Nonperturbative Analysis of the Electroweak Phase Transition in the Two Higgs Doublet Model

Andersen, Jens O.; Gorda, Tyler; Helset, Andreas; Niemi, Lauri; Tenkanen, Tuomas V., I; Tranberg, Anders; Vuorinen, Aleksi; Weir, David J.

Published in: Physical Review Letters

DOI: 10.1103/PhysRevLett.121.191802

Publication date: 2018

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Nonperturbative Analysis of the Electroweak Phase Transition in the Two Higgs Doublet Model

Jens O. Andersen,1,4 Tyler Gorda,2,3,† Andreas Helset,1,4,‡ Lauri Niemi,2,§ Tuomas V. I. Tenkanen,2,¶
Anders Tranberg,2,¶ Aleks Vuorinen,2,** and David J. Weir2,††

1Department of Physics, Faculty of Natural Sciences, Norwegian University of Science and Technology,
Høgskoleni, 5, N-7491 Trondheim, Norway
2Department of Physics and Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland
3Department of Physics, University of Virginia, 382 McCormick Road, Charlottesville, Virginia 22904-4714, USA
4Niels Bohr Institute, University of Copenhagen, Blegdamsvæj 17, 2100 Copenhagen
5Department of Mathematics and Physics, University of Stavanger, 4036 Stavanger, Norway

(Received 15 January 2018; revised manuscript received 21 September 2018; published 7 November 2018)

We perform a nonperturbative study of the electroweak phase transition (EWPT) in the two Higgs doublet model (2HDM) by deriving a dimensionally reduced high-temperature effective theory for the model, and matching to known results for the phase diagram of the effective theory. We find regions of the parameter space where the theory exhibits a first-order phase transition. In particular, our findings are consistent with previous perturbative results suggesting that the primary signature of a first-order EWPT in the 2HDM is \( m_A > m_H^0 + m_Z \).

DOI: 10.1103/PhysRevLett.121.191802

Introduction.—Accounting for the baryon asymmetry in the present Universe is a major unsolved problem in cosmology. One of the leading candidates for a viable mechanism, electroweak baryogenesis (EWBG) [1], suggests that the asymmetry originates from the electroweak phase transition (EWPT) in the early Universe. According to the Sakharov conditions [2] the transition would have to be first order, accompanied by a sizable violation of CP symmetry. Unfortunately, these conditions immediately rule out EWBG within the minimal standard model (SM), as it was demonstrated that the SM EWPT is a crossover [3–6], and that SM CP-violating effects are heavily suppressed at high temperatures [7–9].

Independently of the question of baryon asymmetry, a host of beyond the standard model (BSM) theories have been proposed to solve open problems in physics. Determining whether BSM theories can produce a first-order EWPT and thus facilitate EWBG is nontrivial: quantitatively reliable conclusions about the phase transition typically require a nonperturbative approach, deemed unmanageable for large parameter spaces. Because of this difficulty, analyses based on the finite-temperature effective potential have become standard, particularly for physical observables: in one study [17], errors in excess of 10% in the critical temperature and 50% in the latent heat were found, compared to nonperturbative studies.

In contrast, a more reliable approach uses dimensionally reduced effective theories, originally applied to the SM in Refs. [3–5,18,19], and recently applied to the SM accompanied by a real singlet [20]. In this Letter, we use this method to treat a widely studied BSM model, the two Higgs doublet model (2HDM), where the SM is augmented with an additional Higgs doublet [see Ref. [21] for a review, and Refs. [22–24] for earlier work on dimensional reduction (DR) in the 2HDM]. We derive a three-dimensional high-T effective theory, studying regions of parameter space where this theory has the same form as that of the standard model, similar to Ref. [25]. This reduces determining the phase diagram of the theory to mapping its parameter space to that of the SM effective theory. Equipped with the analysis of Refs. [3–5], we discover interesting and phenomenologically viable regions of parameter space where the EWPT is first order, corroborating key findings of perturbative studies of EWBG in the 2HDM.

Dimensional reduction of the 2HDM.—Our four-dimensional starting theory can be described by the schematic action

\[
S = \int d^4x [\mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{Yukawa}}],
\]

suppressing counterterm and ghost contributions. The field content includes SU(3)_c, SU(2)_L, and U(1)_Y gauge fields, two scalar doublets \( \phi_1 \) and \( \phi_2 \), as well as all fermions...
present in the SM. In our present treatment, we will consider only one quark flavor in the Yukawa sector, namely the top, since it has the largest coupling to the Higgs field. The top quark couples to one doublet only (by convention $\phi_2$), and we have not yet committed to a specific type of 2HDM (I or II).

The extended scalar sector of our model reads

$$L_{\text{scalar}} = \sum_{i=1}^{2} (D_\mu \phi_i)\overline{(D_\mu \phi_i)} + V(\phi_1, \phi_2),$$

with usual covariant derivative $D_\mu$ and the potential

$$V(\phi_1, \phi_2) = \mu_1^2 \phi_1^\dagger \phi_1 + \mu_2^2 \phi_2^\dagger \phi_2 + \mu_3^2 \phi_1^\dagger \phi_2 + \mu_4^2 \phi_2^\dagger \phi_1 + \lambda_1 (\phi_1^\dagger \phi_1)^2 + \lambda_2 (\phi_2^\dagger \phi_2)^2 + \lambda_3 (\phi_1^\dagger \phi_2)(\phi_2^\dagger \phi_1) + \lambda_4 (\phi_1^\dagger \phi_2)^2 + \lambda_5 (\phi_2^\dagger \phi_1)^2.$$

In general, $C(P)$ symmetry is broken when $\lambda_4$ or $\mu_{2,3}^2$ are complex; we have discarded so-called hard $CP$-breaking terms, often parametrized by $\lambda_{6,7,\gamma}$; cf. Refs. [21,26].

The first three-dimensional effective theory, obtained by integrating out the “superheavy” hard scale $\pi T$ (see, e.g., Ref. [20] for details of the procedure), has schematic form

$$S = \int d^3 x [L_{\text{gauge}}^{(3)} + L_{\text{scalar}}^{(3)} + L_{\text{temporal}}^{(3)}],$$

again suppressing ghost and counterterm contributions. The field content is now $SU(2)_L$ and $U(1)_Y$ gauge fields, two Higgs doublets, and temporal scalar fields $A_0^0, B_0, C_0^0$. The fermions are integrated out and the SU(3)$_c$ gauge fields can be neglected [20]. The fundamental scalar sector remains of the form

$$L_{\text{scalar}}^{(3)} = (D_r \phi_1)^\dagger (D_r \phi_1) + (D_r \phi_2)^\dagger (D_r \phi_2) + V(\phi_1, \phi_2),$$

where $r = 1, 2, 3$ is summed over. In the second step of DR, the heavy temporal scalar fields are integrated out.

Although the theory in Eq. (4) is already suitable for lattice simulations, it can be further simplified by noticing that $\phi_1$ and $\phi_2$ mix when $\mu_{1,2}^2 \neq 0$, and near the phase transition there typically exists a hierarchy between the mass eigenvalues. This observation—specific to the 2HDM—allows us to integrate out the heavy mode and study the phase transition with only one scalar field coupled to the gauge fields. Our final effective theory becomes

$$S = \int d^3 x \left[ \hat{L}_{\text{gauge}}^{(3)} + \hat{L}_{\text{scalar}}^{(3)} \right],$$

$$\hat{L}_{\text{scalar}}^{(3)} = (D_r \phi)^\dagger (D_r \phi) + \hat{\mu}_3^2 \phi^\dagger \phi + \hat{\lambda}_3 (\phi^\dagger \phi)^2. (6)$$

Here, $\phi$ is the remaining light $\phi_1$-$\phi_2$ mode, and the parameters of the theory include $\hat{\mu}_3^2$, $\hat{\lambda}_3$, and the 3D gauge couplings $\hat{g}_1$ and $\hat{g}_3$ for the $U(1)_Y$ and $SU(2)_L$ interactions. As in the analysis of Refs. [3,18], we omit all non-perturbative effects related to the $U(1)_Y$ field.

The main task of DR is to perturbatively match the parameters of the original 4D theory, Eq. (2), to those of the final effective theory, Eq. (6). This is accomplished by demanding that the effective theory reproduces the static Green functions of the original theory at large distances $R \gg 1/T$. This results in a number of matching relations from which the effective theory parameters are solved. This procedure is presented in Ref. [26] and summarized in Supplemental Material [27].

As discussed above, the effective theory of Eq. (6) has the same form as that of the SM, studied in Refs. [3–5], but with different matching relations. This allows us to adopt existing numerical results for the strength of the phase transition and study the phase diagram through our matching procedure alone.

The validity of DR can be quantified by estimating the effect of neglected dimension-six operators. While it is difficult to comprehensively gauge their effect, one can evaluate the change in the vacuum expectation value (VEV) of the Higgs field in the effective theory caused by the $(\phi^3 \phi)$ operators. In Eq. (201) of Ref. [18], it was shown that in the SM the dominant neglected contribution comes from the top quark; its effect is about 1%. In the first DR step where the superheavy fields are integrated out, we estimate the effect of new BSM contributions by comparing their magnitude to the contribution from the top quark [see Eqs. (34) and (35) in Supplemental Material [27]]. However, in many cases, the operator $O_{B}^{(6)} = \hat{A}_6 (\phi^3 \phi)^3_{3D}$ generated when the heavier doublet is integrated out dominates over the six-dimensional operators of the first step, denoted $\{O_{A}^{(6)}\}$. We discuss these operators in detail below.

Finally, although the parameter matching is perturbative, the study of the 3D phase diagram is nonperturbative and—within the limitations of lattice methods—exact. The main advantage of our approach lies in proper handling of the infrared physics, which causes trouble in traditional perturbative studies of the EWPT. Resummations are performed when the superheavy and heavy scales are integrated out perturbatively, and the problematic light modes are treated nonperturbatively on the lattice. However, the mapping to precise values of the 4D parameters, where this phase transition occurs in the 2HDM, is limited by the accuracy of the perturbative truncation. We organize the expansion in terms of the gauge coupling $g$, and perform the DR to $O(g^4)$. Thus the calculation is carried out at the one-loop level for quartic couplings, and two-loop level for mass parameters. This exceeds the accuracy used in the perturbative calculations of, e.g., Ref. [28] (see, however, Ref. [29] for a recent two-loop perturbative calculation in the inert doublet
model). The uncertainty in the effective theory due to the choice of renormalization scale is discussed in Supplemental Material [27].

Scanning the parameter space.—The phase diagram of the dimensionally reduced theory can be mapped using the dimensionless parameters \( x \equiv \lambda_3 / g_3^2 \), \( y \equiv \tilde{\mu}_3^2 / g_4^2 \). It is known that within this theory the EWPT occurs near \( y \simeq 0 \), where the Higgs mass parameter becomes negative. In Refs. [3–5], it was found that the transition is first order for \( x \lesssim 0.11 \), and strongly so for \( x \lesssim 0.04 \). In this Letter we are focused on finding where the crossover turns into a first-order transition.

We search for areas of 2HDM parameter space that map onto regions of the 3D effective theory with \( x < 0.11 \) and \( y \simeq 0 \). Since there are ten real parameters in the 4D theory and only three in the 3D one, inverting the mapping process is not unique. We perform scans of the 2HDM parameter space, guided by the results of Ref. [30] that combine phenomenological constraints with a one-loop resummed perturbative determination of the effective potential. Other recent treatments are found in Refs. [31,32].

A uniform scan through a 10-dimensional space is computationally expensive; we must therefore make some simplifying assumptions. We take all parameters of the 2HDM to be real, setting \( \text{Im}(\lambda_3) = 0 \), \( \text{Im}(\mu_{12}^2) = 0 \). This eliminates extra CP violation in the model, which would be crucial for baryogenesis. However, the effect of these imaginary parts on the strength of the transition is expected to be negligible; the CP-violating phase must necessarily be small due to electric dipole moment (EDM) constraints [33–35].

Next, we reparametrize the model following Ref. [30], applying tree-level relations between the \( \overline{\text{MS}} \) parameters and physical quantities; accounting for (possibly sizable) loop effects from vacuum renormalization is left for future work. The masses of the CP-even scalars are denoted by \( m_h = 125 \text{ GeV} \) and \( m_{H_0} \), that of the CP-odd scalar by \( m_{A_0} \), and that of the charged scalar by \( m_{H^\pm} \). We also employ two angles \( \alpha \) and \( \beta \): \( \alpha \) parametrizes mixing between the CP-even states, while \( \beta \) is related to the ratio of the VEVs \( \tan(\beta) \equiv (\nu_2 / \nu_1) \). Here, \( \nu_1 \) and \( \nu_2 \) are the VEVs for \( \phi_1 \) and \( \phi_2 \), respectively, with \( \nu_1^2 + \nu_2^2 = \nu^2 \) and \( \nu = 246 \text{ GeV} \). Finally, there is the squared mass scale \( M^2 \equiv \mu^2 [\tan(\beta) + 1 / \tan(\beta)] \), where we treat \( \mu^2 \equiv \text{Re} \mu_{12}^2 \) as an input parameter. The relations between the physical states and gauge eigenstates can be obtained from Ref. [30].

We also fix \( m_{H^\pm} = m_{A_0} \), since EW precision tests require the mass of the charged Higgs boson to be roughly degenerate with either \( H_0 \) or \( A_0 \) [36,37]. Furthermore, we work in the alignment limit, setting \( \cos(\beta - \alpha) = 0 \). In this limit, the CP-even scalar \( h \) couples to SM particles exactly like the SM Higgs boson. We investigate relatively few values for \( \tan(\beta) \), whereas we perform a more exhaustive scan in a three-dimensional parameter space spanned by \( m_{H_0} \), \( m_{A_0} \), and \( \mu^2 \). At each point, we require that tree-level stability and unitary constraints be satisfied; for details, see Ref. [26]. Furthermore, for the scaling assumptions of DR to be valid, the tree-level mass parameters \( \mu_{11}, \mu_{22} \), and \( \mu_{12} \) should be comparable to the Debye mass \( m_D \sim gT \) near the phase transition. This sets an upper bound for the input parameter \( \mu \lesssim 200 \text{ GeV} \). Finally, we verify that in the effective theory the other doublet really is heavy near the phase transition, so it is justified to integrate it out.

Results.—Following our scanning protocol outlined above, we fix \( \tan(\beta) \) and scan in the two scalar masses \( m_{H_0} \) and \( m_{A_0} \) between 137.5 and 562.5 GeV at spacings of 6.25 GeV, a total of 4624 points. A dense scan in \( \mu \) is then carried out for each pair, from 10 to 150 GeV at intervals of 2.5 GeV for a total of 56 values. In all, each of our fixed-\( \tan(\beta) \) plots results from scanning approximately 260 000 points. The upper limit on \( \mu \) is chosen to ensure that the DR is valid, as explained above.

We first check whether each point is physical, according to our criteria. If so, we then perform the DR for evenly spaced temperatures between 80 and 200 GeV, at intervals of 20 GeV. This allows us to find the value of \( y \) when \( y = 0 \)—on the critical line—by interpolation. We then use \( x \) to characterize the phase transition. We take \( 0 < x < 0.11 \) as an indicator of a first-order EWPT, the upper limit coming from previous lattice work.

FIG. 1. Heat maps with fixed \( \tan(\beta) \), showing regions of first-order EWPT (\( 0 < x < 0.11 \) and \( y \simeq 0 \)) in the alignment limit. The dotted lines correspond to \( m_{A_0} = m_{H_0} \) and \( m_{A_0} = m_{H_0} \pm m_Z \).
Combining different values of \( \mu \), we indicate the relative number of points with a first-order phase transition as a heat map in Fig. 1, for three separate values of \( \tan(\beta) \). The majority of our points reside in the region \( m_{A_0} > m_{H_u} + m_{Z'} \), in accordance with Refs. [28,30] (see, however, Refs. [31,32]). In our framework, sufficiently strong interactions with the second doublet are necessary to bring \( x \) down from its SM value of \( x > 0.11 \). Although the relation between the 4D inputs and \( x \) is complicated by the diagonalization, a mass hierarchy between \( H_0 \) and \( A_0 \) generically results in large portal couplings \( \lambda_{3-5} \) and small values of \( x \) in the upper region. However, at small \( \tan(\beta) \) we also see a considerable number of points in regions where this does not hold.

In Fig. 2, we show a breakdown of the heat map plot with fixed \( \tan(\beta) = 2.0 \) for two values of \( \mu \). We include here an estimate of the effect of two of the neglected six-dimensional operators \( O_{A,1}^{(6)} \) and \( O_{A,2}^{(6)} \) produced when the superheavy scale is integrated out. Generally, decreasing values of \( x \) correspond to increasing importance of dimension-six terms: when the effect of these terms becomes large, the DR breaks down. These plots also show how the lower first-order region disappears as \( \mu \) increases.

We have found by explicit computation that the negative-\( x \) region at large \( m_{A_0} \) is due to the omission of the sixdimensional operator \( O_B^{(6)} \) in the last DR step that, although inversely proportional to the heavy doublet mass, obtains sizable contributions from the large couplings. We estimate its effect by computing the dominant tree-level diagram contributing to the operator coefficient (see Ref. [38] and Supplemental Material [27]) and determining the two-loop effective potential in the final effective theory with this operator included (cf. Refs. [18,39]). We stress, however, that the effective potential is only a tool for estimating errors from omitted six-dimensional operators; our results concerning the phase transition are obtained using the nonperturbative phase diagram of Refs. [3,5].

In Fig. 3, the effective potential is depicted at two values of \( x \), both with and without the effects of the sixdimensional operator \( O_B^{(6)} \). The field \( \phi \) is the 3D background field, defined via \( \langle \phi \rangle_3 = (\phi/\sqrt{2})(0,1)^T \) and related to 4D fields, as described in Supplemental Material [27]. The figure demonstrates how at \( x = 0.108 \)—near the crossover boundary—the six-dimensional operator \( O_B^{(6)} \) has a negligible impact on the potential, while for \( x = 0.068 \) (which corresponds to \( \phi_c/T_c \approx 0.7 \)) the effect is already sizable, continuing to grow as \( x \) decreases. Hence, integrating out the heavier doublet is expected to be a valid approximation when the transition is of weakly first order, but becomes increasingly challenged near the strong transition limit of \( x \lesssim 0.04 \). While we expect our results to be qualitatively robust even there, reaching quantitatively accurate results for very small \( x \) clearly calls for simulations with two dynamical doublets, which we leave for future work.

Experimental constraints on the 2HDM parameter space depend strongly on the way in which fermions couple to the Higgs doublets. With the exception of the top quark, other Yukawa couplings have little effect on our EWPT analysis,
and we have still to indicate whether we are considering type I (all quarks couple to \(\phi_2\)) or type II (up-type quarks couple to \(\phi_2\), down-type quarks to \(\phi_1\)) 2HDM. The most stringent constraints come from flavor physics, where \(B\) decays set the bound \(m_{H^0} \gtrsim 580\) GeV for the charged Higgs mass in type II [40]. Assuming that \(m_{H^0}\) is degenerate with \(m_{A_0}\), in accordance with EW precision tests, this rules out our regions of first-order EWPT in type II, but no such lower bound exists in type I for \(\tan\beta \gtrsim 2\) [40,41].

Additional restrictions come from direct searches for neutral Higgs bosons at the LHC [42]. For type I, the \(H_0 \rightarrow \tau\tau\) cross section is suppressed by \(\cot^2\beta\), and our choices of \(\tan\beta\) are within current experimental bounds. Finally, we have verified that the mass range we scan in is allowed by measurements of the \(h \rightarrow \gamma\gamma\) decay [43], as well as the relatively recent search for \(A_0 \rightarrow Zh\) processes [44]. Having not scanned in the hidden-Higgs region where constraints from charged-scalar searches become important [45], we conclude that our first-order EWPT regions are currently not ruled out by experiments if a type I 2HDM is assumed.

Discussion.—It is a shortcoming of present-day particle cosmology that it is still impossible to reliably determine the nature and strength of the EWPT for a given BSM scenario. This information would be valuable not only for EWBG, but also for gravitational-wave physics, as a first-order EWPT would leave an imprint in the sensitivity range of the LISA mission and other proposed gravitational-wave detectors [46]. We have taken a step towards remedying the situation by studying the mapping of the phase diagram of one viable BSM theory, the 2HDM. Our results concern the EWPT in the alignment limit \(\cos(\beta - \alpha) = 0\). Our work so far supports the idea that the primary signature of a first-order transition in this theory is indeed \(m_{A_0} > m_{H^0} + m_Z\), as suggested by Refs. [28,30].

The techniques discussed in this Letter can be applied, with suitable modifications, to a host of other models where a substantial region of parameter space can be mapped onto the three-dimensional theory of the minimal standard model. In the future, our aim is to perform a thorough comparison of perturbative and nonperturbative results in the 2HDM by keeping both doublets dynamical in the effective theory. Similar projects to study the EWPT and benchmark the accuracy of perturbation theory are already under way in the standard model augmented by a real singlet [47] or triplet field [48]; the EWPT has been perturbatively analyzed for the former in Refs. [49,50], and for the latter in Ref. [51].

The authors would like to thank Tomáš Brauner, Mark Hindmarsh, Stephan J. Huber, Kimmo Kainulainen, Keijo Kajantie, Venus Keus, Mikko Laine, Jose M. No, and Kari Rummukainen for discussions. T. V. I. T. has been supported by the Vilho, Väinö and Kalle Väisälä Foundation. T. G., L. N., T. V. I. T., and A. V. have been supported by the Academy of Finland Grant No. 1303622, as well as by the European Research Council Grant No. 725369. L. N. was also supported by the Academy of Finland Grant No. 308791. D. J. W. (ORCID ID 0000-0001-6986-0517) was supported by Academy of Finland Grant No. 286769.


