Higgs boson couplings to SM particles using the signal-strength data from the LHC and Tevatron experiments [27–29]. First, we impose the constraint from the invisible width of the Higgs boson from LHC at $2\sigma$ level:

$$R_I = \frac{\Gamma_I}{\Gamma_I + \Gamma_{SM,\text{tot}}} < 0.020, \quad (3.2)$$

which here implies $\Gamma_{h_0 \to SS} \lesssim 1.0 \text{ MeV}$. Second, we allow for modifications to the Higgs couplings and parametrize the deviations from the SM with parameters $a_f$ and $a_V$,

$$\mathcal{L}_{\text{eff}} = a_V \left( \frac{2M_W^2}{v_0} h_0 W^+ W^- + \frac{M_Z^2}{v_0} h_0 ZZ \right) - a_f \sum \frac{m_\psi}{v_0} h_0 \bar{\psi} \psi. \quad (3.3)$$

We accept points which are within $2\sigma$ from the best fit values

$$a_V = 0.993, \quad a_f = 0.968, \quad (3.4)$$

which we obtain by performing a $\chi^2$ fit to the Higgs boson signal strength data. We have neglected the imaginary part of $a_f$ and the $h_0 H^+ H^-$ coupling in the fit, since we have checked that these couplings are very small.\footnote{As a comparison, we have performed the $\chi^2$ fit of $a_f$ and $a_V$ with maximum values of the imaginary part of $a_f$ and $h_0 H^+ H^-$ coupling for out data points, and we find that the best-fit value does not change significantly.}

3.2 Dark matter abundance and direct detection limits

We compute the relic abundance of the $S$ bosons for all models passing the experimental constraints described above. We assume that $S$ is a thermal relic, i.e. we assume that at
least some the portal couplings $\lambda_{S1}$, $\lambda_{S2}$ and $\lambda_{S12}$ are sufficiently large (larger than about $10^{-7}$ [30]). We then apply the standard freeze-out formalism [31], employing the accurate approximation scheme introduced in [32]. The relevant annihilation channels of our WIMP candidate are

$$SS \to h_0h_0, H_0H_0, A_0A_0, h_0H_0, h_0A_0, H^+H^-, W^+W^-, ZZ, \bar{f}f.$$  (3.5)

Cross-sections for all these processes are given in a very compact form in appendix C. To treat cases where $S$ is a subdominant DM candidate, we define the ratio [33]

$$f_{\text{rel}} = \frac{\Omega_S h^2}{0.12},$$  (3.6)

which expresses how large fraction of the observed DM abundance is in form of $S$ bosons. All annihilation channels are directly proportional to the couplings between the $S$ boson and the Higgs fields, although the precise dependence is rather complicated (see appendix C). Since the relic abundance is roughly proportional to the inverse of the annihilation cross section, large (small) couplings corresponds to a small (large) relic abundance.

Direct search limits for $S$ follow from the bound on the spin-independent cross section for $S$ scattering off nucleons. It is given by

$$\sigma_{\text{SI}} = \frac{\lambda_{\text{eff}}^2 f_N^2 \mu m_N^2 v_0^2}{4\pi m_h^4 M_S^2},$$  (3.7)

where $m_N = 0.939$ GeV is the nucleon mass, $\mu = m_N M_S/(m_N + M_S)$ is the reduced mass of the nucleon-scalar system, and $f_N \approx 0.30$ [32] gives the strength of the Higgs-nucleon coupling: $g_{h_0 N\bar{N}} = f_N m_N/v_0$. Finally, the effective $SSh_0$ coupling is given by

$$\lambda_{\text{eff}} \equiv \frac{1}{2} \left[ (R_{N_{44}} \cos \beta \sin \theta - R_{N_{42}} \cos \beta \cos \theta + R_{N_{41}} \sin \beta) \lambda_{S12L} \\ + (R_{N_{42}} \cos \beta \sin \theta - R_{N_{44}} \cos \beta \cos \theta - R_{N_{43}} \sin \beta) \lambda_{S12R} \\ + \cos \beta (R_{N_{43}} \cos \theta + R_{N_{41}} \sin \theta) \lambda_{S1} - \sin \beta R_{N_{44}} \lambda_{S2} \right],$$  (3.8)

where $\beta$ is the vacuum mixing angle and $\theta$ the phase between the doublet vevs, and $R_{N_{ij}}$ are the components of the $4 \times 4$-mixing matrix between the neutral scalar fields given in eq. (A.8).

Currently the most stringent limit for $\sigma_{\text{SI}}$ come from the LUX experiment [34]. However, in the case of a subdominant DM, the LUX bound is not directly related to $\sigma_{\text{SI}}$ shown in eq. (3.7). Instead, assuming that all DM components cluster similarly, the actual signal strength from $S$ bosons is suppressed by the fraction $f_{\text{rel}}$. The relevant quantity to compare with the LUX limit then is

$$\sigma_{\text{eff}} = f_{\text{rel}} \sigma_{\text{SI}}.$$  (3.9)

We show the effect of this limit on the models passing all experimental bounds on figure 1. Note that the scatter in $\sigma_{\text{eff}}$ as a function of $m_S$ is much larger here than in the singlet extension of the SM [33]. Note how also the models with relatively low abundance are constrained by the LUX data.
Figure 1. Shown is the LUX bound on a set of models passing all experimental constraints described in section 3.1 and for which \( f_{\text{rel}} \geq 0.01 \).

### 3.3 Electroweak phase transition

One of the necessary conditions for a successful baryogenesis scenario is the departure from thermal equilibrium. In EWBG this is provided by a strong first-order phase transition. In pure 2HDMs, with only leading thermal corrections (i.e. with thermal masses), the phase transition is of second order. While full one-loop effects may induce a first-order transition, it still tends to be rather weak [8, 9]. In a model with a singlet scalar, a strong transition may take place with just the leading corrections (2.25), given a specific two-step symmetry-breaking pattern [35]. This means that, when passing from high towards low temperatures, a minimum of the potential is generated first along the singlet direction. At lower temperature, then, the potential develops a minimum where the doublets have non-zero vevs while the singlet symmetry is restored, \( (\langle h_1 \rangle, \langle h_2 \rangle; \langle s \rangle) = (0, 0; 0) \rightarrow (0, 0; w(T)) \rightarrow (v_1(T), v_2(T); 0) \), ensuring the actual electroweak phase transition. This pattern of minima should of course develop in such a way that the true ground state at \( T = 0 \) is given by eq. (3.1).

Including the thermal mass corrections of eq. (2.25) to the potential, we find the transition temperatures. In particular the electroweak transition temperature, \( T_c \), where the two minima are degenerate is determined by

\[
V(0, 0, w_c, T_c) = V(v_{1c}, v_{2c}, 0, T_c),
\]

where \( w_c = w(T_c) \) and \( v_{ic} = v_i(T_c) \). To determine when the transition is strong enough, we require, as usual, that

\[
v_c/T_c \geq 1,
\]

where \( v_c = \sqrt{v_{1c}^2 + v_{2c}^2} \). It turns out not to be possible to have simultaneously strong transition and dominant DM candidate in the model with \( f_{\text{rel}} \approx 1 \). This is because having a strong transition requires that at least some of the mixing couplings \( \lambda_S \) are large. However, as we noted above, large mixing introduces large annihilation cross sections for \( S \) bosons and hence small relic abundances. Qualitatively the behaviour is the same as in the pure singlet model [32, 33]: a strong first-order EWPT and large \( f_{\text{rel}} \approx 1 \) are realized in different portions of the parameter space.
Figure 2. Scanned data points which give a strong first-order EWPT. Yellow points are excluded by direct DM searches.

Left panel of figure 2 shows the correlation between the DM particle mass $M_S$ and the effective coupling $\lambda_{\text{eff}}$ for models with strong EWPT. Accepted values of $\lambda_{\text{eff}}$ increase as a function of $M_S$ because both quantities are linearly proportional to the mixing couplings $\lambda_{SI}$. Right panel of figure 2 shows the correlation between $M_S$ and the relative relic abundance $f_{\text{rel}}$ for the same set. Note that for $M_S$ between the Higgs resonance and $M_W$, both a strong transition and a relatively large DM abundance, $f_{\text{rel}} \approx 0.15$ could be obtained, but this region is now excluded by LUX.

The fact that LUX most strongly constrains models with small $\lambda_{\text{eff}}$ is because $f_{\text{rel}}$ is roughly inversely proportional to the square of $\lambda_{\text{eff}}$, and thus to large extent cancels the direct $\lambda_{\text{eff}}$ dependence in $\sigma_{SI}$. Larger mass region beyond the Higgs resonance is less constrained because the DM-number density falls with increasing mass.

3.4 Electron EDM constraint

The non-observation of electric dipole moments (EDMs) of electrons, neutrons and atoms gives stringent bounds on CP-violating interactions in multi-Higgs models. As shown in [36], currently the most stringent bound for 2HDMs arises from the electron EDM, for which the ACME experiment gives an upper limit

$$|d_e| < 8.7 \times 10^{-29} \text{e cm},$$

with 90% confidence level [37]. We calculate $d_e$ for the points which give a strong first-order EWPT using the results from ref. [38], where Barr-Zee type contributions to fermionic EDM were calculated in 2HDM. These results are directly applicable here as well, because the singlet scalar $S$ does not directly couple to gauge fields. In figure 3, we show the distribution of models passing all previous cuts as a function of $d_e$ and the neutral scalar mixing matrix element $R_{N42}$, which expresses the projection of $h_0$ to complex part of the second doublet:

$$R_{N42} \equiv \langle H^0_{21} | h_0 \rangle \equiv \sin \Delta_{\text{CP}}.$$
Figure 3. Scatter plot of all models with strong enough EWPT as a function of the mixing parameter $\sin \Delta_{CP}$ and the electron EDM $d_e$. The red region is excluded by the eEDM limit.

3.5 Baryogenesis

The actual baryogenesis mechanism in our model relies on CP-violating interactions of the top quark with the expanding phase-transition walls. The CP violation comes directly from the spatial evolution of the complex phases of the Higgs field $H_2$, which renders the top mass a complex-valued function of the spatial coordinate across the wall. The first step for us is then to work out the evolution of the scalar fields over the bubble wall.

We shall approximate the true phase-transition-wall profile in the usual way, by the stationary path that extremizes the Euclidean one-dimensional action

$$ S_1 = \int dz \left( \sum_i \frac{1}{2} (\partial_z h_i)^2 + \frac{1}{2} (\partial_z S)^2 + \frac{1}{2} \frac{h_1^2 h_2^2}{h_1^2 + h_2^2} (\partial_z \varphi)^2 + V(h_1, h_2, S, \varphi, T_c) \right) . $$

The invariance of the potential under the change of the total phase $\varphi_1 + \varphi_2$ implies a conservation law, which in the $Z_\mu = 0$ gauge allows us to work out the phase $\varphi_2$ in terms of the relative phase $\varphi$ [8]:

$$ \partial_z \varphi_2 = -\frac{h_1^2}{h_1^2 + h_2^2} \partial_z \varphi . $$

The complex, spatially-varying top mass can now be constructed from the phase $\varphi_2(x)$ and the modulus $h_2(z)$:

$$ m_t(z) = \frac{y_t}{\sqrt{2}} h_2(z) e^{i \varphi_2(z)} . $$

In fact, one does not need to solve for the top phase, since only its derivative, given by eq. (3.16), appears in the source term for the diffusion equations for chemical potentials:

$$ S_t = \xi_w \left( K_{8,t} \left( x_1^2 \varphi_2' \right)' - K_{9,t} x_1^2 x_1^2 \varphi_2' \right) . $$

---

[Note: The text continues with further details and equations, but is not fully transcribed due to the nature of the task.]
Here $\xi_w$ is the wall velocity, primes denote $\partial_z T$ and $K_{n,t}$ are dimensionless functions of $x_t \equiv |m_t|/T$ arising from phase-space averaging of certain kinematic variables defined in [39].

Given the source, one can calculate chemical potentials $\mu_j(z)$ for top, bottom, anti-top and Higgs by solving a set of transport equations defined in [7]. Finally the baryon-to-entropy ratio $\eta_B \equiv n_B/s$ is given by

$$\eta_B = \frac{405}{4\pi^2\xi_w g_* T_c} \int_0^\infty dz \Gamma_{\text{sph}}(z) \mu_{B_L}(z) e^{-45 \Gamma_{\text{sph}}(z)/\xi_w}.$$  

(3.19)

We take $\xi_w = 0.1$ for the wall velocity and $g_* = 106.75$ for the number of degrees of freedom in the plasma. The left-chiral baryon chemical potential is

$$\mu_{B_L} = \frac{1}{2} (1 + 4K_{1,t}) \mu_t + \frac{1}{2} (1 + 4K_{1,b}) \mu_b - 2K_{1,t} \mu_c.$$  

(3.20)

For the sphaleron rate we use a formula interpolating between the symmetric and the broken phase [8],

$$\Gamma_{\text{sph}}(z) = \min(10^{-6} T_e, 2.4 T_e e^{-40 v(z)/T_e}),$$  

(3.21)

where $v(z)^2 = h_1(z)^2 + h_2(z)^2$.

In the left panel of figure 4 we show how the baryon-to-entropy ratio relative to the observed value $\eta_B^{\text{obs}} = 8.7 \times 10^{-11}$ [40] correlates with the CP-violation-sensitive parameter $\sin \Delta CP$ defined in (3.13). Shown are only the points which survive the eEDM bound. As expected, the size of $\sin \Delta CP$ correlates with the size of $\eta_B$. This trend is similar to the correlation between $\sin \Delta CP$ and $d_e$ shown in figure 3. However, a large $\eta_B$ does not always imply a large $d_e$, as is clear from the right panel of figure 4, where we show the correlation between $d_e$ and $\eta_B$, again for points that pass the EDM bound. Apparently, while both quantities are sensitive to the CP-violating parameters in the model, they can be sensitive to different linear combinations of them, so that large $\eta_B$ may be obtained simultaneously with a small enough $d_e$.

Figure 5 shows the distributions of various physical parameters in our parametric scan. Orange colour refers to models that pass all experimental cuts described in sections 3.1 and 3.2, and give a strong EWPT, blue to models that in addition satisfy EDM constraint and green to models which also give large baryon-to-entropy ratio $\eta_B/\eta_B^{\text{obs}} \in [0.5, 2]$. These plots
must be interpreted with care, since our scans were partly tuned by hand. Nevertheless, we see that none of the vevs can be very large and in particular both $v_1$ and $v_2$ need to be nonzero. Also the critical temperature is bounded from above: $T_c \lesssim 100$ GeV. Finally for the models with large $\eta_B$, the new scalar masses are in general bound from above: $m_{H^\pm}, m_A, m_S \lesssim 1.4$ TeV and $m_S \lesssim 400$ GeV, which is encouraging from the point of view of experimental verifiability of the model.

3.6 Bubble nucleation

So far we have implicitly assumed that the bubble nucleation takes place at a temperature not too different from the critical temperature. This is typically the case in models where the first-order phase transition is effected by cubic corrections to potential from infrared modes, which leads to rather mild supercooling and small latent heat release. Here the situation is different, because the barrier between the degenerate minima is essentially due to a tree-level term. Thus a stronger supercooling and more latent heat release may be expected, or even a possibility of a formation of a metastable vacuum where the electroweak breaking never takes place.

We study the nucleation problem in the thin-wall limit [41]. The bubble nucleation rate is given by

$$\Gamma \sim T^4 \left( \frac{S_3(T)}{2\pi T} \right)^{3/2} \exp \left( -\frac{S_3(T)}{T} \right), \quad (3.22)$$

where $S_3(T)$ is the three-dimensional action for an O(3)-symmetric bubble. In the thin-wall limit, it is given by

$$S_3(T) = \frac{16\pi}{3} \frac{\sigma^3}{\Delta V(T)^2}, \quad (3.23)$$

where $\Delta V(T)$ is the potential energy difference between the electroweak-symmetric and electroweak-broken minima and $\sigma$ is the surface tension,

$$\sigma = \int d\phi \sqrt{2V}, \quad (3.24)$$

integrated along the path from the symmetric to the broken minimum at temperature $T = T_c$. The bubble nucleation temperature $T_n$ is defined as the temperature at which creating at
least one bubble per horizon volume is of order one. This condition can be written as

\[ \frac{S_3(T_n)}{T_n} = -\log \left( \frac{3}{4\pi} \left( \frac{H(T_n)}{T_n} \right)^4 \left( \frac{2\pi T_n}{S_3(T_n)} \right)^{3/2} \right). \] (3.25)

We show the results of the nucleation-temperature calculation for our data set on the left panel of figure 6. Obviously, a large number of points displayed in figure 4 are missing in figure 6. The reason is that for these models, indicated by red dots in figure 4, no solution to eq. (3.25) was found. In these cases, in the thin-wall approximation, the fields were trapped in the false vacuum. Moreover, of the surviving models only four give large \( \eta_B \).

The situation is actually more dire than this: another implicit assumption in our baryon asymmetry calculation is that the transition walls are subsonic deflagrations, which is required for efficient diffusion of particle asymmetries across the bubble wall. However, with a large latent heat release the walls tend to be supersonic detonations instead. A full microscopic analysis of wall dynamics including a computation of the wall friction is beyond the scope of this paper. However, a deflagration wall must necessarily satisfy a condition [42]

\[ \alpha \equiv \frac{\Delta V(T_n)}{\rho(T_n)} < \frac{1}{3} (1 - \xi_w)^{-13/10} \equiv \alpha_{\text{max}}, \] (3.26)

where \( \rho(T_n) \) is the radiation energy density in the symmetric phase. We show this condition in the right panel of figure 6 (blue dots) for the set of models shown in the left panel. As expected, for the models with the lowest nucleation temperatures, deflagrations are not possible. This applies in particular to all four surviving models with a large baryon-to-entropy ratio.

There are two issues that ameliorate the situation. First, the validity of the thin-wall limit actually requires a small latent heat and/or a large surface tension, which is often not the case here. When not applicable, thin-wall limit tends to overestimate the action \( S_3(T) \), and hence underestimate the nucleation rate and eventually \( T_n \). Accurate calculation of the nucleation rate is quite complicated in the full model, however, and we do not pursue it here. Instead, we compare the nucleation temperatures found in the thin-wall limit and in the full calculation in the simpler, singlet extension of the SM, studied for example in [32, 33]. In
practice we minimize the action
\[ S_3(T) = 4\pi \int r^2 dr \left( \frac{1}{2} \left( \frac{dh}{dr} \right)^2 + \frac{1}{2} \left( \frac{dS}{dr} \right)^2 + V_{\text{SSM}}(h, S, T) \right), \] (3.27)
where \( h \) is the SM-Higgs field, and \( V_{\text{SSM}}(h, S, T) \) is the singlet-model potential.

The results of this analysis are shown in figure 7. The blue dots show the nucleation temperature in the thin-wall approximation and the yellow dots the same quantity found from the solving minimizing the action (3.27). As expected, the thin-wall approximation underestimates nucleation temperatures significantly. We find that the true nucleation temperature \( T_n \) and the thin-wall value \( T_{nw} \) are related by \( T_n = T_c - \kappa (T_c - T_{nw}) \), where the coefficient \( \kappa \) to some extent depends on \( M_S \) and \( \lambda_S \), but is less than 0.7. We anticipated this result in the deflagration limit shown in the right panel of figure 6, where the yellow dots were found by redefining all thin-wall nucleation temperatures by the above equation with \( \kappa = 0.7 \). We believe that this scaling conservatively represents the effect of going beyond thin-wall approximation in the full 2HDM and singlet model, and hence shows that most parameter sets may in fact be deflagrations.

We can also tune our search to prefer models with a higher critical temperatures. Indeed, as is clearly seen from figure 7, the nucleation temperature approaches the critical temperature when \( T_c \) gets higher in the singlet extension of the SM and this feature persists also in the full 2HDSM. Hence, we made a new parametric scan, where we accept only models with \( T_c > 80 \) GeV. The result of this scan is seen in figure 8, where the left panel again shows the baryon-to-entropy ratio and in the right panel the deflagration bound. We now found more points with large asymmetry and, in particular after one rescales the thin-wall nucleation temperatures as explained above, these points are now well below the deflagration bound eq. (3.26).

4 Conclusions and outlook

We have studied the viability of a two-Higgs-doublet and inert-singlet model for EWBG and for DM, taking into account also all existing observational and collider constraints. Our model
Figure 8. Left: shown is the baryon asymmetry $\eta_B$ as a function of $T_n/T_c$ for the new scan with $T_c > 80$GeV (here $T_n$ is the thin-wall nucleation temperature). Right: the deflagration bound for the set displayed in the left plot as a function of $T_n/T_c$. We again used $\alpha_{\text{max}} = 0.38$.

is based on the maximal $\text{GL}(2,\mathbb{C})$ reparametrization symmetry. This implies a universal Yukawa-alignment scheme, where both Higgs fields couple similarly to all fermions and there are no FCNCs. Exploiting the $\text{GL}(2,\mathbb{C})$ symmetry, the model can, in a particular basis, be written with a type-I Yukawa sector combined with the most general CP-violating potential.

Following [19–21] we implemented a novel way to construct potentials with a tree-level stability and to study the symmetry breaking patterns at finite temperatures. This construction was based on the Lorentz symmetry induced by the reparametrization symmetry on bilinears formed from Higgs doublets. These techniques are applicable to all 2HDM models, and they proved extremely useful when performing large-scale parametric scans over the multidimensional phase space of the model.

Dark matter and the strength of the electroweak phase transition in the model follow a similar pattern to the pure singlet extension: in accordance with [33], we find that strong two-step phase transitions are easily found, but they are consistent only with a subleading DM. Likewise, we find that experimental and observational constraints are fairly easy to avoid, with the outstanding exception of the electron-EDM bound, which strongly constrains the CP-violating parameters on the model. EDM constraints are particularly important because creating a large baryon asymmetry during the electroweak phase transition requires large CP-violating parameters; we found that the electron EDM indeed strongly constrains the phase space consistent with baryogenesis. Yet the bounds are not as strong here as in the pure 2HDM case [8], and we found a number of models consistent with all requirements.

Finally, we observed that two-step transitions may suffer from an unexpected problem of providing a too strong phase transition. We found that fields may get trapped in the metastable minimum, and transition walls may not be subsonic as required by a successful baryogenesis scenario. However, our analysis in the full model was restricted to the thin-wall approximation. We then studied the bubble nucleation in full generality in the case of a pure singlet extension of the SM. While the generic problem of too strong transitions remained, we found that thin-wall limit tends to overestimate the strength of the transition. Based on these results we argued that, when corrected for the thin-wall bias, the walls may well remain subsonic in the full 2HDSM. In a revised scan concentrating to models with a large critical temperature, we found many models potentially consistent with all available constraints.

While our results are not a definite proof, they provide a strong indication of a success of baryogenesis in the 2HDM and an inert singlet model. Settling the issue beyond any doubt
would require two significant improvements. First is a detailed analysis of the bubble wall dynamics, including a microscopic computation of the friction on the wall. Second is going beyond the $T_c$-bounce solution, when solving the scalar field profiles over the bubble wall to compute the top quark mass profile and eventually the baryogenesis source. These are both very interesting topics that deserve to be studied in detail in the future.

Acknowledgments

We thank Jim Cline for useful comments and discussions and collaboration at the early stages of this work. KK would like to thank Stephan Huber, Mikko Laine and Jose M. No for useful discussions during the 2016 MIAPP programme “Why is there more Matter than Antimatter in the Universe?”. This work was financially supported by the Academy of Finland projects 278722 and 267842. The CP$^3$-Origins centre is partially funded by the Danish National Research Foundation, grant number DNRF90. VV is supported by the Magnus Ehrnrooth foundation. TA acknowledges partial funding from a Villum foundation grant.

A Diagonalization of mass matrices

In this appendix we show in detail the diagonalization of the scalar mass matrices. Following [43], we write

$$r^\mu = \phi_x \Sigma^{\mu}_{xy} \phi_y,$$

(A.1)

where

$$\phi = (H_{1I}^+, H_{2I}^+, H_{1R}^+, H_{2R}^+, H_1^0, H_2^0, H_{1R}^0, H_{2R}^0),$$

(A.2)

where the subscripts R and I again denote the real and imaginary parts, respectively, and the matrices $\Sigma^\mu$ are block diagonal matrices consisting of $4 \times 4$ elements

$$\Sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 0 & -i\sigma_2 \\ i\sigma_2 & 0 \end{pmatrix}, \Sigma_3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}.$$

(A.3)

Now the $(a,b)$ element of the mass matrix is

$$\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \bigg|_{\phi = \langle \phi \rangle} = -\Sigma^{\mu}_{ab} M_\mu + \Sigma^{\mu}_{ab} \lambda_\mu \langle r^\mu \rangle + 2\Sigma^{\mu}_{ax} \langle \phi_x \rangle \lambda_\mu \langle \phi_y \rangle \Sigma^\nu_{yb},$$

(A.4)

and the minimum conditions are

$$\frac{\partial V}{\partial \phi_a} \bigg|_{\phi = \langle \phi \rangle} = (-\Sigma^{\mu}_{ax} M_\mu + \Sigma^{\mu}_{ax} \lambda_\mu \langle r^\mu \rangle) \langle \phi_x \rangle = 0.$$

(A.5)

Let us consider the neutral part of the mass matrix. The eigenstate corresponding to eigenvalue zero is $G_0 = N_G \Sigma^5 \langle \phi^0 \rangle$, where $\Sigma^5 = \Sigma^0 \Sigma^1 \Sigma^2 \Sigma^3$ and $N_G$ is the normalization factor. Now

$$G_0 = \cos \beta \cos \theta H_{1R}^0 + \sin \beta H_{1I}^0 - \cos \beta \sin \theta H_{2R}^0,$$

(A.6)

where

$$\beta = \arctan \frac{v_2}{v_1}.$$

(A.7)
Hence we can diagonalize the neutral mass matrix by

\[
R_N = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & c_\beta c_\theta & -c_\beta s_\theta & s_\beta \\
0 & s_\beta c_\theta & c_\beta s_\theta & s_\beta \\
0 & s_\beta c_\theta & c_\beta s_\theta & c_\beta s_\theta
\end{pmatrix} \begin{pmatrix}
c_\beta c_\theta & s_\beta & c_\beta s_\theta & 0 \\
-s_\beta c_\theta & c_\beta & s_\beta s_\theta & 0 \\
-s_\beta & 0 & c_\theta & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

(A.8)

where we have used the short-hand notations \( s_\alpha \equiv \sin \alpha, \ c_\alpha \equiv \cos \alpha \), and the neutral mass eigenstates are

\[
\begin{pmatrix}
G_0 \\
A_0 \\
H_0 \\
h_0
\end{pmatrix} = R_N \begin{pmatrix}
H^0_{11} \\
H^0_{12} \\
H^0_{1R} \\
H^0_{2R}
\end{pmatrix}.
\]

(A.9)

On the charged sector there are two eigenstates with eigenvalue zero: \( G_1 = N_{G_1} \Sigma^5 \langle \phi^0 \rangle \) and \( G_2 = N_{G_2} \langle \phi^0 \rangle \). Now

\[
G_1 = \cos \beta \sin \theta H^+_{1R} + \cos \beta \cos \theta H^+_{1I} + \sin \beta H^+_{2I},
\]

\[
G_2 = \cos \beta \cos \theta H^+_{1R} - \cos \beta \sin \theta H^+_{1I} + \sin \beta H^+_{2R}.
\]

(A.10)

Hence charged Goldstone bosons are \( G^\pm = \cos \beta e^{\pm i \theta} H^\pm_1 + \sin \beta H^\pm_2 \). The charged mass matrix in the basis \( \{ H^+_1, H^+_2 \} \) can be diagonalized by

\[
R_C = \begin{pmatrix}
c_\beta e^{i \theta} & s_\beta \\
-s_\beta e^{i \theta} & c_\beta
\end{pmatrix},
\]

(A.11)

and charged mass eigenstates are

\[
\begin{pmatrix}
G^+_+ \\
H^+_+
\end{pmatrix} = R_C \begin{pmatrix}
H^+_1 \\
H^+_2
\end{pmatrix}.
\]

(A.12)

B 1-loop beta functions for scalar couplings

For completeness we show the 1-loop beta functions,

\[
\beta_{\lambda_i} \equiv \frac{d\lambda_i}{d \ln \mu},
\]

(B.1)
we used to compute the perturbativity limits on the model parameters:

\begin{align}
16\pi^2\beta_{\lambda_1} &= 24\lambda_1^2 + 2\lambda_3^2 + 2\lambda_3\lambda_4 + \lambda_4^2 + 4|\lambda_5|^2 + 12|\lambda_6|^2 \\
&+ \frac{1}{2}\lambda_{S1}^2 - 3(3g_L^2 + g_Y^2)\lambda_1 + \frac{3}{8}(3g_L^2 + 2g_Y^2 g_Y^2 + g_Y^4), \\
16\pi^2\beta_{\lambda_2} &= 24\lambda_2^2 + 2\lambda_3^2 + 2\lambda_3\lambda_4 + \lambda_4^2 + 4|\lambda_5|^2 + 12|\lambda_7|^2 + \frac{1}{2}\lambda_{S2}^2 \\
&- 3(3g_L^2 + g_Y^2)\lambda_2 + \frac{3}{8}(3g_L^2 + 2g_Y^2 g_Y^2 + g_Y^4) + 12g_Y^2\lambda_2 - 6g_Y^4, \\
16\pi^2\beta_{\lambda_3} &= (\lambda_1 + \lambda_2)(12\lambda_3 + 4\lambda_4) + 4\lambda_3^2 + 2\lambda_4^2 + 8|\lambda_5|^2 + 4|\lambda_6|^2 \\
&+ 16\text{Re}(\lambda_6^*\lambda_7) + 4|\lambda_7|^2 + \lambda_{S1}\lambda_{S2} \\
&- 3(3g_L^2 + g_Y^2)\lambda_3 + \frac{3}{4}(3g_L^2 - 2g_Y^2 g_Y^2 + g_Y^4) + 6g_Y^2\lambda_3, \\
16\pi^2\beta_{\lambda_4} &= 4(\lambda_1 + \lambda_2)\lambda_4 + 8\lambda_3\lambda_4 + 4\lambda_4^2 + 32|\lambda_5|^2 + 10(|\lambda_6|^2 + |\lambda_7|^2) \\
&+ 4\text{Re}(\lambda_6^*\lambda_7) + |\lambda_{S12}|^2 - 3(3g_L^2 + g_Y^2)\lambda_4 + 3g_Y^2 g_Y^2 + 6g_Y^4\lambda_4, \\
16\pi^2\beta_{\lambda_5} &= (4\lambda_1 + 4\lambda_2 + 8\lambda_3 + 12\lambda_4)\lambda_5 + 5\lambda_5^2 + 2\lambda_6\lambda_7 + 5\lambda_7^2 \\
&+ \frac{1}{2}\lambda_{S12}^2 - 3(3g_L^2 + g_Y^2)\lambda_5 + 6g_Y^2\lambda_5, \\
16\pi^2\beta_{\lambda_6} &= (24\lambda_1 + 6\lambda_3 + 8\lambda_4)\lambda_6 + (6\lambda_3 + 4\lambda_4)\lambda_7 + (20\lambda_6^2 + 4\lambda_7^2)\lambda_5 \\
&+ \lambda_{S1}\lambda_{S2} - 3(3g_L^2 + g_Y^2)\lambda_6 + 3g_Y^2\lambda_6, \\
16\pi^2\beta_{\lambda_7} &= (6\lambda_3 + 4\lambda_4)\lambda_6 + (24\lambda_2 + 6\lambda_3 + 8\lambda_4)\lambda_7 + (4\lambda_5^2 + 20\lambda_7^2)\lambda_5 \\
&+ \lambda_{S2}\lambda_{S12} - 3(3g_L^2 + g_Y^2)\lambda_7 + 9g_Y^2\lambda_7, \\
16\pi^2\beta_{S} &= 18\lambda_5^2 + 2\lambda_{S1}^2 + 2\lambda_{S2}^2 + 4|\lambda_{S12}|^2, \\
16\pi^2\beta_{\lambda_{S1}} &= (12\lambda_1 + 6\lambda_5)\lambda_{S1} + (4\lambda_3 + 2\lambda_4)\lambda_{S2} + 4\lambda_{S1}^2 + 12\text{Re}(\lambda_6^*\lambda_{S12}) \\
&+ 4|\lambda_{S12}|^2 - \frac{3}{2}(3g_L^2 + g_Y^2)\lambda_{S1}, \\
16\pi^2\beta_{\lambda_{S2}} &= (4\lambda_3 + 2\lambda_4)\lambda_{S1} + (12\lambda_2 + 6\lambda_5)\lambda_{S2} + 4\lambda_{S2}^2 + 12\text{Re}(\lambda_7^*\lambda_{S12}) \\
&+ 4|\lambda_{S12}|^2 - \frac{3}{2}(3g_L^2 + g_Y^2)\lambda_{S2} + 6g_Y^2\lambda_{S2}, \\
16\pi^2\beta_{\lambda_{S12}} &= (2\lambda_3 + 4\lambda_4 + 6\lambda_S + 4\lambda_{S1} + 4\lambda_{S2})\lambda_{S12} + 12\lambda_5\lambda_{S12} \\
&+ 6\lambda_6\lambda_{S1} + 6\lambda_7\lambda_{S2} - \frac{3}{2}(3g_L^2 + g_Y^2)\lambda_{S12} + 3g_Y^2\lambda_{S12}. \\
\end{align}

\section{Singlet scalar annihilation cross sections}

Here we give compact expressions for singlet pair annihilation rates used in the computation of the DM abundances. In the mass eigenbasis, the relevant Lagrangian may be written as

\begin{equation}
\mathcal{L} = \sum_{j,k} \left( c_{0j} S_j H_j H_k + \sum_i c_{ij} H_i H_j H_k \right) + \sum_i c_i S_i^2 H_i + \sum_{i,j} \left( g_{ij} H_i^T R_f L + \text{h.c.} \right) + \sum_i \left( g_{W1} W^+ W^- H_i + g_{Zi} Z Z H_i \right),
\end{equation}
where $H = (A_0, H_0, h_0, H^+)$. Now the singlet scalar annihilation cross section to scalars is given by:

$$
\sigma(SS \rightarrow H_j H_k) = \frac{1}{16\pi s^2 v_s} \left( A_{jk}^2 I_{0,j,k} + 2 A_{jk} B_{jk} I_{1,j,k} + B_{jk}^2 I_{2,j,k} \right) .
$$

The rate into fermion-antifermion pairs is given by:

$$
\sigma(SS \rightarrow \bar{f} f) = \frac{I_{0,f,f}}{4\pi s^2 v_s} \sum_{i,j} \frac{c_i c_j}{(s - m_i^2)(s - m_j^2)} \times \left( (y_{fR} y_{jR} - y_{iR} y_{jR}) \frac{s}{2} - 2 y_{fR} y_{jR} m_f^2 \right) ,
$$

and the annihilation to vector bosons is:

$$
\sigma(SS \rightarrow VV) = \frac{I_{0,V,V}}{4\pi s^2 v_s} \sum_{i,j} c_i c_j g_{V,i} g_{V,j} \left( 3 + \frac{s(s - 4 M_V^2)}{4 M_V^2} \right) \left( 1 - \frac{\delta_{V,Z}^2}{2} \right) ,
$$

where

$$
I_{0,j,k} = \sqrt{s^2 + (m_j^2 - m_k^2)^2 - 2s(m_j^2 + m_k^2)} ,
$$

$$
I_{1,j,k} = \frac{4I_{0,j,k}}{m_j^2 + m_k^2 - s} ,
$$

$$
I_{2,j,k} = \frac{2I_{0,j,k}}{m_j^2 m_k^2 - M_S^2(2m_j^2 + 2m_k^2 - s) + \frac{M_S^2}{s}(m_j^2 - m_k^2)^2} + \frac{8I_{0,j,k}}{(m_j^2 + m_k^2 - s)^2} ,
$$

and

$$
A_{jk} = 2(1 + \delta_{jk}) c_0 + 2 \sum_i (1 + \delta_{ij} + \delta_{jk} + \delta_{ki} + 2\delta_{ij} \delta_{jk}) \frac{c_i c_{ij}}{s - m_i^2} ,
$$

$$
B_{jk} = 4c_j c_k .
$$

References


