Incompatibility of fine-structure constant variations at recombination with local observations

Léo Vacher[1,*,†] and Nils Schöneberg[2,†]

[1]*International School for Advanced Studies (SISSA), Via Bonomea 265, 34136, Trieste, Italy*
[2]*Institut de Ciències del Cosmos (ICCUB), Facultat de Física, Universitat de Barcelona (IEEC-UB),
Martí i Franqués, 1, E08028 Barcelona, Spain*

Some attempts of easing the critical Hubble tension present in modern cosmology have resorted to using variations of fundamental constants, such as the fine-structure constant, at the time of recombination. In this article we demonstrate that there are critical hurdles to construct such viable models using scalar fields, due to the striking precision of local constraints on the fine-structure constant stability. These hurdles demonstrate that in single-field models one has to extremely fine-tune the shape of the potential and/or the initial conditions. Indeed, for single-field models in a potential that is not fine-tuned we can put a generic bound at recombination of $\Delta\alpha/\alpha < 5 \times 10^{-4}$ (95% CL).

I. INTRODUCTION

Thanks to the incredible experimental and theoretical effort undertaken within the last decades, the precision of the measurement and understanding of the underlying cosmological model has steadily increased. However, this increased precision has uncovered new tensions between different measurements of cosmological parameters—such as the Hubble constant (which specifies the current expansion rate)—within the cosmological standard model (ΛCDM, involving a cosmological constant and cold collisionless dark matter). This Hubble tension between local distance ladder measurements using Cepheids and determinations from the *Planck* satellite measuring the cosmic microwave background (CMB) anisotropies have now reached a significance of beyond 5σ [1–3].

It has been recently proposed that an early variation of the fundamental constants could ease this Hubble tension by delaying the time at which recombination occurs [4–8]. Such primordial variations can be constrained directly using the CMB anisotropies and spectral distortions, but these constraints remain loose enough to allow for the existence of significant deviations with respect to the local values measured on Earth [8–11]. However, to remain consistent from a high-energy physics perspective, a space-time-dependent fundamental constant—as the fine-structure constant α, quantifying the intensity of the electromagnetic force—must be induced by a fundamental field implemented at the Lagrangian level (for a review see for example [12]). Such fields are for example unavoidable

in string theory, with the scalar dilaton field, partner mode of the graviton, inducing a variation of all the standard model's gauge couplings [13].

In general, any model that shows a displacement of such a coupled field causes a variation of the fine-structure constant in the early Universe. However, as we will detail within this work, if a relatively large relative variation of the fine-structure constant is desired, this displacement should occur as early as possible. As such, investigating scalar fields that become dynamical around the time of recombination are of particular interest. The prime example of such fields is an axionlike particle (ALP) with a decay constant such that it becomes dynamical around recombination. Indeed, such a field has been used to supply an era of early dark energy (EDE) contribution [14,15], which has also been shown to provide a successful path in order to ease the Hubble tension [16–18], though not without caveats [19–25]. These ALPs are deeply motivated both from high energy, as solution to the strong *CP* problem and modes from string theory [26,27]. Moreover, such ALPs could produce some parity-violating signal in the cosmic microwave background which claims to have been detected in Planck data [28–30]. Note, however, that this claim of detection has been contested by recent analyses as [31]. As such, an important question to address is whether such an ALP could be also coupled to electromagnetism and induce a variation of the fine-structure constant, further alleviating the H_0 tension (and/or overcoming the shortcomings of EDE). As proposed in [32], such a field fully coupled to electromagnetism through both a scalar and a pseudo-scalar term is expected to be present in general scalar-tensor theories.

We will see that such a scenario would be in conflict with the extremely tight constraints imposed by laboratory

*lvacher@sissa.it
†L. V. and N. S. contributed equally to this work.

103520-1

data on Earth today. Along with our exploration of the swampland conjectures in [33], we will derive an "almost-no-go" theorem, stating that a simple early varying fine-structure constant scalar field model cannot possibly provide fine-structure constant variations large enough to be cosmologically relevant without an extreme level of fine-tuning in either the potential or the field initial conditions.

We will start by presenting the theoretical background for varying fine-structure constant models in Sec. II A, followed by the derivation of an almost-no-go theorem for models with early fine-structure constant variation in Sec. II B. The validity of this theorem will then be demonstrated in two models: axionlike particles coupled to electromagnetism in Sec. III A and a toy model with a hyperbolic tangent potential in Sec. II B. Finally, we will discuss the extent and the limits of our study followed by our conclusions in Sec. IV.

II. THEORETICAL MOTIVATION

In order to remain consistent with the most basic principles of physics, any variation of the fundamental constants of nature must be implemented at the Lagrangian level, promoting the constants to scalar fields. In Sec. II A, we will present the formalism allowing for a consistent variation of the fine-structure constant. From this theoretical background, we will show in Sec. II B using simple derivations that such scalar field models cannot allow for an early variation of the fine-structure constant while remaining compatible with local data without an extraordinary amount of fine-tuning.

A. Fields coupled to electromagnetism

The fine-structure constant α is a dimensionless gauge coupling quantifying the intensity of the electromagnetic force. It represents a favorable observable in order to investigate the possible variation of the fundamental constants, as its impact on physics is well known and it is possible to measure its value with a great accuracy throughout cosmic history using multiple independent probes [34].[1] As far as we know, the only self-consistent path to promote α as a dynamical quantity is to introduce new fields at the Lagrangian level which are responsible for its variation. If $\alpha \to \alpha(\phi)$, then the electromagnetic kinetic Lagrangian must be modified as

$$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \to -\frac{1}{4} B_F(\phi) F_{\mu\nu} F^{\mu\nu} \qquad (1)$$

in order to preserve the U(1) gauge invariance of the theory, such that the fine-structure constant evolves as

$$\frac{\Delta\alpha}{\alpha_0} = \frac{\alpha(\phi) - \alpha_0}{\alpha_0} = B_F(\phi)^{-1} - 1, \qquad (2)$$

where $\alpha_0 \sim 1/137$ is the value of the fine-structure constant measured in laboratory.[2] Including a kinetic and potential term for the scalar field, its full Lagrangian can be written as

$$\mathcal{L} = -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) - \frac{1}{4} B_F(\phi) F_{\mu\nu} F^{\mu\nu} + \cdots. \qquad (3)$$

Knowing that the possible fine-structure constant variations allowed by experiments are extremely restricted, the scalar coupling can be typically linearized as

$$B_F(\phi) \simeq 1 + \zeta(\phi - \phi_0), \qquad (4)$$

where $\zeta = \partial_\phi B_F|_{\phi=\phi_0}$ (with the sign convention of [36]). This approximation appears to accurately model the time variation of the fine-structure constant through cosmic history for a wide range of models. While we will use this linearization in the remainder of the main text, we also discuss how generic this approach is in Appendix B.

From the expression of B_F, one can express the variations of the fine-structure constant as

$$\frac{\Delta\alpha}{\alpha_0} \simeq -\zeta(\phi - \phi_0). \qquad (5)$$

The amplitude of variations of such a scalar field through cosmological times, and hence the allowed variations of α, are sharply restricted by the atomic clocks measurement of [39] providing a bound of

$$\frac{1}{\alpha_0} \frac{d\alpha}{dt}\bigg|_{z=0} = (1.8 \pm 2.5) \times 10^{-19}/\text{yr.} \qquad (6)$$

As such, in the absence of screening mechanisms, any significant variations of the field from the CMB responsible for a different fine-structure constant at recombination must brutally rapidly decrease to match local constraints.

In addition, direct astrophysical measurements of $\Delta\alpha/\alpha_0$ itself (which are made using quasars) lead to $\Delta\alpha/\alpha_0(1 \le z \le 2.5) \sim 10^{-6}$ (see item 2). For comparison, these measurements can be converted to give an approximate estimation of the fine-structure constant drift rate of $(d\alpha/dt)/\alpha_0|_{z \simeq 1.5} \sim 10^{-16}h/\text{yr}$. Such estimation can be done by comparing the measurement of quasars at different redshifts (here we used a polynomial regression of the data

---

[1]Additionally, α is dimensionless, and only the stability of dimensionless constant can be investigated unambiguously. For a discussion see for example [35].

[2]This formalism for fine-structure constant variations was presented in [36] as a generalization of the models of Bekenstein [37] and Sandvik, Barrow, and Magueijo [38], which are the limiting cases where the scalar field is proportional to the electron charge [leading through a change of variables to $\alpha(\phi) = \alpha_0 e^{2\phi}$ and $B_F(\phi) \propto e^{-2\phi}$ for a canonically normalized field ϕ].

up to various orders). Hence, local data on $(d\alpha/dt)/\alpha_0$ are more stringent than local data by around two and a half orders of magnitude.

B. Almost a no-go theorem

Let us now investigate further how the experimental bounds can constrain the presence of varying α during recombination. For this, let us look at the two ingredients that are present for such a model:

  (i) a relevantly large $\Delta\alpha/\alpha_0$ at the time of recombination (by definition) and
  (ii) a small $d\ln\alpha/d\ln a$ today in order to avoid atomic clock constraints.

As we are going to see, in a standard formalism of coupling based on a single scalar field, these two conditions are highly incompatible.

First, we must consider what we mean by relevantly large $\Delta\alpha/\alpha_0$ during recombination. For our purpose, we will choose a benchmark value of $\Delta\alpha/\alpha_0(z_{\mathrm{cmb}}) = 10^{-2}$, which would in principle completely resolve the Hubble tension (this value was computed with a simple code such that it leads to the same angular sound horizon while raising H_0 to 73.04 km/s/Mpc, the value preferred by [1]).

However, it is already clear that such high values of $\alpha(z_{\mathrm{cmb}})$ are not compatible with the Planck 2018 angular power spectra measurements, since it is possible to infer a constraint of [9]

$$\frac{\Delta\alpha}{\alpha_0}(z_{\mathrm{cmb}}) = (-0.7 \pm 2.5) \times 10^{-3}. \tag{7}$$

However, we note that this constraint is dependent on the geometrical degeneracies present in the CMB and as such might be eased when introducing additional model parameters such as curvature, variation of the neutrino mass, dark energy equation of state, etc. Instead, the bound we derive below will be mostly independent of the specifics of the evolution of the various cosmic energy densities and their perturbations. We also note that it has been derived assuming typically a constant shift of the fine-structure constant. Finally, we will see that the bound we derive is even tighter than that obtained from the CMB itself.

Now, let us introduce the two quantities

$$\mathfrak{D} \equiv \frac{\Delta\alpha}{\alpha_0}(z_{\mathrm{cmb}}) = -\zeta\Delta\phi = -\zeta\int_{\ln a_{\mathrm{cmb}}}^{0} \frac{d\phi}{d\ln a} d\ln a, \tag{8}$$

$$\epsilon \equiv \frac{d\ln\alpha}{d\ln a}\bigg|_{z=0} = -\zeta\frac{d\phi}{d\ln a}\bigg|_{z=0}. \tag{9}$$

In this notation we can succinctly summarize our constraints as $\epsilon \simeq (1.76 \pm 2.44) \times 10^{-9}/h$ from Eq. (6), where we use the usual definition of $h = H_0/[100 \text{ km/s/Mpc}]$.

Comparing with a benchmark value of $\mathfrak{D} \simeq 10^{-2}$ gives us a ratio of $\mathfrak{D}/\epsilon \sim 10^7$. If we reparametrize the field speed evolution through some function $U(z)$ such that

$$\frac{d\phi}{d\ln a} = U(z)\frac{d\phi}{d\ln a}\bigg|_{z=0}, \tag{10}$$

we can immediately relate the two expressions as

$$\mathfrak{D} = \epsilon\int_{0}^{\ln 1+z_{\mathrm{cmb}}} U(z)d\ln(1 + z). \tag{11}$$

Importantly, this relation is *independent* of the precise value of ζ, since it simultaneously rescales $\mathfrak{D}$ and ϵ. As such, the combination of a given desired $\mathfrak{D}$ at early times and a given constraint on ϵ from atomic clock observations gives a coupling-independent constraint on the necessary evolution of the field speed.

Looking at the rough order of magnitude of this constraint, the problem becomes quickly apparent. In only roughly seven e-folds (from recombination to now), the relative field speed $U(z)$ has to vary by approximately a factor of $\mathfrak{D}/\epsilon \sim 10^7$. It turns out that with mild assumptions this combination is almost impossible. Below, we will put a bound on $\mathfrak{D}/\epsilon$ from simple considerations and translate it into a bound on $\mathfrak{D}$. To show this, let us slightly rewrite the equations of motion for the scalar field speed[3]

$$\frac{d\dot\phi}{d\ln a} + 3\dot\phi = -\frac{1}{H}\frac{dV}{d\phi} \tag{12}$$

and explicitly integrate to obtain

$$\dot\phi = Ca^{-3} - a^{-3}\int \frac{a^3}{H}\frac{dV}{d\phi}d\ln a. \tag{13}$$

Let us, for a moment, focus on the homogeneous equation. We are going to come back to the inhomogeneous part in a bit. The homogeneous part $\dot\phi = Ca^{-3}$ tells us that from Hubble drag the field speed can only decay at a rate of $a^{-3} = (1+z)^3$. This immediately places a tight constraint on $U(z) = \dot\phi/H \cdot H_0/\dot\phi_0$. Indeed, since $v(0) = 1$ by definition, we would have $U(z) = (1+z)^3 \cdot H_0/H$ in this case. It is now easy to put a conservative bound on the integral as

---

[3]We neglect the possible couplings of the field to the other sectors of the Universe, expected to be present in such models in the Klein-Gordon and the Friedmann equations through terms of the form $\zeta\Delta\phi\rho_i$ [38]. Given the allowed values for ζ, the impact on the field evolution is expected to be largely subdominant; see also Sec. IV.

$$\mathfrak{D}/\epsilon \approx \int_0^{\ln 1+z_\Lambda} (1+z)^3 \frac{H_0}{H} \mathrm{d}\ln(1+z) + \frac{1}{\sqrt{\Omega_m}} \int_{\ln 1+z_\Lambda}^{\ln 1+z_{\mathrm{cmb}}} (1+z)^3 (1+z)^{-3/2} \mathrm{d}\ln(1+z)$$

$$= \int_0^{\ln 1+z_\Lambda} (1+z)^3 (1+z)^{-\frac{3}{2}(1+w)} \mathrm{d}\ln(1+z) + \frac{2}{3\sqrt{\Omega_m}} [(1+z_{\mathrm{cmb}})^{3/2} - (1+z_\Lambda)^{3/2}]$$

$$< \int_0^{\ln 1+z_\Lambda} (1+z)^3 (1+z)^{-\frac{3}{2}(1+(-1))} \mathrm{d}\ln(1+z) + \frac{2}{3\sqrt{\Omega_m}} [(1+z_{\mathrm{cmb}})^{3/2} - 0]$$

$$= \frac{1}{3} [(1+z_\Lambda)^3 - 1] + \frac{2}{3\sqrt{\Omega_m}} [(1+z_{\mathrm{cmb}})^{3/2}]$$

$$< \frac{1}{3} [(1+10)^3 - 1] + \frac{2}{3 \cdot 0.3} [(1+1200)^{3/2}] \approx 9.3 \times 10^4. \tag{14}$$

If we require our benchmark value of $\mathfrak{D}/\epsilon \sim 10^7$, then this is obviously a contradiction. Indeed, we can use this bound to impose limits on $\mathfrak{D} = \Delta\alpha(z_{\mathrm{cmb}})/\alpha_0$ as we show below. However, we first aim to give a few details on the calculation.

In the first line of Eq. (14) we have split the integral into one part where we assume matter domination to hold (above some z_Λ) and one part where the dark energy might be important.

In the second line we have combined the first and second Friedmann equations

$$\frac{2(\dot{H} + H^2)}{H^2} = -\frac{\frac{8\pi G}{3}(\rho + 3P)}{\frac{8\pi G}{3}\rho} = -(1 + 3w), \quad (15)$$

with $w = P/\rho$ being the total equation of state in the Universe, which can quickly be simplified to

$$\frac{\mathrm{d}\ln H}{\mathrm{d}\ln a} = \frac{\dot{H}}{H^2} = -\frac{1+3w}{2} - 1 = -\frac{3}{2}(1+w), \quad (16)$$

which gives the relation of $H/H_0 = (1+z)^{3/2(1+w)}$ used here. Note that w at this point can and does depend on time and is simply defined as $w(z) = P(z)/\rho(z)$. Now, for any normal contents of the Universe, we have $w \in [-1, 1]$ and this allows us to quickly bound the first summand in the third line of Eq. (14).[4] Let us stress again that this derivation is intended to propose a conservative bound rather than an accurate estimate. In the same direction, we also bounded $1 + z_\Lambda > 0$ for the second summand. In the fourth line we simply evaluate the remaining integral. Finally, we bound $z_\Lambda < 10$ and $z_{\mathrm{cmb}} < 1200$ as well as $\Omega_m > 0.1$ (to get

$\sqrt{\Omega_m} > 0.3$) in the fifth line to give the final upper bound. Note that these are very conservative upper bounds, since CMB data typically require a smaller recombination redshift and measurements of baryonic acoustic oscillations (BAO) or supernovae of type Ia (SNIa) require $z_\Lambda \sim 0.33 \ll 10$ and $\Omega_m \sim 0.3$; see below.

Since the computation is always dominated by the matter-dominated part (due to much larger integration range), we can make the following statement: Given the constraints on Ω_m from Pantheon + SNIa of [40] (0.334 ± 0.018), we find that we can exclude ratios of $\mathfrak{D}/\epsilon < 4.8 \times 10^4$ at 95% CL. This would translate for the current bounds from atomic clocks into $\mathfrak{D} < (2.8 \times 10^{-4})/h$ at 95% CL [using the value from BAO of [41] with $\Omega_m = 0.299 \pm 0.016$ would give instead $\mathfrak{D} < (3.0 \times 10^{-4})/h$ at 95% CL].[5]

This bound is expected to apply maximally for physically motivated and consistent models emerging from unification theories (as grand unified theories or string theories), in which the potential contribution is usually subdominant and in which the presence of couplings with radiation freeze the field during radiation domination and typically just lead to a logarithmically slow variation in matter domination, leaving even less time for the field to decelerate. As we will see below, the presence of a potential that is not fine-tuned does not weaken these bounds.

C. The issue of potentials

We have considered only the homogeneous part of the field evolution of Eq. (13) so far, and one could wonder if the inhomogeneous part involving the potential slope $\mathrm{d}V/\mathrm{d}\phi$ could impart sufficient deceleration to allow the field to evade this issue. Indeed, with complete freedom over initial conditions and shape of the potential, this is always trivially possible. One can simply construct such a potential where $\dot{\phi} \to 0$ toward $z = 0$. As we will argue

---

[4]In principle a $w < -1$ could break this argument, but this is practically irrelevant for two reasons: (i) the first summand turns out to have a negligible 0.5% contribution to the total sum, and for it to have a 50% contribution one would need $w \simeq -2.4$, which is strongly excluded by any late-time data, and (ii) the bound assumed for z_Λ of around 10 is very conservative, and even if a nontrivial dark energy model with $w \ll -1$ would be used, also in this case the bounds could be made tighter by choosing a less conservative z_Λ.

[5]To obtain these constraints we used the corresponding Gaussian probability densities for Ω_m and ϵ from the respective measurements and propagated them according to the laws of transforming and multiplying independent random variables.

below, this requires a fine-tuned and "unnatural" combination of initial conditions and potential shape.

Let us consider the energy conservation equation [or equivalently Eq. (12)], which can be written as

$$\frac{dV}{d\ln a} + \frac{1}{2}\frac{d\dot{\phi}^2}{d\ln a} = -3\dot{\phi}^2 < 0. \qquad (17)$$

This equation corresponds to motion in a potential with a dissipative force,[6] and as such the field must eventually obey $dV/d\ln a < 0$ (as the squared field speed trivially cannot decrease below 0). Indeed, at any location within the potential a field would naturally start rolling downhill [$\dot{\phi} = 0 \Rightarrow \ddot{\phi} = -dV/d\phi$ and thus $\delta[\dot{V}] = -(dV/d\phi)^2\delta t$]. As such, we have a strong expectation that $\dot{V} < 0$ except in a transitory phase where the field has been accelerated before and suddenly encounters an uphill slope and has not sufficiently decelerated. We stress here that it is not impossible to generate $\dot{V} > 0$; it is simply not a "natural" state for a field with arbitrary initial conditions in an arbitrary potential.

However, the problem is slightly worse than just requiring $\dot{V} > 0$ during part of the evolution. This is because Eq. (13) weighs the contributions of a given $dV/d\phi$ with the current Hubble expansion rate as well as a factor of a^3, making the weight a function that is quite spiked toward $a \to 1$. Indeed, during matter domination we have $a^3/H \propto a^{4.5}$. This implies that the deceleration has to happen mostly toward the end of the evolution and, as such, be sudden. Since we are essentially requiring the field to come to an abrupt halt, it does not move over great distances of the potential, thus requiring essentially the potential slope at a singular point (of the current-day field position) to balance exactly to cancel out any remaining velocity. This is the fine-tuning issue for nonoscillatory solutions. In Sec. III B we show an example with a constant slope $dV/d\phi$ to decelerate the field and note that the final field slope has to be extremely fine-tuned.

Another way to quickly decelerate the field might in principle use strong oscillations. However, it can quickly be shown that in a potential of type $V(\phi) = A \cdot \phi^{2n}$ the *envelope* of the oscillations only decays as $a^{-3/(n+1)}$ (see for example [16]), and the envelope of the oscillation speed decays as $a^{-3n/(n+1)}$. We reproduce for convenience the argument of [42] in Appendix A and extend it to the field speed. In any case, this behavior $a^{-3n/(n+1)}$ is *slower* than a^{-3} for all $n \in \mathbb{N}$, implying that the problem for oscillating potentials is equal or even slightly worse, since they are slightly more inefficient at dissipating energy and they

cannot use a slope $dV/d\phi$ to decelerate, since it will be traversed in both directions during any oscillation. For this reason even oscillatory potentials cannot create a large $\mathfrak{D}$ for a given small bound on ϵ using the same reasoning as in Eq. (14). However, here too there is an extremely fine-tuned way to avoid a straight no-go theorem. If the oscillation happens to reach its maximum just at $z = 0$, then the field speed there will be zero, meaning that any bound in ϵ is trivially satisfied. Given that the field speed needs to be about 1–2 orders of magnitude smaller today than the expected speed from the decay of the field speed envelope, this requires even more fine-tuning in the field position. Indeed, one can quickly show that for the aforementioned potential a field offset of $\phi = \phi_{\max}(1 - \varepsilon)$ gives a relative velocity offset of $\sqrt{2n\varepsilon}$. Restricting the velocity to be smaller than the velocity envelope by 2 orders of magnitude [to reach $\mathfrak{D}/\epsilon \sim 10^7$, which is 2 orders of magnitude larger than the bound of Eq. (14)] then would bound $\varepsilon \lesssim 10^{-4}/(2n)$, which especially for large n is quite a fine-tuned position for the final field. We discuss such fine-tuning of oscillatory potentials in an example case in Sec. III A.

In summary, if the potential is oscillatory, an extreme degree of fine-tuning is required for a given value of the fine-structure constant variation $\mathfrak{D} = \Delta\alpha/\alpha$ to be compatible with the strict observations of $\epsilon \ll 1$, namely that the field happens to exactly end up at the turnaround of the oscillation today. If the potential is not oscillatory, one needs to construct a potential that happens to have exactly the correct slope toward $z \to 0$ in order to decelerate the field by just the right amount at the last moment. In either case, the potential and/or the field initial conditions need to be fine-tuned in order to avoid the tight constraints. We show such fine-tuning explicitly for a few examples in Sec. III.

Statement 1. We have demonstrated that one has to either give up large fine-structure constants at recombination or face an extreme degree of fine-tuning of the potential or the initial conditions for any model where the fine-structure constant variation is generated from the variation of a single scalar field.

III. EXAMPLE STUDIES

The aim of this section is to illustrate and strengthen the argument made in Statement 1. For this, we investigate two example cases. One displays a highly oscillatory behavior (due to an axionlike potential) and is investigated in Sec. III A, whereas the other shows a fine-tuned potential that aims to decelerate the field rapidly toward the end of its movement and is investigated in Sec. III B.

A. Axionlike particles coupled to electromagnetism

Let us now verify the validity of the discussion above by considering the canonical case of ALPs. Originally invoked

---

[6]This might become more obvious by replacing $\phi \to \sqrt{m}x$, which then immediately gives the normal kinetic term of a free particle ($\rho = V + \frac{1}{2}m\dot{x}^2$). The dissipative force in this case would simply be $F_{\text{diss}} = -3H\dot{x}$ (whereas the normal force generated by the potential is still $F_{\text{pot}} = -\partial_x V$).

to solve the strong CP problem in quantum chromodynamics [26], the existence of ALPs is commonly proposed in cosmology as a dark matter candidate [43], a source of early dark energy [16] or at the origin of cosmic birefringence in the CMB [28–30].

The potential $V_n(\phi)$ of an ALP is typically modeled by the function [14,15]

$$V_n(\phi) = (m_a f_a)^2 \left[1 - \cos\left(\frac{\phi}{f_a}\right)\right]^n, \qquad (18)$$

which allows for a domination of the field energy density at early times followed by a brutal damping in successive era. As such, ALP are well motivated candidates to implement an early variation of the fine-structure constant, assuming their full coupling to electromagnetism as proposed in [32]. As such, ALPs are well-motivated candidates to implement an early variation of the fine-structure constant.

The eventuality of coupling between EDE and electromagnetism was already investigated in a different context by [44], and in some sense we are generalizing here this work for axionic fields and with new datasets.

One can further note that such a phenomenology could be well motivated within the framework of string theory [45], where both EDE and the early variation of the fine-structure constant could be caused by the existence of a dilaton field coupled to an axion. We will however not discuss this case here, focusing only on single scalar field models for now. As such, we postpone such an investigation for future work.

The "canonical" ALP early dark energy model that has been proven to have a significant impact on the Hubble tension [6,16–18] has four fundamental parameters describing the model. These are $\{m_a, f_a, n, \phi_i\}$, where the first three parameters describe the potential of Eq. (18) and the fourth parameter is the initial condition of the field displacement (the initial field velocity is typically irrelevant due to the strong Hubble drag at early times). These parameters are typically further replaced by parameters carrying more physical meaning, such as $f_{\rm ede}$ (the largest fraction of energy density that the ALP EDE field reaches, related to f_a) and z_c (the critical redshift at which the field begins oscillating due to the weakening Hubble drag, related to m_a).

Furthermore, for the sake of simplicity, the initial field displacement is replaced by the initial position along the cosine curve of the potential, $\Theta_i = \phi_i/f_a$. In addition to this parameter space, we introduce the linear coupling ζ of the fine-structure constant to the field displacement as in Eq. (5), self-consistently propagating the effect of a varied fine-structure constant throughout recombination processes as described for example in [9]. As such, the parameter base can be written as $\{z_c, f_{\rm ede}, n, \Theta_i, \zeta\}$ in this notation.

To investigate the behavior of an ALP EDE coupled to electromagnetism, we used a modified version of the CLASS

software [46] coupled to MontePython [47,48], along the line of previous works using a similar setup [33,49,50] (following also the idea behind the implementation of [16,51,52], but adapting it to the new coding standard of CLASS v3.2.0, which also involves variations of fundamental constants [53]). For the plots we use LIQUIDCOSMO [54]. For simplicity, we first focus on the $n = 3$ case, which was favored by observational data in order to address the H_0 tension [17].

As further discussed in Appendix C, the model acts as desired for a model implementing an early variation of the fine-structure constant, inducing almost no variation of the fine-structure constant at low redshift, until it suddenly reaches a plateau associated with a different value of $\alpha \neq \alpha_0$ after an oscillatory transition close to the CMB epoch. The time at which this transition occurs is directly given by the critical redshift (z_c). The parameters $f_{\rm ede}$ and ζ impact the amplitude of the oscillations and the magnitude of the plateau in $\alpha(z)$ while Θ_i impacts the number of oscillations during the transition phase and the final speed of the field.

In order to confront this model with data, we use likelihoods associated to various datasets[7] also used in [50].

(1) *Cosmological.* We use CMB angular power spectra and lensing reconstruction data from the *Planck* satellite [56,57], baryonic acoustic oscillation data from BOSS DR12 [58], the Pantheon SNIa sample [59], and cosmic chronometers from [60]. We also use a prior on the H_0 value from [61], implemented as discussed in [62] as a prior on the supernovae absolute magnitude.

(2) *Fine-structure.* We use the spectroscopic measurement points of quasars (QSO) from various datasets [63–66] as well as a prior from the Oklo natural nuclear reactor [67], and the updated bound on the time variation of α of Eq. (6) coming from measurements of atomic clocks [39], as already used in [33].

The results of this investigation can be found in Fig. 2, where we show the constraints on the underlying model parameters from only cosmological observations (item 1) or also including fine-structure constant observations (item 2). An enlargement of the constraints on ζ can be found in Fig. 1. The immediate conclusion is that without experiments probing the fine-structure constant, a large range of ζ values is allowed (even slightly favoring a positive value). For example, a point on the edge of the 2σ exclusion limit with $\zeta \simeq 0.06$ can be associated with a value of $\Delta\alpha(z_{\rm cmb})/\alpha_0 \sim 10^{-3}$.[8]

---

[7]We do not discuss the constraints coming from big bang nucleosynthesis here (which would further reinforce the point we want to make), as these are expected to be strongly model dependent. On this topic, see e.g. [55].

[8]These suggestive values of $\Delta\alpha(z_{\rm cmb})/\alpha_0$ naturally depend on the other parameters of the model (such as Θ_i or z_c). Here we use the best-fit values of the corresponding run—rescaling only ζ—to give a qualitative estimate.
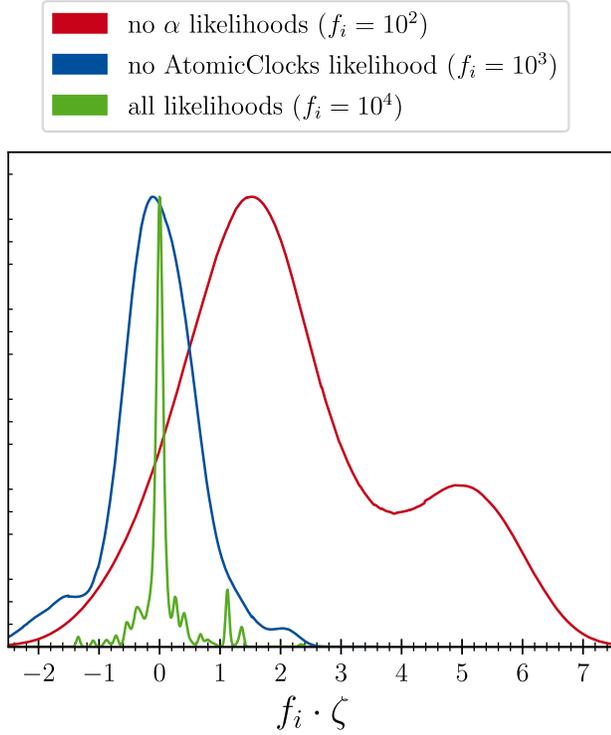
FIG. 1. Contour plots of the parameter space of the ALP scenario for three different datasets: no α data (red curve), no atomic clock data (blue curve) and all likelihoods (green curve). f_i is a rescaling factor allowing one to easily compare the curves.

Instead, once likelihoods probing the fine-structure constant are added, no such points are available anymore, and the ζ is restricted to lead to cosmologically irrelevant $\alpha(z)$ at the recombination times, in accordance with our almost-no-go theorem of Statement 1. Put otherwise, while it would be legitimate to think that local constraints on $\alpha(z)$ at $z \sim 0$ have no bearing on the possible behavior of the field during recombination era at $z \sim 1100$, quite the opposite is true for a well-defined single-field model. With the full likelihood set, the maximum values allowed for the electromagnetic coupling are of the order of $\zeta \sim 10^{-4}$, which can be associated to a value of $\Delta\alpha/\alpha_0(z_{\rm cmb}) \sim 10^{-6}$.

Overall, we report constraints on ζ and the associated suggestive values of $\Delta\alpha/\alpha_0(z_{\rm cmb})$ in Table I.[9]

Importantly, it is possible to circumvent the relatively narrow bounds on ζ reported in Table I entirely if one is allowed to fine-tune their initial conditions (Θ_i) drastically, such that the final field movement will be exactly in a

---

[9]We note that the ALP EDE model is always what drives the higher value of H_0 in this combined model (and there is no correlation between ζ and $f_{\rm ede}$) since the EDE is so much more efficient at resolving the Hubble tension in a way that is statistically preferred from the CMB data compared to the $\alpha(z)$ introduced by the coupling (ζ).

TABLE I. Mean and standard deviation of ζ for the various data combinations in the ALP EDE case, corresponding to the posteriors shown in Fig. 1 and corresponding variation of the fine-structure constant at the emission of CMB associated to the mean.

	No α data	No atomic clocks	All likelihoods
ζ	$(2.2 \pm 1.9) \times 10^{-2}$	$(-0.06 \pm 0.77) \times 10^{-3}$	$(0.06 \pm 0.45) \times 10^{-4}$
$\frac{\Delta\alpha}{\alpha_0}(z_{\rm cmb})$	2.5×10^{-4}	-6.8×10^{-6}	6.8×10^{-8}

turnover of an oscillation. Since these fine-tuned initial conditions take up a tiny negligible prior volume, they are not visible in the posterior contours of Fig. 2. This behavior of fine-tuned initial conditions or potential parameters occupying a vanishing amount of prior space is expected to be generic and not just related to this particular model.

This can lead us to conclude another interesting corollary statement.

Statement 2. Even if fine-tuned initial conditions and/or potential shape can reconcile large variations of the fine-structure constant at recombination with the local bounds, these typically occupy a negligible fraction of the prior volume and might thus still be ruled out in a Bayesian analysis.

However, in our ALP EDE example case we can force the field into such fine-tuning by minimizing the likelihood for a fixed $\zeta = 10^{-3}$ (more than an order of magnitude larger than the allowed range). In this case, we can reach a similarly good maximum likelihood as in the full run, differing only by $\Delta\chi^2 \simeq 23.1$ (largely driven by the QSO and Oklo likelihoods), whereas the original best fit of the full run would result in a $\Delta\chi^2 = 129\,000$ if simply rescaled to have $\zeta = 10^{-3}$. The reason is that with this minimization strategy the fine-tuning of the initial conditions can be performed such that the final field configuration is almost perfectly at the turnover of a single oscillation. Such a behavior is displayed in Fig. 3, comparing the minimized best fit of the full run and the minimized best fit of the run with $\zeta = 10^{-3}$ fixed. While the former reaches the small final field velocity through appropriately small ζ to lead to cosmologically irrelevant $\alpha(z)$ at recombination times, the latter reaches it by fine-tuning the Θ_i such that the final field velocity is almost zero by virtue of the oscillation ending at exactly $z = 0$ (see also the field displacement being at a peak in the same Fig. 3).

Another interesting corollary to the main Statement 1 can be derived from the fact that we cannot find a region of $\Delta\chi^2 \simeq 0$: since this "trick" of fine-tuning the deceleration of the field at late times cannot lead to vanishing $\Delta\alpha(z)$ for a wide range of redshifts in an oscillatory potential, the QSO and Oklo likelihoods at $z > 0$ necessarily impart a reasonably large likelihood penalty.

Statement 3. As such, even when there is an extreme degree of fine-tuning in the design of the potential, other
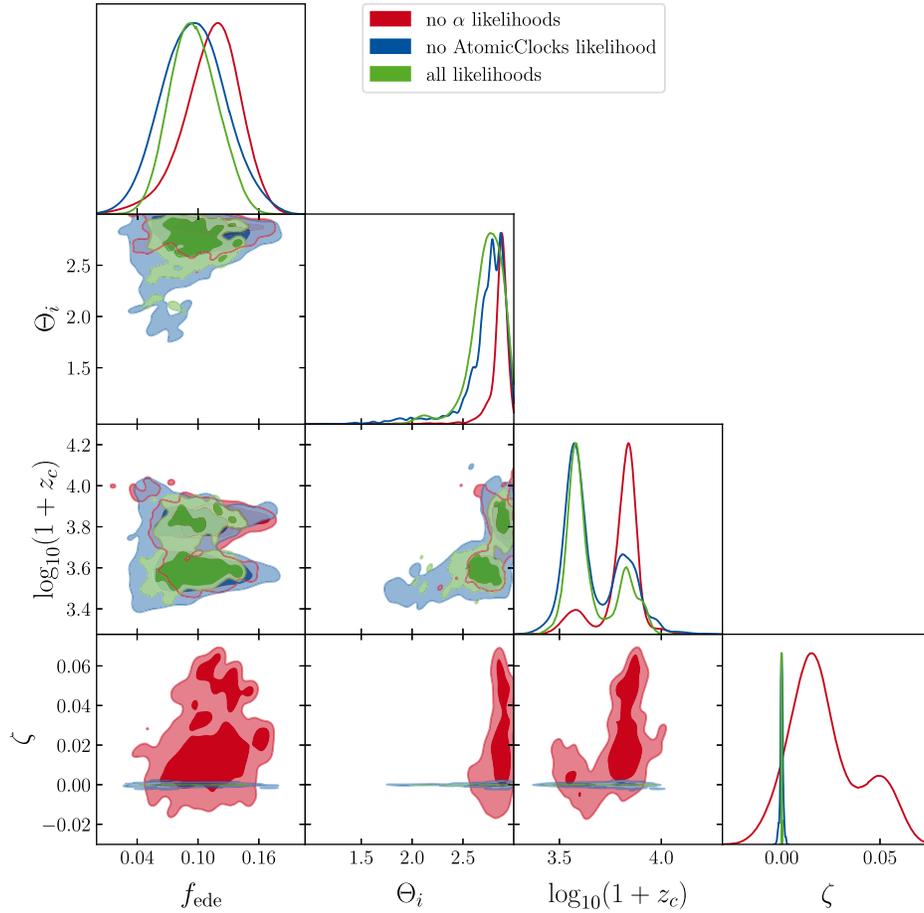
FIG. 2.   Contour plots of the parameter space of the ALP scenario for three different datasets: no α data (red), no atomic clock data (blue) and all likelihoods (green).

data on the fine-structure constant in the late Universe can put even further constraints against a relevant fine-structure at recombination.

B. Toy model

In this section we attempt to build a simple toy model of a case in which the field does not oscillate around a minimum in the potential but instead displays a continuous motion. As such, according to Statement 1 it can only have a significant impact at the time of recombination and avoid the current fine-structure constraints by virtue of fine-tuning of initial conditions or parameter values. For this, we introduce a potential based on the hyperbolic tangent function as

$$V(\phi) = 10^{-10} \cdot [fa\phi + (1-f)(s\phi + \kappa)] + V_0 \quad (19)$$

with

$$f = \frac{1}{2}\left(\tanh\left(\frac{\phi}{\Sigma}\right) + 1\right), \quad (20)$$

with the free parameters $\{V_0, \kappa, a, s, \Sigma\}$.[10] The factor 10^{-10} is used purely for numerical reasons related to the implementation in CLASS. We show one example of such a potential in Fig. 4. The idea of this potential is to first accelerate the field down a slope, which is accomplished by the term $s\phi$ with a negative slope $s < 0$, relevant as long as $\phi \ll -\Sigma$. As soon as the field crosses the inflection point ($\phi \gg \Sigma$), the second part of the potential becomes important and decelerates the field through another linear slope $a\phi$, now with positive slope $a > 0$. The value of Σ quantifies the width of the transitional region. We additionally allow for an offset κ, which could in principle further accelerate or decelerate the field only within the transitional region.

As further discussed in Appendix C, this toy model provides a great example of early fine-structure constant variation, with a single brutal transition of the value of α around z_{cmb}, very close to a Heaviside function. The two slopes s and a together drive the time and the amplitude of

---
[10]The offset parameter κ was introduced here for the sake of generality but does not display any interesting behavior.
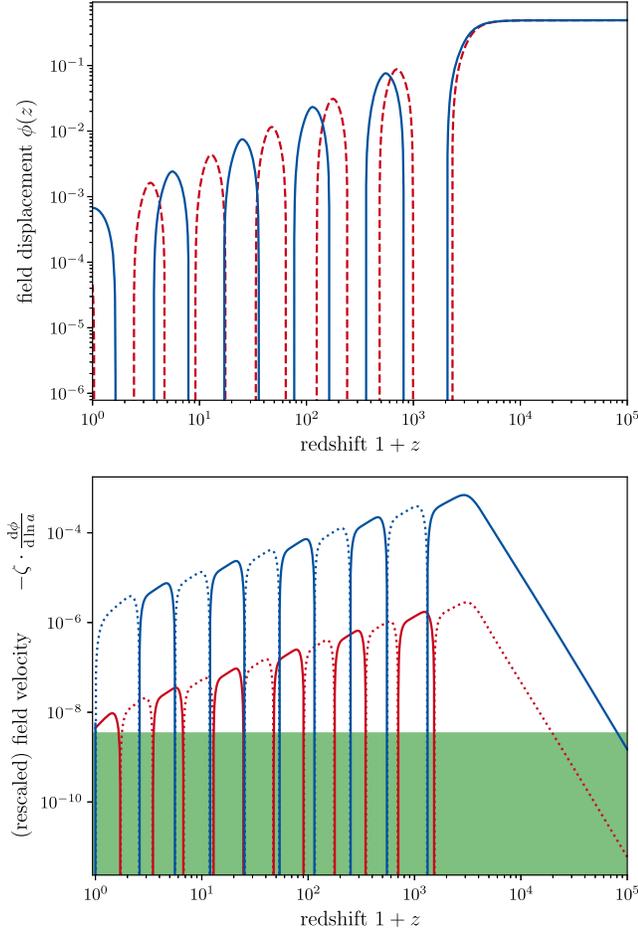
FIG. 3. Axion EDE field (top) and field velocity (bottom) as a function of redshift for the best fit of the full model (red) and for the best fit of the model with $\zeta = 10^{-3}$ fixed (blue). While the red model reaches a small rescaled field velocity due to a small value of ζ, the blue model reaches it only through the fine-tuning of the initial conditions allowing for the exact cancellation of the field speed today in the turnover of an oscillation.
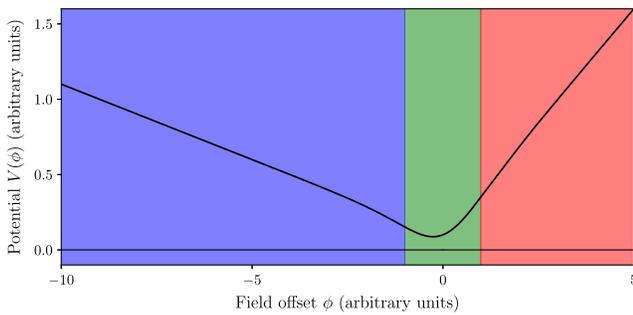

FIG. 4. Example of a potential of the form of Eq. (19) with parameters $a = 0.3$, $s = -0.1$, $\Sigma = 1$, $\kappa = 0$, and $V_0 = 10^{-11}$. The blue range marks the accelerating part of the potential dominated by the downhill slope $s < 0$ for $\phi \ll -\Sigma$, while the red marks the sudden deceleration regime dominated by the uphill slope $a > 0$ for $\phi \gg \Sigma$. The intermediate region is marked in green.

the transition. As expected, ζ simply rescales the whole evolution of the fine-structure constant.

For our toy experiment, we fix the cosmological parameters to the best-fitting parameters of Planck [2] and investigate only the extension of the parameter space, using flat priors on $\log_{10}(-s) \in [1, 15]$, $\log_{10}(\kappa) \in [-5, 5]$, $\log_{10}(a) \in [-10, 5]$, and $\zeta \in [-10, 10]$ and we fix the parameters $\Sigma = 10^{-5}$ and $V_0 = 10^{-8}$. The initial field value is also fixed to $\phi(z \to \infty) = -10^{-2}$. Since the parameter space is highly degenerate and difficult to explore with traditional Markov chain Monte Carlo methods, we make use of the POLYCHORD sampler [68] in this case.

We only put two requirements on the field, namely that it should generate a cosmologically relevant variation of the fine-structure constant at recombination, $\Delta\alpha(z_{\rm cmb}) = 10^{-2} \pm 10^{-3}$ (thus avoiding the potential issue discussed in Statement 2) and that it should avoid constraints by the fine-structure observations detailed in item 2. These two limitations alone are sufficient to force the model parameters into a tight fine-tuned degeneracy, which can be observed in Fig. 5 (lower panel). The degeneracy manifests as a one-to-one relation between the slope for acceleration $(-s)$ and the slope needed for deceleration (a). This slope for deceleration has to balance exactly in such a way as to make the field decelerate entirely until today. Examples of the field trajectories are shown in Fig. 6 where this becomes very apparent as a sudden and rapid stop or drop in velocity toward $z \to 0$ or equivalently $a = 1/(1 + z) \to 1$.

It should also be noted that by construction the field speed typically reaches $\mathcal{O}(10^{-1})$ for this type of potential, thus requiring smaller values of ζ when the field decelerates later, which is typically the case for smaller slopes. This
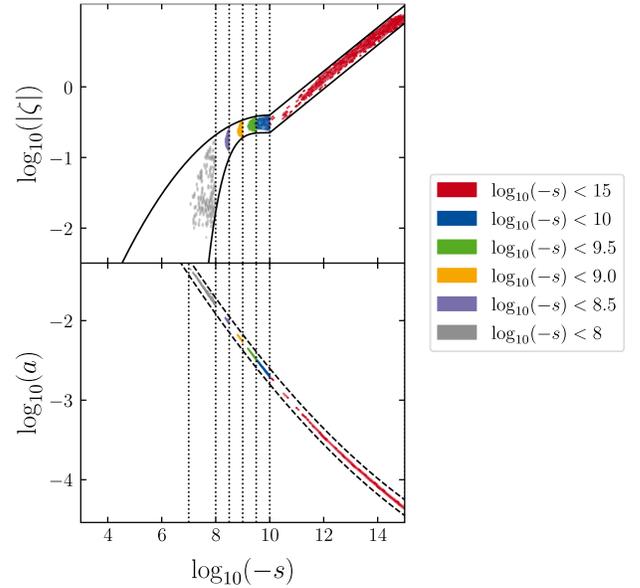

FIG. 5. Constraints on the parameter space of the toy model with the hyperbolic tangent potential. Contours are derived for different bounds on $\log_{10}(-s)$.
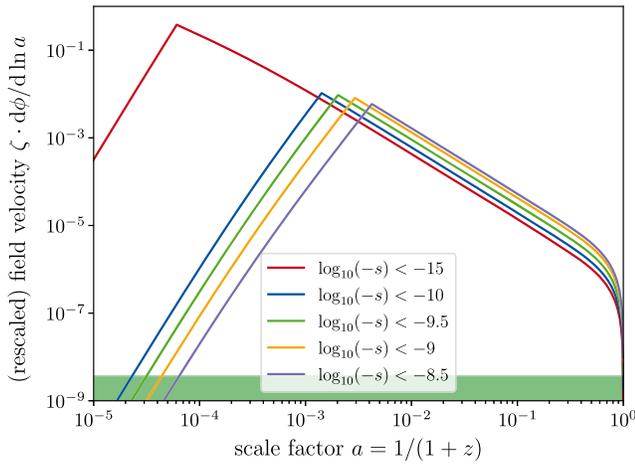
FIG. 6. Rescaled field speeds for the best-fitting models for the various cases. Note that cases with lower $\log_{10}(-s)$ need a smaller ζ to reach the same low rescaled field velocity [the peak of the nonrescaled velocity is always at around $\mathcal{O}(10^{-1})$]. All of the models show a very drastic deceleration toward $z \to 0$ or equivalently $a \to 1$.

imparts an additional (though not as strong) degeneracy between $\log_{10}(|\zeta|)$ and $\log_{10}(-s)$ that can be seen in Fig. 5. As such, even in this case we can conclude that one has to fine-tune the potential such that the field suddenly and rapidly decelerates exactly at $z \to 0$ if one requires cosmologically relevant $\Delta\alpha(z)$ at the time of recombination as per Statement 1.

IV. DISCUSSION AND CONCLUSIONS

In this work we have theoretically derived an almost-no-go theorem (Statement 1) about the impossibility to have a large fine-structure constant at recombination while remaining compatible with late time observations in a single-field model without resorting to extreme fine-tuning of the potential or its initial conditions.

We have given two examples in Sec. III that show the impact of this almost-no-go theorem. Furthermore, we have derived two corollary statements to this main statement, which further restrict the fine-tuned models that avoid Statement 1. Statement 2 states that the prior volume of such fine-tuned models in a Bayesian analysis is typically negligible, while Statement 3 states that other late-Universe data probing the fine-structure data can still be an issue in this case (especially for oscillating potentials). As such, it appears that the often-investigated possibility of a fine-structure constant variation around the time of recombination that has been proposed in the past to ease the Hubble tension [8] does not seem feasible with current late-time and laboratory observations of the fine-structure constant, at least for single-field models.

Indeed, using fairly simple arguments one can derive an upper bound on the variation of the fine-structure constant at recombination of $\Delta\alpha(z_{\rm cmb})/\alpha_0 < 3 \times 10^{-4}/h$ (95% CL),

which can be further generalized to $\Delta\alpha(z_{\rm cmb})/\alpha_0 < 5 \times 10^{-4}$ using a weak prior on the Hubble parameter.

We note that there are a few caveats to this discussion. First and foremost, the argument can be avoided if screening mechanisms exist that would either allow the fine-structure constant α to take on a different value depending on the scale at which it is measured (such as the chameleon mechanism; see for example [69]) or that would act differently on the values of the fine-structure constant in the late and early Universe.[11] Second, we perform a linearization in the field displacement, the validity of which is further discussed in Appendix B. Third, our model does not consider possible couplings of the field to the other sectors of the Universe (such as baryonic matter), as is the case in many theories involving a varying $\alpha(z)$. The existence of such couplings, as the Bekenstein-type couplings $\zeta_i\rho_i$ [36], should impact slightly cosmic evolution of the field but we do not expect it to drastically change our conclusions (the field deceleration still needs to be fine-tuned in terms of the additional coupling parameters).

We also want to point out that the same argument can in principle be made for the variation of other fundamental constants, such as the electron-to-proton mass ratio, which has been shown in the past to be even more successful in easing the Hubble tension [6,7,70]. Ultimately, such an investigation would have to consider the combined variation of the fundamental constants that is expected in most physically motivated models or define a well-defined theory of varying only the electron mass.

Let us also note here that any varying constant model must induce a violation of the Einstein equivalence principle at some level (see for example [12,71]), such that it should also be sharply constrained locally by stringent tests of the universality of free fall performed by experiments such as MICROSCOPE [72]. We did not consider such a constraint here as its relation to the underlying parameters is strongly model dependent and the accurate bound provided by atomic clocks is in any case stringent enough for our statements.

The improvement brought by future datasets is expected to put further stress on coupled scalar field models. The stringent bounds on the stability of fundamental constants from laboratory data will be sharpened by the upcoming nuclear clocks experiments [73] while early Universe constraints will be updated by future generation of CMB

---

[11]Screening mechanisms would typically imply a change in the effective value of α with the local density and hence discard laboratory measurement such as atomic clocks, the density on Earth's surface being much greater than the one of the primordial plasma at recombination. This would however leave the measurements of QSO available, which derive from absorption lines in low-density clouds. However, using the bounds of $(d\alpha/dt)/\alpha_0|_{z\simeq1.5} \sim 10^{-16}h/{\rm yr}$ from QSO measurements and using the same reasoning as above only lead to the mild upper bound of $\Delta\alpha(z_{\rm cmb})/\alpha_0 \lesssim 0.1$.

missions such as Simons Observatory [74] and the combination of CMB-S4 [75] and Lite-BIRD [76], which will measure the anisotropies of the last scattering surface at a cosmic variance accuracy.[12] The combination of these datasets will rule out any cosmologically interesting variation of the fundamental constants at recombination, and thus condemn simple and consistent solutions of the Hubble tension based on such variations.

While the scope of the almost-no-go theorem has to be clearly defined, we stress that the impressive work that has gone into tightening the local laboratory constraints on variations of fundamental constants have yielded such high precision that now even seemingly unrelated cosmological epochs are constrained by these data.

APPENDIX A: OSCILLATORY DYNAMICS

Note that this is a simplified write-up of [42]. The interested reader is encouraged to visit the original source.

In order to determine the behavior of the field envelope, the idea is that the field oscillates between the two critical points of the potential. These are the points where simultaneously $\dot{\phi} = 0$ and $V(\phi_{max}) = V_{max}$, with V_{max} being relatively constant over a single oscillation (as long as the oscillation frequency is much larger than the Hubble frequency).

Then the evolution of V_{max} is directly a tracer of the total energy density ρ [since $\rho(\phi_{max}) = V_{max}$ by definition]. The

advantage of using ρ instead of V is that ρ will typically not decrease drastically during the oscillation, as the potential energy is turned into kinetic energy. To find the density evolution, we decompose $\dot{\phi}^2 = \rho + P \equiv (\gamma + \gamma_p)\bar{\rho}$ with some constant part γ and some oscillatory part γ_p and some mean evolution $\bar{\rho}$. Note that the energy density is *not* conserved [see Eq. (17)], though the overall stress-energy is. Indeed, we use the energy conservation equation $d\rho/d\ln a = -3(\rho + P) = -3(\gamma + \gamma_p)\bar{\rho}$ [which is equivalent to Eq. (17)] to find an equation involving $\bar{\rho}$. On the long timescales of interest we neglect γ_p (average it out, giving also $\bar{\rho}$ from the average of ρ), and we are left with $\bar{\rho} \propto a^{-3\gamma}$. To explicitly find γ, one can average $\dot{\phi}^2/\bar{\rho} \equiv \gamma + \gamma_p$ over any given single oscillation (which will be the same in any other oscillation) and neglect for this the slow variation of the field speed due to the Hubble friction. In this limit $\dot{\phi} = \sqrt{2(\rho - V)}$. This gives a period of oscillation of

$$T = \int \frac{dt}{d\phi}\,dt = \int_{-\phi_{max}}^{\phi_{max}} \frac{1}{\sqrt{2(\rho - V)}}\,d\phi. \qquad (A1)$$

Plugging everything in then leaves us with

$$\gamma \approx \frac{1}{T}\int_0^T \dot{\phi}^2/\bar{\rho}\,dt = \frac{1}{T}\int_{-\phi_{max}}^{\phi_{max}} \dot{\phi}/\bar{\rho}\,d\phi \qquad (A2)$$

$$= \left[\int_{-\phi_{max}}^{\phi_{max}} \sqrt{2(\rho - V)}/\sqrt{\bar{\rho}}\,d\phi\right] \Big/ \left[\int_{-\phi_{max}}^{\phi_{max}} \frac{\sqrt{\bar{\rho}}}{\sqrt{2(\rho - V)}}\,d\phi\right]. \qquad (A3)$$

For the final step, we approximate $\rho \approx \bar{\rho}$ during the oscillation and use $V/\bar{\rho} \approx V/V_{max} \approx (\phi/\phi_{max})^{2n}$ to obtain

$$\gamma = \frac{2n}{n+1}, \qquad (A4)$$

which can then easily be plugged into the equation for

$$V_{max} \propto \bar{\rho} \propto a^{-3\gamma} \propto a^{-6n/(n+1)}, \qquad (A5)$$

and finally using $V_{max} \propto \phi_{max}^{2n}$ we get the evolution of the envelope as

$$\phi_{max} \propto a^{-3/(n+1)}. \qquad (A6)$$

Note that $d[\phi_{max}]/dt \approx -\frac{3H}{n+1}\phi_{max}$ and is not equal to $[d\phi/dt]_{max}$, the envelope of the oscillation speed. For the latter, we simply recall that $\dot{\phi} = \sqrt{2(\rho - V)}$ is maximal when $V \to 0$, and there gives $\dot{\phi} \propto \sqrt{\bar{\rho}} \propto a^{-3n/(n+1)}$ (the same can be found using the virial theorem).

---

[12]See Refs. [5,11] as well as [10] for quantitative forecasts on constraints on varying constants coming from measurements of anisotropies and spectral distortions with future CMB experiments.

APPENDIX B: LINEARIZATION

One may wonder about the legitimacy of the linearization $B_F \simeq 1 + \zeta \Delta\phi$ assumed throughout all this work, especially for the early time evolution of the fields.

To test this linearization, we compare it with the evolution predicted by other well-motivated models. The canonical Bekenstein models [38,77] predict a variation of the form

$$\frac{\Delta\alpha}{\alpha_0} = e^{-\zeta\Delta\phi} - 1 \qquad (B1)$$

with $\zeta = -2$. However, such a high value of ζ is clearly excluded by data for the potential under consideration here, so we generalize it to smaller values of the electromagnetic coupling.

Another relevant model is the runaway dilaton model [13], in which the fine-structure constant evolution is given by [78]

$$\frac{\Delta\alpha}{\alpha_0} = 1 - \zeta(1 - e^{-\Delta\phi}). \qquad (B2)$$

These two models are expected to be well described by our linearization, at least near $\phi = \phi_0$.

The redshift evolutions of $\Delta\alpha/\alpha_0$ in the linearized, the Bekenstein, and the dilaton scenarios are displayed in Fig. 7 using the axionic potential, $\zeta = 10^{-3}$ and best-fit values for all the other parameters. We can see that the Bekenstein-type model is well described by the linearization throughout all of the cosmic history with difference never getting greater than one permille. The runaway dilaton prediction, however, has the same overall behavior but diverges compared to the linearization (up to ∼20%) after it reaches a plateau during radiation domination at very high redshift.

However, this is not a concern for several reasons. First, at the time of recombination ($z \sim 1100$) the typical deviation is still in the permille range. Second, even at earlier times, it can be shown that the leading-order correction causes a smaller variation of $\Delta\alpha(z)$ at equal ζ, and thus our fine-tuning argument would be even stronger for these types of couplings. Third, even in other modified coupling scenarios that differ for a constant ζ by order-unity factors in $\Delta\alpha(z)$, the argument is only slightly weakened but otherwise remains intact
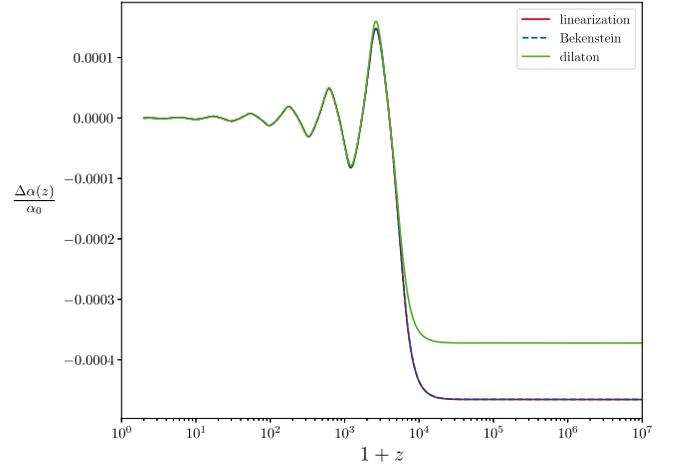

FIG. 7. Comparison of $\Delta\alpha(z)/\alpha_0$ for $\zeta = 10^{-3}$ using different expressions: linearization (red curve), dilatonlike coupling (green curve) and exponential coupling (blue curve).

due to the large order of magnitude between the constraint of Eq. (14) and the typical $\mathfrak{D}/\epsilon$ required to be cosmologically relevant.

One could even imagine stronger deviations in cases where $B_F(\phi)$ has an even more exotic shape that could not even be linearized. It might thus be possible to counter the argument of this paper by using such a well-designed choice of $B_F(\phi)$. However, this choice would have to be motivated from an underlying high-energy theory; otherwise, it would just amount to displace the extreme level of fine-tuning at the level of the initial conditions or potential shape toward the choice of the $B_F(\phi)$ function itself.

We will not discuss here possible kinetic coupling of the field of the form $B_F(\phi, \partial_\mu\phi)$ as in [79] and leave such an investigation for future works.

APPENDIX C: EVOLUTION OF THE FINE-STRUCTURE CONSTANT IN THE MODELS

In this appendix we display several examples of evolutions of $\Delta\alpha(z)/\alpha_0$ and their variation with the relevant underlying model parameters to facilitate the understanding of the main text. In Fig. 8 we show the evolutions for the ALP EDE model, while in Fig. 9 we show the evolution for the tanh toy model of Sec. III B.
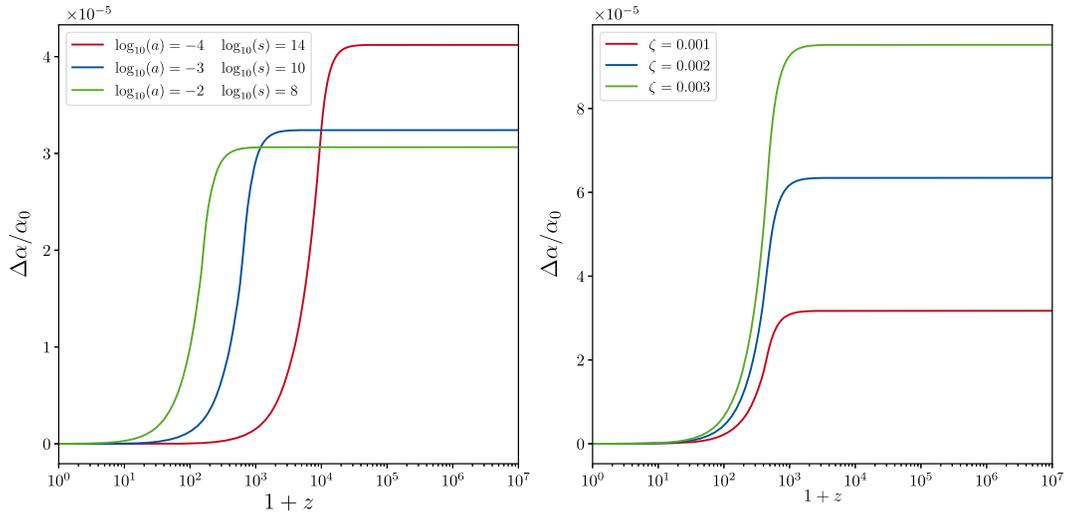
FIG. 8. Impact of the different parameters of the ALP model on the cosmic evolution of $\Delta\alpha(z)/\alpha_0$. We show only the additional model parameters, and note that the usual cosmological ΛCDM parameters do not impart large changes on $\Delta\alpha(z)/\alpha_0$. Both f_{ede} and ζ rescale the amplitude of the fine-structure constant variation, while $a_c = 1/(1+z_c)$ changes the redshift of the transition from constant (due to Hubble drag) to oscillating field. Finally, Θ_i changes the phase of the oscillations (as well as very slightly changing the initial redshift).
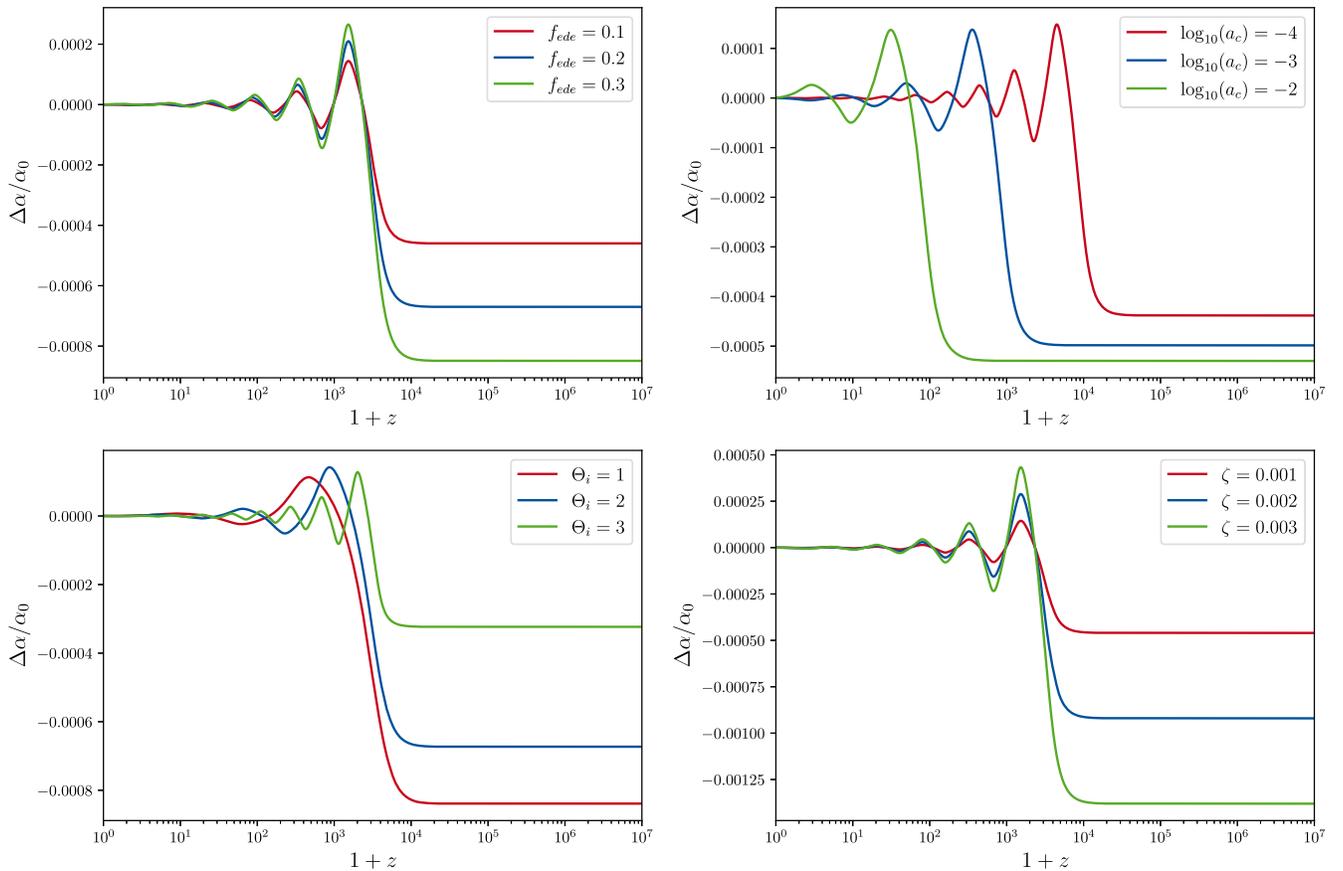

FIG. 9. Impact of the different parameters of the hyperbolic tangent model on the cosmic evolution of $\Delta\alpha(z)/\alpha_0$. For purpose of visualization we follow the degeneracy line of a and s. We see that simultaneously the redshift of the transition and to a lesser degree the amplitude of the transition is modified. For ζ we note the expected behavior of only rescaling the overall amplitude of the $\Delta\alpha(z)/\alpha_0$. We do not show the impact of the offset κ since the variation in terms of $\Delta\alpha/\alpha_0$ is very small.

[1] A. G. Riess *et al.*, A comprehensive measurement of the local value of the Hubble constant with $1 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$ uncertainty from the Hubble space telescope and the SH0ES Team, Astrophys. J. Lett. **934,** L7 (2022).

[2] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro *et al.*, Planck2018 results, Astron. Astrophys. **641,** A1 (2020).

[3] L. Verde, N. Schöneberg, and H. Gil-Marín, A tale of many H_0, arXiv:2311.13305.

[4] Planck Collaboration, Planck intermediate results. XXIV. Constraints on variations in fundamental constants, Astron. Astrophys. **580,** A22 (2015).

[5] L. Hart and J. Chluba, Varying fundamental constants principal component analysis: Additional hints about the Hubble tension, Mon. Not. R. Astron. Soc. **510,** 2206 (2022).

[6] N. Schöneberg, G. F. Abellán, A. Pérez Sánchez, S. J. Witte, V. Poulin, and J. Lesgourgues, The H_0 olympics: A fair ranking of proposed models, Phys. Rep. **984,** 1 (2022).

[7] N. Lee, Y. Ali-Haïmoud, N. Schöneberg, and V. Poulin, What it takes to solve the Hubble tension through modifications of cosmological recombination, Phys. Rev. Lett. **130,** 161003 (2023).

[8] J. Chluba and L. Hart, Varying fundamental constants meet Hubble, arXiv:2309.12083.

[9] L. Hart and J. Chluba, New constraints on time-dependent variations of fundamental constants using Planck data, Mon. Not. R. Astron. Soc. **474,** 1850 (2018).

[10] L. Hart and J. Chluba, Using the cosmological recombination radiation to probe early dark energy and fundamental constant variations, Mon. Not. R. Astron. Soc. **519,** 3664 (2023).

[11] H. Tohfa, J. Crump, E. Baker, L. Hart, D. Grin, M. Brosius, and J. Chluba, A cosmic microwave background search for fine-structure constant evolution, arXiv:2307.06768.

[12] J.-P. Uzan, Varying constants, gravitation and cosmology, Living Rev. Relativity **14,** 2 (2011).

[13] T. Damour and A. M. Polyakov, The string dilation and a least coupling principle, Nucl. Phys. **B423,** 532 (1994).

[14] D. J. E. Marsh, Axion cosmology, Phys. Rep. **643,** 1 (2016).

[15] S. Gasparotto and I. Obata, Cosmic birefringence from monodromic axion dark energy, J. Cosmol. Astropart. Phys. 08 (2022) 025.

[16] V. Poulin, T. L. Smith, D. Grin, T. Karwal, and M. Kamionkowski, Cosmological implications of ultralight axionlike fields, Phys. Rev. D **98,** 083525 (2018).

[17] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, Early dark energy can resolve the Hubble tension, Phys. Rev. Lett. **122,** 221301 (2019).

[18] V. Poulin, T. L. Smith, and T. Karwal, The ups and downs of early dark energy solutions to the Hubble tension: A review of models, hints and constraints circa 2023, Phys. Dark Universe **42,** 101348 (2023).

[19] 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).

[20] M. M. Ivanov, E. McDonough, J. C. Hill, M. Simonović, M. W. Toomey, S. Alexander, and M. Zaldarriaga, Constraining early dark energy with large-scale structure, Phys. Rev. D **102,** 103502 (2020).

[21] G. D'Amico, L. Senatore, P. Zhang, and H. Zheng, The Hubble tension in light of the full-shape analysis of large-scale structure data, J. Cosmol. Astropart. Phys. 05 (2021) 072.

[22] R. Murgia, G. F. Abellán, and V. Poulin, Early dark energy resolution to the Hubble tension in light of weak lensing surveys and lensing anomalies, Phys. Rev. D **103,** 063502 (2021).

[23] A. Amon and G. Efstathiou, A non-linear solution to the S_8 tension?, Mon. Not. R. Astron. Soc. **516,** 5355 (2022).

[24] S. Goldstein, J. C. Hill, V. Iršič, and B. D. Sherwin, Canonical Hubble-tension-resolving early dark energy cosmologies are inconsistent with the Lyman-α forest, Phys. Rev. Lett. **131,** 201001 (2023).

[25] R. Gsponer, R. Zhao, J. Donald-McCann, D. Bacon, K. Koyama, R. Crittenden, T. Simon, and E.-M. Mueller, Cosmological constraints on early dark energy from the full shape analysis of eBOSS DR16, arXiv:2312.01977.

[26] R. D. Peccei, The strong *CP* problem and axions, in *Axions*, edited by M. Kuster, G. Raffelt, and B. Beltrán (Springer, New York, 2008), Vol. 741, p. 3.

[27] P. Svrcek and E. Witten, Axions in string theory, J. High Energy Phys. 06 (2006) 051.

[28] Y. Minami and E. Komatsu, New extraction of the cosmic birefringence from the Planck 2018 polarization data, Phys. Rev. Lett. **125,** 221301 (2020).

[29] P. Diego-Palazuelos, J. R. Eskilt, Y. Minami, M. Tristram, R. M. Sullivan, A. J. Banday, R. B. Barreiro, H. K. Eriksen, K. M. Górski, R. Keskitalo, E. Komatsu, E. Martínez-González, D. Scott, P. Vielva, and I. K. Wehus, Cosmic birefringence from Planck public release 4, arXiv:2203.04830.

[30] P. Diego-Palazuelos, E. Martínez-González, P. Vielva, R. B. Barreiro, M. Tristram, E. de la Hoz, J. R. Eskilt, Y. Minami, R. M. Sullivan, A. J. Banday, K. M. Górski, R. Keskitalo, E. Komatsu, and D. Scott, Robustness of cosmic birefringence measurement against Galactic foreground emission and instrumental systematics, J. Cosmol. Astropart. Phys. 01 (2023) 044.

[31] G. Zagatti, M. Bortolami, A. Gruppuso, P. Natoli, L. Pagano, and G. Fabbian, Planck constraints on Cosmic Birefringence and its cross-correlation with the CMB, arXiv:2401.11973.

[32] V. Flambaum, S. Lambert, and P. Maxim, Scalar-tensor theories with pseudoscalar couplings, Phys. Rev. D **80,** 105021 (2009).

[33] N. Schöneberg, L. Vacher, J. D. F. Dias, M. M. C. D. Carvalho, and C. J. A. P. Martins, News from the swampland—constraining string theory with astrophysics and cosmology, J. Cosmol. Astropart. Phys. 10 (2023) 039.

[34] C. J. A. P. Martins, F. P. S. A. Ferreira, and P. V. Marto, Varying fine-structure constant cosmography, Phys. Lett. B **827,** 137002 (2022).

[35] M. J. Duff, Comment on time-variation of fundamental constants, arXiv:hep-th/0208093.

[36] K. A. Olive and M. Pospelov, Evolution of the fine structure constant driven by dark matter and the cosmological constant, Phys. Rev. D **65,** 085044 (2002).

[37] J. D. Bekenstein, Fine-structure constant: Is it really a constant?, Phys. Rev. D **25**, 1527 (1982).

[38] H. B. Sandvik, J. D. Barrow, and J. Magueijo, A simple cosmology with a varying fine structure constant, Phys. Rev. Lett. **88**, 031302 (2002).

[39] M. Filzinger, S. Dörscher, R. Lange, J. Klose, M. Steinel, E. Benkler, E. Peik, C. Lisdat, and N. Huntemann, Improved limits on the coupling of ultralight bosonic dark matter to photons from optical atomic clock comparisons, Phys. Rev. Lett. **130**, 253001 (2023).

[40] D. Brout *et al.*, The Pantheon + analysis: Cosmological constraints, Astrophys. J. **938**, 110 (2022).

[41] S. Alam *et al.*, Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory, Phys. Rev. D **103**, 083533 (2021).

[42] M. S. Turner, Coherent scalar-field oscillations in an expanding universe, Phys. Rev. D **28**, 1243 (1983).

[43] C. B. Adams *et al.*, Axion dark matter, arXiv:2203.14923.

[44] E. Calabrese, E. Menegoni, C. J. A. P. Martins, A. Melchiorri, and G. Rocha, Constraining variations in the fine structure constant in the presence of early dark energy, Phys. Rev. D **84**, 023518 (2011).

[45] S. Alexander and E. McDonough, Axion-dilaton destabilization and the Hubble tension, Phys. Lett. B **797**, 134830 (2019).

[46] J. Lesgourgues, The cosmic linear anisotropy solving system (class) i: Overview, arXiv:1104.2932.

[47] https://github.com/brinckmann/montepython_public.

[48] T. Brinckmann and J. Lesgourgues, MontePython 3: Boosted MCMC sampler and other features, Phys. Dark Universe **24**, 100260 (2019).

[49] L. Vacher, J. D. F. Dias, N. Schöneberg, C. J. A. P. Martins, S. Vinzl, S. Nesseris, G. Cañas-Herrera, and M. Martinelli, Constraints on extended Bekenstein models from cosmological, astrophysical, and local data, Phys. Rev. D **106**, 083522 (2022).

[50] L. Vacher, N. Schöneberg, J. D. F. Dias, C. J. A. P. Martins, and F. Pimenta, Runaway dilaton models: Improved constraints from the full cosmological evolution, Phys. Rev. D **107**, 104002 (2023).

[51] https://github.com/PoulinV/AxiCLASS.

[52] 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).

[53] Our implementation of aninstantaneous transition of the fine-structure constant is publicly available on the software's repository: https://github.com/lesgourg/class_public.

[54] Available at https://github.com/schoeneberg/liquidcosmo based on getdist https://github.com/cmbant/getdist.

[55] M. T. Clara and C. J. A. P. Martins, Primordial nucleosynthesis with varying fundamental constants. Improved constraints and a possible solution to the lithium problem, Astron. Astrophys. **633**, L11 (2020).

[56] Planck Collaboration, Planck 2018 results. V. CMB power spectra and likelihoods, Astron. Astrophys. **641**, A5 (2020).

[57] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VIII. Gravitational lensing, Astron. Astrophys. **641**, A8 (2020).

[58] T. B. 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).

[59] D. Scolnic *et al.*, The Pantheon + analysis: The full data set and light-curve release, Astrophys. J. **938**, 113 (2022).

[60] M. Moresco *et al.*, Unveiling the Universe with emerging cosmological probes, Living Rev. Relativity **25**, 6 (2022).

[61] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, Cosmic distances calibrated to 1% precision with Gaia EDR3 parallaxes and Hubble space telescope photometry of 75 Milky Way cepheids confirm tension with ΛCDM, Astrophys. J. Lett. **908**, L6 (2021).

[62] D. Camarena and V. Marra, On the use of the local prior on the absolute magnitude of Type Ia supernovae in cosmological inference, Mon. Not. R. Astron. Soc. **504**, 5164 (2021).

[63] C. J. A. P. Martins, The status of varying constants: A review of the physics, searches and implications, Rep. Prog. Phys. **80**, 126902 (2017).

[64] J. K. Webb, J. A. King, M. T. Murphy, V. V. Flambaum, R. F. Carswell, and M. B. Bainbridge, Indications of a spatial variation of the fine structure constant, Phys. Rev. Lett. **107**, 191101 (2011).

[65] M. T. Murphy and K. L. Cooksey, Subaru telescope limits on cosmological variations in the fine-structure constant, Mon. Not. R. Astron. Soc. **471**, 4930 (2017).

[66] M. T. Murphy *et al.*, Fundamental physics with ESPRESSO: Precise limit on variations in the fine-structure constant towards the bright quasar HE 0515 − 4414, Astron. Astrophys. **658**, A123 (2022).

[67] Y. V. Petrov, A. I. Nazarov, M. S. Onegin, V. Y. Petrov, and E. G. Sakhnovsky, Natural nuclear reactor at Oklo and variation of fundamental constants: Computation of neutronics of a fresh core, Phys. Rev. C **74**, 064610 (2006).

[68] https://cobaya.readthedocs.io/en/latest/sampler_polychord .html.

[69] K. A. Olive and M. Pospelov, Environmental dependence of masses and coupling constants, Phys. Rev. D **77**, 043524 (2008).

[70] T. Sekiguchi and T. Takahashi, Early recombination as a solution to the H_0 tension, Phys. Rev. D **103**, 083507 (2021).

[71] T. Damour, F. Piazza, and G. Veneziano, Violations of the equivalence principle in a dilaton-runaway scenario, Phys. Rev. D **66**, 046007 (2002).

[72] P. Touboul *et al.*, MICROSCOPE mission: First results of a space test of the equivalence principle, Phys. Rev. Lett. **119**, 231101 (2017).

[73] P. Fadeev, J. C. Berengut, and V. V. Flambaum, Sensitivity of ^{229}Th nuclear clock transition to variation of the fine-structure constant, Phys. Rev. A **102**, 052833 (2020),

[74] The Simons Observatory collaboration, The Simons Observatory, in *BAAS*, Vol. 51 (2019), p. 147.

[75] CMB-S4 Collaboration, CMB-S4 Science Case, Reference Design, and Project Plan, arXiv:1907.04473

[76] LiteBIRD Collaboration, Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey, Prog. Theor. Exp. Phys., **11443** (2022), 114432F.

[77] J. D. Bekenstein, Fine-structure constant variability, equivalence principle, and cosmology, Phys. Rev. D **66**, 123514 (2002).

[78] C. J. A. P. Martins, P. E. Vielzeuf, M. Martinelli, E. Calabrese, and S. Pandolfi, Evolution of the fine-structure constant in runaway dilaton models, Phys. Lett. B **743**, 377 (2015).

[79] B. J. Barros and V. da Fonseca, Coupling quintessence kinetics to electromagnetism, J. Cosmol. Astropart. Phys. 06 (2023) 048.