

Resolving the Hubble tension with new early dark energy

Florian Niedermann^{*} and Martin S. Sloth[†]

*CP³-Origins, Center for Cosmology and Particle Physics Phenomenology,
University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*

 (Received 4 July 2020; accepted 27 August 2020; published 23 September 2020)

New early dark energy (NEDE) is a component of vacuum energy at the electron volt scale, which decays in a first-order phase transition shortly before recombination [F. Niedermann and M. S. Sloth, New early dark energy]. The NEDE component has the potential to resolve the tension between recent local measurements of the expansion rate of the Universe using supernovae (SN) data and the expansion rate inferred from the early Universe through measurements of the cosmic microwave background (CMB) when assuming Λ CDM. We discuss in depth the two-scalar field model of the NEDE phase transition including the process of bubble percolation, collision, and coalescence. We also estimate the gravitational wave signal produced during the collision phase and argue that it can be searched for using pulsar timing arrays. In a second step, we construct an effective cosmological model, which describes the phase transition as an instantaneous process, and derive the covariant equations that match perturbations across the transition surface. Fitting the cosmological model to CMB, baryonic acoustic oscillations, and SN data, we report $H_0 = 69.6_{-1.3}^{+1.0}$ km s⁻¹ Mpc⁻¹ (68% C.L.) without the local measurement of the Hubble parameter, bringing the tension down to 2.5σ . Including the local input, we find $H_0 = 71.4 \pm 1.0$ km s⁻¹ Mpc⁻¹ (68% C.L.) and strong evidence for a nonvanishing NEDE component with a $\simeq 4\sigma$ significance.

DOI: [10.1103/PhysRevD.102.063527](https://doi.org/10.1103/PhysRevD.102.063527)

I. INTRODUCTION

Over the past decades, our cosmological standard model has been put to the test with ever-increasing precision. Thanks to measurements of the cosmic microwave background (CMB), its parameters are now constrained at the percent level and provide us with a consistent description of our Universe starting shortly after the big bang and ending with a phase of accelerated expansion where length scales are stretched at an exponential rate. It is a remarkably successful model, because it resolves physics between extreme high and low curvature regimes with only a small handful of parameters. One of them is the Hubble parameter H_0 , denoting the present expansion rate of space. Recently, its measurement has given rise to a profound observational and theoretical crisis that is threatening the success of the Λ CDM model. In this paper (and a recent companion paper [1]), we propose a novel and simple resolution to the Hubble crisis in terms of a first-order phase transition that takes place in a dark sector before recombination.

A. The Hubble tension

The problem arises due to two competing sets of measurements that infer different values of H_0 . First, there are

early-time measurements. The most precise value comes from the European Space Agency satellite, *Planck*, that measures the temperature fluctuations, polarization, and lensing in the CMB radiation, with the newest data release finding [2] $H_0 = 67.36 \pm 0.54$ km s⁻¹ Mpc⁻¹.¹ This observation depends on the expansion history before *and* after recombination and, thus, is highly sensitive to our assumed cosmological model. For example, increasing H_0 while keeping all other Λ CDM parameters fixed will shift the whole expansion history to higher values of $H(z)$ and, therefore, affects not only the photon distribution in the primordial plasma but also the free-streaming of photons toward us after they have decoupled from the plasma. Another example of an early-time measurement is provided by observations of baryonic acoustic oscillations (BAO), which are sensitive to physics *before* recombination (as well as H_0 directly) as they depend on the sound horizon that has been imprinted in the visible matter distribution briefly before photons decouple.

On the other hand, there is a qualitatively different set of measurements that is operative at *late times* in a low-redshift patch around our position in space. These *local* observations are almost entirely insensitive to changes in the expansion history and, hence, our cosmological model

^{*}niedermann@cp3.sdu.dk
[†]sloth@cp3.sdu.dk

¹For this value, the *Planck* TT, TE, EE, and lensing dataset have been combined. Here and henceforth, if not stated otherwise, the uncertainties are at 68% C.L.

[3]. They use methods like the cosmic distance ladder to reach out to redshifts of about $z \simeq 0.4$. For example, the SH₀ES Collaboration [4] measured $H_0 = 74.03 \pm 1.42 \text{ km s}^{-1} \text{ Mpc}^{-1}$ by using type Ia supernovae (SN) calibrated through Cepheids as standard candles. Comparing both measurements yields a difference which is 4.4σ in significance. The H0LiCoW experiment, which uses strong-lensing time delays, finds [5] $H_0 = 73.3_{-1.8}^{+1.7} \text{ km s}^{-1} \text{ Mpc}^{-1}$, corresponding to a similar discrepancy. When combined with SH₀ES, it raises the tension with the early-time measurement above the 5σ level [6]. A special role is played by the analysis in Refs. [7,8]. Unlike SH₀ES, it uses the tip of the red giant branch method (rather than Cepheids) to calibrate the distance ladder and reports a value which is still compatible with both early- and late-time probes, $H_0 = 69.6 \pm 1.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (see, however, [9]). There are many other measurements which up to this point have larger uncertainties, though. For example, interferometric and spectroscopic observations of water megamasers in accretion disks of active galactic nuclei yield [10] $H_0 = 73.9 \pm 3 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and gravitational wave measurements give rise to [11] $H_0 = 68_{-7}^{+14} \text{ km s}^{-1} \text{ Mpc}^{-1}$. The fact, however, remains that there is a statistically significant tension between direct late-time and Λ CDM-inferred early-time measurements of H_0 , which is also known as the Hubble tension (for reviews, see [12,13]).

Of course, a frequent concern is that the local probes have systematic errors that are hard to control, as they rely on notoriously difficult distance measurements in our small- z neighborhood. Many potential sources for errors have been studied in the past. For example, in Ref. [14], it was found that redshift uncertainties cannot account for the discrepancy, and, in Refs. [15–17], a similar conclusion was drawn with regard to the local Hubble flow. In Refs. [18,19], the Cepheid calibration of the SH₀ES Collaboration was criticized; however, these corrections have been argued not to be sufficient to resolve the tension in later work [20,21]. In the meanwhile, the Cepheid discussion in the observational community continues with vigor [22–24]. The hope, of course, is that this controversy will be resolved once other local measurements, like the ones based on gravitational waves, will have reduced their uncertainties.

The upshot we can take home from this debate is that systematic errors might still have a say in the matter; however, as of yet, there is no clear indication that they could resolve the tension. We believe this makes it imperative from a theory standpoint to also think about alternative ways of addressing the problem. The crucial observation is that the CMB and BAO measurement have a strong dependence on our cosmological model through their dependence on acoustic oscillations in the photon-baryon plasma. This means that by changing the expansion history we can hope to be able to raise the BAO and CMB

inferred value up to the locally measured one, which itself remains unaffected by the change (see, for example, [13,25]).

This reasoning has already set in motion a lot of activity in the theory community, and different proposals can be divided in two classes. First, there are attempts to modify the background evolution at a late stage after recombination. This is usually done by modifying the expansion history at low redshift by introducing an *ad hoc* modification of $H(z)$, for example, parametrized in terms of a spline model [13,26,27]. However, so far, all of these attempts have been plagued by creating a tension with BAO measurements (or by introducing a lot of free parameters) [13,28,29]. The sound horizon r_s , which characterizes oscillations in the primordial plasma before recombination, is key to understanding this failure [12,30]: BAO measurements are anchored at early times through their dependence on r_s . Used in combination with the Pantheon dataset, they constrain the combination $r_s H_0$ and, therefore, require us to lower r_s if we want to increase H_0 in order to have agreement with the local measurement. However, if we use Λ CDM to describe early-time physics, lowering the sound horizon r_s has been found to be incompatible with raising H_0 , as this would lead to an unacceptable deformation of the CMB power spectrum [12]. In fact, the authors in Ref. [28] argued that any solution to the Hubble tension has most likely to be operative *before* recombination to avoid this BAO trap. This, of course, cannot exclude the possibility that late-time modifications play a partial role in resolving the tension. Also, as pointed out in Ref. [31], new low-redshift data might shed further light on the relevance of late-time approaches.

This brings us to the second direction which attempts to modify Λ CDM at early times. At first sight, this seems to be the least favorable option, as the CMB is a very clean probe. After all, changing the dynamics in the primordial plasma is expected to directly influence the CMB power spectrum, which is highly constrained by *Planck* data. So it is quite intriguing that a promising resolution of the Hubble tension modifies the Λ CDM model exactly where we would have thought it is constrained the most: closely to matter-radiation equality where the composition of the energy budget affects the relative and absolute heights of the first peaks in the power spectrum. This is done by introducing an *early dark energy* (EDE) component [12,32–40], which contributes a sizable fraction to the energy budget around matter-radiation equality, just before it starts to decay subsequently in order not to overclose the Universe.² So far, this has been realized in terms of a single scalar field that transitions from a slow-roll to an oscillating

²To be more precise, here we refer to specific axion setups that have been devised to address the Hubble tension. The general idea of having an early dark energy component is much older though and was, for example, studied in Refs. [41–44] and more recently in Ref. [45].

(or fast-roll) phase via a *second-order phase transition*. During the slow roll, the field energy (playing the role of EDE) mimics a cosmological constant, and, during the oscillating phase, it decays as a power law which is controlled by the shape of the field's *anharmonic* potential. This mechanism, therefore, employs the single-field approach to slow-roll inflation at much lower energies set by the eV rather than a high-energy scale close to grand unified theory physics. It is safe to say that until now EDE is one of the more promising ways to address the Hubble tension, as other early-time proposals have suffered from introducing new tensions not present in the Λ CDM model. One such alternative example is *dark radiation* (DR), which relies on introducing additional relativistic species—often modeled as sterile neutrinos—to lift H_0 to higher values and has long been regarded a consistent resolution [46–49]. However, increasingly precise CMB measurements have recently disfavored this possibility [2,50] (not completely ruled out, though).

The previous implementation of EDE is not free of problems either, as it requires shallow *anharmonic* potentials to be efficient in addressing the Hubble tension. While monomial potentials, ϕ^{2n} with $n \geq 2$, were claimed not to be sufficient, mainly due to a worsening of the fit to high- ℓ polarization data [51], specific terms of the nonperturbative axion potential, $\propto [1 - \cos(\phi/f)]^n$ with $n = 3$, yield a significantly better fit to data [32–34,52]. This, however, constitutes a tuning where leading terms in an instanton approximation have been set to zero [36]. It is unclear whether these potentials can be obtained from a UV theory, and, in particular, from the “string axiverse” as suggested by the authors, in a technically natural way. After all, the generic expectation in axion dark energy physics is to obtain a shallow *harmonic* potential [53–55].

B. New early dark energy

In this work, we discuss a novel solution to the Hubble tension, called *new early dark energy* (NEDE), which is based on a *first-order phase transition* that occurs shortly before recombination in a dark sector at zero temperature [1]. Effectively, it describes a jump in the vacuum energy from the eV down to the meV scale, in accordance with the spacetime curvature observed today. It builds up on the phenomenological success of the EDE proposal but uses different physics to do so. In fact, we think NEDE provides a more natural framework to resolve the Hubble tension while also providing new unique experimental signatures.

Both single-field EDE and NEDE share two defining properties, which are crucial for their phenomenological success. First, there is an additional energy component, not present in Λ CDM, which comes to contribute an $\sim 10\%$ fraction to the energy budget at some point t_* close to matter-radiation equality. Second, that component starts to decay at least as fast as radiation after t_* . How can this be realized in a first-order phase transition of NEDE at zero

temperature? The main challenge is to prevent the phase transition from happening too early, in which case the sound horizon and, hence, also H_0 would be largely unaffected. This, however, is not enough. We also need the phase transition to occur on a timescale $1/\bar{\beta}$ that is short compared to the Hubble expansion, i.e., $H/\bar{\beta} \ll 1$. This avoids the premature nucleation of bubbles of true vacuum that would grow too large before they collide with their smaller cousins. This would lead to large-scale anisotropies which would have imprinted themselves in the CMB. These challenges are overcome by NEDE within a *two-field* scalar model in a dark sector that features a built-in trigger mechanism. This is an adaption of an old idea that has been studied in the context of hybrid inflation at much higher energies [56,57].³ Here, the field starts its evolution deep inside a potential valley with an almost flat bottom, corresponding to an ultralight mass $m \sim 10^{-27}$ eV, and very steep walls of mass $M \sim$ eV. Initially, the field is stuck at its initial sub-Planckian field value due to the large Hubble friction, and its false vacuum energy provides the NEDE component. It is, however, protected from tunneling toward the true minimum of the potential by the high walls of the false vacuum valley. When $H \lesssim m$, the Hubble drag is being released and the field starts rolling down the valley. Now things move fast. Within less than a Hubble time, the valley opens up, effectively shrinking the barrier between the field and its true minimum, which then naturally triggers the phase transition (a sketch is provided in Fig. 2). This leads to the sudden nucleation of a large number of bubbles of true vacuum that start to expand and quickly approach the speed of light. As this process fills space very rapidly, bubbles start to collide before they reach cosmological scales. This condensate of colliding bubble walls is complicated to describe; however, at large observable scales, we expect it to be dominated by kinetic energy that redshifts away even quicker than radiation. This resonates with the observation that anisotropic stress, which comprises a sizable fraction of the colliding bubble condensate on short scales $\lesssim \bar{\beta}^{-1}$, is known to redshift like $1/a^6$ in the simpler case of a homogeneous system. Another part of the condensate is radiated away through gravitational waves and other model-dependent decay channels. In other words, the decay of NEDE at large scales is driven by the short-scale dynamics of colliding bubble walls and their microscopic decay. The corresponding model details, including a derivation of the percolation timescale $\bar{\beta}$ in terms of fundamental parameters as well as a first discussion of gravitational wave signatures, is provided in Sec. II.

Before we can test our model against data, we have to find an effective description valid on cosmological scales.

³In a nonzero temperature scenario, the phase transition could also be triggered due to the cooling of the Universe. For a corresponding late Universe scenario, see [58].

This is done in Sec. III in terms of an ideal fluid with a discontinuous equation of state parameter. It starts out in the very early Universe as -1 , describing the vacuum energy of our scalar field, and then at redshift z_* (~ 5000) it jumps upward to $w_{\text{NEDE}}^* > 1/3$, in accordance with a quickly decaying fluid. Again, this cosmological model is based on the assumption that the phase transition occurs on a short timescale $1/\bar{\beta} \ll 1/H$ and, therefore, can be described as an instantaneous process when solving cosmological equations. The trigger mechanism is implemented in terms of a single ultralight scalar field $\phi(t)$ with mass m . In terms of our microscopic two-field model, it parametrizes the almost flat direction along the bottom of the potential valley and promotes the tunneling probability to a function $\Gamma(\phi)$. The transition is then triggered when the ratio H/m , and, hence, ϕ/ϕ_{ini} , drops below a certain threshold value, which for a generic choice of fundamental model parameters falls in the range $0.18 < H_*/m \lesssim 0.21$, where the lower bound corresponds to the point of maximal tunneling probability and the upper bound makes sure that oscillations of the trigger field around the true vacuum are suppressed. In other words, for a given mass m , the effective parameter H_*/m (implicitly) determines the decay time t_* (or redshift z_* , equivalently). Things become more interesting at the perturbation level. Adiabatic perturbations $\delta\phi$ of the trigger field lead to spatial variations of the decay time. These, in turn, provide the initial conditions for the fluctuations of the decaying fluid after the transition. We use Israel’s junction conditions [59] to covariantly match perturbations across the transition surface. If the phase transition were to occur much earlier, we would not need to worry about this subtlety, because all observable modes would safely reside outside the horizon uniquely determined as adiabatic modes. However, in our case, modes that are observable in the CMB on short scales have already entered the particle horizon, and, as a result, we find that they are sensitive to the details of the phase transition, in particular, the ratio H_*/m .

In Sec. IV, we will finally perform a full cosmological parameter extraction. Our base model introduces two additional parameters (beyond the six Λ CDM parameters), the mass of the trigger field, $\log_{10}(m/m_0)$, and the fraction of NEDE at decay time, f_{NEDE} . The other parameters such as the “trigger parameter” H_*/m , the equation of state parameter of the NEDE fluid, w_{NEDE} , or its rest-frame sound speed c_s^2 , are fixed based on our theory expectations but allowed to vary when we investigate extensions of our base model subsequently. Our joint analysis includes CMB data from *Planck* 2018 (including temperature, polarization, and lensing data), BAO data, constraints on the growth of large-scale structure, the Pantheon dataset, and the locally determined value of H_0 . The latter is excluded from a subset of our runs in order to infer the evidence for NEDE without using “the statistical pull” of H_0 . Special care is taken to avoid sampling volume effects that are

known to lead to unphysical artifacts within EDE-type models when not properly accounted for [34,35]. We also provide a detailed comparison of NEDE with its early-time competitors dark radiation (DR), acoustic dark energy (ADE) [35], and single-field early dark energy (EDE) [33]. On a technical level, we implement each model in a Boltzmann code and use a *Monte Carlo Markov chain* (MCMC) algorithm to perform the parameter extraction.

C. Summary of results

Here we list our most important results.⁴

- (i) There is a consistent and radiatively stable choice of microscopic parameters that leads to a very quick first-order phase transition with percolation time-scale $\bar{\beta}^{-1} \ll H^{-1}$ around $z = 5000$. For a radiatively stable choice of coupling $\tilde{\lambda}$ between the two scalar fields, the two-field tunneling problem can be effectively reduced to a one-field tunneling problem, making our scenario computationally accessible.
- (ii) Taking the limit $\bar{\beta}^{-1} \rightarrow H^{-1}$ renders the expected gravitational wave signal produced by the phase transition marginally compatible with the claimed peak sensitivity of the *Square Kilometre Array* [61–63] for frequencies at around $f = 10^{-9}$ Hz. At the same time, this limit leads to the prediction of distinct anisotropic structures in the CMB arising from the colliding bubble wall condensate.
- (iii) We worked out the full covariant set of cosmological matching equations across a generic spacelike hypersurface valid for sub- and superhorizon modes. We then used it to study both curvature and isocurvature modes within NEDE and identified the region in parameter space where the latter are sufficiently suppressed to be neglected.
- (iv) The base model MCMC analysis without the local measurement of H_0 results in $f_{\text{NEDE}} = 7.7^{+3.8}_{-4.0}\%$ corresponding to a 1.9σ evidence for NEDE, provided we account for volume sampling effects in the limit $f_{\text{NEDE}} \rightarrow 0$. Moreover, the model introduces an approximate degeneracy in the $r_s - H_0$ plane satisfying $H_0 r_s \simeq \text{const}$ as required by BAO measurements. Correspondingly, the uncertainties in the determination of $H_0 = 69.6^{+1.0}_{-1.3} \text{ km s}^{-1} \text{ Mpc}^{-1}$ become relatively large, bringing the tension down to 2.5σ (from 4.3σ within Λ CDM when the same dataset combination is chosen).
- (v) When the local measurement is included, we report for our base model an increased Hubble parameter of $H_0 = 71.4 \pm 1.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and evidence for NEDE at 4.3σ [$f_{\text{NEDE}} = 12.6^{+3.2}_{-2.9}\%$]. The overall fit

⁴We work units in which $\hbar = c = 1$. For our cosmological perturbation theory, we widely adopt the notation and conventions of Ref. [60].

improves by $\Delta\chi^2(\text{total}) = -15.6$. At the same time, the so-called S_8 tension increases slightly to the 2.7σ level but not by much compared to ΛCDM for which we find 1.9σ (with SH_0ES) and 2.3σ (without SH_0ES).

- (vi) Different extensions of our base model confirm the expectations arising from our microscopic description. In particular, we obtain $H_*/m = 0.206^{+0.013}_{-0.022}$ in good agreement with the theoretically viable range, as well as $w_{\text{NEDE}}^* = 0.70^{+0.12}_{-0.16}$, consistent with a fluid dominated by kinetic energy and small-scale anisotropic stress. On the other hand, we find no evidence for large-scale anisotropies.
- (vii) We compare NEDE with single-field EDE. While we find that both models share their phenomenological success, they use different physics to do so. Most notably, NEDE relies on a k -independent rest-frame sound speed and is compatible with a wider range of initial field values $\phi_{\text{ini}}/M_{\text{pl}} > 10^{-4}$. Moreover, its potential adheres to the usual low-energy rules and does not require a fine-tuning.

II. NEW EARLY DARK ENERGY

In the following, we will outline a field theoretical model of NEDE. To that end, we will briefly review how a first-order phase transition depends on the shape of a generic quartic potential. In a second step, we consider a generalized two-field model with a clock field that triggers the transition at a given time t_* in the early expansion history somewhat close to recombination.

A. One-field vacuum decay

Here we briefly review the vacuum decay of a single scalar field at zero temperature. We are particularly interested in how the decay rate depends on the shape of the potential, which will become crucial in devising a dynamical two-field model in the next section. We will broadly follow the discussion in Ref. [64].

We consider a scalar field ψ with general renormalizable potential $V(\psi)$:

$$V(\psi) = \frac{\lambda}{4}\psi^4 - \frac{1}{3}\alpha M\psi^3 + \frac{1}{2}\beta M^2\psi^2 + \gamma M^3\psi + d, \quad (1)$$

where M is a generic mass scale and α , β , and γ are dimensionless constants. We assume the potential to be bounded from below, implying $\lambda > 0$. The constant d corresponds to an irrelevant shift that can be absorbed in a cosmological constant term. Furthermore, we can perform a field translation to set $\gamma = 0$ and switch the sign of ψ if $\alpha < 0$. We can therefore make the choice $d = \gamma = 0$ and $\alpha > 0$ without any loss of generality. After rewriting the potential in terms of the dimensionless field variable $\bar{\psi} = 3\lambda\psi/(\alpha M)$, we have

$$\bar{V}(\bar{\psi}) = \frac{1}{4}\bar{\psi}^4 - \bar{\psi}^3 + \frac{\delta}{2}\bar{\psi}^2, \quad (2)$$

where we introduced the dimensionless potential $\bar{V} = 81\lambda^3 V/(\alpha^4 M^4)$ as well as

$$\delta = 9\frac{\lambda\beta}{\alpha^2}. \quad (3)$$

So we see that the properties of the potential are conveniently captured by the dimensionless parameter δ . For $0 < \delta < 9/4$, the potential exhibits two minima at

$$\bar{\psi}_1 = 0 \quad \text{and} \quad \bar{\psi}_2 = \frac{1}{2}(3 + \sqrt{9 - 4\delta}), \quad (4)$$

as well as a maximum at $\bar{\psi}_{\text{max}} = (3 - \sqrt{9 - 4\delta})/2$. We now assume $\bar{\psi} = 0$ initially, which for $0 < \delta < 9/4$ corresponds to the false vacuum. There are two interesting limiting cases.

1. Thin-wall limit

For $\delta = 2 - \epsilon$ and $\epsilon \ll 1$, both minima are degenerate, corresponding to the thin-wall limit (see solid, blue line in Fig. 1). Specifically, the true vacuum sits at

$$\bar{\psi}_2 \simeq 2, \quad (5a)$$

and the difference in energies is

$$\bar{V}(\bar{\psi}_1) - \bar{V}(\bar{\psi}_2) \simeq 2\epsilon. \quad (5b)$$

The barrier height becomes maximal with $\bar{V}(\bar{\psi}_{\text{max}}) - \bar{V}(\bar{\psi}_1) \simeq 1/4$.

2. Deep-well limit

For $\delta = \epsilon \ll 1$, on the other hand, the height of the barrier becomes tiny, $\bar{V}(\bar{\psi}_{\text{max}}) - \bar{V}(\bar{\psi}_1) \simeq \epsilon^3/54$, and the true vacuum at

$$\bar{\psi}_2 \simeq 3 \quad (6a)$$

is deepest (see dashed, green line in Fig. 1),

$$\bar{V}(\bar{\psi}_1) - \bar{V}(\bar{\psi}_2) \simeq \frac{27}{4}. \quad (6b)$$

We will see that the case most interesting for our application lies between these extremes with $0 < \delta \lesssim 1.5$.

⁵The case $\delta \leq 0$ (or $\beta \leq 0$, equivalently), for which the maximum is at $\bar{\psi} = 0$, is more akin to scenarios of hybrid inflation where the field undergoes a second-order phase transition. We provide a preliminary discussion of such a scenario in Sec. III C 3.

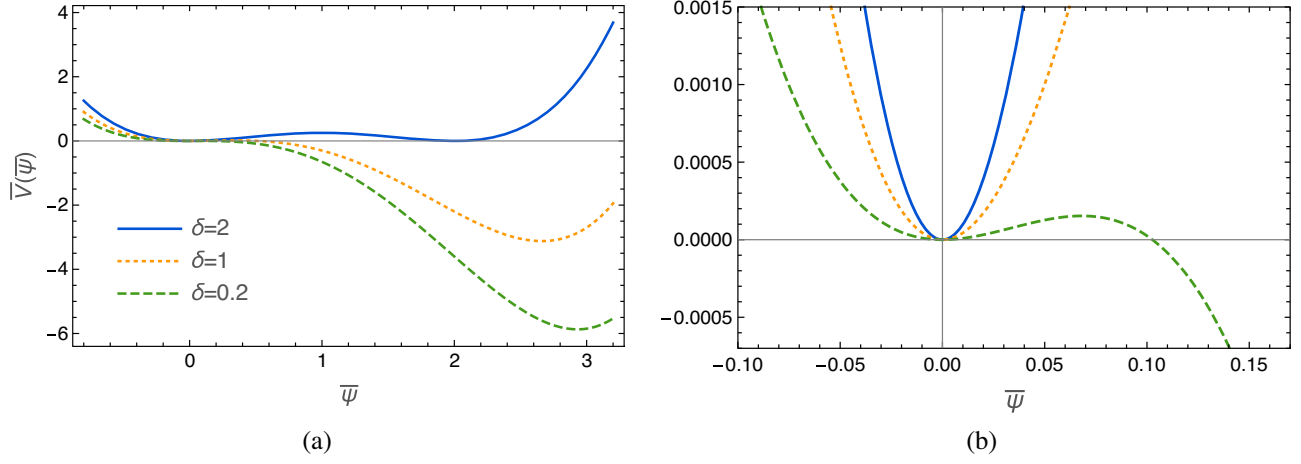


FIG. 1. Dimensionless potential \bar{V} for different choices of δ . The vacua become degenerate as $\delta \rightarrow 2$. For $\delta \rightarrow 0$, the barrier height vanishes and the true vacuum is deepest. (a) Both vacua: $\bar{\psi}_1$ (left) and $\bar{\psi}_2$ (right). (b) Zoomed in on vacuum $\bar{\psi}_1 = 0$.

In order to infer the corresponding tunneling probabilities, we need the Euclidian action evaluated at the $O(4)$ symmetric “bounce solutions” $\bar{\psi}_\delta(r)$, where r is a four-dimensional radial coordinate in Euclidian space. These configurations are defined as stationary points of the Euclidian action

$$\begin{aligned} S_E &= 2\pi^2 \int_0^\infty dr r^3 \left[\frac{1}{2} \left(\frac{d\psi}{dr} \right)^2 + V(\psi) \right] \\ &= \frac{2\pi^2}{\lambda} \int_0^\infty d\bar{r} \bar{r}^3 \left[\frac{1}{2} \left(\frac{d\bar{\psi}}{d\bar{r}} \right)^2 + \bar{V}(\bar{\psi}) \right], \end{aligned} \quad (7)$$

where $\bar{r} = \alpha M r / (3\sqrt{\lambda})$. Correspondingly, they fulfill the Euclidian equations of motion

$$\bar{\psi}''(r) + \frac{3}{r} \bar{\psi}'(r) = \bar{\psi}^3 - 3\bar{\psi}^2 + \delta \bar{\psi} \quad (8)$$

and are subject to the boundary conditions $d\bar{\psi}_\delta/d\bar{r} = 0$ at $\bar{r} = 0$ (to ensure regularity at the origin) and $\bar{\psi}_\delta = 0$ as $\bar{r} \rightarrow \infty$ (for the false vacuum to lie outside the bubble). A fully analytic solution exists only in the thin-wall limit [65]. In general, a numerical shooting method needs to be applied. For the model in Eq. (2), a sufficiently accurate approximation for S_E was derived in Ref. [64], yielding⁶

$$S_E|_{\bar{\psi}=\bar{\psi}_\delta} \simeq \frac{4\pi^2}{3\lambda} (2-\delta)^{-3} (\alpha_1 \delta + \alpha_2 \delta^2 + \alpha_3 \delta^3), \quad (9)$$

where

⁶We checked the result with our own implementation of the shooting method. Moreover, in the limit $\delta \ll 1$, we derive $\lambda S_E \simeq 91/4\delta$, which agrees with the result in Ref. [66] when we identify $\delta_{\text{there}}^2 = 9\lambda M^2/\delta$. And in the opposite limit, it agrees with the analytical result obtained from the thin-wall approximation.

$$\alpha_1 = 13.832, \quad \alpha_2 = -10.819, \quad \alpha_3 = 2.0765. \quad (10)$$

As a result, S_E diverges for $\delta \rightarrow 2$ (thin-wall limit) and vanishes linearly as $\delta \rightarrow 0$ (deep-well limit). For moderate values in the range $\delta \in [0.01, 0.5]$, we find $\lambda S_E \in [0.2, 17.4]$. The tunneling rate per unit volume is then defined as

$$\Gamma = K \exp(-S_E). \quad (11)$$

In Refs. [66,67], it was argued that the determinant factor $K \sim M^4$, provided the dimensionless parameters in Eq. (1) do not introduce large hierarchies. This formula is applicable as long as $S_E/\hbar \gg 1$ for the underlying semiclassical calculation to be valid. This can be satisfied for a sufficiently weak coupling, at least $\lambda < 0.1$, provided δ is not tuned too small. As a numerical example, for $\lambda = 0.01$ and $\delta = 0.1$, we obtain $S_E \simeq 245 \gg 1$. In short, the tunneling rate vanishes in the thin-wall limit and monotonically increases for decreasing δ .

B. Triggered vacuum decay

If the nucleation rate is constant, we will encounter an analogy of the “big bubble” problem, which was discussed already in the context of old inflation [56]. Even if a constant nucleation rate is initially small compared to the growing Hubble rate, there will be a small but nonvanishing probability to form bubbles early in the expansion history of the Universe, and, since these bubbles will expand with the speed of light and rapidly become very big, the true vacuum volume will be dominated by a small number of very large bubbles [56]. This is problematic for two reasons: First, it will lead to unacceptably large inhomogeneities in the Universe. Second, the transition would happen too slow to have the desired impact on the sound horizon. To solve this problem, one needs to make sure that the bubble nucleation is turned on quickly. This can be

achieved by introducing an additional field, which acts as a trigger for the bubble nucleation.

The aim of this section is, therefore, to find a field theoretical model that promotes δ to a function of another field ϕ , i.e., $\delta \rightarrow \delta_{\text{eff}}(\phi)$. By controlling the slow-roll dynamics of ϕ , we can scan over different values of δ_{eff} and switch on the phase transition in ψ (and ϕ) at any given time. The difference between the false and the true vacuum energy stored in the ψ field then provides most of the NEDE component (with only a subdominant contribution arising from ϕ). As NEDE is stable at early times, we must construct a model that initially guarantees $\delta_{\text{eff}} \gg 2$ in order to sufficiently suppress bubble nucleation. At a later stage, however, we want NEDE to start its decay. This means that we look for a dynamical mechanism that sends $\delta_{\text{eff}} \rightarrow \delta_{\text{eff}}^* < 2$ and, in turn, switches on the nucleation of bubbles of true vacuum. This is then followed by a coalescence phase, where the colliding bubble wall condensate redshifts away while also decaying to radiation. This idea has been studied in the context of hybrid inflation in Refs. [56,57], where the authors employed a field theoretical trigger mechanisms to end inflation by tunneling (and overcome the “graceful exit” problem in earlier first-order single-field constructions). We will apply a similar setup as a simple way of realizing our NEDE scenario. To that end, we generalize the single-field model to

$$V(\psi, \phi) = \frac{\lambda}{4}\psi^4 + \frac{1}{2}\beta M^2\psi^2 - \frac{1}{3}\alpha M\psi^3 + \frac{1}{2}m^2\phi^2 + \frac{1}{2}\tilde{\lambda}\phi^2\psi^2, \quad (12)$$

where again everything is expressed in terms of a single mass scale M and positive, dimensionless parameters λ , $\tilde{\lambda}$, α , and β . The kinetic terms for ψ and ϕ are canonically normalized. This potential can be recovered from the previous one in Eq. (1) if we set $\gamma = 0 = d$, replace β by a function of ϕ ,

$$\beta \rightarrow \beta + \tilde{\lambda} \frac{\phi^2}{M^2}, \quad (13)$$

and add a mass term for ϕ .

The general mechanism is simple: Initially, $\psi = 0$ and ϕ is frozen at $\phi \simeq \phi_{\text{ini}}$. Once the Hubble drag is released, i.e., when $H \lesssim m$, ϕ rolls down the potential. Now, if we require

$$\alpha^2 > 4\beta\lambda, \quad (14)$$

or $\delta < 9/4$, equivalently, a second (global) minimum exists. As a result, when ϕ drops below a certain threshold ϕ_* on its way to $\phi = 0$, the field configurations with $\psi = 0$ (representing a valley in the two-dimensional potential) become unstable against quantum tunneling, and NEDE starts to decay. For a schematic plot of the potential, see Fig. 2.

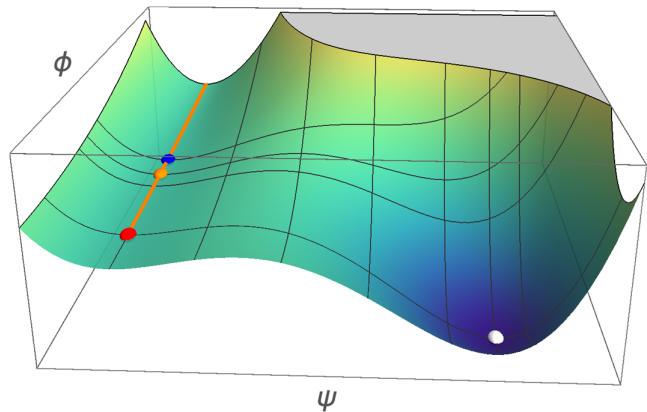


FIG. 2. Illustration of the potential in (12). At early times, the field is frozen high up the false vacuum valley (above the blue dot) and protected against tunneling to the true vacuum (white dot) by a high potential barrier. Tunneling is turned on when the field rolls past the orange dot. Shortly after that, bubble percolation overcomes the expansion of space and reaches its maximal efficiency at the red dot.

We are interested in the tunneling rate as a function of the rolling field $\phi(t)$ and, hence, cosmological time t . To that end, we derive a semianalytic expression for the Euclidian action in the limit where the evolution of ϕ is sufficiently slow and $\tilde{\lambda}/\lambda \ll 1$. As before, it is useful to introduce a dimensionless potential

$$\bar{V}(\bar{\psi}, \bar{\phi}) = \frac{1}{4}\bar{\psi}^4 - \bar{\psi}^3 + \frac{\delta_{\text{eff}}(\bar{\phi})}{2}\bar{\psi}^2 + \frac{1}{2}\kappa^2\bar{\phi}^2, \quad (15a)$$

where we used the dimensionless variables

$$\bar{V} = \frac{81\lambda^3}{\alpha^4 M^4} V, \quad \bar{\psi} = \frac{3\lambda}{\alpha M} \psi, \quad \bar{\phi} = \frac{3\sqrt{\lambda\tilde{\lambda}}}{\alpha M} \phi \quad (15b)$$

and introduced

$$\kappa = \frac{3\lambda}{\alpha\sqrt{\tilde{\lambda}}} \frac{m}{M}. \quad (15c)$$

As advertised, the parameter δ in (3) is promoted to a function of $\bar{\phi}$,⁷

$$\delta_{\text{eff}}(\bar{\phi}) = \delta + \bar{\phi}^2, \quad (16)$$

controlling the shape of the potential for $\bar{\psi}$. In particular, it shows that for $\delta < 0.2$ we cycle through the different configurations depicted in Fig. 1 when $\bar{\phi} \rightarrow 0$.

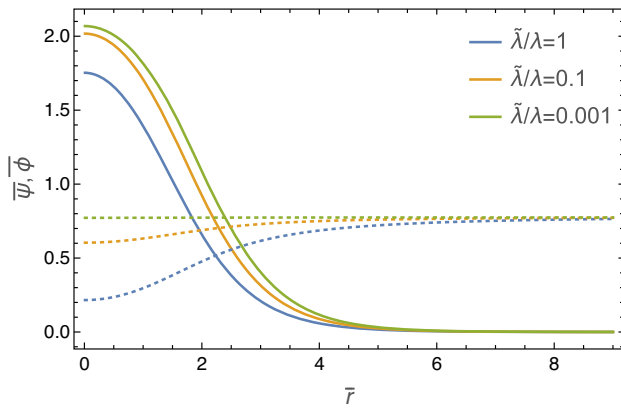
⁷We assume as before $\beta > 0$, which ensures the positivity of δ and, hence, δ_{eff} . It would also be interesting to study the case $\beta < 0$, which for a suitable choice of parameters gives rise to a second-order phase transition like in hybrid inflation (see, for example, [56]). As far as we know, this has not yet been studied in the context of EDE proposals.

The problem of calculating the tunneling probability is rather complicated in the two-field case. First, finding the bounce solution requires us to solve the coupled system of Euclidian equations of motion for ϕ and ψ . Second, due to the rolling of ϕ , the boundary condition in Euclidian time is not $O(4)$ symmetric (or, equivalently, the initial configuration before the bubble nucleation is not Lorentz invariant in real time). This means that the bounce solutions, in general, depend on t_E and the three-dimensional radial coordinate $\mathbf{x}^2 = \delta_{ij}x^i x^j$ explicitly [rather than the $O(4)$ symmetric combination $r = \sqrt{t_E^2 + \mathbf{x}^2}$]. We would therefore need to solve a coupled system of partial (rather than ordinary) differential equations to infer the evolution of the bubbles after their nucleation. This difficulty can, however, be avoided if we assume $\kappa \ll 1$ (we will check the validity of this assumption *a posteriori*). In that case, we can neglect the last term in Eq. (15a), and the corresponding Lorentz-invariant vacuum solution becomes $\bar{\psi} = 0$ and $\bar{\phi} \simeq \text{const}$. This limit corresponds to the case where the field ϕ evolves very slowly on scales $1/M$ (relevant for the bounce solution).

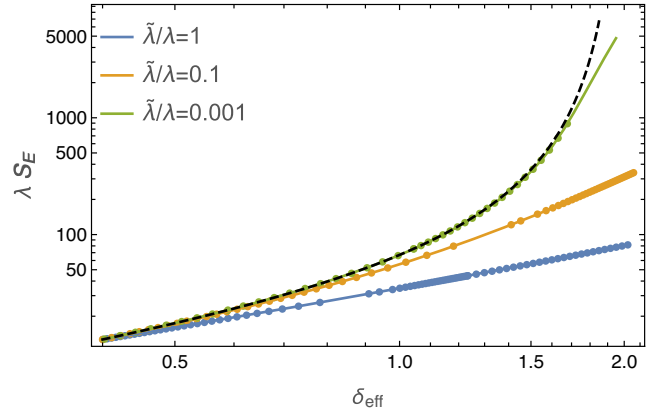
The Euclidian action for $O(4)$ -symmetric configurations then reads

$$S_E = \frac{2\pi^2}{\lambda} \int_0^\infty d\bar{r} \bar{r}^3 \left[\frac{1}{2} \left(\frac{d\bar{\psi}}{d\bar{r}} \right)^2 + \frac{1}{2} \frac{\lambda}{\tilde{\lambda}} \left(\frac{d\bar{\phi}}{d\bar{r}} \right)^2 + \bar{V}(\bar{\psi}) \right], \quad (17)$$

leading to the equations of motion



(a)



(b)

FIG. 3. Result of the two-field shooting method. Each dot corresponds to one numerical integration. Different colors correspond to different choices of $\tilde{\lambda}/\lambda$. (a) Radial profiles for $\bar{\psi}$ (solid lines) and $\bar{\phi}$ (dotted lines) for $\delta_{\text{eff}} = 1$. As $\tilde{\lambda}/\lambda \rightarrow 0$ the profile of $\bar{\phi}$ approaches a constant. (b) Euclidian action as a function of δ_{eff} for $\delta = 0.4$. The dashed line plots the semi-analytic result (9) derived in the one-field case. We find that it is approached as $\tilde{\lambda}/\lambda \rightarrow 0$.

$$\bar{\psi}'' + \frac{3}{\bar{r}} \bar{\psi}' = \bar{\psi}^3 - 3\bar{\psi}^2 + \delta_{\text{eff}}(\bar{\phi}) \bar{\psi}, \quad (18a)$$

$$\bar{\phi}'' + \frac{3}{\bar{r}} \bar{\phi}' = \frac{\tilde{\lambda}}{\lambda} [\bar{\psi}^2 \bar{\phi} + \mathcal{O}(\kappa^2)]. \quad (18b)$$

They are subject to the boundary conditions

$$\lim_{\bar{r} \rightarrow \infty} \bar{\psi} = 0, \quad \lim_{\bar{r} \rightarrow \infty} \bar{\phi} = \bar{\phi}(t_*), \quad (19a)$$

$$\lim_{\bar{r} \rightarrow 0} \bar{\psi}' = 0, \quad \lim_{\bar{r} \rightarrow 0} \bar{\phi}' = 0, \quad (19b)$$

where $\bar{\phi}(t_*)$ is the value of $\bar{\phi}$ at the time of tunneling. The important observation is that in the limit $\tilde{\lambda}/\lambda \ll 1$ both equations decouple and (18b) is trivially solved by $\bar{\phi}(\bar{r}) = \bar{\phi}(t_*) = \text{const}$. In that case, Eq. (18a) becomes formally equivalent to Eq. (8). As a result, the Euclidian action is simply given by its one-field expression (9) subject to the replacement $\delta \rightarrow \delta_{\text{eff}}(\bar{\phi})$:

$$S_E \simeq \frac{4\pi^2}{3\lambda} (2 - \delta_{\text{eff}})^{-3} (\alpha_1 \delta_{\text{eff}} + \alpha_2 \delta_{\text{eff}}^2 + \alpha_3 \delta_{\text{eff}}^3). \quad (20)$$

The time dependence of $\bar{\phi}(t)$ (relevant on cosmological timescales) then determines the time dependence of the tunneling rate $\Gamma(t)$, given by (11).

We also studied the system away from the limit $\tilde{\lambda}/\lambda \ll 1$ by solving the coupled system of equations numerically, yet still assuming $\kappa \ll 1$. The result is depicted in Fig. 3 for different values of $\tilde{\lambda}/\lambda$ and $\delta = 0.4$. We see in Fig. 3(a) that $\bar{\phi}$ indeed approaches a constant profile as $\tilde{\lambda}/\lambda \rightarrow 0$. Moreover, Fig. 3(b) shows that the result in Eq. (9) is

applicable at least up to values $\lambda S_E \simeq 1000$ provided $\tilde{\lambda}/\lambda < 0.001$. This is enough to study our tunneling mechanism, which, as we will see, is triggered only when S_E drops below ~ 250 and bubble percolation becomes more efficient than the expansion of space. For larger values, on the other hand, it is extremely unlikely that a bubble is either nucleated within our Hubble volume or enters from outside the horizon in case it has been nucleated in a different patch before. In addition, we also find that our trigger mechanism would still work if $\tilde{\lambda} \sim \lambda$. In particular, percolation would cease for sufficiently large values of δ_{eff} . Unfortunately, we could not find an analytic expression for S_E in this more general case, but later we will anyway find that it is disfavored by naturalness considerations.

Finally, for $\delta_{\text{eff}} > 2$, all three curves in Fig. 3(b) asymptote to a power law. This is different from the one-field case (corresponding to the dashed line), where S_E diverges for $\delta \rightarrow 2$. The reason is that there is always a nonvanishing probability that the field tunnels along a diagonal direction in the $\phi - \psi$ plane, albeit highly suppressed.

C. Bubble percolation

We can now use the expression in Eq. (20) to study the percolation phase within our model. In particular, we will derive the range of model parameters for which percolation happens quickly on cosmological timescales to avoid phenomenological problems with bubbles growing to cosmological scales.

Before we do so, we derive estimates for the two mass scales M and m controlling the ψ and ϕ sector, respectively. The scale M is fixed by the requirement that NEDE gives an $\sim 10\%$ contribution around the time of the phase transition. Otherwise, its effect on the expansion history would be negligible. To that end, we introduce the NEDE fraction

$$f_{\text{NEDE}} \equiv \frac{\bar{\rho}_{\text{NEDE}}^*}{\bar{\rho}_*} = \frac{V(\psi_*, \phi_*) - V(\psi_2, 0)}{\bar{\rho}_*}, \quad (21)$$

where $\bar{\rho}$ is the total background energy density and we normalized $\bar{\rho}_{\text{NEDE}}^* = V(\psi_*, \phi_*) - V(\psi_2, 0)$ with respect to the true vacuum at $\psi = \psi_2$ and $\phi = 0$. Here and henceforth, we use an asterisk as a shorthand for evaluation at percolation time t_* , i.e., $A_* = A^* \equiv A(t_*)$. We can use the results of the last section to further evaluate (21):

$$\bar{\rho}_{\text{NEDE}}^* = \frac{c_\delta \alpha^4 M^4}{12 \lambda^3} + \frac{1}{2} m^2 \phi_*^2, \quad (22)$$

where

$$c_\delta = \frac{1}{216} (3 + \sqrt{9 - 4\delta})^2 (3 - 2\delta + \sqrt{9 - 4\delta}), \quad (23)$$

evaluating to $c_\delta \in [1, 0.41]$ for $\delta_{\text{eff}} \in [0, 1]$. In addition, we require the trigger field ϕ to be subdominant before the phase transition. A sufficient condition reads

$$m^2 \phi_*^2 < m^2 \phi_{\text{ini}}^2 \ll \frac{c_\delta \alpha^4 M^4}{6 \lambda^3}. \quad (24)$$

We then obtain from (22)

$$M^4 \simeq \frac{12 \lambda^3 f_{\text{NEDE}}}{c_\delta \alpha^4} \bar{\rho}_*. \quad (25)$$

Provided α is not hierarchically small, we see that the mass scale M is bounded by the energy scale of the cosmic fluid at decay time t_* . Assuming that we are still in the radiation-dominated epoch when the transition takes place, we have $\bar{\rho}_* \simeq \bar{\rho}_{\text{rad}}^*/(1 - f_{\text{NEDE}})$, where the radiation density is

$$\rho_{\text{rad}}^* = \frac{\pi^2}{30} g_*(T) (1 + z_*)^4 \left(\frac{g_s(T_0)}{g_s(T_*)} \right)^{4/3} T_0^4. \quad (26)$$

Here, $g_*(T)$ and $g_s(T)$ are the total number of relativistic and the effective number of entropy carrying degrees of freedom at temperature T , respectively. As our transition occurs rather late in the early Universe, we have $g_s(T_*) = g_s(T_0)$, where T_0 is the CMB temperature today. Plugging in numerical values, we find

$$M^4 \simeq (0.4 \text{ eV})^4 \frac{1}{c_\delta} \left(\frac{\lambda^3 \alpha^{-4}}{0.01} \right) \left(\frac{f_{\text{NEDE}}/(1 - f_{\text{NEDE}})}{0.1} \right) \times \left(\frac{g_*(T_*)}{3.9} \right) \left(\frac{z_*}{5000} \right)^4. \quad (27)$$

We will use the estimate $M \sim \text{eV}$ throughout the remainder of this work.

At late times, after bubble percolation has been completed, we have to make sure that the remaining constant energy density stored in the ψ vacuum is not overclosing the Universe, explicitly

$$\frac{c_\delta \alpha^4}{12 \lambda^3} M^4 + \rho_\lambda \lesssim (\text{meV})^4. \quad (28)$$

Here, we introduced the cosmological constant contribution to the energy density, $\rho_\lambda = M_{\text{pl}}^2 \Lambda$. This relation can always be satisfied by tuning the value of Λ , which is nothing else but the usual cosmological constant problem (see, for example, [68]).

Next, we turn to the mass of the trigger field ϕ , which is frozen at early times at $\phi \simeq \phi_{\text{ini}}$. From (24) and (25), we derive the upper bound

$$\frac{\phi_{\text{ini}}}{M_{\text{pl}}} \ll \sqrt{6f_{\text{NEDE}}} \frac{H_*}{m} \lesssim 1, \quad (29)$$

which is sufficient to ensure that the contribution of ϕ to the total energy density is subdominant and smaller than the one arising from ψ (at least during its slow-roll regime before the phase transition). It also shows that ϕ is always sub-Planckian. Once $m \sim H_*$, slow roll ends and $\phi \rightarrow 0$ for the first time, which, in turn, triggers the phase transition as $\delta_{\text{eff}} \rightarrow \delta_{\text{eff}}^*$. Using $\rho_* = 3M_{\text{pl}}^2 H_*^2$ together with (25) and (27), this yields an expression for the mass:

$$m = 1.8 \times 10^{-27} \text{ eV} (1 - f_{\text{NEDE}})^{-1/2} \left(\frac{g_*(T_*)}{3.9} \right)^{1/2} \times \left(\frac{z_*}{5000} \right)^2 \left(\frac{0.2}{H_*/m} \right), \quad (30)$$

where we introduced the ‘‘trigger parameter’’ H_*/m . Later we will show that for a typical parameter choice the decay is triggered at $H_*/m \simeq 0.2$, close to the zero crossing of ϕ . As an example, for the transition to take place at redshift $z \sim 5000$, we need an extremely tiny scale of the order of $m \sim 10^{-27} \text{ eV} \ll M \sim \text{eV}$. We will see that this enormous hierarchy between m and M is key in making the bubble percolation efficient (and avoiding potential problems with large inhomogeneities arising from colliding bubble wall dynamics), and in Sec. II E we demonstrate that the hierarchy is technically natural.

We quantify the efficiency of the bubble nucleation in terms of the percolation parameter $p = \Gamma/H^4$. Provided $p > 1$, at least one bubble can be expected to be nucleated within one Hubble patch and Hubble time. To make the phase transition an instantaneous event on cosmological timescales and avoid phenomenological problems with large bubbles, we impose the stronger condition $p \gg 1$ during bubble percolation. On the other hand, if $p \ll 1$, the percolation cannot keep up with the expansion of space and a typical Hubble patch does not contain any bubble. This is the condition we want to realize before the transition.

In order to express p in terms of our model parameters, we need an estimate for the determinant factor K introduced in Eq. (11), which is notoriously difficult to calculate. However, as we are interested only in the magnitude of p and our ψ potential is controlled by a single scale M —recall that the dimensionless parameters α , β , and λ can be taken to be $\mathcal{O}(0.1-1)$ and the ψ - ϕ interaction term is also of the order of $\sim M$ when nucleation becomes possible due to the vacuum expectation value (VEV) of ϕ —we can use dimensional analysis to identify $K \sim M^4 (\gg H_*^4)$. In other words, bubble nucleation in this regime is a one-scale problem; a corresponding argument is, for example, provided in Refs. [66,67]. In any case, we have

$$p \sim \frac{M^4}{H_*^4} \exp(-S_E) \sim \exp(250 - S_E), \quad (31)$$

where we assumed $m \sim H_* \sim 10^{-27} \text{ eV}$ (corresponding to a phase transition around $z \sim 10^4$) and $M \sim \text{eV}$. We distinguish two phases.

1. Quasistable phase ($p \ll 1$).

Bubble nucleation is irrelevant. This is fulfilled throughout the slow-roll phase when $H \gg m$ and ϕ is frozen close to its initial value ϕ_{ini} ; it still holds when $m \sim H$ and ϕ starts to roll toward zero, provided $S_E > S_E^* \simeq 250$ [corresponding to an exponential suppression of the right side of (31)]. The discussion of the last section shows that this is guaranteed if $\delta_{\text{eff}} \geq 2$ (irrespective of the value of $\tilde{\lambda}$). Because of (16), a sufficient condition for the stability of the false vacuum at early times then reads

$$\frac{\lambda \tilde{\lambda} \phi_{\text{ini}}^2}{\alpha^2 M^2} > \frac{2}{9}. \quad (32)$$

Note that this condition is also sufficiently suppressing bubble nucleation at the earliest possible times when $H \sim M$. *A priori*, these very old bubbles are problematic, because they would have had a lot of time to grow and could have entered our horizon before the time t_* , defying the whole idea of NEDE and leading to large-scale anisotropies. To avoid this problem, we simply demand that there is a sufficiently suppressed probability that a single bubble would have been nucleated into *any* of the Hubble patches that will have entered our particle horizon by the time t_* . This requires that the percolation parameter at that early time is $p|_{H \sim M} < 10^{-56}$, which translates to $S_E \gtrsim 130$, which is indeed a weaker condition than the one derived before.

2. Percolation phase ($p \geq 1$).

Bubbles of true vacuum are being nucleated. This happens for our numerical example when the condition $S_E < 250$ is met. The corresponding percolation time t_* is then implicitly determined through $p(t_*) \simeq 1$, which also defines the critical value $\delta_{\text{eff}}^* = \delta_{\text{eff}}(t_*) (< 1.5)$. This phase should be short on a scale $1/H_*$ to prevent bubbles from growing to cosmological size. In other words, p quickly increases to values $p \gg p(t_*) \simeq 1$, describing a strong burst of nucleation events. For example, for $S_E \rightarrow 10$ as $\phi \rightarrow 0$, the percolation rate could grow as large as $p \sim 10^{104}$, which is an extremely high rate.⁸ However, we will see that in our case the percolation is completed before that maximal value is reached.

⁸This huge percolation parameter sets our proposal apart from similar scenarios in an inflationary context where the scales M and $H_* \simeq m$ are less clearly separated [56,57].

In order to obtain a more quantitative picture, we first derive the time evolution of ϕ . Since its dynamical equation

$$\ddot{\phi} + 3H\dot{\phi} + m^2\phi = 0 \quad (33)$$

is controlled by a single timescale, $1/H \sim 1/m$, its first zero crossing takes place roughly within one Hubble time after it has dropped out of slow roll. In fact, (33) can be solved analytically for a radiation-dominated universe in terms of Bessel functions:

$$\phi(x) = \sqrt{2}\Gamma(5/4)\phi_{\text{ini}}x^{-1/4}J_{1/4}(x/2), \quad (34)$$

where we introduced $x = m/H$ and Γ is the Gamma function. Here, we have already chosen the attractor branch which we identified by expanding $\phi(x)$ in the limit $x \ll 1$. The normalization is such that $\phi \rightarrow \phi_{\text{ini}}$ as $x \rightarrow 0$. A quick numerical evaluation of (34) yields the zero of ϕ at

$$H_c/m \simeq 0.18, \quad (35)$$

where $H_c \lesssim H_*$ is the value of H at zero crossing time t_c . There is a critical value $\delta_{\text{eff}}^* (< 1.5)$ for which the phase transition becomes efficient, implicitly defined by $p|_{\delta_{\text{eff}}^*} \simeq 1$ [corresponding to $S_E(\delta_{\text{eff}}^*) \simeq 250$ in our numerical example]. We can use (16) to infer the corresponding value of ϕ :

$$\phi_*^2 = M^2 \frac{\alpha^2}{9\lambda\tilde{\lambda}} (\delta_{\text{eff}}^* - \delta). \quad (36)$$

Using (32), we derive an upper bound on ϕ_*/ϕ_{ini} :

$$\frac{\phi_*^2}{\phi_{\text{ini}}^2} < \frac{1}{2} (\delta_{\text{eff}}^* - \delta) < 1. \quad (37)$$

The corresponding value of H_* can then be inferred numerically from (34). In particular, for $0 < (\delta_{\text{eff}}^* - \delta) \lesssim 1.5$, we find $0.18 < H_*/m \lesssim 0.6$. Later we will further tighten the upper bound to ensure a short percolation phase. The slope of ϕ at t_* varies between $-2.2\phi_{\text{ini}} < \dot{\phi}_*/H_* < -0.5\phi_{\text{ini}}$. In this work, we are mostly interested in the regime where $\phi_* \simeq 0$. This then translates to $\dot{\phi}_*/H_* \simeq -2.2\phi_{\text{ini}}$, which will be used for our later estimates. It would be straightforward to discuss the more general case, albeit being of minor significance for our order one estimates.

Now we have gathered all ingredients to calculate the duration of the percolation phase. Approximating $\Gamma(t)$ as a linear exponential, $\Gamma(t) \propto \exp(\bar{\beta}t)$, around $t = t_*$, we obtain an estimate for the inverse duration:

$$\bar{\beta} \equiv -\frac{dS_E}{dt} \simeq \frac{\dot{\Gamma}}{\Gamma}, \quad (38)$$

where we assumed that we can neglect the time dependence of K . We further use that S_E is only a function of δ_{eff} for $\tilde{\lambda}/\lambda \ll 1$, which allows us to further evaluate $\bar{\beta}$:

$$\begin{aligned} H_*\bar{\beta}^{-1} &\simeq -H_* \left(\frac{dS_E}{d\delta_{\text{eff}}} \frac{d\delta_{\text{eff}}}{dt} \right)^{-1} \\ &\simeq \left[\frac{d(\lambda S_E)}{d\delta_{\text{eff}}} \right]^{-1} \frac{\alpha^2}{36\tilde{\lambda}} \left(\frac{M}{\phi_{\text{ini}}} \right)^2 \frac{\phi_{\text{ini}}}{\phi_*}. \end{aligned} \quad (39)$$

This expression can be simplified by replacing ϕ_* defined in (36) and using the inequality (32):

$$H_*\bar{\beta}^{-1} < \left[\frac{d(\lambda S_E)}{d\delta_{\text{eff}}} \right]^{-1} \frac{1}{4\sqrt{2}} (\delta_{\text{eff}}^* - \delta)^{-1/2} \lambda. \quad (40)$$

Physically, this constitutes a bound on the maximal time bubbles have to grow before they start colliding. As we want the phase transition to be completed at least within one Hubble time, we demand $H_*\bar{\beta}^{-1} < 1$. A sufficient condition for this to be true is then

$$\lambda < (4\sqrt{2}) \sqrt{\delta_{\text{eff}}^* - \delta} \left[\frac{d(\lambda S_E)}{d\delta_{\text{eff}}} \right]. \quad (41)$$

This constitutes only a very mild restriction on the allowed values of λ . In fact, a numerical evaluation shows that $d(\lambda S_E)/d\delta_{\text{eff}} \in [10, 10^3]$ for $\delta_{\text{eff}}^* \in [0, 1.5]$. This implies that, unless $\sqrt{\delta_{\text{eff}}^* - \delta}$ is not hierarchically suppressed, the bound (41) can be fulfilled for any weakly coupled theory with $\lambda < 1$.

As our phase transition occurs in a dark sector at zero temperature, there is no plasma which would slow down the propagation of the bubble walls. We can therefore assume that they expand at the speed of light. As a result, $\bar{\beta}^{-1}$ is the maximal size a bubble can reach. When this maximal bubble collides and dissipates, it will lead to perturbations within the dark sector of the same size. These structures can leave an imprint on the photon fluid through their gravitational interactions and that way—at least in principle—lead to directly observable structures in the CMB. These structures subtend an angle

$$\theta_{\text{NEDE}} = \frac{1}{\bar{\beta} a_* D_{\text{rec}}}, \quad (42)$$

where we introduced the comoving angular diameter distance at the time of recombination,

$$D_{\text{rec}} = \int_0^{z_{\text{rec}}} dz \frac{1}{H(z)}. \quad (43)$$

We can approximate it based on a fiducial cosmology (more precision is not needed at this stage):

$$D_{\text{rec}} \simeq \frac{1}{H_0} \int_0^{1100} dz [0.3(1+z)^3 + 0.7]^{-1/2} \simeq 3.2/H_0, \quad (44)$$

which then yields

$$\theta_{\text{NEDE}} \simeq 0.4^\circ \times H_* \bar{\beta}^{-1} \left(\frac{g_*(T_*)}{3.9} \right)^{-1/2} \left(\frac{5000}{z_*} \right) \left(\frac{h}{0.7} \right) \times \sqrt{1 - f_{\text{NEDE}}}, \quad (45)$$

where $H_0 \equiv 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$. In particular, for $H_* \bar{\beta}^{-1} \simeq 1$ and $z_* \simeq 5000$, the bubbles give rise to structures⁹ that correspond to an angle of half a degree, which is just larger than the maximal angular resolution of the CMB observation of $l^{10} \sim 2\pi/2500 \simeq 0.14^\circ$. We leave the question as to whether the amplitude of the corresponding structures would be large enough to leave an observable imprint for future work. After all, we can impose the bound

$$H_* \bar{\beta}^{-1} < 0.4 \times \left(\frac{z_*}{5000} \right) \left(\frac{g_*(T_*)}{3.9} \right)^{1/2} \left(\frac{0.7}{h} \right) (1 - f_{\text{NEDE}})^{-1/2}, \quad (46)$$

which then prevents bubbles from growing to observable size.¹¹ This also tightens the upper bound on λ in (41):

$$\lambda < 2.3 \times \left(\frac{z_*}{5000} \right) \left(\frac{g_*(T_*)}{3.9} \right)^{1/2} \left(\frac{0.7}{h} \right) \sqrt{\frac{\delta_{\text{eff}}^* - \delta}{1 - f_{\text{NEDE}}}} \left[\frac{d(\lambda S_E)}{d\delta_{\text{eff}}} \right], \quad (47)$$

which evidently is still compatible with any value $\lambda < 1$ and, hence, does not constitute a meaningful parameter restriction.

In fact, we will mostly impose the even stronger bound $H_* \bar{\beta}^{-1} < 10^{-2}$, which can be easily achieved if $\lambda < 0.1$. We do this for different reasons. First, a percolation phase that is significantly shorter than $1/H_*$ justifies the use of an effective description where the phase transition happens instantaneously on cosmological timescales (we will discuss such an effective model in Sec. III). Second, it justifies the approximation of $\Gamma(t)$ as a linear exponential, used to derive $\bar{\beta}$ in (39). Third, as argued after (11), it also ensures the applicability of the semiclassical approximation used to calculate the tunneling rate Γ . In the following, we will

⁹The CMB cold spot is about 5° on the sky for comparison. It is an intriguing possibility that such a feature in the CMB could be related to NEDE [69,70].

¹⁰This value is compatible with Ref. [71], where it is argued that the CMB is sensitive only to structures with comoving size greater than about $10h^{-1} \text{ Mpc}$.

¹¹Note that NEDE still influences the CMB *indirectly*, because it changes the background evolution and supports adiabatic perturbations that source the gravitational potential also “felt” by the photons.

therefore assume that $\lambda < 0.1$. However, we also have to make sure that λ is not hierarchically small to still be compatible with both our estimate of K used to derive (31) and the condition $\tilde{\lambda} \ll \lambda$.

Finally, having a small value of λ also simplifies the expression for S_E . To be precise, provided $\lambda < 0.02$, we have $S_E \lesssim 250$ (as needed for an efficient percolation) for values of δ_{eff} below $\simeq 0.2$, which then justifies using a linear approximation for (20):

$$\lambda S_E \simeq \frac{\pi^2}{6} \alpha_1 \delta_{\text{eff}}. \quad (48)$$

This, in turn, allows us to approximate δ_{eff}^* as

$$\delta_{\text{eff}}^* \simeq \frac{6}{\alpha_1 \pi^2} \lambda S_E^* \simeq 0.11 \times \left(\frac{\lambda}{0.01} \right) \left(\frac{S_E^*}{250} \right). \quad (49)$$

In short, percolation becomes very efficient for $\lambda < 0.1$ lasting only a fraction of a Hubble time before it has covered the entire space with bubbles of true vacuum. We can therefore describe the phase transition on cosmological scales as an instantaneous process at time t_* implicitly determined by $p(t_*) \simeq 1$. During this short burst, the percolation parameter p increases exponentially until bubble percolation is completed after the time $\bar{\beta}^{-1}$ (before p would reach its peak value at $\phi = 0$).

D. Bubble coalescence and decay

Before we turn to the bubble dynamics, there is an additional decay channel for the energy stored in the (false) vacuum of ψ besides the bubble wall condensate. Part of it goes into oscillations of ϕ around the true vacuum. The reason is that ϕ after the phase transition does, in general, not have a vanishing VEV, which becomes apparent from Fig. 3(a) in the limit $r \rightarrow 0$. If we want to calculate this part, we have to take into account that the mass of ϕ increases due to the nonvanishing VEV of ψ . In other words, oscillations in the ϕ direction around the true vacuum experience a steeper potential than the ones around the false vacuum. From (12),

$$\begin{aligned} m^2 &\rightarrow m^2 + \tilde{\lambda} \psi_2^2 \\ &= m^2 + \mathcal{O}(1) \times \frac{\tilde{\lambda} \alpha^2}{\lambda^2} M^2, \end{aligned} \quad (50)$$

where we used $\psi_2 = \mathcal{O}(1) \alpha M / \lambda$, which follows from (4), to derive the second line. We further assume that the VEV of ϕ remains constant during the phase transition, $\phi(r) \simeq \phi_*$. As we have seen, this is a good approximation for $\tilde{\lambda} / \lambda \ll 1$. After the transition, i.e., for $t_* + \bar{\beta}$, we then get

$$\begin{aligned}\bar{\rho}_\phi|_{t_*+\bar{\beta}^{-1}} &= \frac{1}{2} \left(m^2 + \mathcal{O}(1) \frac{\tilde{\lambda}\alpha^2}{\lambda^2} M^2 \right) \phi_*^2 \\ &= \mathcal{O}(1) \times (\delta_{\text{eff}}^* - \delta) f_{\text{NEDE}} \bar{\rho}_*,\end{aligned}\quad (51)$$

where we neglected the term proportional to m^2 and used (36) and (25). This shows that for a generic choice of model parameters the potential energy in ϕ contributes a sizable fraction to the energy budget even though it was negligible before the transition. In this work, however, we focus on the special case where most of NEDE is converted into kinetic and gradient energy of the bubble walls. Explicitly, if we require $\bar{\rho}_\phi/\bar{\rho}_* \lesssim 10^{-2}$, which is enough suppression to ignore oscillations in the ϕ sector, this translates to a rather mild parameter tuning through (51)¹²:

$$\delta_{\text{eff}}^* - \delta \lesssim \frac{10^{-2}}{f_{\text{NEDE}}} \sim 10^{-1}, \quad (52)$$

where $\delta_{\text{eff}} = \mathcal{O}(0.1)$ as in (49) and $\delta < \delta_{\text{eff}}$. In particular, this is only weakly affecting the bound on λ in (47).

While the space is being filled with true vacuum bubbles, their walls quickly accelerate to the speed of light and start to collide. As bubble percolation becomes efficient very quickly—i.e., it raises to values $p \gg p(t_*) \simeq 1$ in a small fraction of a Hubble time—we can describe this as an instantaneous process. The subsequent collision phase is complicated to describe except for the idealized case of two equally sized bubbles [72]. From a phenomenological perspective, this coalescence phase risks polluting the CMB with non-scale-invariant density perturbations arising from the colliding wall structures. For example, exactly these perturbations strongly constrain models that use a first-order phase transition to end inflation, giving rise to an analogy of the “big bubble constraint” (see, for example, [56]). In our case, however, this would be happening in a dark sector, and it is not clear whether the imprint in the CMB would be strong enough. In any case, in the two-field model, as we have argued before, this problem (or feature) can be avoided by imposing the constraint (46), preventing bubbles from growing to observable sizes.

In the next step, the condensate of colliding bubble walls decays. In fact, there are several channels for dissipating its energy. If the phase transition is effectively happening in vacuum in the dark sector decoupled from the visible sector, the energy released by the bubbles is converted into kinetic and gradient energy of the bubble walls. To keep the discussion as general as possible, we will consider three generic ways how the gradient energy in the bubble walls can dissipate. The first two are model independent, and the last one is model dependent.

¹²We intend to relax this condition in future work. This opens the way toward providing a fraction of dark matter (DM) in terms of decaying NEDE.

1. Gravitational waves

Some of the gradient energy will dissipate in gravitational radiation. To estimate the fraction of the gradient energy that dissipates into gravitational radiation, we will assume that the transverse and traceless part of the energy momentum tensor, which sources gravitational waves, is of the same order as the released vacuum energy $\Pi_\psi \sim \rho_{\text{NEDE}}$ [73], suppressing the tensor structure for simplicity. A typical dimensional estimate of the amount of gravitational waves produced is then obtained by observing that, from the linearized Einstein equation, we have for the gravitational wave amplitude $\dot{h} \sim 16\pi G \Pi_\psi$, and, assuming the duration of the phase transition to be parametrized by $\bar{\beta}$ as in the previous subsection, we find $\ddot{h} \sim \bar{\beta} \dot{h} \sim \bar{\beta}^2 h$. The energy density radiated away in gravitational waves is then $\rho_{\text{GW}} \sim \dot{h}^2 / (32\pi G) \sim 8\pi G \Pi_\psi^2 / \bar{\beta}^2$, or, in other words, only a fraction of ρ_{NEDE} is typically radiated away in gravitational waves:

$$\frac{\rho_{\text{GW}}}{\rho_{\text{NEDE}}} \sim 3(H_* \bar{\beta}^{-1})^2 f_{\text{NEDE}} \ll 1. \quad (53)$$

2. Cosmic fluid

The remaining and dominant part of the gradient energy will redshift away. The details of this process are complicated to describe, because the corresponding state of colliding bubbles is highly inhomogeneous and anisotropic, characterized by structures whose density contrast $\delta_\psi(t, \mathbf{x}) = \rho_\psi(t, \mathbf{x}) / \bar{\rho}_\psi(t) - 1$ exceeds unity. The crucial observation is that the scale of these nonlinearities, $\bar{\beta}^{-1} \ll 1/H$, stays below cosmological length scales [this is a consequence of demanding that the bubble condensate does not leave any direct imprint in the observed CMB power spectrum; cf. Eq. (46)]. This raises the question how these structures, which provide a sizable fraction of the energy budget, affect the expansion history. Asked differently, is there an effective fluid description of a colliding bubble condensate valid on cosmological scales $\gg \bar{\beta}^{-1}$? To our knowledge, this question has not been answered before, and a complete discussion would require a fully *relativistic* coarse-graining procedure, which goes beyond the scope of our work.¹³ However, we can sharpen our intuition by looking at the *nonrelativistic* case. There, small-scale nonlinearities are known to lead to a *positive* excess pressure on large scales, corresponding to an ideal fluid with a nonvanishing equation of state parameter [75,76]:

¹³A quantitative statement can be made before bubbles start to collide. Taking the thin-wall limit, it was shown that the effective equation of state parameter approaches 1/3 due to the relativistic movement of the bubble walls [74]. However, within our model, percolation is very efficient, and bubbles start to collide well within a Hubble time, making this result inapplicable. Besides, the thin-wall approximation is not valid for the values of $\delta_{\text{eff}}^* (< 1.5)$ we are considering.

$$w_{\text{eff}} \simeq \frac{1}{3} \langle v_{\psi}^2 (1 + \delta_{\psi}) \rangle, \quad (54)$$

where $v_{\psi}(t, \mathbf{x})$ is the velocity field associated with ψ and $\langle \circ \rangle$ denotes the smoothing procedure over small-scale nonlinearities as introduced in Ref. [76]. We note that (54) assumes a negligible contribution from the gravitational potential, which requires a highly nonvirialized system (for virialized systems, $w_{\text{eff}} = 0$). The important point is that w_{eff} is not zero as one would naively expect for nonrelativistic matter; it rather receives a positive contribution that scales with δ_{ψ} . Moreover, any anisotropic stress Π_{ψ} averages to zero, which resonates with the observation that there is no preferred direction on sufficiently large scales. Of course, these results hold only for $v_{\psi}^2 \ll 1$, whereas for a realistic bubble wall condensate $v_{\psi}^2 \rightarrow 1$ as the walls quickly accelerate to relativistic speeds. Nevertheless, based on the nonrelativistic result, we also expect in the relativistic case the emergence of an excess pressure on large scales, whose value will depend on the precise tunneling potential (which controls the size of nonlinearities within the bubble wall domains). To be precise, rather than an equation of state parameter of $1/3$, valid for a radiation fluid in linear perturbation theory, we expect its value to move toward stiffness, i.e., $1/3 < w_{\text{eff}} < 1$.

We can also provide a heuristic argument, which relies on the observation that on small scales the bubble wall condensate gives rise to a sizable amount of anisotropic stress Π_{ψ} . In homogeneous setups, this component is known to dilute as $1/a^6$ [77,78], mimicking a stiff fluid with $w = 1$. The tendency of the large-scale fluid toward stiffness, $w_{\text{eff}} > 1/3$, can then be understood as a consequence of the anisotropic contributions on small scales.

3. Additional channels

There could be additional model-dependent decay channels for the ψ fluid, so that the remaining fluid, which has not yet been converted into gravitational waves, could decay into radiation. In principle, one could have an almost instant decay of the ψ fluid into some dark or visible radiation component immediately after the phase transition. In addition, one might consider the simple possibility that the ψ condensate decays into ϕ particle excitations. However, this is suppressed by a very stringent bound on the coupling scale $\tilde{\lambda}$ following from naturalness considerations, which we will derive in the next section.

To keep the discussion as general as possible, we will in our phenomenological treatment of NEDE consider all three possible decay channels and model NEDE after the phase transition as an effective fluid with an equation of state parameter $1/3 < w_{\text{NEDE}}^* < 1$. Of course, w_{NEDE}^* is generically time dependent; however, as the impact of NEDE on cosmological observables is sharply localized

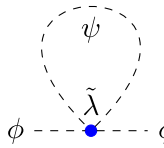
around its decay time, we can approximate it as a constant. For example, in our baseline model, we will fix it to $w_{\text{NEDE}}^* = 2/3$, which we believe to be a natural choice, but we will also allow it to vary within its physical range when we discuss different model extensions of our baseline model.

E. Radiative stability

So far, the only restrictions on our model parameters arose from (14) and (47) and the tuning in (52) (if we want to study a regime for which ϕ oscillations are suppressed after the decay). In particular, α and β can still be $\mathcal{O}(1)$. This means that our model is controlled only by the coupling constants, $\tilde{\lambda}$ and λ , satisfying $\tilde{\lambda} \ll \lambda < 0.1$, as well as the two mass scales, m and M , satisfying $m \ll M \ll M_{pl}$. The hierarchy of the latter is fixed by (25):

$$\frac{M^2}{m^2} = \mathcal{O}(1) \times \frac{\sqrt{\lambda^3 f_{\text{NEDE}}}}{\alpha^2} \frac{M_{pl}}{H_*} \gg 1, \quad (55)$$

which evaluates to $\sim 10^{53}$ if we plug in phenomenologically plausible numbers. This raises the question of how we can stabilize it against quantum corrections. After all, the ultralight mass m will receive quantum corrections from its coupling to the much heavier ψ field. Working perturbatively around the $\psi = 0$ background, the dominant contribution comes from the diagram



$$\delta m^2 = \phi \text{---} \tilde{\lambda} \text{---} \psi \text{---} \tilde{\lambda} \text{---} \phi \quad (56)$$

$$= \frac{\tilde{\lambda}}{2} \frac{1}{(4\pi)^{d/2}} \Gamma(1 - d/2) (\beta M^2)^{d/2-1},$$

where we used dimensional regularization. After properly renormalizing this expression, we find a correction (suppressing logs but keeping factors of π) of the order of $\tilde{\lambda} \beta M^2 / (32\pi^2)$. If we allow for a mild tuning of m^2 of one order of magnitude, we obtain an upper bound on $\tilde{\lambda}$:

$$\tilde{\lambda} \leq \mathcal{O}(1) \times 10^3 \frac{m^2}{\beta M^2} \ll 1. \quad (57)$$

This is an extremely small coupling bounded from above by $\sim 10^{-50}$ when using the estimates suggested in (27) and (30). While this ensures that our assumption $\tilde{\lambda} \ll \lambda$ can be easily justified, it also raises concerns about the effectiveness of our trigger mechanism. In fact, if the coupling scale is too small, it can lead to a violation of condition (32), which would mean that bubble nucleation is not shutting off in the very early Universe. We therefore assume that $\tilde{\lambda}$

saturates its naturalness bound, which then via (25) and (32) yields a lower bound on ϕ_{ini} :

$$\mathcal{O}(1) \times 10^{-9} \frac{\delta}{0.1} \left(\frac{\lambda}{0.01} \right) \left(\frac{H_*/m}{0.2} \right)^2 \left(\frac{f_{\text{NEDE}}}{0.1} \right) < \frac{\phi_{\text{ini}}^2}{M_{pl}^2} \ll 1, \quad (58)$$

where the upper limit followed from condition (29), ensuring that ϕ is subdominant before the decay. We also find that we can easily have $\phi_{\text{ini}}^2 = 10^{-8} M_{pl}^2$ without introducing any additional parameter tunings, making sure that we safely stay away from any quantum gravity regime. Having a small coupling also raises the question as to whether our assumption $\kappa \ll 1$ used to derive the Euclidian action was justified as $\kappa \propto 1/\sqrt{\tilde{\lambda}}$ (see the discussion in Sec. II B). Again, using the upper bound on the coupling, we derive from (15c)

$$\kappa^2 = \mathcal{O}(1) \times 10^{-6} \left(\frac{\lambda}{0.01} \right) \left(\frac{\delta}{0.1} \right), \quad (59)$$

which indeed is sufficiently small. We also note that diagrams with more vortices are suppressed by additional powers of $\tilde{\lambda}$ and, hence, always subdominant even for huge (but sub-Planckian) values of the external momenta.

Moreover, having an explicit expression for $\tilde{\lambda}$ allows us to further simplify the inverse percolation time $\bar{\beta}$ in (39); using (36), it evaluates to

$$H_* \bar{\beta}^{-1} \simeq \left[\frac{d(\lambda S_E)}{d\delta_{\text{eff}}} \right]^{-1} \frac{\alpha \sqrt{\tilde{\lambda}} M}{12 \sqrt{\tilde{\lambda}} \phi_{\text{ini}}} (\delta_{\text{eff}}^* - \delta)^{-1/2}. \quad (60)$$

We again substitute the upper bound on $\tilde{\lambda}$, approximate $S_E(\delta_{\text{eff}})$ as in (48) and use (25), which then yields

$$H_* \bar{\beta}^{-1} = \mathcal{O}(1) \times 10^{-3} \left(\frac{\delta}{0.1} \right)^{1/2} \left(\frac{f_{\text{NEDE}}}{0.1} \right)^{1/2} \left(\frac{0.1}{\delta_{\text{eff}}^* - \delta} \right)^{1/2} \times \left(\frac{M_{pl}/\phi_{\text{ini}}}{10^4} \right) \left(\frac{H_*/m}{0.2} \right) \left(\frac{\lambda}{0.01} \right)^{3/2}, \quad (61)$$

where we assumed parameter values in accordance with our previous discussion and the $\mathcal{O}(1)$ factor accounts for the uncertainty in the upper bound on $\tilde{\lambda}$. The lesson from this is that we can use λ as a dial to realize percolation phases that can be extremely short, lasting $\ll 1/H_*$, when $\lambda \ll 1$ but also rather long, up to $\sim 1/H_*$, when $\lambda \sim 1$. Note, however, that the last limit has to be treated with care, as different approximations we used rely on λ being somewhat small. Alternatively, we can achieve a longer duration by demanding $\delta_{\text{eff}}^* - \delta < 0.1$, which, however, constitutes a parameter tuning. As we will see in the next section, the value of $H_* \bar{\beta}^{-1}$ controls the strength of the gravitational wave

signal, which becomes stronger the longer the phase transition lasts.

We are now able to provide for illustration an explicit example for a consistent set of model parameters corresponding to a phase transition at $z_* = 5000$ with $f_{\text{NEDE}} = 10\%$ and $H_*/m = 0.2$. When we confront our model with data in Sec. IV, scanning over all consistent choices of parameters, we will see that this choice falls within the 95% C.L. of the model's mean cosmological parameters. We can further take $\lambda = 0.01$, $\beta \simeq \alpha \simeq 1$, which, in turn, determines $\delta = 9\lambda\beta/\alpha \sim 0.1$ as well as the mass scales $M = 0.04$ eV and $m \sim 2.0 \times 10^{-27}$ eV from (27) and (30), respectively. From (31), we then obtain the percolation condition as $S_E^* \simeq 230$, which via (49) fixes $\delta_{\text{eff}}^* \simeq 0.1$. By imposing a very mild one-digit tuning of β (or α), we can then ensure $\delta - \delta_{\text{eff}}^* \simeq 0.1$ in accordance with (52). In order to keep loop corrections under control, we choose $\tilde{\lambda} \simeq 10^3 m^2/M^2 \simeq 2.3 \times 10^{-48}$ saturating the upper bound in (57). With these choices, the microscopic system is fully determined. It amounts due to (61) to a percolation time $H_* \bar{\beta}^{-1} \sim 10^{-3}$, corresponding to a quick transition on cosmological timescales.

F. Gravitational waves

A unique feature of our model, setting it apart from the second-order scenario, is the production of a stochastic background of gravitational waves. Here, we estimate the corresponding spectrum (for a review, see, e.g., [79]). This task is greatly simplified by the fact that the transition occurs in a dark sector at zero temperature. So there is no plasma slowing down the expansion of the bubble walls, and gravitational waves are produced only during the collision of bubbles of characteristic size $1/\bar{\beta}$. The reason is that colliding bubbles, as opposed to a single expanding bubble, produce a nonvanishing anisotropic stress which, in turn, sources gravitational waves (for seminal papers, see [80,81]). The spectrum in terms of the dimensionless energy density can be calculated by using the ‘‘envelope approximation’’ first introduced in Ref. [82], resulting in¹⁴

$$h^2 \Omega_{GW} = 3.8 \times 10^{-8} (H_* \bar{\beta}^{-1})^2 \left(\frac{f_{\text{NEDE}}}{0.1} \right)^2 \times \left(\frac{3.9}{g_*(T_*)} \right)^{1/3} S_{GW}(f), \quad (62)$$

with spectral shape

¹⁴We cite from Ref. [79]. With regard to their notation, we set $\kappa_\phi = 1$, $\kappa_\nu = 0$, and $v_w = 1$ (as there is no coupling to any plasma) and identified $\alpha = f_{\text{NEDE}}/(1 - f_{\text{NEDE}})$ (assuming that the transition still takes place during the radiation-dominated epoch) as well as $\beta = \bar{\beta}$.

$$S_{GW}(f) \simeq \frac{3.8(f/f_*)^{2.8}}{1 + 2.8(f/f_*)^{3.8}}. \quad (63)$$

The peak frequency as measured today is extremely small:

$$f_* \simeq 2.6 \times 10^{-17} \text{ Hz} \frac{1}{(H_* \bar{\beta}^{-1})} \left(\frac{1+z_*}{5000} \right) \left(\frac{g_*(T_*)}{3.9} \right)^{1/6}. \quad (64)$$

The spectrum grows as f^3 for $f \ll f_*$ and falls off as f^{-1} for $f \gg f_*$. Only in the latter regime do we have the hope of detecting a signal once we reach the typical frequency of *pulsar timing arrays* at $f \sim 10^{-9}$ Hz [83]. Substituting (64) into (62), we find

$$h^2 \Omega_{GW} \simeq 1.3 \times 10^{-15} (H_* \bar{\beta}^{-1}) \left(\frac{10^{-9} \text{ Hz}}{f} \right) \left(\frac{f_{\text{NEDE}}}{0.1} \right)^2 \times \left(\frac{3.9}{g_*(T_*)} \right)^{1/6} \left(\frac{1+z_*}{5000} \right). \quad (65)$$

In Fig. 4, we illustrate that for $(H_* \bar{\beta}^{-1}) \rightarrow 1$, which can be achieved for large enough values of $\lambda (< 1)$ or $\delta_{\text{eff}}^* - \delta < 0.1$, we just reach the claimed peak sensitivity¹⁵ of the *Square Kilometre Array* at $h^2 \Omega_{GW} \sim 10^{-15}$ [61–63]. This limit also forces us into the green shaded region, where the CMB becomes potentially sensitive to structures produced by the bubble collisions (as discussed in Sec. II C). On the other hand, if we want the phase transition to be quick, as we do for the most part of this work, we are forced below the dotted line. In that case, the signal falls short of the experiment's claimed peak sensitivity.

The overall picture can change significantly if we look at different sophistications of our model. For example, by coupling the scalar field to the standard model sector, we expect to enhance the gravitational wave signal due to sound waves produced in the primordial fluid as well as turbulence effects. At the same time, we would be more sensitive to the perturbations produced during the collision phase, because they could be directly imprinted in the photon fluid. Also, the above derivation assumes a purely radiation-dominated universe and does not take into account the admixture of the NEDE component in the cosmic fluid, which will chance the numerical value of the signal (not the order of magnitude, though).

In summary, while we have provided a first discussion of the model's distinct signatures, both possibilities, gravitational waves and bubble wall imprints in the CMB, demand a more complete analysis.

¹⁵This has to be understood as a necessary condition for detection. A more complete analysis needs to consider the signal-to-noise ratio, too.

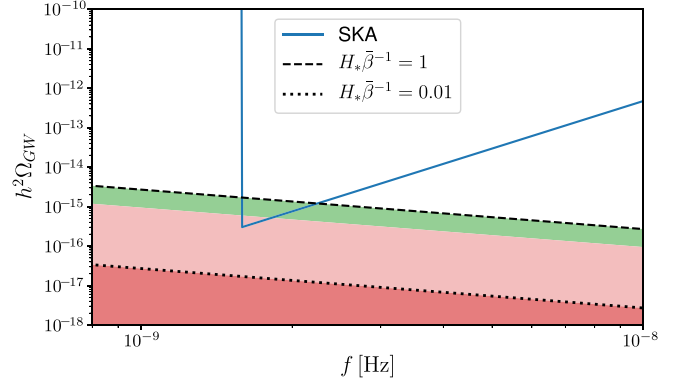


FIG. 4. Sensitivity curve (blue solid) of the *Square Kilometre Array* (SKA) in terms of the dimensionless energy density of gravitational waves. Below the black dotted line (dark red), we can describe the phase transition as an instantaneous process on cosmological scales, making it accessible to our effective description in Sec. III. On the other hand, pushing $\bar{\beta}$ toward cosmological scales, i.e., $\bar{\beta}^{-1} \rightarrow H^{-1}$, might lead to an observable gravitational wave signal. In this limit, we also expect the colliding bubble walls to leave a direct imprint in the CMB (green shaded region). We used $f_{\text{NEDE}} = 0.14$ and $z_* = 5300$ as in Table I.

III. COSMOLOGICAL MODEL

Simulating the percolation of bubbles with their subsequent collision and dissipation phase is a complicated task. We therefore build our analysis on different simplifying assumptions relating to the evolution of the background and the perturbations, separately.

- (1) Bubble nucleation happens instantaneously on cosmological timescales. In the previous section, we have argued that this is satisfied for our two-field model for a wide range of natural parameter choices, requiring only a mild suppression of λ . For example, for $\lambda \simeq 0.01$ we found in (61) that $H_* \bar{\beta}^{-1} \sim 10^{-3}$ [this choice, due to (46), also ensures that the CMB observation is not resolving spatial structures created by the largest bubbles].
- (2) The condensate of colliding vacuum bubbles can be described as a fluid with effective equation of state parameter $> 1/3$ on large scales $\gg \bar{\beta}^{-1}$ (see the discussion in Sec. II D). There is also a radiation component arising from its decay. As the impact of both fluids on cosmological observables is localized around the decay time, we can approximate the two-component fluid as a single fluid with a constant equation of state parameter. This assumption could be relaxed in future work by explicitly modeling a two-component fluid.
- (3) Motivated by our microscopic model (12), we assume that the time t_* at which the phase transition starts is determined by the mass m of a subdominant trigger field ϕ . It is initiated according to (35) around

the time where ϕ enters its oscillatory stage and the percolation parameter $p > 1$ for the first time.

- (4) The fluid perturbations start with vanishing shear stress σ_{NEDE} directly after the transition. This is motivated by the fact that perturbations created due to the collision dynamics have a wavelength too short to leave an imprint on the CMB power spectrum, provided condition (46) is fulfilled. Otherwise, the perturbation dynamics is fully characterized by the background equation of state parameter w_{NEDE} , the rest-frame sound speed c_s , and the viscosity parameter c_{vis} .

In accordance with these assumptions, we do not have to resolve the ψ dynamics explicitly. We will rather introduce a single fluid that exhibits a *sudden* change in its equation of state, switching from -1 to w_{NEDE}^* at time t_* . We note that we will refer to both phases of matter as the NEDE fluid. *Before* the transition it models the (classical) vacuum energy of the field ψ , and *after* the transition it describes the energy density stored in the bubble wall condensate and its decay products. In contrast, we will track the evolution of ϕ explicitly. As we will see, this is necessary to keep track of spatial fluctuations of the transition time. While these assumptions will allow us to relatively quickly arrive at a first phenomenological check of our model, we intend to relax and scrutinize them further in our subsequent work.

A. Instantaneous transition

Our effective model parametrizes the NEDE equation of state parameter as

$$w_{\text{NEDE}}(t) = \begin{cases} -1 & \text{for } t < t_*, \\ w_{\text{NEDE}}^* & \text{for } t \geq t_*. \end{cases} \quad (66)$$

This describes a sudden change at time t_* from a constant energy density to a decaying fluid with equation of state parameter w_{NEDE}^* . The energy density then scales as

$$\bar{\rho}_{\text{NEDE}}(t) = \bar{\rho}_{\text{NEDE}}^* \left(\frac{a_*}{a(t)} \right)^{3[1+w_{\text{NEDE}}(t)]}, \quad (67)$$

where $\bar{\rho}_{\text{NEDE}}^*$ has been introduced in (21) denoting the energy density of NEDE directly before the phase transition. It can be related to the potential energy of our two-field model via (22) (strictly speaking, this definition also contains a minuscule energy contribution arising from ϕ). Also note that $\bar{\rho}_{\text{NEDE}}$ is continuous across the transition. From a microscopic perspective, this is justified, because the transition happens instantaneously and all of the field's potential energy is ultimately converted to kinetic and gradient energy of the bubble walls. Later, for our parameter extraction, we will use the fraction of NEDE before its decay, f_{NEDE} , as defined in (21), rather than $\bar{\rho}_{\text{NEDE}}^*$ as a free parameter, because it gives us a more intuitive idea about the amount of NEDE.

We denote the jump of a function $f(t)$ across the phase transition at time t_* as

$$[f]_{\pm} = \lim_{\epsilon \rightarrow 0} [f(t_* - \epsilon) - f(t_* + \epsilon)] \equiv f^{(-)} - f^{(+)}. \quad (68)$$

The jump of the Friedmann equations then yields

$$[H]_{\pm} = 0, \quad (69a)$$

$$[\dot{H}]_{\pm} = 4\pi G(1 + w_{\text{NEDE}}^*)\bar{\rho}_{\text{NEDE}}^*, \quad (69b)$$

where we used the continuity of the background energy density,

$$[\bar{\rho}]_{\pm} = 0, \quad (70)$$

and assumed the continuity of all other equation of state parameters (except for NEDE), i.e., $[w_i]_{\pm} = 0$. In Appendix A, we present an alternative but equivalent derivation of (69a) based on Israel's matching equations [59].

We also introduce the field ϕ to trigger the phase transition. Its on-shell dynamics before the transition is controlled by (33). As it is a subdominant field component, its backreaction on the geometry can be neglected, and it therefore quickly locks into the attractor solution valid for a radiation-dominated universe [as derived later in (100)]. On the background level, its single purpose is to provide a time switch triggered by its first zero crossing at (35). Since its VEV does not change during the (almost instantaneous) phase transition, we have $[\phi]_{\pm} = 0$. Integrating its dynamical equation in (33) over an infinitesimal time interval yields $[\dot{\phi}]_{\pm} = 0$. As argued before, we are studying a regime where the scalar is also subdominant after the phase transition. In terms of our field theory model, this corresponds to the limit (52). A summary of the background evolution is provided in Fig. 5 (with values corresponding to our best-fit base model discussed in Sec. IV). The left panel shows the composition of the background energy density, $\bar{\rho} = \sum_i \bar{\rho}_i$, and the right panel depicts the discontinuity in the effective equation of state parameter.

B. Perturbation matching

Before the phase transition, $\rho_{\text{NEDE}} = \text{const}$ and NEDE fluctuations can be consistently assumed to vanish. After the transition, on the other hand, they are necessarily produced due to the gravitational sourcing of their equations. We will also see that they are crucial for the phenomenological success of our first-order phase transition. However, their implementation involves a subtlety due to the fact that the transition occurs during the CMB epoch where observable modes have already entered the horizon. As a result, we cannot use the usual superhorizon approximation to initialize perturbations in the NEDE fluid after

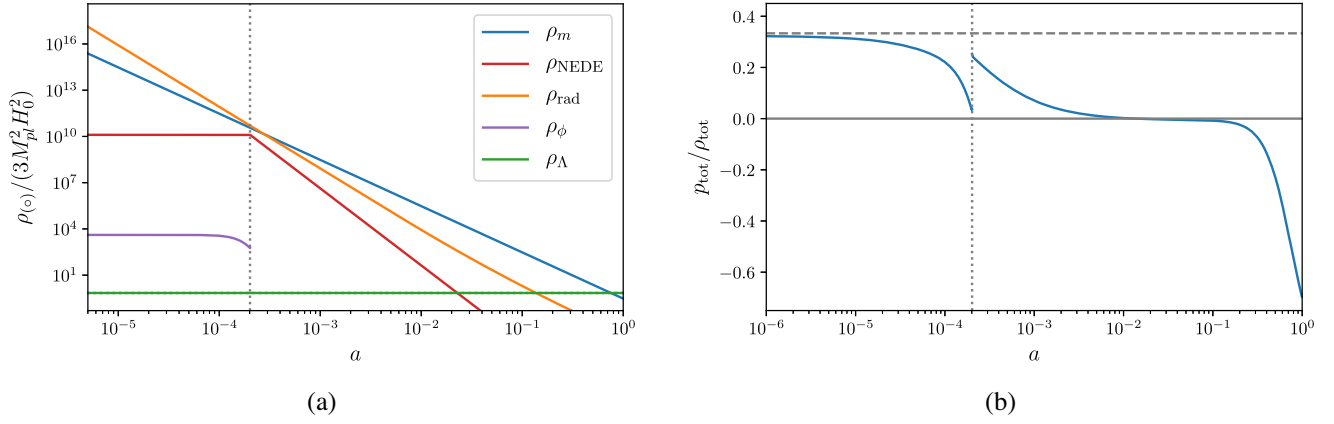


FIG. 5. Background evolution for the best-fit model in Table I (with SH_0ES) as a function of the scale factor a . The transition is characterized by a jump in the effective equation of state parameter (dotted line). $\rho_m = \rho_b + \rho_{\text{cdm}}$ and ρ_{rad} are the matter and radiation density, respectively, and ρ_Λ is the cosmological constant contribution. (a) Different fluid components. ρ_{NEDE} decays quickly after the transition (dotted line), and ρ_ϕ is subdominant before and vanishing after the transition. (b) Effective e.o.s. parameter, $\omega_{\text{eff}} = p_{\text{tot}}/\rho_{\text{tot}}$. During radiation and matter domination the dashed and the solid line is approached, respectively.

the phase transition. Instead, we have to derive the full set of matching equations describing how cosmological perturbations evaluated directly before and after the transition at time t_* can be related. As the central result of this section, we will express the initial NEDE density perturbation $\delta\rho_{\text{NEDE}}^*(\mathbf{x}) \equiv \rho_{\text{NEDE}}(t_*, \mathbf{x}) - \bar{\rho}_{\text{NEDE}}(t_*)$ and the initial divergence of the fluid velocity $\theta_{\text{NEDE}}^*(\mathbf{x})$ in terms of fluctuations of the spacelike transition surface Σ for both sub- and superhorizon modes.

On a technical level, we introduce a function $q(t, \mathbf{x})$ to define Σ :

$$q(t_*, \mathbf{x})|_\Sigma = \text{const.} \quad (71)$$

Later, we will identify $\phi(t, \mathbf{x}) = q(t, \mathbf{x})$, implying that the transition happens across a surface of constant trigger field. Fluctuations in ϕ will therefore lead to small spatial fluctuations of the decay time around the background value t_* , explicitly $\delta q^*(\mathbf{x}) \equiv q(t_*, \mathbf{x}) - \bar{q}(t_*) = \delta\phi(t_*, \mathbf{x})$, which, in turn, will give rise to adiabatic fluctuations in the EDE fluid (provided we choose adiabatic initial conditions for ϕ).

In the following, we first derive the general matching equations that are needed to consistently implement a generic transition in cosmological perturbation theory. After that, we apply the result to our particular two-field model. We work in the synchronous gauge, for which the metric perturbations in momentum space take the form

$$ds^2 = -dt^2 + a(t)^2(\delta_{ij} + h_{ij})dx^i dx^j, \quad (72)$$

where

$$h_{ij} = \frac{k_i k_j}{k^2} h + \left(\frac{k_i k_j}{k^2} - \frac{1}{3} \delta_{ij} \right) 6\eta \quad (73)$$

and $h = \delta^{ij} h_{ij}$. This choice leaves a residual gauge freedom, $x^0 \rightarrow x^0 + \bar{\xi}^0$ and $x^i \rightarrow x^i + g^{ij} \partial_j \bar{\xi}$, with

$$\bar{\xi} = -D_2 a^2 \int dt \frac{1}{a^2} + D_1 a^2, \quad (74a)$$

$$\bar{\xi}_0 = D_2, \quad (74b)$$

where we introduced $D_1 = D_1(\mathbf{k})$ and $D_2 = D_2(\mathbf{k})$, which are time-independent functions of the spatial momenta. The metric perturbations then transform as

$$\Delta_{\bar{\xi}} h = 2 \frac{k^2}{a^2} \bar{\xi} + 6H \bar{\xi}_0, \quad (75a)$$

$$\Delta_{\bar{\xi}} \eta = -H \bar{\xi}_0. \quad (75b)$$

The initial perturbations in the NEDE fluid can be derived from the perturbed Einstein equations that are first order in time derivatives [60]:

$$\frac{1}{2} H \dot{h} - \frac{k^2}{a^2} \eta = 4\pi G \delta\rho, \quad (76a)$$

$$\frac{k^2}{a^2} \dot{\eta} = 4\pi G (\bar{\rho} + \bar{p}) \frac{\theta}{a}. \quad (76b)$$

We introduced the total divergence of the fluid velocity, $(\bar{\rho} + \bar{p})\theta = \sum_i (\bar{\rho}_i + \bar{p}_i)\theta_i$, as well as the total energy density perturbation, $\delta\rho = \sum_i \delta\rho_i$, and employed the same definitions as in Ref. [60], in particular, $ik^i \delta T_i^0 \equiv (\bar{\rho} + \bar{p})a\theta$ and $\delta\rho \equiv -\delta T_0^0$.¹⁶ Under the residual gauge freedom, they transform as

¹⁶The factor a arises from the conversion to cosmological time t .

$$\Delta_{\bar{\xi}}\delta\rho = \dot{\bar{\rho}}\bar{\xi}_0, \quad (77a)$$

$$\Delta_{\bar{\xi}}\delta\theta = \frac{k^2}{a}\bar{\xi}_0. \quad (77b)$$

Using (75) and (77), it is then straightforward to show that the Einstein equations in (76) are indeed invariant. In order to obtain a closed system, we need to impose Israel's matching conditions [59], which we derive in full generality in Appendix A. They express the metric field after the transition, e.g., $\eta^{(+)}$, in terms of the fields right before it, e.g., $\eta^{(-)}$. In the synchronous gauge, they read

$$[\eta]_{\pm} + H_* \left[\frac{\delta q}{\bar{q}} \right]_{\pm} = 0, \quad (78a)$$

$$[\dot{\eta}]_{\pm} + \left[\dot{H} \frac{\delta q}{\bar{q}} \right]_{\pm} = 0, \quad (78b)$$

$$[h]_{\pm} + 6[\eta]_{\pm} = 0, \quad (78c)$$

$$[\dot{h}]_{\pm} + 6[\dot{\eta}]_{\pm} + 2k^2 \left[\frac{\delta q}{\bar{q}} \right]_{\pm} = 0, \quad (78d)$$

where $\delta q(t, \mathbf{k})$ parametrizes fluctuations in the transition surface around some homogeneous background $\bar{q}(t)$. In Appendix A, we also derive the matching for matter fluctuations of a given fluid component i with constant equation of state parameter during the transition (hence, not applicable to NEDE):

$$[\delta_i]_{\pm} + 3H_*(1 + w_i) \left[\frac{\delta q}{\bar{q}} \right]_{\pm} = 0, \quad (79a)$$

$$a_*[\theta_i]_{\pm} - k^2 \left[\frac{\delta q}{\bar{q}} \right]_{\pm} = 0, \quad (79b)$$

where as usual $\delta_i = \delta\rho_i/\bar{\rho}_i$. Moreover, we show that higher momenta of the distribution functions are continuous across the surface; specifically, for the anisotropic stress we have $[\sigma_i]_{\pm} = 0$, and similarly for the higher moments in the Boltzmann hierarchy, $[F_{\nu l}]_{\pm} = [F_{\gamma l}]_{\pm} = 0$ for $l \geq 3$ (relevant only for neutrinos ν and photons g).

In general, our residual gauge freedom can be fixed independently before and after the transition, corresponding to four *a priori* independent constants $D_1^{(\pm)}$ and $D_2^{(\pm)}$. However, the derivation of the above relations made the assumption that spatial diffeomorphisms $\partial_i \xi$ are continuous across the transition surface.¹⁷ This requires $[\xi]_{\pm} = 0$, which due to (74a) fixes one of the four constants. Using the transformation of δq ,

¹⁷It is straightforward to relax this assumption, but it does not add any useful freedom to the calculation.

$$\Delta_{\bar{\xi}}\delta q = \dot{\bar{q}}\bar{\xi}_0, \quad (80)$$

together with the definition in (74), it is then straightforward to show that the matching equations are invariant under the residual gauge freedom provided $[\xi]_{\pm} = 0$; in particular, we find

$$\Delta_{\bar{\xi}} \left[\frac{\delta q}{\bar{q}} \right]_{\pm} = [D_2]_{\pm} \equiv D_2^{(-)} - D_2^{(+)}. \quad (81)$$

We use this freedom to impose $\left[\frac{\delta q}{\bar{q}} \right]_{\pm} = 0$, which simplifies the matching equations significantly:

$$[h]_{\pm} = [\eta]_{\pm} = [\delta_i]_{\pm} = [\theta_i]_{\pm} = 0, \quad (82a)$$

$$[\dot{h}]_{\pm} = -6[\dot{\eta}]_{\pm} = 6[\dot{H}]_{\pm} \frac{\delta q_*}{\bar{q}_*}, \quad (82b)$$

where $[\dot{H}]_{\pm}$ is specified in (69b).¹⁸

In summary, all perturbations (except for NEDE perturbations) are continuous, while their first derivatives, in general, are not. In particular, the discontinuity of $\dot{\eta}$ and \dot{h} is set by the spatial variations of the transition surface $\delta q_* = \delta q(t_*, \mathbf{k})$. The choice of δq_* is a physical one, and we will discuss different constructions later.

The derivation of the junction conditions assumed that the equation of state of a particular matter component i is not changing during the transition, so it cannot be applied to NEDE. In fact, $\bar{\rho}_{\text{NEDE}} = \text{const}$ before the decay, which implies that, unlike with ordinary fluids, the gravitational source term is decoupled from its perturbations [see the fluid perturbation equation in (113)], and it is consistent to assume

$$\delta_{\text{NEDE}} = \theta_{\text{NEDE}} = \sigma_{\text{NEDE}} = F_{\text{NEDE}l} = 0 \quad \text{for } t < t_*, \quad (83)$$

where σ_{NEDE} denotes the anisotropic stress in the NEDE fluid and we allowed for the possibility of higher multipoles $F_{\text{NEDE}l}$ (as relevant when going beyond a fluid description). This then allows us to further evaluate the discontinuity of the NEDE perturbations

$$[\delta_{\text{NEDE}}]_{\pm} = -\delta_{\text{NEDE}}^{(+)} \equiv -\delta_{\text{NEDE}}^*, \quad (84a)$$

$$\begin{aligned} [(\bar{\rho}_{\text{NEDE}} + \bar{p}_{\text{NEDE}})\theta_{\text{NEDE}}]_{\pm} &= -\bar{\rho}_{\text{NEDE}}^*(1 + w_{\text{NEDE}}^*)\theta_{\text{NEDE}}^{(+)} \\ &\equiv -\bar{\rho}_{\text{NEDE}}^*(1 + w_{\text{NEDE}}^*)\theta_{\text{NEDE}}^*. \end{aligned} \quad (84b)$$

The initial NEDE perturbations δ_{NEDE}^* and θ_{NEDE}^* can now be derived from Einstein's equations. To be precise, the

¹⁸The remaining two constants are then fixed as usual by requiring adiabatic superhorizon perturbations in the metric to scale as $h = Ck^2\tau^2$ (with τ the conformal time and C constant) when fixing initial conditions during radiation domination (see, for example, [60]).

discontinuity of (76) together with (82), (69b), and (84) results in our main result of this section:

$$\delta_{\text{NEDE}}^* = -3(1 + w_{\text{NEDE}}^*)H_* \frac{\delta q_*}{\dot{q}_*}, \quad (85a)$$

$$\theta_{\text{NEDE}}^* = \frac{k^2 \delta q_*}{a_* \dot{q}_*}. \quad (85b)$$

Physically, these equations tell us that for a general transition it would be inconsistent to set the perturbations in the NEDE sector to zero after the transition, as they are linked to δq_* . Note, however, that no such condition applies to the shear σ_{NEDE}^* (or any higher multipoles $F_{\text{NEDE}l>3}$ in the Boltzmann hierarchy). We can therefore consistently impose

$$\sigma_{\text{NEDE}}^* = 0 \quad \text{and} \quad F_{\text{NEDE}l>3}^* = 0. \quad (86)$$

We note that the junction conditions apply to both sub- and superhorizon modes. As mentioned before, this is important, as modes we observe in the CMB today with $\ell > \ell_*$ were subhorizon during the phase transition, where we defined

$$\begin{aligned} \ell_* &\simeq a_* H_* D_{\text{rec}} \simeq 160 \times \frac{1}{\sqrt{1 - f_{\text{NEDE}}}} \left(\frac{g_*}{3.9} \right)^{1/2} \\ &\times \left(\frac{0.7}{h} \right) \left(\frac{1 + z_*}{5000} \right). \end{aligned} \quad (87)$$

Here, we assumed that the transition occurs during radiation domination and approximated D_{rec} as in (43). Specifically, if the phase transition takes place at $z_* = 5000$ and $f_{\text{NEDE}} = 0.1$, modes with $\ell > 180$ are on subhorizon scales during the transition.

Finally, we check that the comoving curvature perturbation is conserved during the phase transition. For different fluid components i , it is defined in the synchronous gauge as

$$\zeta_i = -\eta - H \frac{\delta \rho_i}{\dot{\rho}_i}. \quad (88)$$

Its conservation amounts to the statement

$$\begin{aligned} [\zeta_i]_{\pm} &= -[\eta]_{\pm} - H_* \left[\frac{\delta \rho_i}{\dot{\rho}} \right]_{\pm} \\ &= -[\eta]_{\pm} + \frac{1}{3(1 + w_i)} [\delta_i]_{\pm} \\ &= 0, \end{aligned} \quad (89)$$

where we used the matching conditions (79a) and (78a). Note that the same result was also derived in Ref. [84] in an inflationary context. In the special case of adiabatic

superhorizon perturbations, i.e., $\delta \rho_i / \dot{\rho}_i = \delta \rho_j / \dot{\rho}_j$ for $j \neq i$ and $k \ll aH$, definition (88) implies

$$\zeta_i = \zeta_{j \neq i} \equiv \zeta \quad \text{for } k^2 / (aH)^2 \ll 1, \quad (90)$$

where $\zeta = \text{const} + \mathcal{O}(k^2 / (aH)^2)$. Things are slightly more complicated for NEDE perturbations, though. Before the transition, ζ_{NEDE} is not well defined, as the second term (88) is indefinite, but directly afterward we derive

$$\begin{aligned} \zeta_{\text{NEDE}}^* &= -\eta_* + \frac{1}{3(1 + w_{\text{NEDE}}^*)} \delta_{\text{NEDE}}^* \\ &= -\eta_* - H_* \frac{\delta q_*}{\dot{q}_*}. \end{aligned} \quad (91)$$

The question as to whether NEDE perturbations fulfill the adiabaticity condition in (90) depends on the choice of $q(t, \mathbf{x})$. Their nonconservation would correspond to the presence of entropy perturbations, of which a possible gauge-invariant definition reads¹⁹

$$\begin{aligned} S_{\text{NEDE},i}^* &= \zeta_i^* - \zeta_{\text{NEDE}}^* \\ &= \frac{1}{3(1 + w_i)} \delta_i^* + H_* \frac{\delta q_*}{\dot{q}_*}. \end{aligned} \quad (92)$$

C. Transition surface

In our particular model, the transition surface Σ is fixed by surfaces of constant ϕ . Recall that the ϕ dependence of the Euclidian action in (20) controls the tunneling rate, which becomes efficient when ϕ drops below the threshold value ϕ_* , implicitly determined by $p(\phi_*) \simeq 1$. We therefore have

$$q(t, \mathbf{x}) = \phi(t, \mathbf{x}). \quad (93)$$

First, note that this is a gauge-invariant statement and, hence, valid in any frame. This can be seen easily by considering the transformation of δq under general diffeomorphisms:

$$\Delta_{\xi} \delta q = \dot{q} \xi_0 = \dot{\phi} \xi_0 = \Delta_{\xi} \delta \phi, \quad (94)$$

where we used that $\dot{\phi} = \dot{q}$. In other words, perturbations in the subdominant field ϕ set the initial conditions for perturbations in the NEDE fluid after the phase transition. In the synchronous gauge, they fulfill

$$\delta \phi'' + 2\mathcal{H} \delta \phi' + (k^2 + a^2 m^2) \delta \phi = -\frac{1}{2} h' \phi', \quad (95)$$

¹⁹This definition is independent of the index i provided there are no entropy perturbations before the transition, implying $\zeta_i = \zeta_{j \neq i}$.

where we found it convenient to work in conformal time τ . We also introduced the conformal Hubble parameter $\mathcal{H} \equiv aH = 1/\tau$ and used the shorthand $' = d/d\tau$. On the right-hand side, we substitute the adiabatic solution for superhorizon metric perturbations [60], $h = -\zeta(k\tau)^2/2$, which is not affected by the presence of the subdominant scalar ϕ . In the limit where $k \rightarrow 0$, the corresponding equation becomes

$$\delta\phi'' + 2\mathcal{H}\delta\phi' + a^2m^2\delta\phi = \mathcal{O}(k^2/\mathcal{H}^2). \quad (96)$$

This limit is useful, as it allows us to extract the adiabatic and entropy modes. Also note that the fluctuation equation becomes formally equivalent to the background equation in (33) expressed in conformal time:

$$\phi'' + 2\mathcal{H}\phi' + a^2m^2\phi = 0. \quad (97)$$

1. Adiabatic case

The adiabatic mode corresponds to the trivial solution, which is vanishing at leading order, i.e., $\delta\phi = \mathcal{O}(k^2/\mathcal{H}^2)$. If we further use²⁰

$$\delta_i/(1+w_i) = \frac{1}{4}\zeta k^2\tau^2 + \mathcal{O}(k^4/\mathcal{H}^4), \quad (98)$$

we can derive from (92) that

$$S_{\text{NEDE},i}^* \xrightarrow{k \rightarrow 0} 0. \quad (99)$$

In other words, there are no superhorizon entropy perturbations created by the transition provided we assume adiabatic initial conditions for ϕ . We also derive the first nonvanishing correction to $\delta\phi$ in the limit $m^2 \ll H^2$. To that end, we further evaluate the source term in (95) by substituting the background solution

$$\phi = \phi_{\text{ini}} \left(1 - \frac{1}{20} \frac{m^2}{H(t)^2} \right) + \mathcal{O}(m^4/H^4), \quad (100)$$

which we obtained by expanding (34). Recall that this expression is valid only during radiation domination. This is consistent, as ϕ is a subdominant field and, thus, we can neglect its gravitational backreaction. There is also a second branch, which we set to zero as it becomes quickly negligible. At leading order, we then obtain

$$\delta\phi = -\frac{1}{420} \phi_{\text{ini}} \frac{m^2 k^2}{a(t)^2 H(t)^4} \zeta + \mathcal{O}(m^4/H^4, k^4/(aH)^4). \quad (101)$$

²⁰This result can also be found in Ref. [60] by identifying $\zeta = -2C + \mathcal{O}(k^2/\mathcal{H}^2)$.

In particular, this expression vanishes in the limit $\zeta \rightarrow 0$, as expected for an adiabatic (or curvature) mode. We will use it to define adiabatic initial conditions when we implement our model in a Boltzmann code.

Next, we can use (100) and (101) to derive an expression for δ_{NEDE}^* on superhorizon scales, provided the transition takes place when $H_*/m > 1$. Even though this is not the most interesting case for our model where a typical scenario suggests $H_*/m \simeq 0.2$, it will still turn out to give a good approximation on superhorizon scales. Plugging the perturbative expressions back into (85) yields at leading order in $k^2/(aH)^2$

$$\delta_{\text{NEDE}}^* = -\frac{1+w_{\text{NEDE}}^*}{28} k^2 \tau_*^2 \zeta \quad \text{and} \quad \theta_{\text{NEDE}}^* = \frac{1}{84} k^4 \tau_*^3 \zeta, \quad (102)$$

which we can compare to adiabatic superhorizon perturbations in the neutrino fluid at the same time [60]:

$$\delta_\nu^* = \frac{1}{3} k^2 \tau_*^2 \zeta \quad \text{and} \quad \theta_\nu^* = \frac{1}{36} \frac{23 + 4R_\nu}{15 + 4R_\nu} k^4 \tau_*^3 \zeta, \quad (103)$$

where $R_\nu = \rho_\nu/(\rho_\gamma + \rho_\nu) = \text{const}$ and ζ is conserved at leading order. This shows us that, even if we have a radiation fluid in both cases, i.e., $w_{\text{NEDE}}^* = 1/3$, perturbations are different already at the order of k^2 . While this difference is negligible for superhorizon modes, it becomes important when modes enter the horizon ($k\tau \geq 1$). For a transition that happens around $z = 5000$, these are modes with $k \gtrsim 0.015/\text{Mpc}$ (or $\ell \gtrsim 180$). This is exactly the fingerprint left by the trigger field when setting the initial conditions for δ_{NEDE} and θ_{NEDE} . In this more general case, we will use our Boltzmann code, which will be introduced in Sec. IV, to numerically determine the evolution of $\delta\phi$. Using our best-fit cosmology, we obtain the curves in Fig. 6, where we plot $\delta_{\text{NEDE},k}^*/\zeta_{k=0}$ and $\theta_{\text{NEDE},k}^*(k)/\zeta_{k=0}$ as a function of H_*/m for different values of k . Note that we keep the decay redshift z_* fixed, so variations in the initial fluctuations result only from the fact that the transition happens at a different stage of the trigger field evolution [recall that $\phi(t)$ depends only on the ratio H/m]. For example, the dotted line corresponds to the zero of $\phi(t)$, and its first minimum sits toward the left at $H_*/m \simeq 0.12$. We did not include it in the plot because the fluctuations diverge at this point due to the vanishing of $\dot{q}(t) = \dot{\phi}(t)$ in the denominator of (85). The plot shows that the analytic result in (102) constitutes a good approximation only for superhorizon modes. In particular, we see that the initial fluctuations after the phase transition become sizable on subhorizon scales.

In summary, we find that the dynamics of the trigger field on both the background and perturbation level control the size of initial perturbations in the NEDE fluid after the phase transition. The crucial parameter is given by the

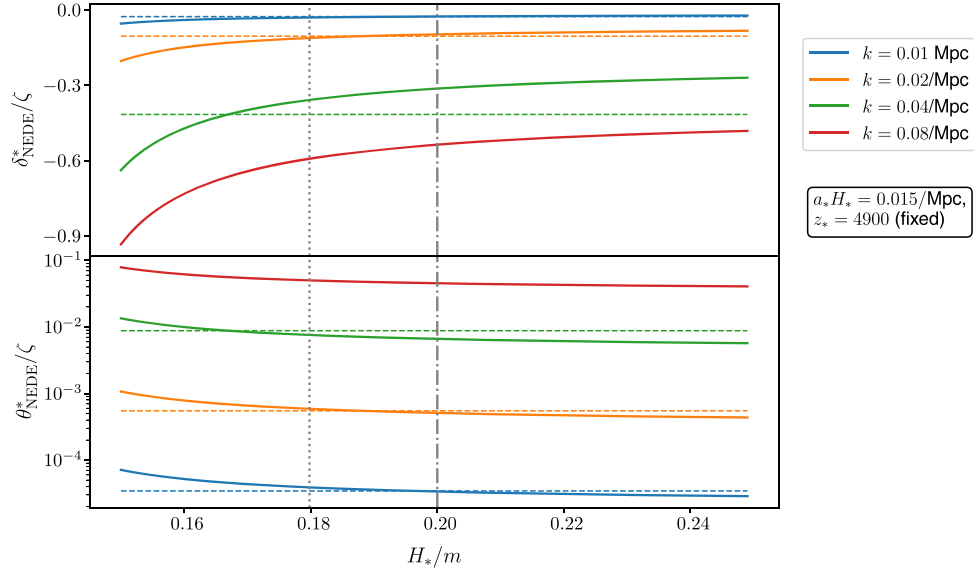


FIG. 6. Dependence of initial fluctuations in NEDE fluid on trigger parameter H_*/m for best-fit NEDE as in Table I (with SH₀ES) and z_* fixed. The dotted line correspond to the case where the phase transition coincides with the zero of ϕ . The dash-dotted line indicates $H_*/m = 0.2$ as used in our base model. The dashed horizontal lines depict the analytical approximation (102) valid for superhorizon modes with $k < a_*H_*$ (orange and blue).

ratio H_*/m to which we refer as the trigger parameter. From (36), we derive that it is related to our microscopic model via

$$H_*/m = \phi^{-1} \left(\frac{1}{3} M \alpha \frac{\sqrt{\delta_{\text{eff}}^* - \delta}}{\sqrt{\lambda \tilde{\lambda}}} \right), \quad (104)$$

where ϕ^{-1} is obtained by inverting (34) on the interval before its first minimum, and, in accordance with our previous discussion in Sec. II C, δ_{eff}^* is implicitly determined by the percolation condition $p|_{\delta_{\text{eff}}^*} \simeq 1$. To get an idea of its actual value, we can expand $\phi(x)$ around its zero at $x_c \simeq 5.56$, which results in

$$\phi_* = \frac{d\phi_c}{dx} (x_* - x_c), \quad (105)$$

where we remind that $x = m/H$. A quick numerical evaluation of the derivative of (34) gives $d\phi_c/dx \simeq -0.28\phi_{\text{ini}}$. The bound in (37) together with (105) then translates to the upper bound

$$\frac{H_*}{m} < \frac{H_c}{m} \left(1 + 0.14 \sqrt{\frac{\delta_{\text{eff}}^* - \delta}{0.1}} \right), \quad (106)$$

where we assumed $\delta_{\text{eff}}^* - \delta_* \leq 0.1$ for the expansion to be valid. In particular, if we saturate the upper bound in (52), i.e., $\delta_{\text{eff}}^* - \delta_* \simeq 0.1$, we obtain $H_*/m < 0.21$. At the same time, the transition has to happen before the tunneling probability becomes maximal when $\phi \rightarrow 0$, leaving only a rather narrow window:

$$0.18 < H_*/m \lesssim 0.21. \quad (107)$$

The only way to relax the upper bound would be to allow for $\delta_{\text{eff}} - \delta = \mathcal{O}(1)$, which would, however, lead to sizable oscillations of the trigger field around the true vacuum, incompatible with the assumptions made in our previous analysis. As a result, the range (107) constitutes a distinct prediction of our model, which we will confront with data in Sec. IV.

2. Nonadiabatic case

Evidently, (101) is only a particular solution of (95) to which we can still add a homogeneous one, corresponding to an entropy (or isocurvature) mode. After all, entropy modes are expected to be generated during inflation, because ϕ is a subdominant field component.

In the $k \rightarrow 0$ limit, the isocurvature solution can be derived from (96). In fact, we have solved the formally equivalent background equation (97) before when we derived (34). We therefore have (valid during radiation domination)

$$\delta\phi = \sqrt{2}\Gamma(5/4)\delta\phi_{\text{ini}}x^{-1/4}J_{1/4}(x/2) + \mathcal{O}(k^2/\mathcal{H}^2). \quad (108)$$

Recall that $x = m/H$ and $\delta\phi \rightarrow \delta\phi_{\text{ini}}$ as $x \rightarrow 0$ (corresponding to early times). Again, we have picked the attractor branch of the more general (two-parameter) family of solutions. We therefore find

$$\delta\phi = \frac{\delta\phi_{\text{ini}}}{\phi_{\text{ini}}} \phi, \quad (109)$$

where the adiabatic case corresponds to $\delta\phi_{\text{ini}} = 0$. The phase transition occurs when $\phi \rightarrow 0$ (as $H/m \rightarrow 0.18$). As $\delta\phi$ tracks the same evolution (up to a rescaling), we also have $\delta\phi \rightarrow 0$ at the same time, which, in turn, leads to a suppression of S_{NEDE}^* in (92) depending on how close ϕ_* (and, hence, $\delta\phi_*$) is to its zero. More explicitly,

$$S_{\text{NEDE}}^* = H_* \frac{\delta\phi_*}{\dot{\phi}_*} = \frac{1}{2} \frac{H_*}{m} \frac{\delta\phi_{\text{ini}}}{\phi_{\text{ini}}} \frac{\phi_*}{\frac{d}{dx}\phi_*}, \quad (110)$$

where we used (109) to derive the last line. Approximating $d\phi_c/dx \simeq d\phi_*/dx \simeq -0.28\phi_{\text{ini}}$ and using (105), we find

$$|S_{\text{NEDE}}^*| \simeq 10^{-5} \left(\frac{m(H_c^{-1} - H_*^{-1})}{0.1} \right) \left(\frac{H_*/m}{0.2} \right) \left(\frac{\delta\phi_{\text{ini}}/\phi_{\text{ini}}}{10^{-3}} \right). \quad (111)$$

For the above suggested parameter choice, we would get an entropy mode of the order of $\sim 10^{-5}$, almost matching the size of the adiabatic mode $\zeta \simeq 5 \times 10^{-5}$. However, as NEDE provides only a fraction ~ 0.1 of the energy budget, the effect of the entropy mode on the CMB power spectrum is further suppressed. To still get an idea of a reasonable phenomenological bound on S_{NEDE}^* , we can look at the CMB bound on neutrino density isocurvature, which is found to be of the order of $\mathcal{P}_{\nu\nu}/(\mathcal{P}_{\nu\nu} + \mathcal{P}_{\zeta\zeta}) \simeq 0.1$ [85], where \mathcal{P}_{xx} denotes the power spectrum for the mode x . Somewhat naively, it can be translated to $S_\nu/\zeta \lesssim \sqrt{0.1} \simeq 0.3$. In our case, this bound will be even weaker by a factor of $f_{\text{NEDE}} \sim 0.1$, resulting in $S_{\text{NEDE}}^*/\zeta \lesssim \mathcal{O}(1)$. This, in turn, might allow modes as large as $S_{\text{NEDE}}^* \sim 10^{-4}$ without leaving an observable imprint in the CMB. A more precise statement would require that we explicitly track the evolution of NEDE entropy modes in a Boltzmann code when doing a cosmological parameter fit. However, the message here is that we can easily go to a regime where entropy perturbations are phenomenologically irrelevant. This can be achieved by either having $\delta\phi_{\text{ini}}/\phi_{\text{ini}} \ll 1$ or imposing $m(H_c^{-1} - H_*^{-1}) \ll 1$. Because of (106), the latter condition requires $(\delta_{\text{eff}}^* - \delta) \ll 1$. Provided $\delta\phi_{\text{ini}}/\phi_{\text{ini}} < 10^{-3}$, the bound derived in (52) is already strong enough to guarantee a sufficient suppression of (111), as it ensures $m(H_c^{-1} - H_*^{-1}) \lesssim 0.1$ due to (106).

Alternatively, we can also enhance the entropy perturbation by increasing $\delta\phi_{\text{ini}}$, at least as long $\delta\phi_{\text{ini}}/\phi_{\text{ini}} \lesssim 0.1$ to make the perturbative treatment applicable. In fact, it might be possible that including entropy perturbations to our model even helps with resolving the Hubble tension. We leave a more detailed discussion of their phenomenology within NEDE to future work.

Finally, if we assume a standard inflationary scenario, the amplitude of the isocurvature perturbation in our trigger field is set by the Hubble scale during inflation, i.e., $\delta\phi_{\text{ini}} \sim H_I$. The absence of entropy perturbations

then provides an upper bound on the scale of inflation; explicitly,

$$\frac{H_I}{M_{pl}} < \mathcal{O}(1) \times 10^{-6} \left(\frac{0.1}{m(H_c^{-1} - H_*^{-1})} \right) \left(\frac{0.2}{H_*/m} \right) \times \left(\frac{\phi_{\text{ini}}/M_{pl}}{10^{-4}} \right), \quad (112)$$

where the $\mathcal{O}(1)$ factor accounts for the uncertainty in the phenomenological bound on S_{NEDE}^* . Given the before mentioned freedom in tuning the phase transition to be close to $\phi = 0$ and, hence, $(H_* - H_c) \ll 1$, this is not a particularly strong bound.

In summary, our model allows for both natural scenarios with and without a sizable contribution of entropy modes. Whether they are favored by the data remains to be seen. For the remainder of this paper, we simply assume they are sufficiently suppressed and set $\delta\phi_{\text{ini}} = 0$.

3. Hybrid early dark energy

It is also worth pointing out that our cosmological model can capture the physics of a second-order phase transition within a hybrid scenario. For example, this can be achieved in our two-field model (12) by setting $\beta = 0$ (this is only for simplicity, as $\beta = \gamma \leq 0$ and $\gamma \neq 0$ would also be compatible with this scenario). A schematic picture of such a potential is provided in Fig. 7. It gives rise to a smooth crossover from an early phase dominated by the vacuum energy stored in ψ to another phase where the field oscillates around its true vacuum. As before, this second phase is triggered when the Hubble drag is released and the subdominant field ϕ passes the inflection point (blue dot). The crucial observation is that a hierarchy of scales $m \ll M \sim \text{eV}$ is again required to have the transition occur around matter-radiation equality. As a result, the transition takes places on a cosmologically short timescale $1/M \ll 1/H \sim 1/m$, which then justifies the use of our cosmological model. Also in this case, (85) can be used to initialize fluctuations in the ψ fluid after the transition. The main difference with NEDE is that the oscillations around the true vacuum correspond to a relatively weak decay. In fact, when averaging over oscillation cycles, the field condensate gives rise to a fluid with a vanishing effective equation of state parameter, which cannot help with the Hubble tension. The only way of avoiding that conclusion and achieving a quicker decay would consist in tuning the potential such that close to its minimum ψ_2 there is an anharmonic direction approximately described by a monomial potential with $(\psi - \psi_2)^{2n}$ and $n \geq 2$. This, however, would constitute a fine-tuning similar to the one imposed in single-field EDE (see the discussion in Sec. IV C 3). We therefore believe that the first-order scenario is more natural, as it explains the quick decay of NEDE in terms of radiation and anisotropic stress.

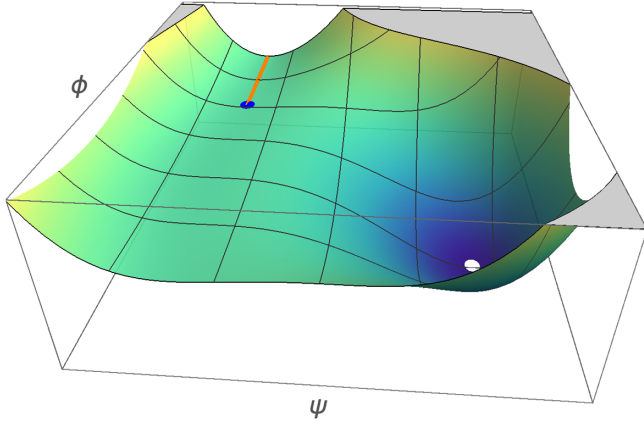


FIG. 7. The two-field potential in (12) can also give rise to a second-order phase transition. Similar to the first-order case, the field is initially frozen high up in a potential valley. When the Hubble drag is released around matter-radiation equality, it rolls along the almost flat bottom of the valley (orange line) until it crosses the inflection point (blue dot). Here its trajectory gets deflected toward a deep potential sink, around which it will oscillate while dissipating its potential energy.

D. Perturbation equations

We will model NEDE after the phase transition as a generic fluid, which at the level of perturbations is characterized by its density perturbation δ_{NEDE} , velocity divergence θ_{NEDE} , and anisotropic stress σ_{NEDE} (higher moments in the Boltzmann hierarchy we set to zero: $F_{\text{NEDE}/\geq 3} = 0$). For a time-dependent equation of state parameter $w(t)$, the perturbation equations in the synchronous gauge read [32,86] (for vanishing spatial curvature)

$$\begin{aligned} \dot{\delta}_{\text{NEDE}} = & -(1+w) \left(\frac{\theta_{\text{NEDE}}}{a} + \dot{h}/2 \right) - 3(c_s^2 - w)H\delta_{\text{NEDE}} \\ & - 9(1+w)(c_s^2 - c_a^2)H^2 \frac{\theta_{\text{NEDE}} a^2}{k^2}, \end{aligned} \quad (113a)$$

$$\dot{\theta}_{\text{NEDE}} = -(1 - 3c_s^2)H\theta_{\text{NEDE}} + \frac{c_s^2 k^2 / a}{1+w} \delta_{\text{NEDE}} - \frac{k^2}{a} \sigma_{\text{NEDE}}. \quad (113b)$$

Here, we introduced the effective sound speed in the fluid rest frame, $c_s^2 = \delta p_{\text{NEDE}}^{(\text{rest})} / \delta \rho_{\text{NEDE}}^{(\text{rest})}$, as well as the adiabatic sound speed

$$c_a^2 = w - \frac{1}{3} \frac{\dot{w}}{1+w} \frac{1}{H}. \quad (114)$$

With our assumption $w(t) \equiv w_{\text{NEDE}}^* = \text{const}$, we obtain $c_a^2 = w_{\text{NEDE}}^*$. In the general approach of Ref. [86], these equations are supplemented by (valid for $w \neq -1$)

$$(\dot{\sigma}_{\text{NEDE}} + 3H\sigma_{\text{NEDE}}) = \frac{8}{3} \frac{1}{1+w} c_{\text{vis}}^2 \left(\frac{\theta_{\text{NEDE}}}{a} + \frac{\dot{h}}{2} + 3\dot{\eta} \right), \quad (115)$$

describing the evolution of the anisotropic stress σ_{NEDE} , defined as

$$(\bar{\rho} + \bar{p})_{\text{NEDE}} \sigma_{\text{NEDE}} = - \left(\frac{k_i k_j}{\mathbf{k}^2} - \frac{1}{3} \delta_i^j \right) \Pi_{\psi,j}^i, \quad (116)$$

where $\Pi_{\psi,j}^i$ is the traceless part of the energy momentum tensor associated with NEDE. The above system is then subjected to the initial conditions specified in (85) and (86). Here, we will assume the vanishing of the viscosity parameter, $c_{\text{vis}} = 0$, which together with the initial condition (86) implies the trivial solution $\sigma_{\text{NEDE}} = 0$. Note that this is not in contradiction with having a nonvanishing anisotropic stress arising from the bubble condensate. The point is that the latter contribution is present only on scales that cannot be probed by CMB measurements provided condition (46) on the maximal bubble size $\simeq \bar{\beta}^{-1}$ is satisfied. As a possible avenue for future work, introducing anisotropic stress by having $c_{\text{vis}} \neq 0$ and setting the initial value to $\sigma_{\text{NEDE}}^* \neq 0$ will enable us to probe the parameter range of our model for which bubble signatures are imprinted in the observable CMB signal. In fact, we will provide a preliminary discussion of this possibility later in Sec. IV B 2 when we allow $c_{\text{vis}} \neq 0$ (but still have $\sigma_{\text{NEDE}}^* = 0$).

Finally, as discussed in Ref. [86], this parametrization can also mimic the behavior of free-streaming particles for the special choice $(w_{\text{NEDE}}^*, c_s^2, c_{\text{vis}}^2) = (1/3, 1/3, 1/3)$. However, as we will briefly discuss later, this possibility is not favored by the data, and, therefore, we will mostly explore the parameter regime for which $c_{\text{vis}} = 0$ and $w_{\text{NEDE}}^* = c_s^2 = 2/3$. As discussed in Sec. II D, this is motivated by our expectation that the scalar field condensate is dominated by kinetic energy (rather than potential energy) and small-scale anisotropic stress, which drives the effective equation of state of the NEDE fluid toward stiffer values. Also, we expect c_s^2 and w_{NEDE}^* to be well described by a constant, because the NEDE fluid is relevant only for a narrow redshift window around z_* , making cosmological observables rather insensitive to its time dependence.

E. NEDE phenomenology

The primary effect of NEDE is to lift $H(z)$ to higher values before recombination. This changes the anchoring of the inverse distance ladder through a reduction of the (comoving) sound horizon:

$$r_s(z_{\text{rec}}) = \int_{z_{\text{rec}}}^{\infty} dz \frac{v(z)}{H(z)}, \quad (117)$$

where $v(z)$ denotes the sound speed in the photon-baryon fluid. In order to keep the angular scales we observe through CMB and BAO measurements fixed, the corresponding distance scales have to be reduced, which is achieved by increasing H_0 . To be explicit, the CMB observation constrains the angular position of the first peak in the CMB power spectrum:

$$\theta_{\text{rec}} = \frac{r_s(z_{\text{rec}})}{D_{\text{rec}}}. \quad (118)$$

A smaller value of the sound horizon at recombination, $r_s(z_{\text{rec}})$, then needs to be balanced by reducing the comoving distance D_{rec} between us and the surface of last scattering, which due to (43) is achieved by anchoring the expansion history at a higher value of H_0 . This, in turn, alleviates the Hubble tension (see also [12,33,35] for different implementations of the same mechanism).

A similar argument holds in the case of BAO measurements. There we observe the imprint the sound horizon at radiation drag time $z_d \lesssim z_{\text{rec}}$ leaves in the visible matter spectrum in our low-redshift environment ($0.1 \lesssim z \lesssim 1$). Again, the NEDE modification together with an increased value of H_0 lowers both $r_s(z_d)$ and the comoving distance of the observed galaxies in a way which leaves the angular position of the peak in the matter power spectrum approximately invariant. It is this balanced interplay between the shortening of the sound horizon and the reduction in the angular distances where late-time modifications fall short. As they modify only the expansion history after recombination, they leave the sound horizon invariant (or change it too little by changing the inferences we draw on other Λ CDM parameters), which leads to unwanted shifts in the

angular positions of both CMB and BAO peaks (for a more detailed discussion of this BAO obstruction, see, for example, [12,13,87–89]).

The above early-time mechanism is shared by different modifications of Λ CDM, most notably DR (or extra sterile neutrinos), EDE, and ADE as well as our proposed model. However, this is not enough to explain the phenomenological success of these models. After all, the CMB measurement is constraining not only the position of the first angular peak but also the shape of the angular power spectrum in a wide k range. As photon-baryon acoustic oscillations are sensitive to the evolution of the new energy component at both the background and perturbation level through their gravitational coupling, any modification at that time risks to alter the CMB power spectrum beyond what is phenomenologically allowed. To understand how new early dark energy and similar proposals get around this obstacle, we follow the discussion in Ref. [35] and consider the Weyl potential:

$$\psi + \phi = \eta + \frac{1}{2k^2} \left[\frac{d^2 h}{d\tau^2} + 6 \frac{d^2 \eta}{d\tau^2} \right], \quad (119)$$

where η and h are the metric fluctuations in the synchronous gauge introduced in (72). We plot it in the upper two panels in Fig. 8 for a mode which enters the CMB sound horizon before (right) and after (left) the phase transition, indicated by the vertical dashed lines.

The decay of the Weyl potential during the radiation-dominated epoch leads to driving effects in the photon-baryon fluid [90]. Modes that have entered the horizon before matter domination are most sensitive to this effect. This can be seen in the upper two panels in Fig. 8, showing

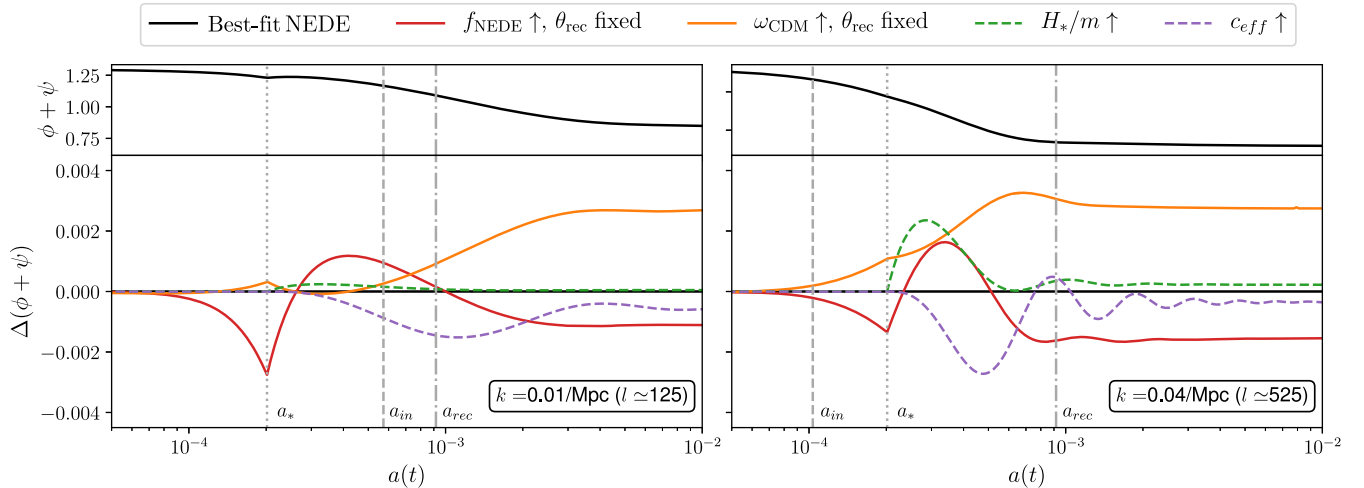


FIG. 8. Relative change in the Weyl potential with our best-fit NEDE as baseline; see Table I. The dashed line corresponds to the time where modes in the photon-baryon fluid enter their sound horizon and become sensitive to changes in the Weyl potential due to driving effects arising from its decay. In the left plot this happens after and in the right one before the decay at a_* . An increase in f_{NEDE} is approximately balanced by an increase in ω_{cdm} . The effect of the trigger parameter H_*/m cannot be fully compensated by a change in a k -independent sound speed and, therefore, mimics the effect of a k -dependent sound speed as in single-field EDE.

a stronger decay for the larger k mode. As NEDE makes its most significant contribution when radiation still dominates (around the dotted line in Fig. 8), it affects the driving of CMB acoustic oscillations through the changes it introduces to the expansion history and through its own acoustic oscillations. In the two bottom panels in Fig. 8, we depict the effect different NEDE parameters have on the Weyl potential. The best-fit NEDE model serves as the baseline. The red line corresponds to a 10% increase in f_{NEDE} . Here, the corresponding shift in the CMB peak is compensated by increasing H_0 such that $\theta_{\text{rec}} = \text{const.}$ That way we make sure that the leading-order degeneracy has been accounted for. Before the transition at $a = a_*$, NEDE behaves as a cosmological constant, and the corresponding repulsive boost leads to a quicker decay of the Weyl potential. This behavior is reverted at the point of the phase transition, because there NEDE is converted from vacuum energy to a decaying fluid with equation of state parameter $w_{\text{NEDE}}^* > 1/3$. This fluid then supports its own acoustic oscillations that themselves source the gravitational potential. As has been discussed in the context of ADE, their subsequent Jeans stabilization then leads to an overall suppression of the potential. This excess decay can be compensated at leading order by increasing ω_{cdm} , which is depicted by the yellow line. In Fig. 9, we show how this cancellation plays out on the level of the CMB power spectrum. While an increase in w_{cdm} is able to balance the additional power on angular scales $\ell < 700$, this no longer works on scales > 700 . Here, the dominant effect arises from changes in the diffusion damping scale r_d . It controls the suppression of the power spectrum at large ℓ , which goes as $\exp[-2(r_d/\lambda)^2]$ [91]. If we replace the wavelength λ of a given mode with its angular scale $\ell \simeq 2\pi D_{\text{rec}}/\lambda$ and use (118), we obtain a diffusion-induced suppression factor:

$$\exp\left[-2\left(\frac{r_d \theta_{\text{rec}}}{r_s 2\pi} \ell\right)^2\right]. \quad (120)$$

As θ_{rec} is directly (and very precisely) fixed by CMB measurements, the damping is solely controlled by the ratio r_d/r_s , where both $r_s[H]$ and $r_d[H]$ are functionals of $H(z)$. Now we use that the mean squared diffusion distance is given by [92–94] $r_d^2 \simeq \int_{z_{\text{rec}}}^{\infty} dz F_d(z)/H$, where $F_d(z)$ depends on the fraction of free electrons, the baryon-photon ratio, and the cross section for Thomson scattering. From (117), we then derive the response to a functional change $H(z) \rightarrow H(z) + \Delta H(z)$:

$$\begin{aligned} \Delta\left(\frac{r_d}{r_s}\right) &\simeq \frac{r_d}{r_s^2} \frac{v}{2} \int_{z_{\text{rec}}}^{\infty} dz \frac{\Delta H}{H^2(z)}, \\ &\simeq -\frac{r_d}{r_s} \frac{\Delta r_s}{2r_s}, \end{aligned} \quad (121)$$

where we assumed that both functions $v(z)$ and $F_d(z)$ are approximately constant in the redshift window in which the integral picks up its main contribution. This shows that increasing $H(z)$ before recombination (or, equivalently, lowering r_s) generically increases the ratio r_d/r_s which, in turn, dampens the power spectrum at large ℓ (in agreement with Refs. [33,95]).²¹ As can be seen in Fig. 9, we cannot compensate this loss of power by increasing w_{cdm} , which has too little influence on small scales. Instead, we can eliminate it by reducing the spectral tilt, corresponding to an increase in n_s toward 1, and slightly adapting the amplitude of primordial perturbations, A_s (dotted line).

In summary, we have identified the leading-order degeneracies in the CMB measurement introduced by NEDE:

- (1) $f_{\text{NEDE}} > 0$ introduces a shift in the angular peaks which is compensated by increasing H_0 .
- (2) Acoustic oscillations in the NEDE fluid lead to an excess decay of the Weyl potential, which, in turn, affects the driving of acoustic CMB modes that enter the sound horizon around the decay time ($\ell \simeq 300$). This can be counteracted by increasing w_{cdm} .
- (3) A shortening of the CMB damping scale r_D , due to the modified expansion history, affects modes on short scales and needs to be compensated by reducing the spectral tilt through $n_s \rightarrow 1$ and increasing the amplitude A_s .

We will confirm these relations when we perform the cosmological parameter extraction in Sec. IV. Of course, these degeneracies are not perfect and, therefore, lead to *Planck* data residuals, which we will discuss when comparing our best-fit model to its EDE-type competitors in Sec. IV C (see also Fig. 17).

Finally, we discuss the effect of the trigger parameter H_*/m , which is a distinctive feature of our model. In Fig. 6, we have seen that varying H_*/m while keeping z_* fixed changes the initial amplitude of the NEDE perturbations. This effect is relevant only for modes that have entered the particle horizon before the decay (for $z_* = 5000$, this corresponds to $\ell \gtrsim 180$). This can also be seen in Fig. 8, where only the larger k mode responds to the increase in H_*/m (green dashed line). The corresponding deformation of the Weyl potential at large k (right panel) can then be balanced by also increasing the effective sound speed c_{eff} . However, this balancing does not work for the smaller k mode (left panel). From this, we conclude that the effect of the trigger parameter cannot be captured by a k -independent sound speed. This resonates with the observation in Refs. [34,35], where it was found that the phenomenological success of single-field EDE relies on having a k -dependent sound speed.

²¹Note that our analytic estimate agrees with the numerical result obtained in Ref. [33] (see their Fig. 2).

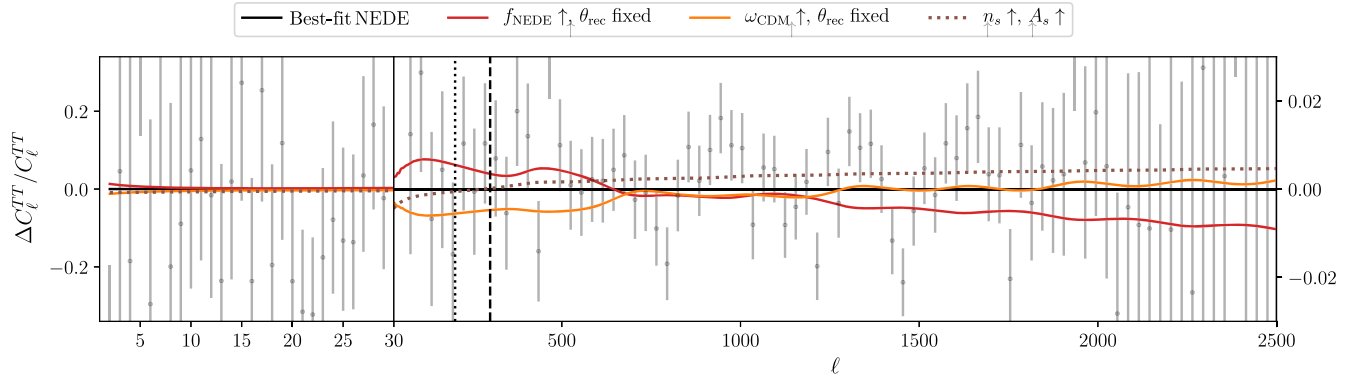


FIG. 9. *Planck* data residuals (gray dots) with respect to the best-fit NEDE model as in Table I (combined analysis with SH₀ES). Increasing f_{NEDE} while keeping the angular position of the first CMB peak fixed leads to more power on large and less power on small scales due to changes in the Weyl potential and stronger damping effects, respectively. The first effect can be balanced by increasing ω_{cdm} and the second one by reducing the spectral tilt of the primordial spectrum. The dashed and dotted vertical lines highlight the modes that, respectively, entered the sound and particle horizon at the time of the NEDE phase transition.

IV. COSMOLOGICAL PARAMETER EXTRACTION

A. Methodology

The cosmological parameter extraction is performed with the Metropolis-Hastings algorithm using its implementation in the Monte Carlo Markov chain (MCMC) code *MontePython* [96,97]. We assume that chains are converged if the Gelman-Rubin criterion [98] yields $R - 1 < 0.01$. To integrate the background cosmology and solve the linear perturbation equations, we use the Boltzmann code *CLASS* (Cosmic Linear Anisotropy Solving System) [99] supplemented by our NEDE modifications, publicly available on GitHub [100]. Also, we add the ADE model to *CLASS* following Ref. [35] and use the publicly available implementation of EDE [101].

We will use different combinations of the following datasets:

- (i) *Local H_0 measurement* (SH₀ES).—The most recent SH₀ES value [4], $H_0 = 74.03 \pm 1.42 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (68% C.L.), using a Gaussian prior.
- (ii) *High- z supernovae* (Pantheon).—The Pantheon dataset [102] comprised of 1048 SNe Ia in a range $0.01 < z < 2.3$. Since the absolute magnitude scale M is treated as a nuisance parameter, this dataset constrains the shape of the expansion history in the corresponding redshift range rather than H_0 itself.
- (iii) *Baryonic acoustic oscillations* (large- z and small- z BAO).—Anisotropic BAO measurements at redshift $z = 0.38, 0.51$, and 0.61 based on the CMASS and LOWZ galaxy samples of BOSS DR 12 [103], as well as small- z , isotropic BAO measurements of the 6dF Galaxy Survey [104] and the SDSS DR7 main Galaxy sample [105] at $z = 0.106$ and $z = 0.15$, respectively.
- (iv) *Large-scale structure* (LSS).—Constraints on the growth of structure, quantified by $f\sigma_8(z)$, from

redshift-space distortions at redshift $z = 0.38, 0.51$, and 0.61 based on the CMASS and LOWZ galaxy samples of BOSS DR 12 [103]. We note that this measurement is implemented in *MontePython* in conjunction with the large- z BAO measurement, taking into account their (strong) covariance. When we calculate the tension with LSS, we use the $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ measurement from a combined tomographic weak gravitational lensing analysis of the Kilo Degree Survey and the Dark Energy Survey [111], $S_8 = 0.762^{+0.025}_{-0.024}$ (68% C.L.).

- (v) *Cosmic microwave background* (*Planck* 2018).—The *Planck* 2018 TT, TE, EE, and lensing likelihood with all nuisance parameters included [106].
- (vi) *Big bang nucleosynthesis* (BBN).—The primordial helium abundance from Ref. [107], $Y_p = 0.2449 \pm 0.0040$ (68% C.L.), via a Gaussian likelihood. This is the conservative choice given the tighter bounds from Ref. [108], $Y_p = 0.2446 \pm 0.0029$ (68% C.L.). As the helium abundance is sensitive to the early expansion history, it mainly constrains the effective number of relativistic degrees of freedom N_{eff} during the BBN epoch. *CLASS* infers the value of Y_p as a function of $(\omega_b, N_{\text{eff}})$ from an interpolation table created with *PARthENoPE* v1.2 [109] with a nuclear rate $d(p, \gamma)^3\text{He}$ extrapolated from observations by Ref. [110].

We will use two different combinations of the above datasets for our parameter extraction:

- (1) combined analysis without SH₀ES: Pantheon, BAO, LSS, *Planck* 2018, and BBN;
- (2) combined analysis with SH₀ES: Pantheon, BAO, LSS, *Planck* 2018, SH₀ES, and BBN.

We introduce two new components to *CLASS*: the trigger field ϕ and the NEDE fluid. The former we describe in terms of a massive scalar field with mass m and

subdominant energy density $\bar{\rho}_\phi$, and the latter is described on the background level as a fluid with time dependence as in (67). Here, we discuss the implementation of both components in turn.

1. Trigger field

This sector requires us to introduce three new input parameters to CLASS, the field's mass m , its initial value ϕ_{ini} as well as the trigger parameter $x_*^{-1} = H_*/m$, defining the transition time with respect to the evolution of ϕ . The background value of the scalar field is initialized at the same time as the other background quantities.²² Specifically, we put ϕ and $\dot{\phi}$ on the attractor (100). We further require ϕ to be subdominant, which due to (29) is guaranteed if $\phi_{\text{ini}}/M_{pl} \ll 1$. To be specific, in our baseline model we will set $\phi_{\text{ini}}/M_{pl} = 10^{-4}$, which is also compatible with the lower bound in (58). In any case, its precise initial value does not matter from a phenomenological perspective—at least if it is subdominant—because ϕ_{ini} drops out of the matching conditions in (85). This is because they are proportional to the ratio $\delta\phi/\dot{\phi}$, which due to (100) and (101) is indeed independent of ϕ_{ini} . We also checked explicitly that the power spectrum shows a negligible dependence on ϕ_{ini} (small compared to numerical uncertainties). The field is then propagated forward in time via the background equation (97) until the decay of NEDE is triggered when the condition $H = m/x_*$ is met. As argued before in Sec. III C, our model predicts a rather narrow range of possible values for $x_*^{-1} = H_*/m$ detailed in (107). For our base model, we will therefore fix it to $H_*/m = 0.2$, which is the most natural expectation. In fact, a smaller value would require an increasingly delicate parameter tuning. We will also perform one analysis where we allow it to vary within its allowed range.

2. NEDE field

The NEDE component is straightforwardly implemented in CLASS by using the parametrization in (67). We use $\bar{\rho}_{\text{NEDE}}^*$, i.e., the constant value of NEDE before its decay, as a new input parameter. The fluctuations in the NEDE fluid are set to zero before the transition and initialized right after by using our matching equations (85) together with the identification $\delta q_*/\dot{q}_* = \delta\phi_*/\dot{\phi}_*$, which closes the dynamical system. To properly resolve the matching dynamics and avoid numerical artifacts, we increase the numerical precision by lowering the step size around the transition. After the transition, the fluctuations are propagated forward in time using the equations (113) describing the dynamics of generic fluid perturbations, fully characterized by the three parameters $(w_{\text{NEDE}}^*, c_s^2, c_{\text{vis}}^2)$. Following the discussion in

Sec. II D, we will make the choice $(2/3, 2/3, 0)$ for our base model. However, we will also consider different extensions that include the case of free-streaming radiation, which according to Ref. [86] can be approximated by $(1/3, 1/3, 1/3)$, and different other cases with freely varying fluid parameters.

With these choices, our base model comes equipped with only two undetermined parameters, $\{\bar{\rho}_{\text{NEDE}}^*, m\}$, which will be inferred from a cosmological parameter extraction. To be specific, we vary $\{f_{\text{NEDE}}, \log_{10}(m/m_0)\}$, where $m_0 = 1/\text{Mpc}$ in order to keep with the CLASS convention, and f_{NEDE} is the fraction of NEDE at the time of decay, related to $\bar{\rho}_{\text{NEDE}}^*$ via (21). Rather than using the CLASS input parameter m , we will present our results in terms of the more intuitive quantity z_* , the redshift at decay time. Its value depends on the expansion history and needs to be inferred numerically from $H(z_*)$ (which is fixed by the product $x_*^{-1} \times m$). During radiation domination, it is approximately given by

$$z_* \simeq 1.6 \times 10^4 (1 - f_{\text{NEDE}})^{1/4} \left(\frac{3.9}{g_*}\right)^{1/4} \times \left(\frac{H_*/m}{0.2}\right) 10^{0.5 \log_{10}(m/m_0) - 1.75}, \quad (122)$$

where we used (30). As expected, the main dependence is due to m , but also f_{NEDE} (and $x_*^{-1} = H_*/m$) go into its determination. The closure of the budget equation is finally ensured by fixing the cosmological constant Λ [using iterative runs to find an accurate estimate for $\bar{\rho}_{\text{NEDE}}(t_0)$].

We will compare NEDE to the ΛCDM model and its most prominent early-time competitors DR, ADE, and single-field EDE. In all cases, we vary the six ΛCDM parameters on which we impose flat priors with standard boundaries, specifically $\{\omega_b, \omega_{\text{cdm}}, h, \ln 10^{10} A_s, n_s, \tau_{\text{reio}}\}$, where $\omega_b = \Omega_b h^2$ and $\omega_{\text{cdm}} = \Omega_{\text{cdm}} h^2$ are the dimensionless baryon and cold dark matter density, respectively, h is the dimensionless Hubble constant defined via $H_0 \equiv h \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, τ_{reio} is the reionization optical depth, and $\ln 10^{10} A_s$ and n_s are the initial amplitude and tilt of the primordial curvature spectrum at $k = 0.05 \text{ Mpc}^{-1}$, respectively. For the neutrino sector, we use the *Planck* convention and model it in terms of two massless and one massive species with $M_\nu = 0.06 \text{ eV}$, where we ensure that the effective number of relativistic degrees of freedom is $N_{\text{eff}} = 3.046$ (except for dark radiation, where N_{eff} is promoted to a free parameter). We impose the flat prior $0 < f_{\text{NEDE}} < 0.3$ where the upper boundary could easily be removed, as it is never probed as shown in Fig. 10.

There is a subtlety related to the mass parameter (or the decay time, equivalently), which applies equally to all models of the EDE type (also see the discussion in Sec. III of Ref. [34]). When $f_{\text{NEDE}} \rightarrow 0$, the ΛCDM model is

²²For concreteness, we use the CLASS standard value $a = 10^{-14}$ as the initial time.

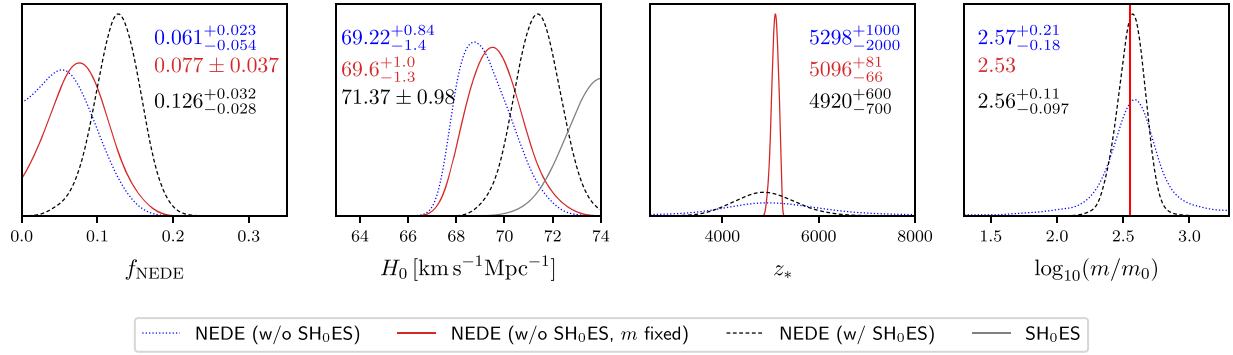


FIG. 10. The marginalized posterior distribution for our combined analysis without SH_0ES . If we allow the mass parameter m , and hence the decay time z_* , to vary, the increased sampling volume as $f_{\text{NEDE}} \rightarrow 0$ leads to an unphysical enhancement for $f_{\text{NEDE}} \ll 1$ (blue dotted line). We avoid this problem by fixing $\log_{10}(m/m_0) = 2.58$ (red solid line), which leads to a Gaussian distribution with 1.9 σ evidence for $f_{\text{NEDE}} > 0$.

recovered and all values of m (or z_*) are exactly equivalent, corresponding to an enhancement of the relevant sampling volume. This, in turn, leads to unphysical artifacts in the posterior distributions which are “pulled” toward the degenerate point at $f_{\text{NEDE}} = 0$ and start to deviate from a Gaussian shape. This can be seen in Fig. 10, where we depict different posteriors for our combined analysis without SH_0ES . In particular, if we vary both parameters, m and f_{NEDE} , corresponding to the blue dotted curve, there is a probability enhancement for $f_{\text{NEDE}} < 0.06$. This is due to the fact that the constraints are no longer mainly driven by data but by the prior volume.

In order to get around this technical obstruction, we perform a two-step analysis where we first vary both m and f_{NEDE} to obtain the best-fit values (which are, of course, not affected by this volume effect). Then we perform a second run where we fix m close to its best-fit value and vary only f_{NEDE} . To be precise, the prior boundaries we impose for the first and second run are²³

$$1.3 < \log_{10}(m/m_0) < 3.3, \quad (123a)$$

$$\log_{10}(m/m_0) = 2.58, \quad (123b)$$

corresponding to the dotted blue and solid red line, respectively. Because of (122), the first prior probes the rather wide range $1.3 < 10^{-3}z_* < 13.1$, whereas the second prior corresponds to a narrow redshift range around 5100. This completely avoids the issue as can be seen from the red curves in Fig. 10, which are nearly Gaussian for both f_{NEDE} and H_0 . Strictly speaking, this analysis assumes that our underlying theory fixes m to a very specific value, and, therefore, it can provide us only with the posteriors valid in this rather special case. Nevertheless, since

²³We use different prior boundaries in the case of free-streaming radiation where the decay happens significantly earlier at $10^{-3}z_* = 35_{-20}^{+8}$.

$f_{\text{NEDE}} = 0$ is still included in the allowed parameter range, we believe it is a consistent way of estimating the statistical preference for our model. We will apply the same procedure when we analyze ADE and EDE without SH_0ES in order to ensure a fair comparison. The more general case, where $\log_{10}[m/m_0]$ is allowed to vary freely, will still be investigated when we include the SH_0ES measurement. In that case, corresponding to the black dashed curve, f_{NEDE} never probes the degenerate parameter region due to the statistical “pull” toward higher values of H_0 , and conventional χ^2 statistics can be used to assess the quality of the NEDE fit.

Finally, we will also discuss three different extensions of our base model where we allow (w_{NEDE}^* , c_s , f_{NEDE} , $\log_{10}[m/m_0]$), (c_{vis} , f_{NEDE} , $\log_{10}[m/m_0]$), and (H_*/m , f_{NEDE}) to vary while keeping the other parameters fixed.

B. Discussion of results

1. Base model

The standard cosmological parameters and their covariances as inferred from our joined analysis (with and without SH_0ES) are presented in Table I and Fig. 11, respectively. In Fig. 12, we depict the two parameters of our baseline model as well as the sound horizon at radiation drag, $r_s^d = r_s(z_d)$, and the growth parameter $S_8 = \sigma_8 \sqrt{\Omega_m}/0.3$. The main effect of NEDE is to increase the mean value of H_0 while also enlarging its uncertainties. As a result, we find an overlap with the local SH_0ES measurement at the 95% confidence level (light gray and red contours); more specifically, we find $H_0 = 71.4 \pm 1.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$ when we include the local measurement, which leads to a total χ^2 improvement of 15.6. When we do not include the local measurement, we find $H_0 = 69.6_{-1.3}^{+1.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$, which corresponds to a reduced tension of 2.5 σ (down from 4.3 σ).

As explained before, the main effect of NEDE is to lower the sound horizon $r_s(z)$, which is balanced by a higher

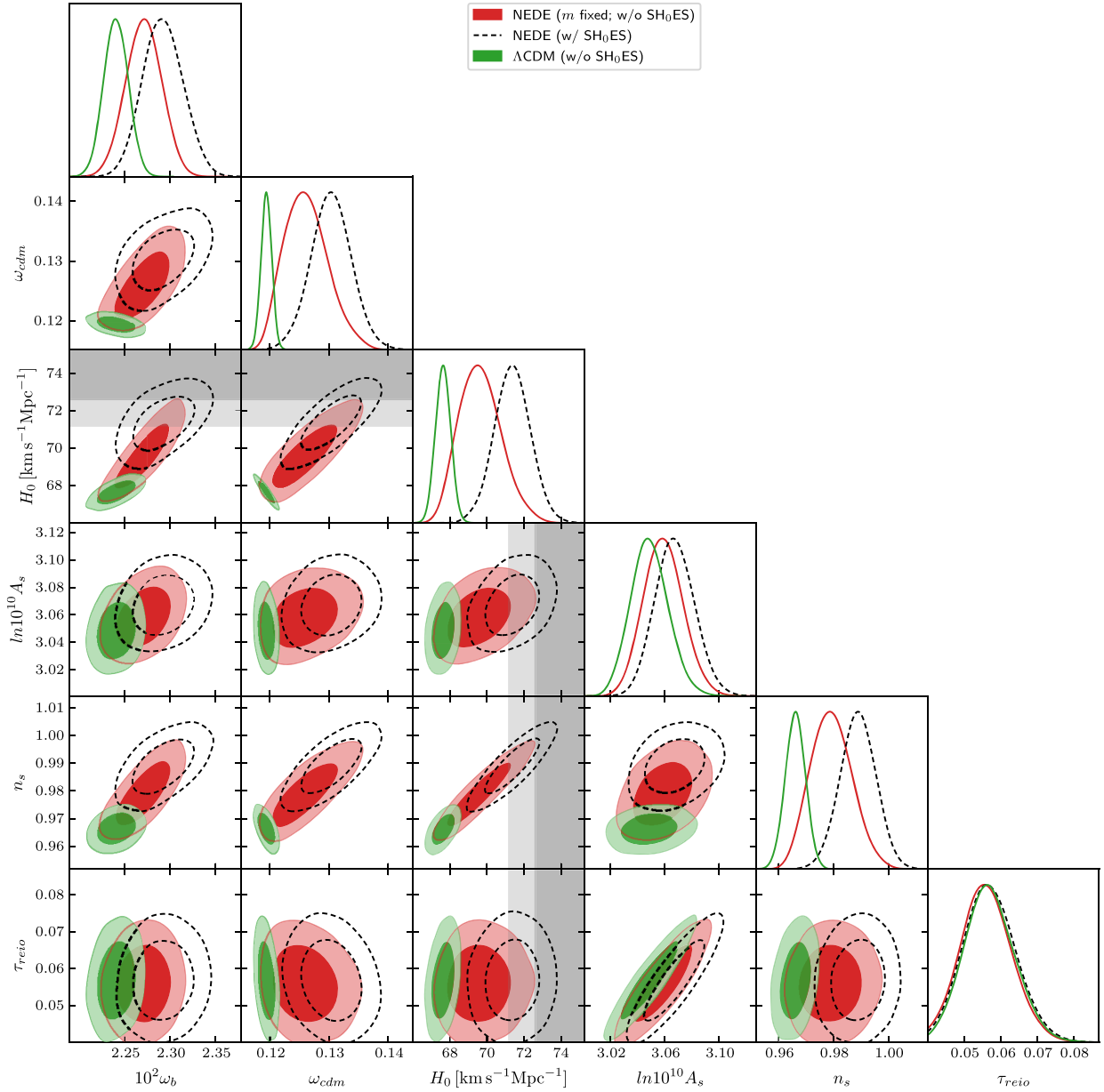


FIG. 11. Covariances and marginalized distributions of standard cosmological parameters for NEDE (red) and Λ CDM (green) from a run to the combined datasets without SH₀ES. The result of the analysis with SH₀ES included is presented as the dashed contours. Here and henceforth, the darker and lighter shades correspond to the 68% C.L. and the 95% C.L., respectively. The gray bands depict the constraints from the SH₀ES measurement. The 95% C.L. contours of SH₀ES and NEDE largely overlap.

value of H_0 . To be specific, we obtain a 4% reduction of r_s^d , from 147.20 ± 0.23 Mpc (Λ CDM without SH₀ES) down to $141.0_{-1.7}^{+1.6}$ Mpc (NEDE with SH₀ES), which is then approximately balanced by a 6% increase in H_0 . We also report a 4.3σ evidence for a nonvanishing NEDE component when we include the local measurement of H_0 , corresponding to $f_{\text{NEDE}} = 12.6_{-2.9}^{+3.2}\%$, and if we do not include the local measurement, the evidence for NEDE still reaches 1.9σ , corresponding to a mild preference for our model driven mostly by *Planck* data. This resonates with our χ^2 results detailed in Table II. They show that the

improvement is almost entirely due to the SH₀ES dataset with $\Delta\chi^2(H_0) = -13.8$. However, the important point is that the quality of the fit to the other datasets does not become worse; in fact, it even improves by $\Delta\chi^2(\setminus H_0) = -1.8$. This beneficial effect is even more pronounced for our analysis without SH₀ES, where we obtain $\Delta\chi^2(\text{total}) = -2.9$. In other words, the evidence for NEDE is mainly but not entirely driven by the local measurement of H_0 . More importantly, none of the fits to other datasets shows any significant deterioration, let alone the emergence of a new tension. On closer inspection,

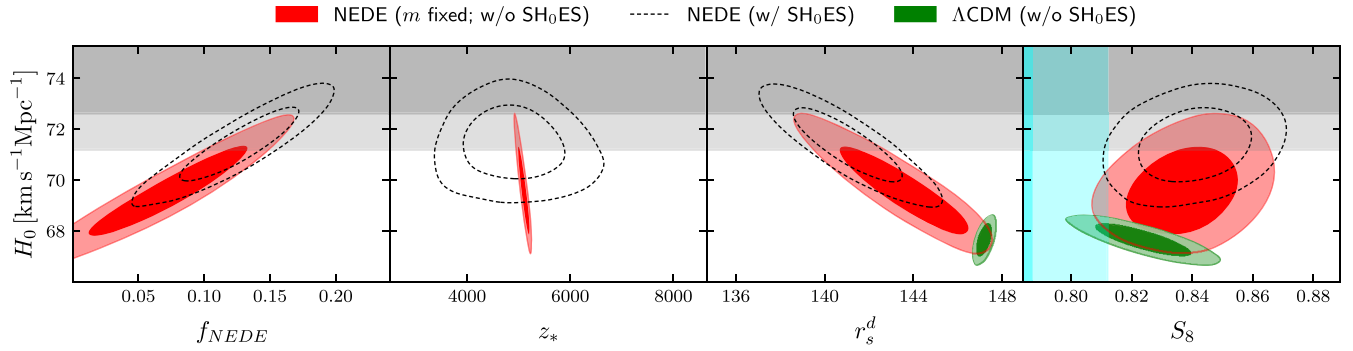


FIG. 12. The Hubble parameter today vs a subset of parameters for the combined analysis with (dashed contour) and without (red filled contour) SH₀ES. The light blue band corresponds to the constraint on S_8 from weak gravitational lensing combining different cosmic shear surveys [111]. The gray band represents the SH₀ES measurement. Without SH₀ES, the decay time is constrained to a narrow redshift range to avoid sampling volume artifacts.

we see that NEDE improves the fit to the CMB temperature and polarization spectrum at both low and high ℓ , specifically $\Delta\chi^2(\text{TT, EE, TE}) = -3.6$, at the prize of slightly worsening the lensing fit with $\Delta\chi^2(\text{lens}) = +0.4$, provided we include SH₀ES. We will see later when discussing Fig. 17 that this improvement can be attributed to specific kinky and oscillatory features in the power spectrum, which also provides a potential means of detecting NEDE in future surveys or discriminating it from its competitors.

The decay time turns out to be a relatively solid prediction with little sensitivity to the prior on H_0 .

Specifically, we obtain $z_* = 4920^{+620}_{-730}$ (with SH₀ES) which is compatible with our assumption that the transition occurs during radiation domination but close enough to recombination that it can still influence the value of the sound horizon [note that the integral in (117) picks up its main contribution at its lower boundary].

Looking at different Λ CDM parameters, we can identify the approximate degeneracies predicted in the last sections: First, there is a strong positive (negative) correlation between H_0 and f_{NEDE} (r_s^d) as expected. The effect of a larger value of H_0 is then balanced by a larger value of ω_{cdm}

TABLE I. The mean value and $\pm 1\sigma$ error (with best-fit value in parentheses) of the cosmological parameters from our combined analyses for Λ CDM and NEDE with and without imposing the SH₀ES prior on H_0 .

Parameter	Λ CDM		NEDE	
	With H_0	Without H_0	With H_0	Without H_0
$100\omega_b$	$2.251^{+0.014}_{-0.013}$ (2.251)	$2.240^{+0.014}_{-0.014}$ (2.242)	$2.292^{+0.022}_{-0.024}$ (2.297)	$2.271^{+0.020}_{-0.020}$ (2.271)
ω_{cdm}	$0.1184^{+0.0009}_{-0.0009}$ (0.1183)	$0.1194^{+0.0009}_{-0.0009}$ (0.1194)	$0.1304^{+0.0034}_{-0.0035}$ (0.1306)	$0.1261^{+0.0033}_{-0.0042}$ (0.1254)
h	$0.6813^{+0.0041}_{-0.0041}$ (0.6816)	$0.6763^{+0.0042}_{-0.0042}$ (0.6764)	$0.714^{+0.010}_{-0.010}$ (0.715)	$0.696^{+0.010}_{-0.013}$ (0.695)
$\ln 10^{10} A_s$	$3.053^{+0.014}_{-0.016}$ (3.053)	$3.049^{+0.013}_{-0.015}$ (3.050)	$3.067^{+0.014}_{-0.015}$ (3.068)	$3.059^{+0.014}_{-0.016}$ (3.058)
n_s	$0.9686^{+0.0037}_{-0.0037}$ (0.9698)	$0.9661^{+0.0038}_{-0.0037}$ (0.9672)	$0.9889^{+0.0067}_{-0.0066}$ (0.9912)	$0.9792^{+0.0073}_{-0.0082}$ (0.9794)
τ_{reio}	$0.0599^{+0.0071}_{-0.0078}$ (0.0598)	$0.0570^{+0.0064}_{-0.0075}$ (0.0572)	$0.0571^{+0.0068}_{-0.0077}$ (0.0572)	$0.0562^{+0.0065}_{-0.0075}$ (0.0558)
f_{NEDE}	$0.126^{+0.032}_{-0.029}$ (0.1296)	$0.077^{+0.038}_{-0.040}$ (0.072)
$\log_{10}(m/m_0)$	$2.56^{+0.12}_{-0.10}$ (2.57)	2.58 (fixed)
H_*/m	0.2 (fixed)	0.2 (fixed)
w_{eff}^*	2/3 (fixed)	2/3 (fixed)
c_s^2	2/3 (fixed)	2/3 (fixed)
σ_8	$0.8090^{+0.0060}_{-0.0065}$ (0.8092)	$0.8104^{+0.0057}_{-0.0061}$ (0.8110)	$0.841^{+0.010}_{-0.010}$ (0.841)	$0.828^{+0.010}_{-0.012}$ (0.826)
S_8	$0.814^{+0.010}_{-0.010}$	$0.824^{+0.010}_{-0.010}$	$0.841^{+0.012}_{-0.012}$	$0.837^{+0.012}_{-0.012}$
r_s^d [Mpc]	$147.40^{+0.23}_{-0.23}$ (147.38)	$147.20^{+0.23}_{-0.23}$ (147.19)	$141.0^{+1.6}_{-1.7}$ (140.9)	$143.5^{+2.1}_{-1.8}$ (143.8)
z_*	4920^{+620}_{-730} (4960)	5100^{+80}_{-70} (5110)
Tension SH ₀ ES	...	4.3σ	...	2.5σ
Tension S_8	1.9σ	2.3σ	2.8σ	2.7σ
$\Delta\chi^2$	0	0	-15.6	-2.9
$f_{\text{NEDE}} \neq 0$	4.3σ	1.9σ

TABLE II. The best-fit $\chi^2 = -2\ln(\mathcal{L})$ for each dataset from combined analysis with and without SH₀ES. The quality of the overall fit is assessed in terms of $\Delta\chi^2 = \chi^2(\text{NEDE}) - \chi^2(\Lambda\text{CDM})$, corresponding to the ratios between the maximum likelihoods \mathcal{L} .

SH ₀ ES Extension ΔDOF	ΛCDM		NEDE		NEDE extensions			
	With H_0	Without H_0	With H_0	Without H_0	With H_0	With H_0	With H_0	With H_0
	H_*/m	w_{NEDE}, c_s	c_{vis}	FS ^a
	0	0	2	1 (2) ^b	2 (3) ^b	4	3	2
<i>Planck</i> high- ℓ TT, TE, EE	2348.9	2347.1	2348.2	2345.9	2348.6	2348.4	2348.7	2352.1
<i>Planck</i> low- ℓ TT	22.7	23.1	20.8	21.7	20.8	20.7	20.8	21.4
<i>Planck</i> low- ℓ EE	397.3	396.7	396.4	396.2	396.1	396.3	396.3	396.8
<i>Planck</i> lensing	9.1	8.8	9.5	9.1	9.5	9.6	9.5	9.3
BAO low- z	1.6	1.2	1.8	1.4	1.8	1.7	1.8	1.8
BAO high- z +LSS	5.9	6.7	6.8	6.5	6.8	6.9	6.8	6.3
Pantheon	1027.0	1027.3	1027.3	1027.1	1027.3	1027.0	1027.0	1027.1
BBN	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
SH ₀ ES	17.1	...	3.3	...	3.1	2.5	3.1	9.4
χ^2	3829.6	3810.7	3814.0	3807.8	3813.9	3813.1	3814.0	3824.3
$\Delta\chi^2$	-15.6	-2.9	-15.6	-16.5	-15.6	-5.3

^aFor this run, NEDE is assumed to be free-streaming radiation with $w_{\text{NEDE}} = c_s^2 = c_{\text{vis}}^2 = 1/3$.

^bWe fixed $\log_{10}(m/m_0)$ close to its base model best-fit value detailed in Table I to avoid prior volume artifacts; see the discussion in Sec. IV A. This leads to a slight underestimate of the fit improvement.

in order to compensate a stronger decay of the Weyl potential. At the same time, the value of n_s moves toward scale invariance, and the amplitude of primordial perturbations, A_s , is slightly enhanced to compensate the stronger damping on small scales. These degeneracies can be identified in both cases with and without SH₀ES. As a peculiar feature of these early-time modifications, note that n_s becomes 2σ compatible with scale invariance.

Finally, NEDE also affects the matter power spectrum. It increases the amplitude of local structures parametrized by $S_8 = \sigma_8 \sqrt{\Omega_m}/0.3$ from $S_8 = 0.824 \pm 0.010$ (ΛCDM without SH₀ES) to $S_8 = 0.841 \pm 0.012$ (NEDE with SH₀ES). This slightly raises the tension with weak lensing measurements, already present within ΛCDM at the 2.3σ level,²⁴ to 2.8σ (see also Table I). This is based on a recent study in Ref. [111] reporting $S_8 = 0.762^{+0.025}_{-0.024}$ from the combined tomographic weak gravitational lensing analysis of the Kilo Degree Survey and the Dark Energy Survey. We depict this constraint as the blue band in Fig. 12. The conclusion we draw from this is that whatever physics or systematics is responsible for this so-called S_8 tension [112], it seems to be equally affecting NEDE and ΛCDM . In other words, while NEDE in its simple form does not help with the S_8 tension, it also does not exacerbate it significantly. In fact, in Ref. [113], it was recently argued that the S_8 tension might be a result of the ‘‘lensing tension,’’ which questions the internal consistency of the

Planck dataset. Taking that into account (essentially by neglecting high- ℓ *Planck* data), they found a 3σ preference for single-field EDE over ΛCDM despite the inclusion of different LSS probes. We also note that we cannot reproduce the somewhat bleak picture given in Ref. [114]. There, the authors use for one of their analyses the same datasets as we do (except for BBN, which has a negligible impact anyhow), corresponding to our joint analysis with SH₀ES. They then find a lower evidence for single-field EDE compared with the 4.3σ evidence we report for NEDE. Whether this tendency toward stronger evidence within NEDE persists when including other LSS probes remains to be seen. Let us stress though that such an analysis needs to carefully take into account the prior volume effects discussed in Sec. IV A when not including SH₀ES (or another local H_0 probe).

2. Minimal extensions

In our base model, we fixed the trigger parameter, $x_*^{-1} = H_*/m = 0.2$, as well as the fluid parameters, $w_{\text{NEDE}}^* = 2/3$, $c_s^2 = 2/3$, and $c_{\text{vis}} = 0$. Their values correspond to generic choices in agreement with our microscopic model and expectations from other EDE-type models. Here, we allow them to vary, which constitutes a nontrivial phenomenological check of NEDE and its underlying microphysics. The MCMC analyses in this section include the local measurement of H_0 to create more constraining power and achieve quicker convergence. In all cases, we found that the mean values are 1σ compatible with the ones obtained in our base model in Table I. In order not to overload this work with data, we

²⁴The tension goes down to 1.9σ when including SH₀ES. The reason is that within ΛCDM ω_{cdm} decreases when H_0 goes up. This is not a valid way of relieving the tension, though, because it relies on combining incompatible datasets.

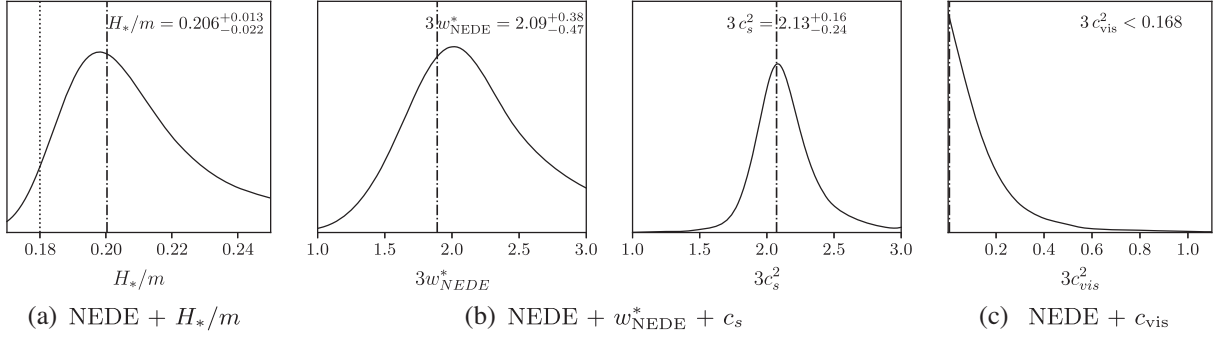


FIG. 13. Posterior distributions of three different NEDE extensions for our joint analysis with SH₀ES. In each case, the listed parameters (and the Λ CDM parameters) are allowed to vary while the other NEDE parameters are fixed. The dash-dotted line represents the best-fit value. The dotted line in (a) corresponds to $\phi_* = 0$.

therefore cite them only if relevant. We do, however, provide the results of our χ^2 analysis in Table II and present the posterior distributions of the additional model parameters in Fig. 13. Finally, we will also study the possibility that NEDE decays completely into free-streaming radiation.

a. Trigger parameter.

The trigger parameter determines where exactly in the evolution of ϕ the phase transition is initiated, corresponding to the ratio $x_*^{-1} = H_*/m$. As illustrated in Fig. 6, it controls the amplitude of initial fluctuations in the NEDE fluid right after the decay. The particular dynamics of our two-field model requires it to be within the narrow range (107), explicitly $0.18 < H_*/m \lesssim 0.21$, provided we prohibit a fine-tuning of the model's fundamental parameters [see the discussion before Eq. (107)]. In order to test this prediction, we allowed the trigger parameter to vary in the slightly larger range $0.17 < H_*/m \lesssim 0.25$.²⁵ As a result of this analysis, we report $H_*/m = 0.206^{+0.013}_{-0.022}$, where the posterior distribution is depicted in Fig. 13(a). This value lies nicely within the predicted interval. In particular, we find that it is larger than 0.18 (dotted vertical line), implying that the transition occurs before the tunneling probability becomes maximal when ϕ crosses zero for the first time, in accordance with a triggered phase transition. We consider this strong additional evidence for our two-field vacuum decay model.

b. Fluid parameters.

The fluid sector has been discussed in Secs. II D and III D. The upshot is that a condensate of colliding bubble walls is dominated by kinetic rather than potential energy on large scales. Together with the radiation arising from the

decay of the condensate, we therefore expect an effective value of the equation of state parameter within the range $1/3 < w_{\text{NEDE}}^* < 1$. Our base model prior $w_{\text{NEDE}} = 2/3$ was then simply chosen to be the central value. Moreover, we assumed the effective sound speed to equal the adiabatic one, implying $c_s^2 = w_{\text{NEDE}} = 2/3$. To check the phenomenological viability of these choices, we allowed both parameters to vary at the same time. The results are depicted in Fig. 13(b), indeed showing that our base model prior was justified.²⁶ We also see that the viable range $w_{\text{NEDE}}^* = 0.70^{+0.12}_{-0.16}$ is pretty large, allowing for a possibly wide range of collision and decay scenarios. The constraints on the rest-frame sound speed $c_s^2 = 0.711^{+0.057}_{-0.081}$ are significantly tighter but still 1σ compatible with $c_s^2 = w_{\text{NEDE}}$. It remains to be seen if such a value can be accommodated in a detailed collision model. We also note that adding two additional parameters to the fit reduces the evidence for NEDE down to 3.4σ , corresponding to $f_{\text{NEDE}} = 11.3^{+3.7}_{-3.3}\%$ and $H_0 = 71.0^{+1.1}_{-1.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$.

c. Viscosity parameter.

Here, we investigate the possibility that a nonvanishing viscosity parameter c_{vis} leads to the buildup of anisotropic stress σ_{NEDE} through the sourcing of velocity and metric shear on the right-hand side of (115). As we have argued in Sec. III D, anisotropic stress arises as a generic component of a perturbed ideal fluid and, as such, needs to be included in a comprehensive analysis. Moreover, we are not discussing the possibility that the initial value $\sigma_{\text{NEDE}}^* \neq 0$, which would be another way of generating anisotropic stress even if $c_{\text{vis}} = 0$ [see (115)]. As is obvious from Fig. 13(c), we find no evidence for $c_{\text{vis}} \neq 0$ when we allow it to vary in the range $0 < c_{\text{vis}}^2 < 1$; instead, we obtain the upper bound $c_{\text{vis}}^2 < 0.56$.

²⁵We fix $\log_{10}(m/m_0) = 2.63$ (close to the base model best fit in Table I) to avoid sampling problems that otherwise occur due to a background degeneracy along $m x_*^{-1} = \text{const}$ (which is broken at the perturbation level, though).

²⁶This is also in agreement with the analysis in [35], where in the case of ADE it was shown that data prefer the choice $c_s^2 \simeq w_{\text{ADE}}^*$.

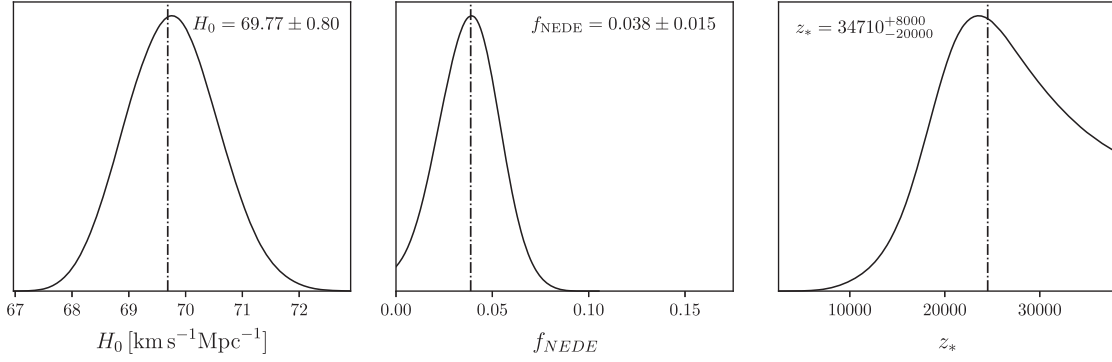


FIG. 14. Posterior distributions of our combined analysis with SH₀ES where we model NEDE after the phase transition as free-streaming radiation, corresponding to the choice $(w_{\text{NEDE}}^*, c_s^2, c_{\text{vis}}^2) = (1/3, 1/3, 1/3)$. Best-fit values are represented by the dash-dotted vertical lines.

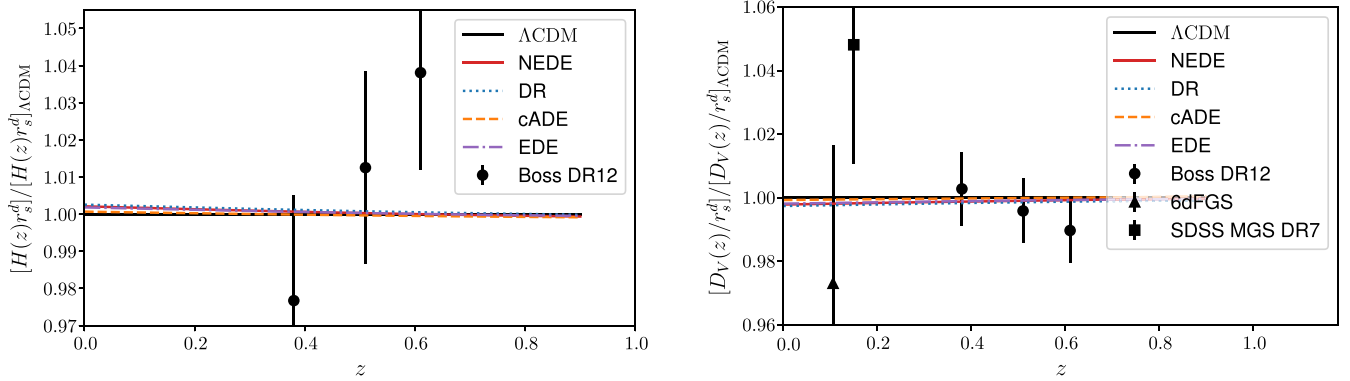


FIG. 15. Different BAO measurements represented as black dots. We compare to the best-fit models obtained from our combined data analysis with SH₀ES. As opposed to late-time modifications, all the early-time modifications studied in this work provide a similarly good fit as ΛCDM .

In the context of our particular model, this result makes it appear less likely that bubbles grow large and, therefore, leave a sizable imprint in the effective fluid in the form of anisotropic stress (which would otherwise average to zero). This, however, is only an indication, as the effect of these large bubbles should be modeled via a k -dependent initial value for $\sigma_{\text{NEDE}}^*(k) \neq 0$. In any case, bubbles never reach cosmological sizes if the bound (46) on $\bar{\beta}$ is satisfied, which is also needed to justify the modeling of the phase transition as an instantaneous event.

d. Free-streaming radiation.

Rather than an extension of the base model, we consider a modification where NEDE after the phase transition is converted to free-streaming radiation. In our model, this can be realized through a quick decay of the bubble wall condensate into light particles. On a purely phenomenological level, this case should also be close to the axion proposal in Ref. [36]. As explained before in Sec. III D, we model the free streaming through the choice $(w_{\text{NEDE}}^*, c_s^2, c_{\text{vis}}^2) = (1/3, 1/3, 1/3)$. This approximation relies on

an early truncation of the Boltzmann hierarchy at the order of $l = 2$ [86].²⁷

Our results are presented in Fig. 14. In particular, we find $f_{\text{NEDE}} = 3.8 \pm 1.5\%$, corresponding to a 2.5σ evidence for free-streaming NEDE. This has to be contrasted with 4.3σ evidence for the base model relying on a stiffer fluid. This is also in accordance with a smaller overall fit improvement of $\Delta\chi^2(\text{total}) = -5.3$. Moreover, the transition occurs significantly earlier for free-streaming radiation, $z_* = 35000^{+8000}_{-20000}$, which implies that changes to the Weyl potential, and, hence, the driving of acoustic CMB oscillations, are limited to rather small scales in the range $420 < \ell < 1300$. We conclude that free-streaming NEDE allows one only to alleviate rather than fully resolve the tension. However, we will see that with an overall improvement of $\Delta\chi^2(\text{total}) = -5.3$ it is still superior to dark radiation (or sterile neutrinos) with $\Delta\chi^2(\text{total}) = -3.2$.

²⁷We also devised an independent code where we solved the Boltzmann hierarchy up to higher order, yielding compatible but slightly stronger results, $H_0 = 70.1 \pm 0.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $f_{\text{NEDE}} = 0.046^{+0.017}_{-0.016}$ and $z_* = 20130^{+3600}_{-3300}$.

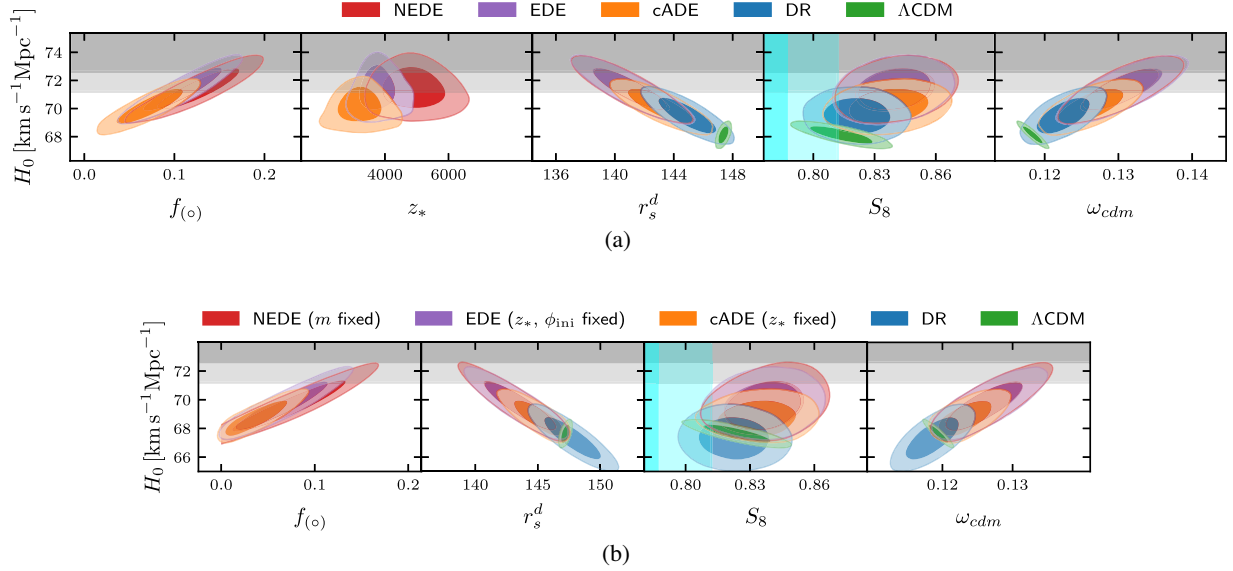


FIG. 16. The Hubble parameter today vs a subset of parameters for NEDE and its competitors. The blue and gray bands corresponds to the 95% C.L. constraints on H_0 (SH₀ES) and S_8 [111], respectively. (a) Combined analysis with SH₀ES. (b) Combined analysis without SH₀ES.

C. Early-time competitors

NEDE competes with other early-time modifications that similarly rely on reducing the sound horizon by adding a new energy component that is relevant before matter-radiation equality. Specifically, we will consider DR, ADE, and single-field EDE. Here, we briefly discuss each proposal and compare it with NEDE. In each case, we perform two combined analyses, one with and one without the local measurement of H_0 . The corresponding results are collected in Appendix B; in Table III we list the mean values together with their 1σ uncertainties, and in Table IV the χ^2 improvements are detailed.

These proposals have a common phenomenological feature which explains their previous success and sets them apart from late-time modifications. They all keep the product $r_s^d H_0 \simeq \text{const}$, which we argued in Sec. III E is key to accommodating BAO measurements. This is demonstrated in Fig. 15, where we see that all best-fit models provide a consistent fit to BAO data even when we use the local measurement to pull H_0 to higher values. The left panel depicts $H(z)r_s^d$ directly, and in the right panel we plot the combination $D_V(z)/r_s^d = [zD(z)^2/H(z)]^{1/3}/r_s^d$, where $D(z)$ is the comoving angular diameter distance at redshift z generalized from (43). Both plots are normalized with respect to the Λ CDM curve.

1. Dark radiation

DR (or extra sterile neutrinos) was an early and widely popular attempt to address the Hubble tension (e.g., [2,46,115]). It relies on introducing new relativistic (or mildly relativistic) degrees of freedom in a dark sector that

behave as free-streaming radiation. It can be parametrized by promoting the effective number of relativistic degrees of freedom, N_{eff} , to a free parameter. The deformation parameter is then given by $\Delta N_{\text{eff}} = N_{\text{eff}} - 3.046$, where we subtracted the effective number or relativistic degrees of freedom corresponding to three neutrino species. This modification has been studied extensively in the literature and in its simple form also been ruled out as a full resolution of the Hubble tension (see, for example, [2,48,50] but also more promising recent attempts based on interacting dark radiation [116,117]). This statement is also backed by our MCMC analysis, which shows only little evidence for $\Delta N_{\text{eff}} \neq 0$ (1.9σ) leading to a small overall improvement of the fit ($\Delta\chi^2 = -3.2$) when the local measurement is included (in agreement with the recent analysis in Ref. [117]). Similarly, without the SH₀ES value, we find that the Hubble tension is still at the $\simeq 4\sigma$ level. In the following, we will mainly use DR as a reference modification to infer the efficiency of the other EDE-type proposals.

From a phenomenological perspective, there are two major differences. First, DR is changing the expansion history not only in a small window around radiation-matter equality but also at much earlier times. This is problematic, because it affects BBN, causing an increase of the primordial helium abundance, Y_p . This worsens the corresponding fit by $\Delta\chi^2(\text{BBN}) = 0.8$ when we include the local H_0 measurement. This argument could be refined by also taking into account the deuterium abundance as done in Ref. [118], with the qualitatively same outcome. As we can also see from Table IV, this problem is completely avoided by modifications of the EDE type which have a

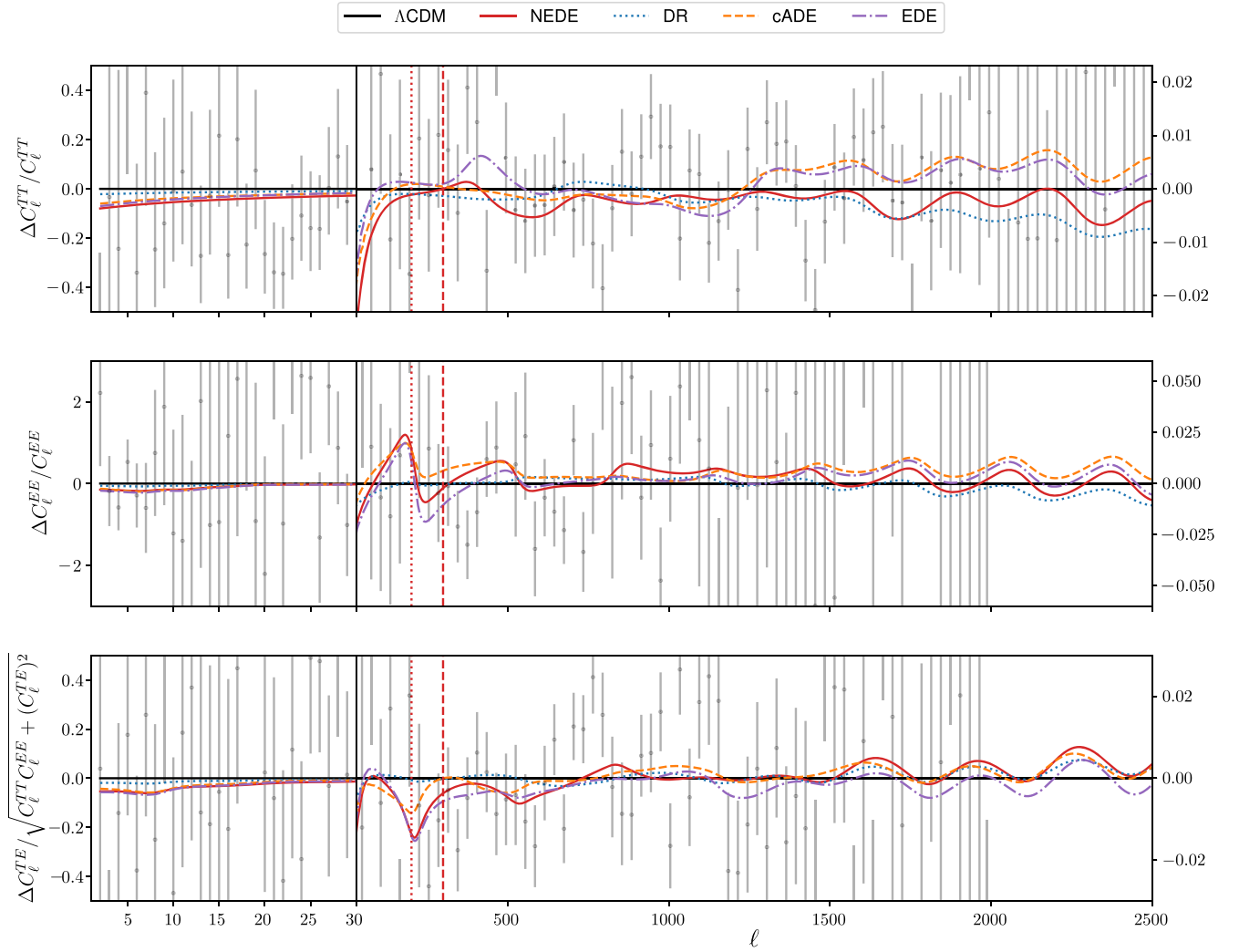


FIG. 17. *Planck* data residuals (gray dots) of NEDE (red solid line), DR (blue dotted line), and cADE (orange dashed line) and EDE (purple dash-dotted line) with respect to Λ CDM. Best-fit values are taken from combined analysis with SH_0ES (Table I). The plots use a different scaling for low ℓ (< 30) and high ℓ (> 30). The red dashed and dotted vertical lines highlight the modes that, respectively, entered the sound and particle horizon at the time of the NEDE phase transition.

negligible effect during BBN. A second background effect of DR is the lowering of r_s , which due to (121) also implies an enhancement of the ratio r_d/r_s . This leads through (120) to stronger diffusion damping effects. So far, this is analogous to NEDE. However, as DR is only a one-parameter extension of Λ CDM, it does not offer enough freedom to fully compensate this effect by, for example, increasing n_s as explained in Sec. III E. This becomes also apparent from the residuals depicted in Fig. 17, where we detect a stronger loss of power at small scales than for the EDE-type models (see also [91]). Accordingly, DR causes the strongest degradation of the high- ℓ *Planck* fit when SH_0ES is included, amounting to $\Delta\chi^2(\text{high-}\ell) = 3.2$. Besides this damping problem, at the perturbation level, DR corresponds to an additional free-streaming radiation component which sources anisotropic stress. This has been shown to lead to a shift in the phase of CMB oscillations

[119] which is also adversely affecting the CMB fit. However, as recently argued in Refs. [117,120], both problems can be improved upon by introducing an interacting radiation component.

In summary, we find that DR cannot resolve the Hubble tension mainly because it is constrained by the BBN probe but also because it does not offer enough flexibility to balance the additional damping and phase shift effects it introduces to the photon-baryon acoustic oscillations.

2. Acoustic dark energy

ADE was developed to mitigate the before-mentioned detrimental CMB effects by introducing acoustic oscillations in a dark fluid (rather than free-streaming particles) [35]. At the background level, it is very similar to other EDE implementations, as it describes a smooth but

relatively sudden transition from an equation of state parameter close to (but slightly above) -1 at early times to $w_{\text{ADE}}^* > 1/3$ at late times, parametrized as

$$w_{\text{ADE}}(a) + 1 = \frac{1 + w_{\text{ADE}}^*}{[1 + (a_*/a)^{6(1+w_{\text{ADE}}^*)}]^{1/2}}, \quad (124)$$

where $a_* \sim 1/3000$ leads to a transition close to matter-radiation equality. As with NEDE, the amount of ADE is implicitly defined via its relative abundance at “decay time”, $f_{\text{ADE}} = \rho_{\text{ADE}}^*/(3M_{\text{pl}}^2 H_*^2)$.

On the perturbation level, ADE is described as a fluid with time-dependent equation of state where the corresponding fluctuations δ_{ADE} and θ_{ADE} are controlled by the system (113). Here, we will focus on its canonical implementation (cADE) for which $(w_{\text{ADE}}^*, c_s^2, c_{\text{vis}}^2) = (1, 1, 0)$. In Ref. [35], it was argued that this can be realized with a canonical scalar field that converts all of its potential to kinetic energy around a_* (without converting it back). An example is provided by a field that after the release of the Hubble drag explores a flat direction in its potential. In particular, this excludes scenarios where the field oscillates around a minimum like the ones discussed in Refs. [33,51]. The parametrization also covers more exotic models with a nonstandard kinetic term, which, in principle, allow for $w_{\text{ADE}}^* = c_s^2 < 1$ [35]. As it is generically harder to justify these models from a fundamental perspective, we will focus here on the stiff fluid case only.

We implemented cADE in CLASS and performed a MCMC likelihood analysis imposing flat priors on the two parameters $-4.2 < \log_{10}(a_*) < -2.6$ and $0 < f_{\text{ADE}} < 0.3$. The resulting parameter covariances are shown as the orange contours in Fig. 16. For the analysis without the local measurement of H_0 , we fixed $\log_{10}(a_*) = -3.46$ (close to the best fit determined from a previous run) to avoid the sampling problems outlined in Sec. IV A. In this case, cADE predicts $H_0 = 68.94_{-0.93}^{+0.68} \text{ km s}^{-1} \text{ Mpc}^{-1}$, which reduces the tension down to $\simeq 3.2\sigma$ (including w_{ADE}^* and c_s^2 as free parameters would presumably further reduce the tension). The evidence for cADE is rather low with $f_{\text{ADE}} = 0.038_{-0.027}^{+0.018}$. If we include the local prior on H_0 , on the other hand, we reach a higher value, $H_0 = 70.15_{-0.81}^{+0.86} \text{ km s}^{-1} \text{ Mpc}^{-1}$, together with a significant fit improvement of $\Delta\chi^2(\text{total}) = -9.9$ and 3.2σ evidence for $f_{\text{ADE}} \neq 0$, which places it somewhere in the middle between DR and NEDE.²⁸ It appears that cADE is held back by the large- ℓ temperature and polarization data, which indeed deteriorate significantly by $\Delta\chi^2(\text{high-}\ell) = 2.9$ if we apply

more statistical pull toward higher values of H_0 by including the SH₀ES value.

What explains the difference between cADE and NEDE? Besides the stronger decay of the background energy density, the main difference lies in the perturbation sector. In contrast to NEDE, the fluid perturbations are initialized long before the decay when they are still frozen outside the horizon, i.e., $aH \gg k$. In the case of NEDE, on the other hand, modes are initialized much later at $z_* \sim 5000$, right after the phase transition, implying that for $\ell \gtrsim 180$ they are inside the particle horizon. As we discussed in Sec. III E, that makes them particularly sensitive to fluctuations in the transition surface, which is a distinct k -dependent feature of NEDE not present in ADE. Another notable difference between cADE and both EDE and NEDE is its later decay time, which roughly coincides with radiation-matter equality; to be specific, $z_* = 3230_{-500}^{+400}$. This is due to the quicker decay of cADE, which prevents adversary driving effects which would otherwise prevent the decay from happening too late.

In summary, cADE shows the same approximate degeneracies as NEDE but is not able to fully resolve the tension, even though it leads to a clear reduction of it by introducing only two new parameters. This general picture is confirmed in Fig. 17, where cADE shows larger residuals in the temperature power spectrum for $600 < \ell < 1000$.

3. Early dark energy

Finally, there is single-field EDE [32–34]. It introduces an ultralight axion field with decay constant $f_a (\sim M_{\text{pl}})$ and potential

$$V^{(n)}(\varphi) = \Lambda_a^4 [1 - \cos(\varphi/f_a)]^n, \quad (125)$$

where $\Lambda_a (\sim eV)$ is a dark sector energy scale. At early times when $H^2 \gg d^2V/d\varphi^2$, the field is stuck at $\varphi \simeq \varphi_{\text{ini}} (< \pi f_a)$. Once the Hubble drag is released around $z = z_*$, implicitly defined via $H_*^2 \simeq d^2V/d\varphi^2$, φ starts to oscillate in a near power-law potential, φ^{2n} , which leads to a quick decay of EDE. By averaging over oscillation cycles, the decay can be described by an ideal fluid with equation of state parameter $w_{\text{EDE}}^* \simeq (n-1)/(n+1)$, where $n = 1$ corresponds to a decay like matter, $n = 2$ to one like radiation, and $n \geq 3$ to an even quicker decay. In fact, the first two cases were shown to be disfavored by data, whereas the model with $n = 3$, where kinetic dominates over potential energy, is most efficient in addressing the Hubble tension, implying $w_{\text{EDE}}^* = 1/2$. Such a shallow potential term, scaling like ϕ^6 , has been argued to arise in axion scenarios when taking into account higher-order instanton corrections [32,121,122]. The problem with this argument, however, is that within the low-energy theory *all* terms are expected to be present, i.e., in general, $V(\varphi) = \sum_n c_n V^{(n)}$ with dimensionless coefficients $c_n = \mathcal{O}(1)$.

²⁸cADE performs slightly better in Ref. [35] with $H_0 = 70.60 \pm 0.85 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We attribute this to the difference in datasets, as their analysis used *Planck* 2015 (rather than *Planck* 2018) data.

The model therefore relies on the assumption $c_1, c_2 \ll c_3$, which from a low-energy perspective appears to be a nongeneric choice [36].

Besides this naturalness issue, the above model constitutes an interesting proposal for resolving the Hubble tension. On a cosmological level, it introduces four new parameters: the decay time z_* , the fraction of EDE at decay time, $f_{\text{EDE}} = \rho_{\text{EDE}}^*/(3M_{\text{pl}}^2 H_*^2)$, the initial field value $\varphi_{\text{ini}} (< \pi f_a)$, and the integer power n . Here, we set $n = 3$ and, therefore, study a three-parameter extension of Λ CDM. Like with ADE, when described as an ideal fluid, EDE shows a very similar behavior to NEDE on the background level, even though it relies on a second- rather than a first-order phase transition. Differences, however, arise on the perturbation level, where the higher-order potential in (125) leads to an effectively k -dependent sound speed $c_s(k)$, which was argued to be crucial for relaxing the Hubble tension [34]. In particular, $c_s(k)$ depends on the initial field value φ_{ini} , which needs to be set close to the Planck scale for the mechanism to work properly, thereby probing the flat part of the potential. Our cosmological model, on the other hand, uses a k -independent sound speed. Also, NEDE can accommodate an initial field value of its trigger field in a wide range (58). The reason is that in NEDE the trigger field ϕ is subdominant throughout its whole evolution; moreover, ϕ_{ini} drops out of the junction conditions, rendering cosmological observables insensitive to its value. Another crucial difference is that the transition in EDE is a smooth process at both the background and perturbation level, whereas it is of a sudden nature in NEDE. To be precise, in NEDE it is initiated when ϕ/ϕ_{ini} drops below a certain threshold determined by the parameter H_*/m alone. As we have argued in Sec. III E, it is this trigger mechanism which can mimic a k -dependent sound speed.

In order to perform our MCMC analysis of EDE, we used an existing CLASS implementation of EDE and scanned in addition to the Λ CDM parameters over $(\log_{10}(a_*), f_{\text{EDE}}, \varphi_{\text{ini}})$. We also tracked the full evolution of φ including its late-time oscillations. Everything else was implemented as detailed in Sec. IV A. The posterior distributions turn out to be slightly bimodal, where the dominant mode is centered around $\log_{10}(a_*) \simeq -3.6$ and the less significant one around $\log_{10}(a_*) \simeq -3.8$. In fact, in Ref. [34], it was argued that this can be attributed to the constraining effect of the high- ℓ polarization data, which is also what we find here. As we want to focus on the more significant mode only, we chose rather tight prior boundaries for our final MCMC analysis, $-3.7 < \log_{10}(a_*) < -3.2$ (and $2.0 < \varphi_{\text{ini}} < 3.14$). When we performed our analysis without SH₀ES, we encountered again the sampling problem detailed in Sec. IV A. In fact, it is even worse compared to NEDE, as two (rather than one) parameters, f_{EDE} and φ_{ini} , become degenerate in the limit $f_{\text{EDE}} \rightarrow 0$, eliminating any evidence for EDE. We were again able to

mitigate this problem by fixing both f_{EDE} and φ_{ini} to their near best-fit values as detailed in Table III.

The resulting covariances of EDE are presented in Fig. 16. It is striking to see that the EDE contours (purple) are largely overlapping with the ones of NEDE (red) despite the differences of both models at the perturbation level. This picture is solidified when we look at the results listed in Tables III and IV. Specifically, EDE achieves $H_0 = 69.6_{-1.2}^{+1.0}$ km s⁻¹ Mpc⁻¹ without SH₀ES, which brings the Hubble tension down to 2.5σ and leads to a weak detection of EDE at the 2σ level. With SH₀ES, our MCMC analysis yields a stronger 4σ detection²⁹ and $H_0 = 71.3 \pm 1.0$ km s⁻¹ Mpc⁻¹,³⁰ corresponding to an improvement of the overall fit of $\Delta\chi^2(\text{total}) = -16.4$, which is almost identical with NEDE. The transition times z_* are quite different though. While EDE prefers to decay at $z = 3760_{-410}^{+320}$, the NEDE phase transition is considerably earlier at $z_* = 4920_{-730}^{+620}$ (see also Sec. IV C).

On the level of the cosmological parameter extraction, NEDE trades the parameter φ_{ini} for the trigger parameter H_*/m and achieves a similarly promising outcome in terms of cosmological observables. However, there is a crucial difference: While φ_{ini} is unconstrained within the axion framework of EDE, we do not have much freedom in choosing the ratio H_*/m . In fact, the trigger dynamics of our underlying model fix it to be $H_*/m \simeq 0.2$, which happens to be the value preferred by data. A closer inspection of Fig. 17 also reveals that NEDE leads to different oscillatory residuals in both the temperature and polarization power spectrum (in particular, for $\ell \gtrsim 1500$). We attribute this mainly to different perturbation dynamics related to the trigger sector. Ultimately, we believe next-generation CMB experiments like the ground-based CMB-S4 [123] or The Simons Observatory [124] will be able to discriminate between both scenarios (notwithstanding that gravitational waves might provide another distinct feature of NEDE) and detect either of them in CMB data. In order to further quantify these statements, an analysis with mock data similar to the one in Ref. [34] would be needed.

V. CONCLUSION

NEDE is an early-time modification of the expansion history that relies on a first-order phase transition in a dark sector to add a short burst of gravitational repulsion around the eV scale. It introduces a two-field mechanism in order to trigger the phase transition at the right time and avoid problems with premature nucleation events that would otherwise risk a contamination of the CMB. By performing

²⁹This value is compatible with the $\simeq 3.5\sigma$ reported in Ref. [34].

³⁰As a cross-check, note that this is compatible with the value found in Ref. [34], $H_0 = 71.49 \pm 1.20$ km s⁻¹ Mpc⁻¹, using the same datasets with the exception of *Planck* 2018 data (instead of 2015 data) and the helium abundance.

a cosmological parameter extraction of an effective cosmological implementation of NEDE, we find that it shares the phenomenological success with its single-field predecessor despite different perturbation dynamics. We report 4.3σ evidence for our NEDE base model and an increased $H_0 = 71.4 \pm 1.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$ with SH₀ES included in our analysis. This corresponds to an improved fit with $\Delta\chi^2(\text{total}) = -15.6$ when including the fraction of NEDE f_{NEDE} and the trigger mass m as additional free parameters. Without SH₀ES, we still find 1.9σ evidence and $H_0 = 69.6^{+1.0}_{-1.3} \text{ km s}^{-1} \text{ Mpc}^{-1}$, which amounts to a reduced Hubble tension of 2.5σ (down from 4.3σ for Λ CDM).

This improvement is achieved for a rather generic choice of microscopic parameters, where the relevant terms in the potential of the NEDE field ψ are set by $M \sim \text{eV}$ and its quartic coupling is weak $\lambda < 0.1$. In fact, only a mild one-digit tuning is applied to suppress isocurvature modes and prevent sizable trigger field oscillations around the true vacuum. The trigger sector, on the other hand, necessitates ultralight physics with a mass scale $m \sim 10^{-27} \text{ eV}$, corresponding to a decay around redshift ~ 5000 . We demonstrated the technical naturalness of this setup. In particular, we derive an upper bound on the coupling between both fields to ensure the radiative stability of the ultralight scale.

The crucial trigger parameter H_*/m , controlling the amplitude of subhorizon modes in the NEDE fluid, is predicted by the model to be in a very narrow range around $H_*/m = 0.2$ for natural, non-fine-tuned, choices of parameters. A fit to data treating it as a free parameter gives a best-fit value of $H_*/m = 2.06^{+0.013}_{-0.022}$, which we take as additional strong evidence for the NEDE model.

We also provide a detailed comparison with single-field EDE and find a similar successful phenomenology. On the theory side, however, EDE is faced by two challenges. First, a steep anharmonic potential is needed to generate a sufficiently quick decay of the background condensate, and, second, the potential has to flatten to avoid detrimental effects in the polarization power spectrum that otherwise arise when the range of subhorizon modes with sound speed < 0.9 is too small [34]. While the last challenge can be naturally satisfied within an axion framework, an anharmonic potential constitutes a nongeneric choice and requires further theory adaptations. NEDE gets around both obstacles in a different way. We expect the stiff fluid needed for a quick decay to be effectively generated by the small-scale anisotropic stress associated with the colliding bubble wall condensate, and the issue with polarization residuals does not arise in our case which we attribute to the momentum-dependent trigger dynamics.

While we expect the high- ℓ signatures of NEDE to be revealed more clearly through future ground-based CMB experiments, a possible detection through gravitational waves is a unique feature of our first-order scenario. As the transition occurs relatively late in the early Universe as compared to the electroweak phase transition, the expected

gravitational wave spectrum is peaked at extremely small frequencies f_* ; e.g., for $H_*\bar{\beta}^{-1} > 10^{-3}$, we have $f_* < 10^{-14} \text{ Hz}$. We still find that the spectrum's large frequency tail can overlap with the claimed sensitivity bound of future gravitational wave experiments provided the duration of the phase transition is long enough. We also expect this signal to be strengthened in extended scenarios where we include a coupling between the dark sector and visible matter due to sound waves in the primordial fluid as well as turbulence effects.

We will end this paper by outlining different directions for future investigations. First, there is the theoretical challenge of finding a more fundamental framework that explains not only the light physics involved in our model but also the extremely weak coupling between both fields needed to keep radiative corrections under control. We observe that the trigger mass lies between a typical mass scale for ultralight DM ($\sim 10^{-22} \text{ eV}$) and quintessence DE ($\sim 10^{-33} \text{ eV}$), both of which can be realized in axion scenarios [125,126], where the mass is protected by a weakly broken shift symmetry. We believe that the two-field mechanism of NEDE can also be obtained from a string-inspired axion setup [127,128] but leave the details to future work.

We have seen that the phenomenological success of NEDE crucially depends on the properties of the fluid after the decay and, in particular, on its equation of state parameter w_{NEDE}^* , as well as its rest-frame sound speed c_s . To be specific, by extending our base model we obtained $w_{\text{NEDE}}^* = 0.70^{+0.12}_{-0.16}$ and $c_s^2 = 0.711^{+0.057}_{-0.081}$, suggestive of a quickly decaying fluid with $c_s^2 = w_{\text{NEDE}}^*$. Adding two more parameters then also diminishes the evidence for a nonvanishing fraction of NEDE down to 3.4σ [$f_{\text{NEDE}} = 11.3^{+3.7}_{-3.3}\%$]. While we indeed expect the relativistic bubble wall condensate to manifest itself on cosmological scales as a fluid with $w_{\text{NEDE}}^* > 1/3$ due to its small-scale anisotropies and kinetic energy domination, we have not yet provided a microscopic calculation. Instead, we have treated both quantities as phenomenological parameters and fixed them in our base run to $c_s^2 = w_{\text{NEDE}}^* = 2/3$, as suggested by our extended analysis and previous studies. We expect to further explore this in future work.

From a more phenomenological perspective, there remains the issue with the growth of structure as inferred from weak lensing measurements. This so-called S_8 tension [112] is slightly increased from 2.3σ to 2.8σ when compared with Λ CDM. This means that if the tension gets worse in the future, it will threaten the success not only of Λ CDM but also of EDE-type models. We therefore believe that further sophistications of NEDE should be driven by the goal to reduce both tensions at the same time. Such possibilities include considering nontrivial interactions between NEDE and the visible sector or including a sizable amount of scalar field oscillations around the new

vacuum after the phase transition (assumed to vanish in the present study) which will change the decay properties of the Weyl potential. In order to test this class of proposal, it will be important to include additional LSS probes in the data analysis.

ACKNOWLEDGMENTS

We thank Guido D’Amico, Edmund Copeland, José Espinosa, Colin Hill, Nemanja Kaloper, Antonio Padilla, Vivian Poulin, Kari Rummukainen, and David Weir for useful discussions. We also thank the SDU eScience Center for providing us with the computational resources for the MCMC analysis. This work is supported by Villum Fonden Grant No. 13384.

APPENDIX A: JUNCTION CONDITIONS IN COSMOLOGICAL PERTURBATION THEORY

In this Appendix, we derive the perturbed matching conditions across the transition surface Σ , formally defined in terms of the function $q(t, \mathbf{x})$ via (71). To that end, we restore the most general form of the metric (72):

$$ds^2 = -(1 + E)dt^2 + 2a(t)\partial_i F dt dx^i + a(t)^2[(1 + A)\delta_{ij} + \partial_i \partial_j B] dx^i dx^j, \quad (\text{A1})$$

with scalar perturbations $A(t, \mathbf{x})$, $B(t, \mathbf{x})$, $E(t, \mathbf{x})$, and $F(t, \mathbf{x})$ (in this work we are not interested in vector and tensor modes). The synchronous gauge corresponds to the choice

$$A = -2\eta, \quad k^2 B^2 = -(h + 6\eta), \quad E = F = 0. \quad (\text{A2})$$

The different fluctuation fields then transform under an infinitesimal coordinate transformation $x^0 \rightarrow x^0 + \xi^0$ and $x^i \rightarrow x^i + g^{ij} \partial_j \xi$ as

$$\Delta_\xi E = 2\dot{\xi}_0, \quad \Delta_\xi F = -\frac{1}{a}(\xi_0 + \dot{\xi} - 2H\xi), \quad (\text{A3a})$$

$$\Delta_\xi A = 2H\xi_0, \quad \Delta_\xi B = -\frac{2}{a^2}\xi. \quad (\text{A3b})$$

In order to simplify our derivation, we consider only spatial diffeomorphisms ξ that are continuous across the surface, i.e., $[\xi]_\pm = 0$; otherwise, ξ and ξ_0 are arbitrary functions. For the function $q(t, \mathbf{x})$, we find

$$\Delta_\xi \delta q = \dot{\hat{q}} \xi_0. \quad (\text{A4})$$

We now use this gauge freedom to fix

$$\delta q \rightarrow \delta \hat{q} = \delta q + \dot{\hat{q}} \xi_0 \stackrel{!}{=} 0, \quad (\text{A5})$$

where $\hat{\xi}_0 = -\delta q / \dot{\hat{q}}$. Working in the gauge $\delta \hat{q} = 0$, which we henceforth indicate by hatted variables, greatly simplifies the derivation of the matching conditions.

Next, we introduce the projection tensor

$$\tilde{g}_{\alpha\beta} = g_{\alpha\beta} + n_\alpha n_\beta,$$

where n is the unit normal vector with respect to the transition surface,

$$n_\alpha = c[g]\partial_\alpha q \quad (\text{A6})$$

with $c[g]$ a normalization constant determined though $n_\alpha n^\alpha = -1$. At the background level, we find $\bar{n}_\alpha = (1, 0, 0, 0)$. The extrinsic curvature tensor is then defined as

$$K_{\alpha\beta} = \frac{1}{2} \mathcal{L}_n \tilde{g}_{\alpha\beta} = \frac{1}{2} n^\gamma \partial_\gamma \tilde{g}_{\alpha\beta} + \tilde{g}_{\gamma(\alpha} \partial_{\beta)} n^\gamma. \quad (\text{A7})$$

Israel’s junction conditions in the absence of matter localized on Σ read [59]

$$[K_{\alpha\beta}]_\pm = 0, \quad (\text{A8})$$

where we use the notation introduced in (68). It is straightforward to show that this reduces to (69a) on the level of background quantities. Perturbing $K_{\alpha\beta}$ and projecting on the spatial components yields

$$\delta K_{ij} = \delta n^0 H g_{ij} - 1/2 \delta \dot{g}_{ij} + a^2 \delta_{k(i} \partial_{j)} \delta n^k, \quad (\text{A9})$$

where we used the background expression for the normal vector. Further perturbing (A6) as well as its normalization condition, we derive

$$\delta n_i = 0, \quad \delta n^i = \frac{1}{a} \delta^{ik} \partial_k \hat{F}, \quad \delta n^0 = \delta n_0 = \frac{1}{2} \hat{E}. \quad (\text{A10})$$

From this, it follows that $\delta g_{ij} = \delta \tilde{g}_{ij}$. With these definitions, we finally have

$$\delta K_{ij} = \left[\frac{1}{2} H \hat{E} - \frac{1}{2} \dot{\hat{A}} - H \hat{A} \right] g_{ij} + a^2 \partial_i \partial_j \left[-\frac{1}{2} \dot{\hat{B}} - H \hat{B} + \frac{1}{a} \hat{F} \right]. \quad (\text{A11})$$

The continuity of the hypersurface-induced metric $[\tilde{g}_{ij}]_\pm = 0$, together with Israel’s matching (A8), requires at the perturbation level

$$[\delta \tilde{g}_{ij}]_\pm = 0, \quad (\text{A12a})$$

$$[\delta K_{ij}]_\pm = 0, \quad (\text{A12b})$$

which, in turn, provides us with the following junction conditions in the “hat” gauge:

$$[\hat{A}]_{\pm} = 0, \quad (\text{A13a})$$

$$[\hat{B}]_{\pm} = 0, \quad (\text{A13b})$$

$$H_*[\hat{E}]_{\pm} = [\dot{\hat{A}}]_{\pm}, \quad (\text{A13c})$$

$$[\dot{\hat{B}}]_{\pm} = \frac{2}{a_*}[\hat{F}]_{\pm}. \quad (\text{A13d})$$

We now use (A3) together with $\hat{\xi}_0 = -\delta q/\dot{\hat{q}}$ to restore δq and, hence, the full gauge invariance³¹:

$$[A]_{\pm} = 2H_0 \left[\frac{\delta q}{\dot{\hat{q}}} \right]_{\pm}, \quad (\text{A14a})$$

$$[B]_{\pm} = 0, \quad (\text{A14b})$$

$$H_*[E]_{\pm} + 2 \left[\dot{H} \frac{\delta q}{\dot{\hat{q}}} \right]_{\pm} = [\dot{A}]_{\pm}, \quad (\text{A14c})$$

$$a_*^2[\dot{B}]_{\pm} - 2 \left[\frac{\delta q}{\dot{\hat{q}}} \right]_{\pm} = 2a_*[F]_{\pm}. \quad (\text{A14d})$$

Using (A2) then yields the junction conditions in the synchronous gauge in (78).

Next, we consider a fluid with energy density $\rho_i = \bar{\rho}_i + \delta\rho_i$ and equation of state parameter $w_i = \text{const}$. In order to derive the matching equations for the density contrast, $\delta_i = \delta\rho_i/\bar{\rho}_i$, and the divergence of the velocity field, θ_i , we follow a slightly different but entirely equivalent approach to the one used before. The conservation of the energy momentum tensor implies at linear order two independent equations:

$$\dot{\delta}_i + (1 + w_i) \left(\frac{\theta_i}{a} + \frac{k^2}{a^2} L + \frac{3}{2} \dot{A} \right) = 0, \quad (\text{A15a})$$

$$a\dot{\theta}_i - \frac{w_i}{1 + w_i} k^2 \delta_i + (1 - 3w_i)aH\theta_i + k^2\sigma_i - k^2 \frac{E}{2} = 0, \quad (\text{A15b})$$

where we introduced $L = a(F - a\dot{B}/2)$ and assumed that the rest-frame sound speed equals the adiabatic sound speed. As a consistency check, we recover (113) in the limit $c_{s,i}^2 = c_{a,i}^2 = w_i = \text{const}$ by using the synchronous gauge condition in (A2). We now impose the hat gauge (corresponding to $\delta\hat{q} = 0$), which “straightens” the transition surface. In order to derive the matching equations, we follow a slightly different but entirely equivalent approach to the one used before: We integrate both equations over an

infinitesimal time interval; specifically, we perform $\lim_{\epsilon \rightarrow 0} \int_{t_* - \epsilon}^{t_* + \epsilon} dt \circ$ and use that the terms proportional to \hat{L} , $\hat{\delta}_i$, $\hat{\theta}_i$, σ_i , and \hat{E} vanish, as they do not contain any distributional part $\propto \delta(t - t_*)$. In other words, we expect perturbations without at least one time derivative acting on them to be regular. This last steps works only because we work in a gauge for which the transition surface is not fluctuating. In any case, we obtain³²

$$[\hat{\delta}_i]_{\pm} = -\frac{3}{2}(1 + w_i)[\hat{A}]_{\pm} = 0, \quad (\text{A16a})$$

$$[\hat{\theta}_i]_{\pm} = 0, \quad (\text{A16b})$$

where we used (A13a) and (A13d) to derive the second equality in the first line. We now use the transformation rule

$$\Delta_{\xi} \delta_i = -3H(1 + w_i)\xi_0, \quad (\text{A17a})$$

$$\Delta_{\xi} \theta_i = \frac{k^2}{a} \xi_0 \quad (\text{A17b})$$

together with $\hat{\xi}_0 = -\delta q/\dot{\hat{q}}$ to restore δq and, hence, the full gauge invariance:

$$\left[\delta_i + H(1 + w_i) \frac{\delta q}{\dot{\hat{q}}} \right]_{\pm} = 0, \quad (\text{A18a})$$

$$\left[\theta_i - \frac{k^2}{a} \frac{\delta q}{\dot{\hat{q}}} \right]_{\pm} = 0. \quad (\text{A18b})$$

In the next step, we can also evaluate the discontinuity of (A15) directly, which after using (A18) and (A14) yields

$$\left[\dot{\delta}_i + \frac{3}{2}(1 + w_i)\dot{A} \right]_{\pm} = 0, \quad (\text{A19a})$$

$$\left[a\dot{\theta}_i + k^2\sigma_i - k^2 \frac{E}{2} \right]_{\pm} = 0. \quad (\text{A19b})$$

This discussion applies directly to cold dark matter perturbations and the first two momenta of the neutrino distribution $F_{\nu l}$, i.e., for $\delta_{\nu} = F_{\nu 0}$ and $\theta_{\nu} = (3/4)kF_{\nu 1}$. There are several special cases, though.

First, for massless neutrinos (and photons), we need to solve, in general, the full Boltzmann hierarchy of higher momenta $F_{\nu l}$, where $F_{\nu 2} \equiv 2\sigma_i$ is the anisotropic stress. In practical terms, this means that Eqs. (A15) have to be supplemented with [60]

³¹Note that this agrees with the result in Ref. [84] subject to the dictionary $E = 2\phi_{\text{there}}$, $F = B_{\text{there}}$, $A = -2\psi_{\text{there}}$, and $B = 2E_{\text{there}}$.

³²As a cross-check, note that we could have obtained Israel's equations in (A14) by performing the same steps with Einstein's equations that are second order in time derivatives.

$$\dot{\sigma}_\nu - \frac{4}{15} \left(\frac{\theta_\nu}{a} + \frac{K^2}{a^2} L \right) + \frac{3}{10} \frac{k}{a} F_{\nu 3} = 0, \quad (\text{A20a})$$

$$\dot{F}_{\nu l} - \frac{k}{a(2l+1)} [lF_{\nu(l-1)} - (l+1)F_{\nu(l+1)}] = 0, \quad l \geq 3. \quad (\text{A20b})$$

Integrating these equations again over an infinitesimal interval implies the continuity $[\sigma_\nu]_\pm = [F_{\nu l}]_\pm = 0$ (for $l \geq 3$) provided the functions $F_{\nu l}$ are sufficiently regular. Taking the jump, we also find $[\dot{\sigma}_\nu]_\pm = [\dot{F}_{\nu l}]_\pm = 0$.

Second, in the case of photons and baryons (and electrons), the dynamical equations (A15) [and (A20) for photons] have to be equipped with interaction terms.

However, it is easy to check that the additional terms do not change the matching conditions provided we make the same regularity assumptions as before. In summary, the matching relations (A18) and (A19) are valid for all fluid components as long as the equation of state is not changing during the transition. In the case of relativistic degrees of freedom, including photons and neutrinos, the higher momenta and their first time derivatives are continuous.

APPENDIX B: COMPETITOR TABLES

In this Appendix, we provide the result tables from our analysis of DR, cADE, and single-field EDE. While Table III lists the means and their uncertainties, Table IV displays the χ^2 improvements with respect to Λ CDM.

TABLE III. The mean value and $\pm 1\sigma$ error of the cosmological parameters from our combined analyses with and without imposing the SH₀ES prior on H_0 for dark radiation, cADE, and single-field EDE.

Parameter	Dark radiation		cADE		EDE ($n = 3$)	
	With H_0	Without H_0	With H_0	Without H_0	With H_0	Without H_0
$100\omega_b$	$2.265^{+0.016}_{-0.016}$	$2.236^{+0.017}_{-0.017}$	$2.287^{+0.023}_{-0.023}$	$2.257^{+0.016}_{-0.017}$	$2.276^{+0.020}_{-0.021}$	$2.263^{+0.017}_{-0.017}$
ω_{cdm}	$0.1225^{+0.0024}_{-0.0024}$	$0.1186^{+0.0025}_{-0.0026}$	$0.1262^{+0.0030}_{-0.0029}$	$0.1234^{+0.0021}_{-0.0030}$	$0.1304^{+0.0037}_{-0.0037}$	$0.1262^{+0.0032}_{-0.0041}$
h	$0.6949^{+0.0084}_{-0.0086}$	$0.6732^{+0.0096}_{-0.0100}$	$0.7015^{+0.0086}_{-0.0081}$	$0.6894^{+0.0068}_{-0.0093}$	$0.713^{+0.010}_{-0.010}$	$0.696^{+0.010}_{-0.012}$
$\ln 10^{10} A_s$	$3.062^{+0.014}_{-0.016}$	$3.047^{+0.015}_{-0.017}$	$3.065^{+0.015}_{-0.015}$	$3.056^{+0.014}_{-0.015}$	$3.069^{+0.014}_{-0.016}$	$3.060^{+0.014}_{-0.016}$
n_s	$0.9761^{+0.0055}_{-0.0055}$	$0.9642^{+0.0062}_{-0.0062}$	$0.9821^{+0.0065}_{-0.0064}$	$0.9731^{+0.0051}_{-0.0061}$	$0.9867^{+0.0066}_{-0.0068}$	$0.9775^{+0.0064}_{-0.0068}$
τ_{reio}	$0.0593^{+0.0066}_{-0.0082}$	$0.0571^{+0.0068}_{-0.0076}$	$0.0574^{+0.0066}_{-0.0078}$	$0.0558^{+0.0063}_{-0.0074}$	$0.0576^{+0.0066}_{-0.0078}$	$0.0568^{+0.0064}_{-0.0077}$
$f_{\text{ADE}}, f_{\text{EDE}}$	$0.073^{+0.026}_{-0.023}$	$0.038^{+0.018}_{-0.027}$	$0.109^{+0.031}_{-0.027}$	$0.066^{+0.032}_{-0.033}$
ΔN_{eff}	$0.25^{+0.13}_{-0.13}$	$-0.05^{+0.14}_{-0.15}$
ϕ_{ini}	$2.705^{+0.13}_{-0.072}$	2.637 (fixed)
σ_8	$0.8204^{+0.0083}_{-0.0091}$	$0.8080^{+0.0088}_{-0.0094}$	$0.833^{+0.010}_{-0.010}$	$0.8226^{+0.0087}_{-0.0100}$	$0.838^{+0.010}_{-0.010}$	$0.8269^{+0.0093}_{-0.0110}$
S_8	$0.821^{+0.011}_{-0.011}$	$0.823^{+0.011}_{-0.011}$	$0.837^{+0.013}_{-0.013}$	$0.834^{+0.012}_{-0.012}$	$0.839^{+0.012}_{-0.012}$	$0.836^{+0.012}_{-0.012}$
r_s^d [Mpc]	$144.9^{+1.3}_{-1.3}$	$147.8^{+1.5}_{-1.5}$	$143.1^{+1.4}_{-1.5}$	$145.0^{+1.6}_{-1.1}$	$141.1^{+1.7}_{-1.8}$	$143.5^{+2.0}_{-1.8}$
z_*	3230^{+400}_{-500}	2880 (fixed)	3760^{+320}_{-410}	3780 (fixed)
Tension SH ₀ ES	...	3.9σ	...	3.2σ	...	2.5σ
Tension S ₈	2.2σ	2.2σ	2.7σ	2.6σ	2.8σ	2.7σ
$\Delta\chi^2$	-3.2	-0.2	-9.9	-1.0	-16.4	-2.9
$f_{(\odot)}, \Delta N_{\text{eff}} \neq 0$	1.9σ	0.4σ	3.2σ	1.4σ	4.0σ	2.0σ

TABLE IV. The best fit $\Delta\chi^2$ with Λ CDM as baseline model from combined analysis with and without SH₀ES. The Λ CDM reference values are provided in Table II.

SH ₀ ES ΔDOF	Dark radiation		cADE		EDE		NEDE	
	With H_0	Without H_0	With H_0	Without H_0	With H_0	Without H_0	With H_0	Without H_0
	1	1	2	1 (2) ^a	3	1 (3) ^a	2	1 (2) ^a
<i>Planck</i> high- ℓ TT, TE, EE	3.2	-0.3	2.9	0.0	-2.3	-1.7	-0.8	-1.2
<i>Planck</i> low- ℓ TT	-0.7	0.2	-1.5	-0.6	-1.5	-1.2	-1.9	-1.4
<i>Planck</i> low- ℓ EE	-0.4	-0.4	-1.0	0.8	-1.0	-0.3	-0.9	-0.5
<i>Planck</i> lensing	0.1	0.0	0.2	0.1	0.4	0.2	0.4	0.3
BAO low- z	0.2	-0.1	0.0	0.0	0.1	0.2	0.1	0.2
BAO high- z + LSS	0.2	0.1	0.8	0.1	1.0	-0.2	0.9	-0.2
Pantheon	-0.1	0.2	0.0	0.1	1.5	-0.2	0.3	-0.1
BBN	0.8	0.0	0.0	0.0	0.2	0.2	0.0	0.0
SH ₀ ES	-6.4	...	-11.3	...	-14.7	...	-13.8	...
Total $\Delta\chi^2$	-3.2	-0.2	-9.9	-1.0	-16.4	-2.9	-15.6	-2.9

^aWe varied $f_{(c)}$ (after setting the other parameters close to their best fit) only to avoid prior volume artifacts; see the discussion in Sec. IV A.

- [1] F. Niedermann and M. S. Sloth, New early dark energy, [arXiv:1910.10739](#).
- [2] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, [arXiv:1807.06209](#).
- [3] S. Dhawan, D. Brout, D. Scolnic, A. Goobar, A. Riess, and V. Miranda, Cosmological model insensitivity of local H_0 from the Cepheid distance ladder, *Astrophys. J.* **894**, 54 (2020).
- [4] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the Hubble constant and stronger evidence for physics beyond Λ CDM, *Astrophys. J.* **876**, 85 (2019).
- [5] K. C. Wong *et al.*, H0LiCOW XIII. A 2.4% measurement of H_0 from lensed quasars: 5.3 σ tension between early and late-Universe probes, [arXiv:1907.04869](#).
- [6] L. Verde, T. Treu, and A. Riess, Tensions between the early and the late universe, *Nat. Astron.* **3**, 891 (2019).
- [7] W. L. Freedman *et al.*, The Carnegie-Chicago Hubble program. VIII. An independent determination of the Hubble constant based on the tip of the red giant branch, [arXiv:1907.05922](#).
- [8] W. L. Freedman, B. F. Madore, T. Hoyt, I. S. Jang, R. Beaton, M. G. Lee, A. Monson, J. Neeley, and J. Rich, Calibration of the tip of the red giant branch (TRGB), [arXiv:2002.01550](#).
- [9] W. Yuan, A. G. Riess, L. M. Macri, S. Casertano, and D. Scolnic, Consistent calibration of the tip of the red giant branch in the large magellanic cloud on the Hubble space telescope photometric system and a re-determination of the Hubble constant, *Astrophys. J.* **886**, 61 (2019).
- [10] D. Pesce *et al.*, The Megamaser cosmology project. XIII. Combined Hubble constant constraints, *Astrophys. J. Lett.* **891**, L1 (2020).
- [11] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO and Virgo, [arXiv:1908.06060](#).
- [12] L. Knox and M. Millea, Hubble constant hunter's guide, *Phys. Rev. D* **101**, 043533 (2020).
- [13] J. L. Bernal, L. Verde, and A. G. Riess, The trouble with H_0 , *J. Cosmol. Astropart. Phys.* **10** (2016) 019.
- [14] T. M. Davis, S. R. Hinton, C. Howlett, and J. Calcino, Can redshift errors bias measurements of the Hubble constant?, *Mon. Not. R. Astron. Soc.* **490**, 2948 (2019).
- [15] R. Wojtak, A. Knebe, W. A. Watson, I. T. Iliev, S. He, D. Rapetti, G. Yepes, and S. Gottlber, Cosmic variance of the local Hubble flow in large-scale cosmological simulations, *Mon. Not. R. Astron. Soc.* **438**, 1805 (2014).
- [16] I. Odderskov, S. Hannestad, and J. Brandbyge, The variance of the locally measured Hubble parameter explained with different estimators, *J. Cosmol. Astropart. Phys.* **03** (2017) 022.
- [17] H.-Y. Wu and D. Huterer, Sample variance in the local measurements of the Hubble constant, *Mon. Not. R. Astron. Soc.* **471**, 4946 (2017).
- [18] M. Rigault *et al.* (Nearby Supernova Factory), Strong dependence of type Ia supernova standardization on the local specific star formation rate, [arXiv:1806.03849](#).
- [19] M. Rigault *et al.*, Confirmation of a star formation bias in type Ia supernova distances and its effect on measurement of the Hubble constant, *Astrophys. J.* **802**, 20 (2015).

- [20] B. M. Rose, P. M. Garnavich, and M. A. Berg, Think global, act local: The influence of environment age and host mass on type Ia Supernova light curves, *Astrophys. J.* **874**, 32 (2019).
- [21] D. O. Jones *et al.*, Should type Ia supernova distances be corrected for their local environments?, *Astrophys. J.* **867**, 108 (2018).
- [22] T. Shanks, L. Hogarth, and N. Metcalfe, Gaia Cepheid parallaxes and ‘Local Hole’ relieve H_0 tension, *Mon. Not. R. Astron. Soc.* **484**, L64 (2019).
- [23] A. G. Riess, S. Casertano, D. Kenworthy, D. Scolnic, and L. Macri, Seven problems with the claims related to the Hubble tension, *Mon. Not. R. Astron. Soc.* **484**, L64 (2019).
- [24] L. Breuval *et al.*, The Leavitt law of Milky Way Cepheids from Gaia DR2 static companion parallaxes, [arXiv:1910.04694](https://arxiv.org/abs/1910.04694).
- [25] E. Mrtzell and S. Dhawan, Does the Hubble constant tension call for new physics?, *J. Cosmol. Astropart. Phys.* **09** (2018) 025.
- [26] R. E. Keeley, S. Joudaki, M. Kaplinghat, and D. Kirkby, Implications of a transition in the dark energy equation of state for the H_0 and σ_8 tensions, *J. Cosmol. Astropart. Phys.* **12** (2019) 035.
- [27] M. Raveri, Reconstructing gravity on cosmological scales, *Phys. Rev. D* **101**, 083524 (2020).
- [28] K. Aylor, M. Joy, L. Knox, M. Millea, S. Raghunathan, and W. L. K. Wu, Sounds discordant: Classical distance ladder & Λ CDM-based determinations of the cosmological sound horizon, *Astrophys. J.* **874**, 4 (2019).
- [29] L. Verde, J. L. Bernal, A. F. Heavens, and R. Jimenez, The length of the low-redshift standard ruler, *Mon. Not. R. Astron. Soc.* **467**, 731 (2017).
- [30] N. Arendse, A. Agnello, and R. Wojtak, Low-redshift measurement of the sound horizon through gravitational time-delays, *Astron. Astrophys.* **632**, A91 (2019).
- [31] C. Krishnan, E. O. Colgin, Ruchika, A. A. Sen, M. Sheikh-Jabbari, and T. Yang, Is there an early Universe solution to Hubble tension?, [arXiv:2002.06044](https://arxiv.org/abs/2002.06044).
- [32] V. Poulin, T. L. Smith, D. Grin, T. Karwal, and M. Kamionkowski, Cosmological implications of ultralight axionlike fields, *Phys. Rev. D* **98**, 083525 (2018).
- [33] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, Early Dark Energy Can Resolve The Hubble Tension, *Phys. Rev. Lett.* **122**, 221301 (2019).
- [34] T. L. Smith, V. Poulin, and M. A. Amin, Oscillating scalar fields and the Hubble tension: A resolution with novel signatures, *Phys. Rev. D* **101**, 063523 (2020).
- [35] M.-X. Lin, G. Benevento, W. Hu, and M. Raveri, Acoustic dark energy: Potential conversion of the Hubble tension, *Phys. Rev. D* **100**, 063542 (2019).
- [36] N. Kaloper and Dark energy, H_0 and weak gravity conjecture, *Int. J. Mod. Phys. D* **28**, 1944017 (2019).
- [37] S. Alexander and E. McDonough, Axion-dilaton destabilization and the Hubble tension, *Phys. Lett. B* **797**, 134830 (2019).
- [38] E. Hardy and S. Parameswaran, Thermal dark energy, *Phys. Rev. D* **101**, 023503 (2020).
- [39] J. Sakstein and M. Trodden, Early Dark Energy from Massive Neutrinos—A Natural Resolution of the Hubble Tension, *Phys. Rev. Lett.* **124**, 161301 (2020).
- [40] M. Braglia, W. T. Emond, F. Finelli, A. E. Gumrukcuoglu, and K. Koyama, Unified framework for early dark energy from α -attractors, [arXiv:2005.14053](https://arxiv.org/abs/2005.14053).
- [41] C. Wetterich, Phenomenological parameterization of quintessence, *Phys. Lett. B* **594**, 17 (2004).
- [42] M. Doran and G. Robbers, Early dark energy cosmologies, *J. Cosmol. Astropart. Phys.* **06** (2006) 026.
- [43] V. Pettorino, L. Amendola, and C. Wetterich, How early is early dark energy?, *Phys. Rev. D* **87**, 083009 (2013).
- [44] E. Calabrese, D. Huterer, E. V. Linder, A. Melchiorri, and L. Pagano, Limits on dark radiation, early dark energy, and relativistic degrees of freedom, *Phys. Rev. D* **83**, 123504 (2011).
- [45] F. Noble Chamings, A. Avgoustidis, E. J. Copeland, A. M. Green, and B. Li, Early dark energy constraints on growing neutrino quintessence cosmologies, *Phys. Rev. D* **100**, 043525 (2019).
- [46] M. Wyman, D. H. Rudd, R. Vanderveld, and W. Hu, Neutrinos Help Reconcile Planck Measurements with the Local Universe, *Phys. Rev. Lett.* **112**, 051302 (2014).
- [47] C. Dvorkin, M. Wyman, D. H. Rudd, and W. Hu, Neutrinos help reconcile Planck measurements with both the early and local Universe, *Phys. Rev. D* **90**, 083503 (2014).
- [48] B. Leistedt, H. V. Peiris, and L. Verde, No New Cosmological Concordance with Massive Sterile Neutrinos, *Phys. Rev. Lett.* **113**, 041301 (2014).
- [49] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIV. Dark energy and modified gravity, *Astron. Astrophys.* **594**, A14 (2016).
- [50] M. Raveri, W. Hu, T. Hoffman, and L.-T. Wang, Partially acoustic dark matter cosmology and cosmological constraints, *Phys. Rev. D* **96**, 103501 (2017).
- [51] P. Agrawal, F.-Y. Cyr-Racine, D. Pinner, and L. Randall, Rock ‘n’ roll solutions to the Hubble tension, [arXiv:1904.01016](https://arxiv.org/abs/1904.01016).
- [52] T. Karwal and M. Kamionkowski, Dark energy at early times, the Hubble parameter, and the string axiverse, *Phys. Rev. D* **94**, 103523 (2016).
- [53] C. T. Hill, D. N. Schramm, and J. N. Fry, Cosmological structure formation from soft topological defects, *Comments Nucl. Part. Phys.* **19**, 25 (1989); **25**, 407 (1988).
- [54] J. A. Frieman, C. T. Hill, A. Stebbins, and I. Waga, Cosmology with Ultralight Pseudo Nambu-Goldstone Bosons, *Phys. Rev. Lett.* **75**, 2077 (1995).
- [55] J. E. Kim and H. P. Nilles, A quintessential axion, *Phys. Lett. B* **553**, 1 (2003).
- [56] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, False vacuum inflation with Einstein gravity, *Phys. Rev. D* **49**, 6410 (1994).
- [57] M. Cortes and A. R. Liddle, Viable inflationary models ending with a first-order phase transition, *Phys. Rev. D* **80**, 083524 (2009).
- [58] S. Dutta, S. D. H. Hsu, D. Reeb, and R. J. Scherrer, Dark radiation as a signature of dark energy, *Phys. Rev. D* **79**, 103504 (2009).
- [59] W. Israel, Singular hypersurfaces and thin shells in general relativity, *Nuovo Cimento B* **44**, 1 (1966); Erratum, *Nuovo Cimento B* **48**, 463 (1967).

- [60] C.-P. Ma and E. Bertschinger, Cosmological perturbation theory in the synchronous and conformal Newtonian gauges, *Astrophys. J.* **455**, 7 (1995).
- [61] C. L. Carilli and S. Rawlings, Science with the square kilometer array: Motivation, key science projects, standards and assumptions, *New Astron. Rev.* **48**, 979 (2004).
- [62] G. Janssen *et al.*, Gravitational wave astronomy with the SKA, *Proc. Sci. AASKA14* (2015) 037 [arXiv:1501.00127].
- [63] A. Weltman *et al.*, fundamental physics with the square kilometre array, *Pub. Astron. Soc. Aust.* **37**, e002 (2020).
- [64] F. C. Adams, General solutions for tunneling of scalar fields with quartic potentials, *Phys. Rev. D* **48**, 2800 (1993).
- [65] S. R. Coleman, The fate of the false vacuum. 1. Semiclassical theory, *Phys. Rev. D* **15**, 2929 (1977); Erratum, *Phys. Rev. D* **16**, 1248 (1977).
- [66] A. D. Linde, Decay of the false vacuum at finite temperature, *Nucl. Phys.* **B216**, 421 (1983); Erratum, *Nucl. Phys.* **B223**, 544 (1983).
- [67] C. G. Callan, Jr. and S. R. Coleman, The fate of the false vacuum. 2. First quantum corrections, *Phys. Rev. D* **16**, 1762 (1977).
- [68] S. Weinberg, The cosmological constant problem, *Rev. Mod. Phys.* **61**, 1 (1989).
- [69] P. Vielva, E. Martinez-Gonzalez, R. Barreiro, J. Sanz, and L. Cayon, Detection of non-Gaussianity in the WMAP 1—year data using spherical wavelets, *Astrophys. J.* **609**, 22 (2004).
- [70] M. Cruz, E. Martinez-Gonzalez, P. Vielva, and L. Cayon, Detection of a non-Gaussian spot in WMAP, *Mon. Not. R. Astron. Soc.* **356**, 29 (2005).
- [71] A. R. Liddle and D. Wands, Microwave background constraints on extended inflation voids, *Mon. Not. R. Astron. Soc.* **253**, 637 (1991).
- [72] S. W. Hawking, I. G. Moss, and J. M. Stewart, Bubble collisions in the very early universe, *Phys. Rev. D* **26**, 2681 (1982).
- [73] P. Binetruy, A. Bohe, C. Caprini, and J.-F. Dufaux, Cosmological backgrounds of gravitational waves and eLISA/NGO: Phase transitions, cosmic strings and other sources, *J. Cosmol. Astropart. Phys.* **06** (2012) 027.
- [74] R. Watkins and L. M. Widrow, Aspects of reheating in first order inflation, *Nucl. Phys.* **B374**, 446 (1992).
- [75] P. Peebles, Phenomenology of the invisible universe, *AIP Conf. Proc.* **1241**, 175 (2010).
- [76] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological non-linearities as an effective fluid, *J. Cosmol. Astropart. Phys.* **07** (2012) 051.
- [77] J. D. Barrow and M. S. Turner, Inflation in the universe, *Nature (London)* **292**, 35 (1981).
- [78] M. S. Turner and L. M. Widrow, Old inflation is not prevented by large amounts of anisotropy, *Nature (London)* **326**, 206 (1987).
- [79] C. Caprini and D. G. Figueroa, Cosmological backgrounds of gravitational waves, *Classical Quantum Gravity* **35**, 163001 (2018).
- [80] A. Kosowsky, M. S. Turner, and R. Watkins, Gravitational radiation from colliding vacuum bubbles, *Phys. Rev. D* **45**, 4514 (1992).
- [81] A. Kosowsky, M. S. Turner, and R. Watkins, Gravitational Waves from First Order Cosmological Phase Transitions, *Phys. Rev. Lett.* **69**, 2026 (1992).
- [82] A. Kosowsky and M. S. Turner, Gravitational radiation from colliding vacuum bubbles: Envelope approximation to many bubble collisions, *Phys. Rev. D* **47**, 4372 (1993).
- [83] C. J. Moore, R. H. Cole, and C. P. L. Berry, Gravitational-wave sensitivity curves, *Classical Quantum Gravity* **32**, 015014 (2015).
- [84] N. Deruelle and V. F. Mukhanov, On matching conditions for cosmological perturbations, *Phys. Rev. D* **52**, 5549 (1995).
- [85] Y. Akrami *et al.* (Planck Collaboration), Planck 2018 results. X. Constraints on inflation, arXiv:1807.06211.
- [86] W. Hu, Structure formation with generalized dark matter, *Astrophys. J.* **506**, 485 (1998).
- [87] J. Evslin, A. A. Sen, and Ruchika, Price of shifting the Hubble constant, *Phys. Rev. D* **97**, 103511 (2018).
- [88] P. Lemos, E. Lee, G. Efstathiou, and S. Gratton, Model independent $H(z)$ reconstruction using the cosmic inverse distance ladder, *Mon. Not. R. Astron. Soc.* **483**, 4803 (2019).
- [89] K. Aylor, M. Joy, L. Knox, M. Millea, S. Raghunathan, and W. L. K. Wu, Sounds discordant: Classical distance ladder and λ cdm-based determinations of the cosmological sound horizon, *Astrophys. J.* **874**, 4 (2019).
- [90] W. Hu, M. Fukugita, M. Zaldarriaga, and M. Tegmark, CMB observables and their cosmological implications, *Astrophys. J.* **549**, 669 (2001).
- [91] Z. Hou, R. Keisler, L. Knox, M. Millea, and C. Reichardt, How massless neutrinos affect the cosmic microwave background damping tail, *Phys. Rev. D* **87**, 083008 (2013).
- [92] M. Zaldarriaga and D. D. Harari, Analytic approach to the polarization of the cosmic microwave background in flat and open universes, *Phys. Rev. D* **52**, 3276 (1995).
- [93] W. Hu and N. Sugiyama, Small scale cosmological perturbations: An analytic approach, *Astrophys. J.* **471**, 542 (1996).
- [94] W. Hu and M. J. White, The damping tail of CMB anisotropies, *Astrophys. J.* **479**, 568 (1997).
- [95] W. Hu and M. White, Acoustic signatures in the cosmic microwave background, *Astrophys. J.* **471**, 3051 (1996).
- [96] B. Audren, J. Lesgourgues, K. Benabed, and S. Prunet, Conservative constraints on early cosmology: An illustration of the Monte Python cosmological parameter inference code, *J. Cosmol. Astropart. Phys.* **02** (2013) 001.
- [97] T. Brinckmann and J. Lesgourgues, MontePython 3: Boosted MCMC sampler and other features, *Phys. Rev. D* **97**, 063506 (2018).
- [98] A. Gelman and D. B. Rubin, Inference from iterative simulation using multiple sequences, *Stat. Sci.* **7**, 457 (1992).
- [99] D. Blas, J. Lesgourgues, and T. Tram, The cosmic linear anisotropy solving system (CLASS) II: Approximation schemes, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [100] <https://github.com/flo1984/TriggerCLASS>.
- [101] <https://github.com/PoulinV/AxiCLASS>.
- [102] D. M. Scolnic *et al.*, The complete light-curve sample of spectroscopically confirmed SNe Ia from Pan-STARRS1

- and cosmological constraints from the combined pantheon sample, *Astrophys. J.* **859**, 101 (2018).
- [103] S. Alam *et al.* (BOSS Collaboration), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological analysis of the DR12 galaxy sample, *Mon. Not. R. Astron. Soc.* **470**, 2617 (2017).
- [104] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, The 6dF Galaxy Survey: Baryon acoustic oscillations and the local Hubble constant, *Mon. Not. R. Astron. Soc.* **416**, 3017 (2011).
- [105] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, The clustering of the SDSS DR7 main Galaxy sample I. A 4 per cent distance measure at $z = 0.15$, *Mon. Not. R. Astron. Soc.* **449**, 835 (2015).
- [106] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. V. CMB power spectra and likelihoods, [arXiv:1907.12875](https://arxiv.org/abs/1907.12875).
- [107] E. Aver, K. A. Olive, and E. D. Skillman, The effects of He I $\lambda 10830$ on helium abundance determinations, *J. Cosmol. Astropart. Phys.* **07** (2015) 011.
- [108] A. Peimbert, M. Peimbert, and V. Luridiana, The primordial helium abundance and the number of neutrino families, *Rev. Mex. Astron. Astrofis.* **52**, 419 (2016).
- [109] O. Pisanti, A. Cirillo, S. Esposito, F. Iocco, G. Mangano, G. Miele, and P. D. Serpico, PArthENoPE: Public algorithm evaluating the nucleosynthesis of primordial elements, *Comput. Phys. Commun.* **178**, 956 (2008).
- [110] E. G. Adelberger *et al.*, Solar fusion cross sections II: The pp chain and CNO cycles, *Rev. Mod. Phys.* **83**, 195 (2011).
- [111] S. Joudaki *et al.*, KiDS + VIKING-450 and DES-Y1 combined: Cosmology with cosmic shear, *Astron. Astrophys.* **638**, L1 (2020).
- [112] M. Raveri and W. Hu, Concordance and discordance in cosmology, *Phys. Rev. D* **99**, 043506 (2019).
- [113] A. Chudaykin, D. Gorbunov, and N. Nedelko, Combined analysis of Planck and SPTPol data favors the early dark energy models, *J. Cosmol. Astropart. Phys.* **08** (2020) 013.
- [114] J. C. Hill, E. McDonough, M. W. Toomey, and S. Alexander, Early dark energy does not restore cosmological concordance, *Phys. Rev. D* **102**, 043507 (2020).
- [115] L. Lancaster, F.-Y. Cyr-Racine, L. Knox, and Z. Pan, A tale of two modes: Neutrino free-streaming in the early universe, *J. Cosmol. Astropart. Phys.* **07** (2017) 033.
- [116] C. D. Kreisch, F.-Y. Cyr-Racine, and O. Dore, The neutrino puzzle: Anomalies, interactions, and cosmological tensions, *Phys. Rev. D* **101**, 123505 (2020).
- [117] N. Blinov and G. Marques-Tavares, Interacting radiation after Planck and its implications for the Hubble tension, [arXiv:2003.08387](https://arxiv.org/abs/2003.08387).
- [118] N. Schöneberg, J. Lesgourgues, and D. C. Hooper, The $\text{bao} + \text{bbn}$ take on the Hubble tension, *J. Cosmol. Astropart. Phys.* **10** (2019) 029029.
- [119] D. Baumann, D. Green, J. Meyers, and B. Wallisch, Phases of new physics in the CMB, *J. Cosmol. Astropart. Phys.* **01** (2016) 007007.
- [120] C. Brust, Y. Cui, and K. Sigurdson, Cosmological constraints on interacting light particles, *J. Cosmol. Astropart. Phys.* **08** (2017) 020.
- [121] M. Kamionkowski, J. Pradler, and D. G. E. Walker, Dark Energy from the String Axiverse, *Phys. Rev. Lett.* **113**, 251302 (2014).
- [122] R. Kappl, H. P. Nilles, and M. W. Winkler, Modulated natural inflation, *Phys. Lett. B* **753**, 653 (2016).
- [123] K. N. Abazajian *et al.* (CMB-S4 Collaboration), *CMB-S4 Science Book*, 1st ed. (2016), [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [124] P. Ade *et al.* (Simons Observatory Collaboration), The Simons Observatory: Science goals and forecasts, *J. Cosmol. Astropart. Phys.* **02** (2019) 056.
- [125] D. J. E. Marsh, Axion cosmology, *Phys. Rep.* **643**, 1 (2016).
- [126] E. G. Ferreira, Ultra-light dark matter, [arXiv:2005.03254](https://arxiv.org/abs/2005.03254).
- [127] P. Svrcek and E. Witten, Axions in string theory, *J. High Energy Phys.* **06** (2006) 051.
- [128] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, String axiverse, *Phys. Rev. D* **81**, 123530 (2010).