

Searching for coupled, hyperlight scalars across cosmic history

Masha Baryakhtar,^{1,*} Olivier Simon^{2,3,†} and Zachary J. Weiner^{1,4,‡}¹*Department of Physics, University of Washington, Seattle, Washington 98195, USA*²*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*³*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*⁴*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

(Received 18 February 2025; accepted 21 May 2025; published 25 June 2025)

Cosmological scalar fields coupled to the Standard Model drive temporal variations in the fundamental constants that grow with redshift, positioning the early Universe as a powerful tool to study such models. We investigate the dynamics and phenomenology of coupled scalars from the early Universe to the present to consistently leverage the myriad searches for time-varying constants and the cosmological signatures of scalars' gravitational effects. We compute the in-medium contribution from Standard Model particles to the scalar's dynamics and identify only a limited range of couplings for which the scalar has an observable impact on the fundamental constants without either evolving before recombination or gravitating non-negligibly. We then extend existing laboratory and astrophysical bounds to the hyperlight scalar regime. We present joint limits from the early and late Universe, specializing to hyperlight, quadratically coupled scalars that modulate the mass of the electron or the strength of electromagnetism and make up a subcomponent of the dark matter today. Our dedicated analysis of observations of the cosmic microwave background, baryon acoustic oscillations, and type Ia supernovae provides the most stringent constraints on quadratically coupled scalars with masses from $10^{-28.5}$ to $\sim 10^{-31}$ eV, below which quasar absorption spectra yield stronger bounds. These results jointly limit hyperlight scalars that comprise a few percent of the current dark matter density to near- or subgravitational couplings to electrons or photons.

DOI: [10.1103/pyls-gvyv](https://doi.org/10.1103/pyls-gvyv)

I. INTRODUCTION

Particle physics often progresses by postulating that the parameters of a theory, hitherto considered fundamental constants, are dynamical functions of an underlying field. The Standard Model Higgs mechanism provides the most successful such example [1–3], while others are proposed solutions to extant issues, such as the Peccei-Quinn mechanism to address the strong CP problem [4–6]. In this work we consider extensions to the Standard Model that introduce new scalar fields whose values control one or more fundamental “constants,” which are then meaningfully and observably *nonconstant* throughout spacetime when the scalar has nontrivial dynamics.

Variations of fundamental constants can be mechanized by new scalar fields that arise, for example, in scalar-tensor

theories of gravity [7–9] and in theories with compactified extra dimensions (like string theory) in the form of the so-called dilaton or moduli fields [7,10–18]. “Bottom-up” perspectives use principles of effective field theory to build Lagrangians for dynamical variations in fundamental constants by promoting Standard Model (SM) parameters to depend on new scalars constrained by symmetry principles [19–22]. These scalars result in three (related) types of observational signatures that have spurred vibrant experimental programs: modulations of fundamental constants with space, time, and environment [7,11,14,20,23,24], observable violations of the weak equivalence principle [12,15,21,22,25,26], and deviations from the $1/r$ law of gravitational interactions [18,27,28].

Scalar fields have cosmological consequences as well—even simply through gravity if they are sufficiently abundant. A cosmological density of the new scalar may be produced in the very early Universe, for instance through the misalignment mechanism [13,29–36]. Depending on its initial conditions and interactions, the scalar makes up at least a subcomponent of the cosmological dark sector today as matter, radiation, or something more exotic. In general, scalar fields redshift with the expansion of the Universe and their observational effects are more pronounced earlier in cosmological history—more so if the leading dependence

* Contact author: mbaryakh@uw.edu† Contact author: osimon@princeton.edu‡ Contact author: zweiner@perimeterinstitute.ca

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

of fundamental constants on the scalar field amplitude is quadratic [23,26,37,38] or higher rather than linear. In an era of precision cosmology, early-Universe observations are particularly apt probes of dynamical scalar fields and may explore parameter space complementary to that from traditional particle physics tests.

Primordial nucleosynthesis and recombination are the two earliest processes with observable relics that are sensitive to SM parameters; measurements of nuclear abundances and the cosmic microwave background (CMB) therefore test for possible evolution of fundamental constants at temperatures around an MeV and eV, respectively. Limits on the variation in the fine-structure constant [39–50] and the electron mass [51–58] at recombination have been placed largely without regard for the microphysics involved. On the other hand, nucleosynthesis provides some of the leading constraints on quadratically coupled scalar dark matter [37,59]. Unlike the yields of primordial elements, which depend only on the conditions while they formed, CMB anisotropies are sensitive not just to physics at the time of recombination but also to intervening dynamics of the Universe up to the present day. Therefore, a consistent CMB analysis must take into account the joint effect of a new scalar’s cosmological abundance as well as its microphysical couplings to matter [60]. Grounding an analysis with a concrete model enables a consistent study of the gravitational interactions of the scalar field, the impact of SM matter on the cosmological evolution of the field, and complementary constraints of fundamental constants across epochs.

In this work, we explore the parameter space of a hyperlight scalar field responsible for variations in the mass of the electron m_e and the electromagnetic fine-structure constant α . We place constraints on the scalar’s cosmological abundance and microphysical coupling strengths with cosmological datasets, utilizing our prior work in which a light scalar field is consistently implemented [60], as well as with complementary late-time astrophysical and terrestrial probes. We target regimes that mimic the phenomenological models of prior study in which α and m_e took on different but time-independent values during recombination, which imposes nontrivial constraints on viable models. In particular, we show that such scalar fields may neither be too heavy nor too strongly coupled to SM matter, for in either case the scalar would become dynamical before photon last scattering. But neither can the scalar be too light, for it must have redshifted via its cosmological dynamics sufficiently to satisfy stringent constraints on fundamental constant variations nearer to the present day. We therefore focus on massive, hyperlight subcomponents of dark matter with particle mass $10^{-32} \text{ eV} \lesssim m_\phi \lesssim 10^{-28} \text{ eV}$. Constraints from late-time searches have to be reinterpreted for such low masses; we find that cosmological constraints are the most stringent down to $10^{-31} - 10^{-30} \text{ eV}$, where they are superseded by measurements of quasar spectra.

This work is divided as follows. In Sec. II, we review the effective field theory of dynamical modulations in fundamental constants by a scalar field, with a focus on α and m_e . We also discuss the influence of the cosmological abundance of SM particles on the scalar’s effective potential and dynamics. Section III recasts searches for varying constants and new scalar fields coupled to the SM from quasar absorption lines, equivalence principle violation, and precision metrology. Finally, in Sec. IV, we detail the physical impact of varying fundamental constants on nucleosynthesis and review the constraints using CMB and other cosmological observables derived in Ref. [60]. We combine the results of the preceding sections in Sec. IV C to set limits on a new scalar’s coupling to the photon or the electron across the hyperlight parameter space, and we conclude in Sec. V.

We use natural units in which $\hbar = c = 1$ and define the reduced Planck mass $M_{\text{pl}} \equiv 1/\sqrt{8\pi G} = 2.435 \times 10^{18} \text{ GeV}$. We employ the Einstein summation convention for space-time indices; repeated spatial indices (Latin characters) are implicitly summed regardless of their placement. We also use upright boldface to denote spatial vectors. Unless otherwise specified, we fix a homogeneous Friedmann-Lemaître-Robertson-Walker spacetime with metric

$$ds^2 = dt^2 - a(t)^2 \delta_{ij} dx^i dx^j, \quad (1.1)$$

with $a(t)$ the scale factor. Dots denote derivatives with respect to cosmic time t , and the Hubble rate is $H \equiv \dot{a}/a$.

II. SCALAR FIELD MODELS FOR VARYING FUNDAMENTAL CONSTANTS

In this section, we review the field-theoretic description of the modulation of fundamental constants by a new scalar field. After some general considerations, we narrow our scope to the mass of the electron and the electromagnetic fine-structure constant. We summarize the fundamental couplings of such fields to matter and radiation (Sec. II A) as well as their effective coupling to composite matter (Sec. II B). We also detail how the cosmological abundance of SM particles contributes to the scalar’s effective potential (Sec. II C) and comment on variations in α and m_e due to such in-medium effects (Sec. II D). Finally, we describe the cosmological dynamics of coupled scalars in Sec. II E, considering in turn the case in which the scalar’s dynamics in an expanding Universe are wholly dictated by its bare potential and where SM matter makes a non-negligible contribution to its effective potential.

A. Varying constants with new scalars

We set the stage by defining notation for a general description of theories featuring new scalars coupled to the SM. We consider a new, real scalar field that couples to the Standard Model in a manner that effectively generates

spacetime dependence of some subset of SM parameters. We take the simplest case of a massive field φ with negligible self-interactions,

$$\mathcal{L} = \frac{1}{8\pi G} (\partial_\mu \varphi \partial^\mu \varphi - m_\varphi^2 \varphi^2) + \mathcal{L}_{\text{SM}}[\varphi, \dots]; \quad (2.1)$$

the ellipsis denotes other SM fields. The scalar's couplings each take the generic form of a function of φ multiplying a dimension-four operator already present in the SM Lagrangian $\mathcal{L}_{\text{SM}}$. The canonical, dimensionful field is $\phi \equiv \varphi/\sqrt{4\pi G} = \sqrt{2}M_{\text{pl}}\varphi$, while φ 's kinetic term is analogous to that of gravity. The normalization of φ is useful for parametrizing its strength relative to dimensionless metric perturbations in cosmological and fifth-force contexts, as the SM contribution to the Klein-Gordon equation for φ resembles the gravitational Poisson equation:

$$\ddot{\varphi}(t, \mathbf{x}) + 3H\dot{\varphi}(t, \mathbf{x}) - \frac{1}{a(t)^2} \partial_i \partial_i \varphi(t, \mathbf{x}) + m_\varphi^2 \varphi(t, \mathbf{x}) = 4\pi G \frac{\partial \mathcal{L}_{\text{SM}}}{\partial \varphi}. \quad (2.2)$$

If the bare mass term dominates both Hubble friction and the interaction term on the right-hand side of Eq. (2.2), the scalar behaves as (a component of) dark matter.

We prescribe the spacetime dependence of a fundamental constant λ , inherited from a slowly varying background field $\varphi(t, \mathbf{x})$, via coupling functions g_λ defined by

$$g'_\lambda(\varphi) \equiv \frac{1}{\lambda(\varphi)} \frac{d\lambda(\varphi)}{d\varphi}. \quad (2.3)$$

Integrating Eq. (2.3) and taking g_λ to vanish at the vacuum expectation value (VEV) $\langle \varphi \rangle$ gives

$$\lambda(\varphi) = \lambda(\langle \varphi \rangle) e^{g_\lambda(\varphi)} \approx \lambda(\langle \varphi \rangle) [1 + g_\lambda(\varphi)], \quad (2.4)$$

the second equality at leading order in small $g_\lambda(\varphi)$. Consequently, the scalar shifts λ from its VEV by

$$\frac{\Delta\lambda}{\lambda(0)} \equiv \frac{\lambda(\varphi) - \lambda(0)}{\lambda(0)} \approx g_\lambda(\varphi) \quad (2.5)$$

to leading order, where we take $\langle \varphi \rangle = 0$ as for Eq. (2.1). Since the value of the cosmological background field today, $\varphi_0 \equiv \varphi(t_0)$, is not its VEV in general, $\lambda(0)$ may not coincide with the coupling $\lambda_0 \equiv \lambda(\varphi_0)$ that would be measured by a contemporary experiment. Tests of equivalence principle violation and other scalar phenomenology beyond the present-day value of fundamental constants alone can, in principle, distinguish between $\lambda(0)$ and λ_0 .

The coupling functions can be expanded as a power series around $\varphi = 0$,

$$g_\lambda(\varphi) = \sum_{n=1} \frac{d_\lambda^{(n)}}{n!} \varphi^n. \quad (2.6)$$

At small φ and still assuming that φ is slowly varying, such that its derivative interactions are relatively suppressed, the coupling functions are parametrically dominated by the lowest powers of the expansion, motivating the existing studies of linear and quadratic couplings. Tests of the equivalence principle stringently limit a light scalar's linear couplings to the SM to far below gravitational strength, independently of its cosmological abundance [26] (see Sec. III D). Furthermore, in the perturbative regime $|g_\lambda|$ monotonically increases with φ , which in turn generally decays in amplitude as the Universe expands. We therefore expect cosmological signatures at high redshift to be larger in magnitude compared to present-day ones, to an extent that is greater for higher-order interactions.

The set of specific “constants” λ that one considers “fundamental” depends on the energy scales under consideration. In this work, we restrict our attention to the fine-structure constant $\alpha \equiv e^2/4\pi$ and the electron mass m_e because of their particular relevance to cosmology and recombination. We thus replace α and m_e with $\alpha(\varphi)$ and $m_e(\varphi)$ in the SM Lagrangian:

$$\mathcal{L}_{\text{SM}}[\varphi, \dots] = -\frac{1}{16\pi\alpha(\varphi)} F^{\mu\nu} F_{\mu\nu} + \bar{e}[i\not{D} - m_e(\varphi)]e + \dots, \quad (2.7)$$

where the ellipsis designates the rest of the SM Lagrangian below the scale of electroweak symmetry breaking. Note that Eq. (2.7) takes the “Yang-Mills” normalization with the (inverse) gauge coupling multiplying the kinetic terms of gauge fields.

In the remainder of this work we aim to treat scalar couplings as generally as is practical; based on the above considerations, however, we typically anchor the discussion to quadratic coupling functions with a vanishingly small linear term,

$$g_\lambda(\varphi) \approx d_\lambda^{(2)} \varphi^2/2, \quad (2.8)$$

and electromagnetically coupled [Eq. (2.7)] scalars, as encoded by the Lagrangian¹

$$\mathcal{L}_{\text{SM}}[\varphi, \dots] \approx \frac{\varphi^2}{2} \left(\frac{d_e^{(2)}}{16\pi\alpha(0)} F^{\mu\nu} F_{\mu\nu} - m_e(0) \bar{e}e \right) + \mathcal{L}_{\text{SM}}[0, \dots], \quad (2.9)$$

¹We replace the subscript α 's with e 's for the photon couplings, e.g., $d_e^{(n)} \equiv d_\alpha^{(n)}$, to conform with prior literature.

where $\mathcal{L}_{\text{SM}}[0, \dots]$ denotes the usual SM Lagrangian. The corresponding φ -dependent electron mass and fine-structure constant are

$$m_e(\varphi) \approx m_e(0) \left(1 + d_{m_e}^{(2)} \varphi^2 / 2 \right), \quad (2.10a)$$

$$\alpha(\varphi) \approx \alpha(0) \left(1 + d_e^{(2)} \varphi^2 / 2 \right). \quad (2.10b)$$

The assumed suppression of terms linear in φ could be achieved with a $\mathbb{Z}_2^{\varphi}: \varphi \mapsto -\varphi$ symmetry; we mention models that feature such a symmetry in the following.

1. Model-building considerations

Before proceeding to the phenomenology of hyperlight scalars in the early Universe, we comment on several aspects of these models in the context of particle theory. A more complete understanding the class of theories we consider calls for a UV completion, which we do not specify here. We focus only on a subset of possible couplings (to α and m_e); the others pertinent at lower energies are the strong confinement scale Λ_{QCD} and the up and down quark masses, m_u and m_d [21,22]. These parameters would ultimately inherit φ dependence from higher-scale (i.e., electroweak-scale) physics—the Higgs’s mass, vacuum expectation value, and Yukawa couplings to fermions along with the three $\text{SU}(3) \times \text{SU}(2) \times \text{U}(1)$ gauge couplings [11,14,61–63]. At yet higher energies, the scalar couplings in Eq. (2.7) are characteristic of so-called “dilaton” and moduli fields found in string theory compactifications and supersymmetric models, although typically with an unsuppressed linear term [14,15,18,21,22].

Many UV completions thus would, in addition to electron and photon couplings, introduce couplings to quarks and gluons. In the Higgs portal model, in which the scalar interacts with the SM Higgs field H via renormalizable operator $\varphi^2 H^\dagger H$, the Yukawa interactions between the Higgs and SM fermions f generate interactions of the form $m_f(\varphi) \bar{f} f$ for all fundamental SM fermions, each with the same φ dependence. In particular, $g'_{m_e}(\varphi) = g'_v(\varphi)$ for the electron. The QCD scale Λ_{QCD} also depends on heavy quark masses via the renormalization of the strong gauge coupling α_s and introduces variations in nucleon masses at the same order. Specifically, the dependence on the Higgs VEV is $g'_{m_q}(\varphi) = g'_v(\varphi)$ for quark masses and $g'_{\Lambda_{\text{QCD}}} = (2/9)g'_v$ for the QCD scale, yielding $g'_{m_N} \approx 0.79g'_{\Lambda_{\text{QCD}}} + 0.24g'_v = 0.42g'_v$ for nucleon masses, with the second term dominated by the strange quark’s contribution to nucleon masses [64]. The “axi-Higgs” model, the cosmology of which is studied in Refs. [65–67], is a recent example of this class of models. References [65–67] claim that, because nucleon masses are parametrically dominated by Λ_{QCD} , their relative variations are suppressed

as $g'_{m_N} \approx g'_{m_{u,d}} / \Lambda_{\text{QCD}} \approx 10^{-3} g'_v$; they therefore neglect variations in baryon masses when modeling the physics of recombination and nucleosynthesis. However, as outlined above, any model derived from Higgs couplings is not well described by variations in the m_e alone at low energies. We confine ourselves to scalar field interactions with the electron Yukawa coupling and the fine-structure constant [Eq. (2.7)] and leave the interesting question of the effect of baryon mass variations on cosmology to future work.

Depending on the assumed symmetries, scalars may interact with fermions and gauge bosons via operators other than those parametrized by Eq. (2.7). For instance, a shift-symmetric pseudovector coupling is characteristic of (pseudo)Nambu-Goldstone bosons such as axionlike particles [68–71]. A pseudoscalar coupling is also allowed by symmetries and is equivalent to the pseudovector coupling alone only in processes whose Feynman diagrams have a single φ per electron [[72], Sec. 90]. Pseudoscalar (PS) and pseudovector (PV) operators effectively contribute to those explicitly included in Eq. (2.7). Consider a background field φ with general electron couplings of the form

$$\begin{aligned} \mathcal{L} \supset i \bar{e} \not{D} e - m_e(0) [1 + G_S(\varphi)] \bar{e} e + i m_e(0) G_{\text{PS}}(\varphi) \bar{e} \gamma^5 e \\ + G_{\text{PV}}(\varphi) \partial_\mu \varphi \bar{e} \gamma^\mu \gamma^5 e + \dots, \end{aligned} \quad (2.11)$$

with the ellipsis designating Lorentz structures with higher derivatives of φ . The scalar functions $G_S(\varphi)$, $G_{\text{PS}}(\varphi)$, and $G_{\text{PV}}(\varphi)$ are not uniquely defined—they mix into one another under chiral rotations. One can show that a chiral rotation $e^{i\gamma^5 \theta(\varphi)/2} = \cos[\theta(\varphi)/2] I_4 + i\gamma^5 \sin[\theta(\varphi)/2]$ of the electron acts as a two-dimensional rotation $[\text{SO}(2)]$ on the two-component vector with entries $1 + G_S(\varphi)$ and $G_{\text{PS}}(\varphi)$. With an appropriate choice of $\theta(\varphi)$, one can therefore rotate away the pseudoscalar or the scalar term (including the bare mass term) entirely, at the cost of introducing a pseudovector coupling $i\theta'(\varphi) \partial_\mu \varphi \bar{e} \gamma^\mu \gamma^5 e / 2$. The correction to the electron mass for slowly varying $\varphi(t, \mathbf{x})$ is

$$g_{m_e}(\varphi) = \sqrt{[1 + G_S(\varphi)]^2 + G_{\text{PS}}(\varphi)^2} - 1, \quad (2.12)$$

which is invariant under chiral rotations. One may alternatively obtain this result, with additional corrections proportional to gradients of the field from the pseudovector coupling, by computing the poles of the electronic two-point function for Eq. (2.11) with φ as a background field.

A similar discussion holds for photon couplings; for example

$$\mathcal{L} \supset -\frac{1 - G_{FF}(\varphi)}{16\pi\alpha(0)} F_{\mu\nu} F^{\mu\nu} + G_{F\tilde{F}}(\varphi) \frac{F_{\mu\nu} \tilde{F}^{\mu\nu}}{16\pi\alpha(0)} + \dots, \quad (2.13)$$

where again G_{FF} and $G_{F\tilde{F}}$ are scalar functions of φ and $\dots$ designates Lorentz structures with higher derivatives of φ .

The phenomenology of the latter coupling has been studied extensively and for instance can source cosmic birefringence [73–79]. Models can also simultaneously feature both couplings, which transform differently under parity [80–82]. For our purposes, the electromagnetic exchange potential (i.e., the amplitude for $ee \rightarrow ee$) receives corrections from all interaction terms, resulting in a coupling function $g_e(\varphi) \equiv g_\alpha(\varphi)$ that depends on $G_{F\tilde{F}}$ in addition to G_{FF} . However, $F_{\mu\nu}\tilde{F}^{\mu\nu}$ by itself is a total derivative term, which holds for $G_{F\tilde{F}}(\varphi)F_{\mu\nu}\tilde{F}^{\mu\nu}$ unless φ varies in space-time. For an Abelian gauge boson like the photon, any physical effect must therefore be proportional to derivatives of φ .

Thus, both the PV coupling in Eq. (2.11) and the $F\tilde{F}$ coupling in Eq. (2.13) are “derivative” couplings, the corrections from which are suppressed for the nonrelativistic, ultralight fields we consider here and the nonderivative, scalar couplings fully captures the low-energy physics.

Finally, we note that a light mass for a scalar with the non-shift-symmetric couplings we consider is not protected from radiative corrections. For instance, if the theory is ultraviolet completed by new physics at some cutoff scale Λ , the electron coupling contributes a one-loop effective potential $m_e(\varphi)^2\Lambda^2/16\pi^2$ at leading order in large Λ .² As discussed in the Introduction, we focus on cosmologies featuring fundamental constants that vary only after recombination and thus m_φ near or below 10^{-28} eV—not merely ultralight, but *hyperlight*. Even for Planckian scalar misalignments, the scale of the bare potential is only of order eV^4 , implying some drastic cancellation of quantum corrections (which naively are at the very least m_e^4 in magnitude) to achieve effective masses so light. For the quadratic couplings we consider, quantum corrections nominally shift the mass by $\sim d_{m_e}^{(2)}m_e(0)^2\Lambda^2/8\pi^2M_{\text{pl}}^2$; an order-unity coupling renders masses $m_\varphi \lesssim 10^{-11}$ eV unnatural for $\Lambda = 100$ GeV [24,37,38,84]. Moreover, radiative corrections generate higher-order self-interactions, since $m_e(\varphi)^2/m_e(0)^2 = e^{2g_{m_e}(\varphi)} \approx 1 + 2g_{m_e}(\varphi) + 2g_{m_e}(\varphi)^2$. Corrections of yet higher order arise beyond one loop and at different orders in Λ but are in general suppressed by powers of $g_{m_e}(\varphi) < 1$. A quadratic coupling, for instance, induces a quartic interaction $\sim (d_{m_e}^{(2)})^2 m_e(0)^2\Lambda^2\varphi^4/8\pi^2$ at one loop, which, while a problem in its own right, requires less tuning [by a factor of $g_{m_e}(\varphi)$] relative to the bare potential.

The cosmological probes we develop here thus test scalars that are nominally un-natural, as is the case for many existing probes of interacting, ultralight scalars. In

principle, the phenomenology we consider only requires the effective theory to be valid at energy scales below 100 eV or so (compared to signatures from nucleosynthesis, for instance, which is sensitive to temperatures up to $\sim \text{MeV}$ scales [85–91]). Whether this translates into a less severely tuned theory (that also remains a suitable description of known physics at moderate energy scales) ultimately can only be assessed in concrete UV models. Apparently unnatural scalars could arise accidentally in models with anthropic solutions for the cosmological constant [92], have naturally light masses from discrete symmetry structures [84] (see also Ref. [38] for technically natural scalars with quadratic couplings), or simply be fine-tuned. Nevertheless, perhaps more important than specific tuning considerations is the fact that an ever-growing set of precision cosmological data allows us to test the existence and unique signatures of hyperlight fields, with abundances consistent with observational constraints [93–97] and couplings as weak as gravitational strength [60].

B. Effective field theory at low energies

In this section, we discuss the effective field theory description of the interaction between scalars and composite matter. Around the time of primordial nucleosynthesis, the scalar interacts with nucleons and bound nuclei. After recombination, some electrons and nuclei are free while others are bound as neutral atoms. An effective description in terms of these composite states is therefore pertinent to study the scalars interactions with the early-Universe plasma and present-day experimental apparatuses. Here we build upon and expand the discussions in Refs. [21,22].

The coupling of φ to composite states is not readily apparent from a theory written in terms of fundamental SM degrees of freedom, as in Eq. (2.7). Consider the fundamental coupling of φ to the electromagnetic tensor $F_{\mu\nu}F^{\mu\nu} \propto \mathbf{E}^2 - \mathbf{B}^2$, for example. As noted in Ref. [20], freely propagating electromagnetic radiation neither sources nor interacts with φ because $\langle \mathbf{E}^2 - \mathbf{B}^2 \rangle$ vanishes for a traveling plane wave. Although $\langle \mathbf{E}^2 - \mathbf{B}^2 \rangle$ can be nonzero for, e.g., a singular standing wave, it vanishes for an *incoherent* superposition of plane waves such as the cosmological blackbody radiation. Conversely, $F_{\mu\nu}F^{\mu\nu}$ can be nonzero at order α in the presence of charged matter, as discussed in Sec. II C. Thus even if the “bare” particle constituents of matter do not directly couple to φ , composite states can nevertheless inherit a coupling to φ , e.g., through their electromagnetic binding energy [20,23,98]. In light of this, we define a new set of φ -dependent parameters—namely, the masses m_X of matter species X and potential new couplings c_i between those species:

$$\mathcal{C}_+ = \{m_e, m_p, m_n, m_H, m_{\text{He}}, m_{\text{He}^{2+}}, \dots\} \cup \{\alpha, c_1, c_2, \dots\}. \quad (2.14)$$

²The photon coupling does not modify the parameters of the free SM Lagrangian and therefore generates no effective potential at one loop [83]; the two-loop contributions for such noncanonical theories are technically challenging to compute and we do not consider them here.

A new coupling function is associated with each such parameter.

A proper treatment of atomic species as an effective field theory associates a unique field and mass parameter to every atomic excited state considered in the theory, labeled by its spectroscopic quantum numbers

$n^{2S+1}L_J$. Since in many applications the distinction between atomic states makes a negligible difference to the interaction of the species with φ , we only consider ground-state atoms, assigning each a fermionic or bosonic field of the appropriate spin representation:

$$\begin{aligned} \mathcal{L}_{\text{SM}}[\varphi, \dots] = & -\frac{1}{16\pi\alpha(\varphi)^2} F^{\mu\nu} F_{\mu\nu} + \sum_f^{\{e,p,n,{}^4\text{He}^+,{}^3\text{He},\dots\}} \bar{f} [i\not{D} - m_f(\varphi)] f \\ & + \sum_b^{\{\text{H},{}^4\text{He},{}^4\text{He}^{2+},\dots\}} \left[|D_\mu b|^2 - m_b^2(\varphi) b^* b \right] + \dots + \mathcal{L}_{\text{int}}[c_1(\varphi), c_2(\varphi), \dots], \end{aligned} \quad (2.15)$$

where the fields f are spinors, b complex scalars, and the first ellipsis denotes the quadratic Lagrangian for fields of higher spin representation. Some species remain charged under electromagnetism, while $\mathcal{L}_{\text{int}}$ allows for short-range interactions between species, such as collisions between neutral species and ionization and recombination processes.

The remaining task is to relate the scalar coupling functions for the set $\mathcal{C}_+$ to the set of coupling functions to fundamental SM parameters $\mathcal{C}$. These relations, the so-called “dilaton charges” [21,22],³ are assembled into a charge matrix $\mathcal{Q}$ with elements

$$(\mathcal{Q}_\sigma)_\lambda = \frac{\lambda}{\sigma} \frac{\partial \sigma}{\partial \lambda}, \quad \lambda \in \mathcal{C}, \sigma \in \mathcal{C}_+, \quad (2.16)$$

which relates the new and old coupling functions via

$$g'_\sigma(\varphi) = \sum_{\lambda \in \mathcal{C}} (\mathcal{Q}_\sigma)_\lambda g'_\lambda(\varphi). \quad (2.17)$$

For a composite species X , the coupling functions g'_{m_X} provide a convenient expression of the dependence of m_X on φ through Eq. (2.4). They also determine the effective potential $V_{\text{matter}}(\varphi)$ which dictates the reaction of φ to matter, whether it be a localized source mass (Sec. III D) or the matter content of the Universe averaged over cosmological scales (Secs. II C 3 and II C 4).

As an example, consider an atomic species X with atomic number Z and mass number A . Somewhat heuristically, its total mass comprises the masses of the constituent protons, neutrons, and electrons as well as nuclear and atomic binding energies,

$$\begin{aligned} m_X \simeq & Zm_p(\alpha) + (A - Z)m_n(\alpha) + Zm_e \\ & + E_{\text{nuclear},X}(\alpha) + E_{\text{atomic},X}(\alpha, m_e), \end{aligned} \quad (2.18)$$

where $E_{\text{nuclear},X}$ is the nuclear binding energy and $E_{\text{atomic},X}$ the atomic binding energy. Here we highlight the dependence on α and m_e relevant to this work; in addition, the proton and neutron masses and nuclear energies themselves depend on Λ_{QCD} and quark masses. The dilatonic charge $(\mathcal{Q}_{m_X})_\lambda$ is roughly the fraction of the mass of X that originates from the electron’s mass or electromagnetic energy. The distinction between atomic excited states is often unnecessary: the difference in their atomic binding energies $E_{\text{atomic},X}(\alpha, m_e) \simeq \mathcal{O}(\alpha^2 m_e)$ is $E_{\text{atomic},X}/m_X \sim 10^{-8}$, yielding a negligible change to the dilatonic charge of the species. An exception is precision measurements of energy level differences, such as in atomic clock systems and astrophysical spectroscopy.

The value of the electron dilatonic charge at leading order in α is the ratio of masses,

$$(\mathcal{Q}_{m_X})_{m_e} \approx Zm_e/m_X. \quad (2.19)$$

The form of the electromagnetic dilatonic charge $(\mathcal{Q}_{m_X})_e$ is less certain due to nuclear effects. Historically, Ref. [20] is the first attempt to estimate $(\mathcal{Q}_{m_X})_e$ in the context of time-varying fundamental constants, arriving at $4(\mathcal{Q}_{m_X})_e \approx 1.3 \times 10^{-2}$ by using a classical charge distribution model of the proton [or equivalently by extrapolating the large Z limit of Eq. (2.20) below to $Z = 1$]. More modern methods yield $(\mathcal{Q}_{m_X})_e \sim 10^{-4}$; see Ref. [99], Sec. II A, for a summary. References [21,22] estimate $(\mathcal{Q}_{m_X})_e$ using the semiempirical mass formula for nuclei (or the Bethe-Weizsäcker formula), as well as the value for the electromagnetic energies of the proton and neutron from [100]. While improvements in nuclear modeling and measurements have further refined these estimates, they nevertheless capture $\mathcal{O}(1)$ of the effects and are sufficient for the

³Our definition differs slightly from that of Refs. [21,22], which define the dilatonic charge only as a property of matter species rather than as a relation between Lagrangian parameters (which would include, e.g., α). Because the concept developed here is so proximate, we retain this nomenclature.

present discussion. Therefore, we use Eqs. (16)–(21) of Ref. [21], extended to ionized species:

$$(Q_{m_X})_{m_e} = F_A \left(5.5 \frac{Z_e}{A} \right) \times 10^{-4}, \quad (2.20a)$$

$$(Q_{m_X})_e = F_A \left(-1.4 + 8.2 \frac{Z_p}{A} + 7.7 \frac{Z_p(Z_p - 1)}{A^{4/3}} \right) \times 10^{-4}, \quad (2.20b)$$

where Z_p is the number of protons, Z_e the number of electrons, A the total number of nucleons, and $F_A \equiv Am_{\text{amu}}/m_X$ with $m_{\text{amu}} = 931$ MeV.

Note that because the dilatonic charge $(Q_{m_X})_e$ is non-zero, Eq. (2.4) implies that a scalar coupled only to the photon still modulates the masses of nucleons and atoms. However, because $(Q_{m_p})_e$ and $(Q_{m_p})_{m_e}$ are both of order 10^{-4} , the relative variations in m_e and in α are much larger than the relative variation in the nucleon and atomic masses for comparable Lagrangian couplings; we therefore take the approximation $m_X(\varphi) = m_X(0)$ for any nucleon, nucleus, or atom X in our treatment of the CMB anisotropies [60]. On the other hand, we take into account the α dependence of nucleon, nuclear, and atomic masses on the potential for φ sourced by matter (Sec. II C), primordial nucleosynthesis (Sec. IV A), and φ -mediated violations of the equivalence principle (Secs. III D and III F), where it is a leading effect. These considerations motivate a more complete study of cosmological dynamics with varying nucleon masses, including φ dependence in m_u , m_d , and Λ_{QCD} , which is beyond the scope of the present work. (See Ref. [81] for exploration of such scenarios in the context of axio-dilaton scalar-tensor theories.)

C. Matter potentials

Having established the effects a coupled scalar can have on the fundamental parameters of the SM, we next consider the effects of SM matter on φ itself. The matter content generates an effective, in-medium potential $\partial\mathcal{L}_{\text{SM}}/\partial\varphi \equiv -V'_{\text{matter}}(\varphi)$ that encodes how φ is sourced or scattered by matter distributions. The magnitude of $V'_{\text{matter}}(\varphi)$ compared to $m_\phi^2 M_{\text{pl}}^2 \varphi$ and $3H\dot{\varphi}$ determines relevance of matter effects to the dynamics of φ at any particular epoch. Even though matter-driven dynamics for φ and the associated variations in α or m_e are not inadmissible *a priori*, our analysis of the CMB in Ref. [60] assumes that the scalar evolution is driven only by the bare mass term $m_\phi^2 \varphi$ and cosmological expansion. We compute in some generality the effective matter potential at various cosmological epochs in order to delineate the parameter space where it can or cannot be neglected. In this section we establish the formalism and important results; we discuss the consequences thereof—namely, for which couplings the scalar field, and therefore

its effect on the fundamental constants, remains effectively frozen until after recombination—in Sec. II E.

A convenient way to compute the scalar’s effective potential due to a background of matter with finite temperature and density is the “method of background fields” [101–103] applied in thermal field theory. This approach evaluates the (thermal) path integral for all coupled degrees of freedom, treating φ as constant and restoring its spacetime dependence after the fact; this approximation is appropriate if the scales most important to the calculation are sufficiently separated from those over which φ varies. For the scalar couplings encoded by $\alpha(\varphi)$ and $m_e(\varphi)$ [Eq. (2.7)], this amounts to computing the thermodynamic pressure $P(\alpha, \{m_X\}, \{T_X\}, \{\mu_X\}, \dots)$ for a collection of species X with nonzero temperatures T_X and chemical potentials μ_X (or equivalently, nonzero temperatures T_X and number densities n_X), treating the fundamental parameters as constant and replacing each with its φ -dependent form at the end. Then $V_{\text{matter}}(\varphi) = -P(\varphi)$, i.e., the matter potential encodes how thermodynamic properties of a medium—whether the primordial thermal bath or a volume of bulk material—vary as φ changes the fundamental properties of its constituents. Scalar-SM interactions thus enter the equation of motion simply as

$$-\frac{\partial V_{\text{matter}}}{\partial\varphi} = \frac{\partial P}{\partial\varphi} = \sum_\lambda \frac{\partial P}{\partial \ln \lambda} g'_\lambda(\varphi). \quad (2.21)$$

In practice, we evaluate Eq. (2.21) at leading order in φ —i.e., with $P(\{\lambda(\varphi)\}) \rightarrow P(\{\lambda(0)\})$ and, for quadratic couplings, $g'_\lambda(\varphi) \rightarrow d_\lambda^{(2)} \varphi$. Note that φ enters in the effective potential only through the SM parameters it modulates. The smallness of variations in SM fundamental constants therefore guarantees that higher-order interactions generated by the quadratic coupling (i.e., terms proportional to higher powers of $d_\lambda^{(2)}$) may be self-consistently neglected.

In what follows, we evaluate the thermodynamic pressure of a multispecies system as a function of φ . In Sec. II C 1, we approximate the cosmological bath as a gas of noninteracting matter and radiation and use the well-known expressions for the total thermodynamic pressure P_{nonint} to obtain the corresponding matter potential $V_{\text{nonint}}(\varphi)$. This noninteracting approximation is sufficient in the late Universe, i.e., after electron-positron annihilation; however, a consistent calculation at $\mathcal{O}(\alpha)$ at earlier times must account for long-range electromagnetic interactions between charged species (i.e., “plasma effects”). In Sec. II C 2 we therefore compute the explicit α -dependent contributions to the pressure to obtain the potential $V_{\text{int}}(\varphi)$ due to interactions. We then apply our results to assess the dominant contributions to $V_{\text{matter}}(\varphi) = V_{\text{nonint}}(\varphi) + V_{\text{int}}(\varphi)$ in the late and early Universe in Secs. II C 3 and II C 4, respectively. Note that our discussion of V_{nonint} and its evolution through the course of cosmological history is similar to that of Ref. [14].

1. Noninteracting gas

We first consider a noninteracting gas of fermions and bosons of the different species X [i.e., neglecting all interaction terms in the effective Lagrangian Eq. (2.15)]. Then [101,102,104,105]

$$P_{\text{nonint}} = \sum_X^{\{\gamma, e, \bar{e}, p, \bar{p}, n, \bar{n}, \text{H}, {}^4\text{He}, \dots\}} P_{\text{nonint}, X} \quad (2.22)$$

with

$$P_{\text{nonint}, X}(\varphi) = \frac{2s_X + 1}{6\pi^2} T_X^4 \int_0^\infty dx \frac{x^4 / \sqrt{x^2 + [m_X(\varphi)/T_X]^2}}{\exp\left(\sqrt{x^2 + [m_X(\varphi)/T_X]^2} - \mu_X/T_X\right) \pm 1}, \quad (2.23)$$

where s_X is the spin of species X , μ_X is its chemical potential, and the plus and minus signs are taken for fermions and bosons, respectively. Even though a free photon gas contributes to the total pressure of the Universe, it does not depend on any scale other than T_γ [i.e., since $m_\gamma(\varphi) = 0$]; the partial pressure $P_{\text{nonint}, \gamma}$ of a free photon gas therefore does not depend on φ and does not contribute to its dynamics.

Next, recall that the spatially averaged energy density of a species X is

$$\bar{\rho}_{\text{nonint}, X}(\varphi) = \frac{2s_X + 1}{2\pi^2} T_X^4 \int_0^\infty dx \frac{x^2 \sqrt{x^2 + [m_X(\varphi)/T_X]^2}}{\exp\left(\sqrt{x^2 + [m_X(\varphi)/T_X]^2} - \mu_X/T_X\right) \pm 1}. \quad (2.24)$$

One can then check that the pressure differential yields

$$\left(\frac{\partial P_{\text{nonint}, X}}{\partial \varphi}\right)_{\mu_X, T_X} = -g'_{m_X}(\varphi) \Theta_{\text{nonint}, X}(m_X(\varphi)) \quad (2.25)$$

where

$$\begin{aligned} \Theta_{\text{nonint}, X}(m_X(\varphi)) &= \bar{\rho}_{\text{nonint}, X}(\varphi) - 3P_{\text{nonint}, X}(\varphi) \quad (2.26) \\ &= \frac{2s_X + 1}{2\pi^2} m_X(\varphi)^2 T_X^2 \mathcal{J}_\pm(m_X(\varphi), T_X, \mu_X) \quad (2.27) \end{aligned}$$

is the trace the (noninteracting) stress-energy tensor of the species, and we defined

$$\mathcal{J}_\pm(m, T, \mu) \equiv \int_0^\infty dx \frac{x^2 / \sqrt{x^2 + (m/T)^2}}{\exp\left(\sqrt{x^2 + (m/T)^2} - \mu/T\right) \pm 1}. \quad (2.28)$$

As observed in Refs. [14,37,59,106], the result that φ effectively couples to $\sum_X \Theta_{\text{nonint}, X}$ can also be obtained by noting that in the interaction Lagrangian φ couples to the bilinear mass operators $m_f \bar{f} f$, $m_b^2 \bar{b}^* b$, etc. In the thermal state these operators correspond to the trace of the thermal energy-momentum tensor of each free species at $\varphi = 0$.

In the relativistic limit $T_X \gg m_X, \mu_X$, the matter potential reduces to

$$\Theta_{\text{nonint}, X}(\varphi) \approx \frac{2s_X + 1}{12} m_X(\varphi)^2 T_X^2 \times \begin{cases} 1, & \text{bosons,} \\ \frac{1}{2}, & \text{fermions.} \end{cases} \quad (2.29)$$

In the nonrelativistic and dilute (nondegenerate) limit, $T_X \ll m_X$ and $T_X \ll m_X - \mu_X$,

$$\Theta_{\text{nonint}, X}(\varphi) \approx m_X(\varphi) n_X, \quad (2.30)$$

where the number densities n_X are averaged over cosmological scales. In summary, the effective source term $V'_{\text{matter}}(\varphi)$ depends on φ simply via the coupling functions $g'_{m_X}(\varphi)$, weighted by the trace of species X 's energy-momentum tensor, Eq. (2.25). Each $g'_{m_X}(\varphi)$ is related to the fundamental couplings $g'_e(\varphi)$ and $g'_{m_e}(\varphi)$ by the dilatonic charges of species X , which heuristically quantify how much of the mass m_X is of electronic or electromagnetic origin.

2. Interacting plasma

As discussed in Sec. II B, the electromagnetic energy content in matter depends on α , meaning the contributions to $g'_{m_X}(\varphi)$ in Eqs. (2.29) and (2.30) from the photon coupling are suppressed by powers of α . In other words, the dilatonic charges $(Q_{m_X})_e$ are really $\mathcal{O}(\alpha)$ quantities, the same order as the leading contributions from electromagnetic interactions. We now compute the contribution P_{int} from interactions to the pressure by treating the matter

system as a quantum electrodynamics (QED) gas or plasma of photons and charged matter.

The pressure of an interacting QED gas is known at $\mathcal{O}(e^2)$ [104] and beyond [107]. Formally, the contribution to the pressure due to the electromagnetic interaction between particles at $\mathcal{O}(\alpha)$ corresponds to the sum of “sunset”-type bubble Feynman diagrams of individual charged matter fields (i.e., single loops of a matter field of a given type with an internal photon line inserted). Because particle-antiparticle pairs are created and annihilated in a QED plasma, the full expression contains products of the phase space distribution functions of particles and antiparticles of the same type, with $\mu_X = -\mu_{\bar{X}}$. In standard cosmological thermodynamics, however, SM species are either relativistic with a negligible matter-antimatter asymmetry ($|\mu_X| \ll T_X$) or nonrelativistic and dilute but with a large matter-antimatter asymmetry. In the former, symmetric case, we denote the total interaction of particles and antiparticles of the same type with

$P_{\text{int},X} + P_{\text{int},\bar{X}}$. In the latter, nonrelativistic case, one can simply take $P_{\text{int},\bar{X}} \approx 0$ for antimatter and $P_{\text{int},X}$ the nonrelativistic limit of the term corresponding to matter-matter interactions.

Taking these assumptions, working at leading order in α , and assuming all species share a common plasma temperature T_p , P_{int} may be written as a sum over the pressure of each matter species:

$$P_{\text{int}} \approx \sum_X^{\{e,\bar{e},p,\bar{p},H^+,{}^4\text{He}^{2+},\dots\}} P_{\text{int},X}. \quad (2.31)$$

To evaluate the relativistic limit while still capturing the exponential suppression at annihilation, we take Eq. (5.58) from Ref. [104] for the contribution of a single type of particle/antiparticle sector to the interaction pressure and substitute $m_X = 0$ everywhere except inside distribution functions:

$$P_{\text{int},X} + P_{\text{int},\bar{X}} \approx \frac{2q_X^2 \alpha T_p^4}{3\pi} \left[\mathcal{J}_+(m_X, T_p, 0) + \frac{3}{\pi^2} \mathcal{J}_+(m_X, T_p, 0)^2 \right], \quad (2.32)$$

where q_X is the electromagnetic charge of X .

As before, the source term in the equation of motion of φ is given by the pressure differential $\sim \partial P_{\text{int}} / \partial \varphi$, replacing α and m_X with their φ -dependent values. The contribution from X particles together with $\bar{X}$ antiparticles is

$$\left(\frac{\partial P_{\text{int},X}}{\partial \varphi} + \frac{\partial P_{\text{int},\bar{X}}}{\partial \varphi} \right)_{T_p, \mu_X} \approx -g'_e(\varphi) \frac{2q_X^2 \alpha(\varphi) T_p^4}{3\pi} \left[\mathcal{J}_+(m_X, T_p, 0) + \frac{3}{\pi^2} \mathcal{J}_+(m_X, T_p, 0)^2 \right]. \quad (2.33a)$$

In the relativistic limit $T_p \gg m_X$, the interaction contribution to the matter potential is then

$$\left(\frac{\partial P_{\text{int},X}}{\partial \varphi} + \frac{\partial P_{\text{int},\bar{X}}}{\partial \varphi} \right)_{T_p, \mu_X} \approx -g'_e(\varphi) \frac{5\pi}{72} q_X^2 \alpha(\varphi) T_p^4. \quad (2.33b)$$

The nonrelativistic ($T_p \ll m_X$) and dilute (nondegenerate) limit with a matter-antimatter asymmetry ($n_X \gg n_{\bar{X}}$) is $T_p \ll m_X - \mu_X \ll m_{\bar{X}} - \mu_{\bar{X}}$. The interaction contribution is given by [[104], see Eq. (5.61)],

$$\left(\frac{\partial P_{\text{int},X}}{\partial \varphi} \right)_{T_p, \mu_X} \approx -g'_e(\varphi) m_X(\varphi) n_X \left[q_X^2 \alpha(\varphi) \left(\frac{4\pi}{12} \frac{T_p^2}{m_X(\varphi)^2} - \frac{n_X \lambda_{T_p,X}^3}{4\sqrt{2\pi}} \sqrt{\frac{T_p}{m_X(\varphi)}} \right) \right], \quad (2.33c)$$

where $\lambda_{T_p,X} = \sqrt{2\pi/m_X T_p}$ is the thermal de Broglie wavelength of species X . The first term in large brackets in Eq. (2.33c) is equal to $\omega_{p,X}^2 T_p^2 / 12$, where $\omega_{p,X}^2 = 4\pi \alpha q_X^2 n_X / m_X$ is the classical plasma frequency of a plasma of the charged species X . Heuristically, photons propagating in a plasma are “dressed” in the field of X and acquire a “plasma mass” $\omega_{p,X}$. As a result, the electromagnetic operator $F_{\mu\nu} F^{\mu\nu} = \mathbf{E}^2 - \mathbf{B}^2$ is nonzero, meaning plasma photons can source φ . The second term in large brackets in Eq. (2.33c) is quantum in origin as it involves the thermal de Broglie wavelength.

3. After electron-positron annihilation

After annihilation and nucleosynthesis, species are nonrelativistic and nondegenerate and antimatter is negligible, so the effective potential in the noninteracting limit is given by Eq. (2.30) summed over species:

$$-\frac{\partial V_{\text{nonint}}}{\partial \varphi} = \frac{\partial P_{\text{nonint}}}{\partial \varphi} \approx \sum_X^{\{e,\bar{e},p,\bar{p},H^+,{}^4\text{He}^{2+},\dots\}} -g'_{m_X}(\varphi) m_X(0) n_X. \quad (2.34)$$

Turning to the potential from interactions, first note that for species with a common temperature, $\partial P_{\text{int}}/\partial\varphi$ [Eq. (2.33c)] is dominated by the lightest species (namely, the electron)—an ubiquitous feature of nonrelativistic plasma physics. Moreover, the explicit photon coupling term in Eq. (2.33c) is multiplied by α , and the implicit photon contribution to $g_{m_X}(\varphi)$ in Eq. (2.34) is proportional to the dilatonic charge $(Q_{m_X})_e$, which is formally $\mathcal{O}(\alpha)$ as well. Thus, the plasma contributions are indeed necessary for a consistent treatment of the effective potential under the photon coupling. However, these terms [in Eq. (2.33c)] are further suppressed by $(T_p/m_X)^2$ or $n_X \lambda_{T_p,X}(0)^3 \sqrt{T_p/m_X(0)}$, which are both small at late times, and also by the hierarchy $\sim m_e/m_p$ between the electron and baryon energy densities. The plasma contribution of electrons is therefore highly suppressed at late times relative to the nonplasma contributions of nuclei. Thus, we conclude that plasma effects in the late Universe are subdominant and take $-V'_{\text{matter}}(\varphi) \approx P'_{\text{nonint}}(\varphi)$ after electron-positron annihilation.

In principle, the sum in Eq. (2.34) distinguishes between neutral atoms, free nuclei, free neutrons, and free electrons. After primordial nucleosynthesis, the Universe contains no free neutrons—the content of the Universe is made up of a fraction $Y_{\text{He}} \equiv \bar{\rho}_{\text{He}}/(\bar{\rho}_{\text{H}} + \bar{\rho}_{\text{He}})$ of helium nuclei by mass, with the remainder almost entirely in free protons. Moreover, as explained in Sec. II B, atomic binding energies make a negligible contribution to dilatonic charges. For nonrelativistic species, the thermal potential then only depends on the *number* of particles of each species, as is evident in Eq. (2.34). Since the total numbers of electrons and nuclei are the same whether or not they are bound into atoms (taking the Universe to be completely charge neutral), Eq. (2.34) may be computed as if all nuclei and electrons were bound into atoms (again when neglecting contributions from binding energy). As a consequence, the effective potential in this limit is unchanged in form throughout recombination and reionization.

An additional subtlety is that, after photon decoupling, the electrons and photons do not have a common temperature as assumed in Eq. (2.33c). Namely, after recombination the temperature of matter redshifts as $1/a^2$ rather than $1/a$. However, we only expect the appropriate extension of the bracketed term in Eq. (2.33c) to remain a suppression factor for any temperatures below the electron mass. We therefore ignore this distinction; determining whether the scalar rolls substantially leading up to recombination does not require treating this regime regardless.

In summary, we may conveniently parametrize the thermal potential in terms of effective dilatonic charges for the total baryonic content, i.e., as a weighted sum of those for hydrogen and helium atoms:

$$(Q_b)_\lambda \equiv (1 - Y_{\text{He}})(Q_{m_{\text{H}}})_\lambda + Y_{\text{He}}(Q_{m_{\text{He}}})_\lambda \approx 10^{-4} \cdot \begin{cases} 4.8, & \lambda = m_e, \\ 6.3, & \lambda = e, \end{cases} \quad (2.35)$$

for a fiducial helium mass fraction $Y_{\text{He}} = 0.25$. The thermal force term is then

$$\frac{\partial V_{\text{matter}}}{\partial\varphi} = [(Q_b)_{m_e} g'_{m_e}(\varphi) + (Q_b)_e g'_e(\varphi)] \bar{\rho}_b. \quad (2.36)$$

Fig. 1 depicts the temperature evolution of the potential [specifically, the coefficients of the coupling functions in Eq. (2.36)]. In Fig. 1, we obtain the noninteracting electron pressure differential (for the electron coupling) by integrating Eq. (2.26) directly with μ_e and $\mu_{\bar{e}} = 0$, and likewise for the contribution to the pressure differential from plasma interactions [Eq. (2.33b)]; see Eq. (2.39) and the surrounding discussion. The nonrelativistic limit [Eq. (2.36) divided by $g'_\lambda(\varphi)$] is simply added to the relativistic result (since the detailed evolution in the brief interval where they are comparable in size is unimportant).

Specializing to quadratic couplings, the gradient of the effective potential is

$$4\pi G \frac{\partial V_{\text{matter}}}{\partial\varphi} = \frac{\bar{\rho}_b}{\bar{\rho}_{\text{crit}}} \frac{3H^2}{2} \left[(Q_b)_{m_e} d_{m_e}^{(2)} + (Q_b)_e d_e^{(2)} \right] \varphi \equiv \delta m_T^2 \varphi, \quad (2.37)$$

where $\bar{\rho}_{\text{crit}} = 3H^2 M_{\text{pl}}^2$ is the critical energy density of a flat Universe. Evidently, a quadratic coupling sources an effective in-medium mass for the scalar, often referred to as a thermal mass:

$$\delta m_T^2 = \sum_{\lambda \in \{e, m_e\}} (5 \times 10^{-36} \text{ eV})^2 d_\lambda^{(2)} a^{-3} \frac{(Q_b)_\lambda}{5 \times 10^{-4}} \frac{\omega_b}{0.0224}, \quad (2.38)$$

where the present-day baryon density $\omega_b \equiv \bar{\rho}_{b,0}/3H_{100}^2 M_{\text{pl}}^2$. Its parametrization in terms of the Hubble rate squared in Eq. (2.37) facilitates identifying under what conditions thermal effects are important to the scalar's dynamics. Since $\bar{\rho}_b/\bar{\rho}_{\text{crit}} \lesssim 1/6$ at all times [108] (with the inequality saturated during the matter-dominated era), matter effects never dominate the scalar's dynamics after electron-positron annihilation if both quadratic couplings satisfy $d_\lambda^{(2)} \lesssim 1/(Q_b)_\lambda \sim 10^4$, as also evident in Fig. 1. In Sec. IV B 1 we discuss more precise conditions under which even subdominant effects can be neglected within the context of our analysis of variations of fundamental constants as relevant to analyses of cosmological data.

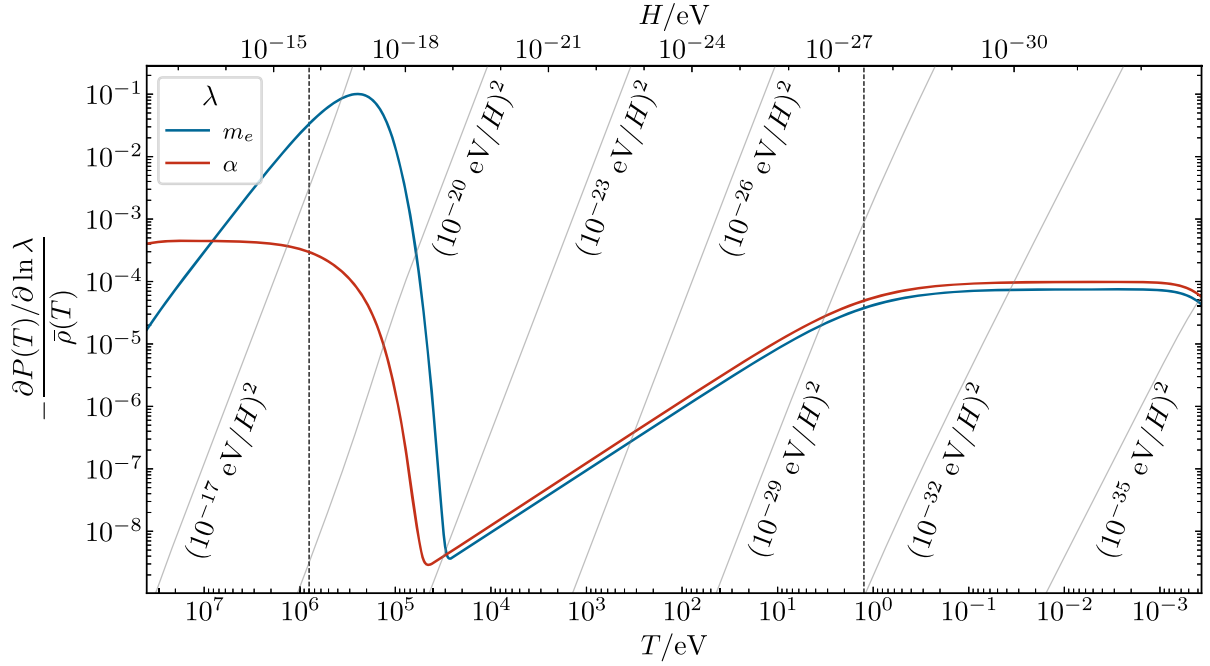


FIG. 1. Contributions to the effective matter potential Eq. (2.21) for a scalar coupled to the electron (blue) or photon (red) relative to the total energy density of the Universe. For a quadratically coupled scalar, the vertical axis is equal to $2\delta m_T^2/3H^2 d_\lambda^{(2)}$ where δm_T is its in-medium or “thermal” mass. Gray lines depict various mass scales relative to the Hubble rate (both squared) as labeled on the figure; where they intersect the red and blue curves indicates when the in-medium mass takes the labeled value (up to a factor of $3d_\lambda^{(2)}/2$). Vertical lines indicate when the photon temperature $T = m_e$ and $T = T_{\text{eq}} \approx 0.80$ eV.

4. Before electron-positron annihilation

Before electrons and positrons annihilate efficiently [i.e., at temperatures $T \gg m_e(0)$], they are relativistic and freely produced in pairs ($\mu_e \approx -\mu_{\bar{e}} \approx 0$), drastically increasing their comoving number density relative to later times. The electron-coupling contribution to the matter potential must be computed directly from Eq. (2.25), whose relativistic limit is $g'_{m_e}(\varphi)m_e(\varphi)^2 T^2/6$. Of course, in numerical solutions one may compute the full matter potential (without a relativistic approximation) by numerically integrating Eq. (2.26), as in Fig. 1.

The potential from the plasma effects of baryons remains significantly suppressed relative to their nonplasma one, but that of electrons and positrons is much enhanced. Namely, electrons and positrons each contribute Eq. (2.33b). Since $\bar{\rho}_b/T^4 \approx 2\eta\zeta(3)m_p/\pi^2/T \approx (T/0.14 \text{ eV})^{-1}$ (where $\eta \approx 6 \times 10^{-10}$ is the baryon-to-photon number ratio), the baryonic contributions are substantially smaller than those from the electron’s plasma interactions at $T > m_e(0)$. On the other hand, the relativistic limit of P_{int} [Eq. (2.32)] should have subleading corrections $\sim \alpha m_e^2 T^2$, which would induce a forcing term proportional to $\alpha(\varphi)g'_{m_e}(\varphi)$. However, this contribution is smaller than the noninteracting result by a factor $\alpha(\varphi)$ and may be neglected.

In principle, $\partial P_{\text{nonint}}/\partial \ln \alpha$ of the nonrelativistic, noninteracting baryons evolves nontrivially as neutrons decay and nucleosynthesis occurs. However, because the

baryon-to-photon ratio is so small, electron-positron annihilation does not complete until temperatures $T \approx 20$ keV [109,110], after nucleosynthesis completes. The matter potential from relativistic electrons (for either the electron or photon coupling) therefore dominate over the baryonic contributions in this epoch in spite of the exponential suppression that begins at $T \approx m_e$, rendering the details of nucleosynthesis irrelevant.

In sum, the full effective potential up to $e^+ - e^-$ annihilation is well approximated as that due just to the electron:

$$\begin{aligned} \frac{\partial V_{\text{matter}}}{\partial \varphi} = -\frac{\partial P}{\partial \varphi} \approx & \frac{2m_e(\varphi)^2 T^2}{\pi^2} \mathcal{J}_+(m_e, T, 0) g'_{m_e}(\varphi) \\ & + \frac{2\alpha(\varphi) T^4}{3\pi} \left[\mathcal{J}_+(m_e, T, 0) + \frac{3}{\pi^2} \mathcal{J}_+(m_e, T, 0)^2 \right] \\ & \times g'_e(\varphi) \end{aligned} \quad (2.39)$$

with $\mathcal{J}_+$ defined in Eq. (2.28). In the limit $T \gg m_e$, $\mathcal{J}_+(m_e, T, 0) \approx \pi^2/12$. Substituting the radiation-era result $H^2 = \pi^2 g_\star(T) T^4 / 90 M_{\text{pl}}^2$, where $g_\star$ is the effective number of relativistic degrees of freedom in the Universe, and again specializing to quadratic couplings, the in-medium mass is

$$\delta m_T^2 = \frac{H^2}{g_\star(T)} \left[\frac{15 d_{m_e}^{(2)}}{2\pi^2} \left(\frac{T}{m_e(0)} \right)^{-2} + \frac{25\alpha(0) d_e^{(2)}}{8\pi} \right] \quad (2.40)$$

in the relativistic limit. Shortly before annihilation, $g_\star \approx 10.75$. As illustrated by Fig. 1, δm_T^2 is largest relative to $H^2 \propto \bar{\rho}$ around the time of electron-positron pair annihilation (i.e., $T \approx m_e$). Requiring the impact of matter to be irrelevant to the scalar's dynamics in the early Universe imposes $d_{m_e}^{(2)} \lesssim 10$ and $d_e^{(2)} \lesssim 10^3$. Again, we defer more precise discussion of subdominant matter effects within the context our analysis of variations of fundamental constants to Sec. IV B 1.

D. Standard Model sources of variations in fundamental constants

The plasma effects studied in Sec. II C 2 induce environment dependence of α and m_e purely via SM processes. As discussed earlier, in thermal field theory photons and charged particles are “dressed” in the surrounding bath of other charged particles and photons. The term $\partial P_{\text{int},X}/\partial\varphi$ due to interactions, Eq. (2.33c), takes the form of the number density times an effective mass. Heuristically, we might interpret the coefficient of n_X in Eq. (2.33c) as an estimate of the mass correction for species X due to plasma effects:

$$\frac{\delta m_X(T_p)}{m_X} = q_X^2 \alpha \left(\frac{4\pi T_p^2}{12 m_X^2} - \frac{n_X \lambda_{T_p,X}^3}{4\sqrt{2\pi}} \sqrt{\frac{T_p}{m_X}} \right). \quad (2.41)$$

For the electron, writing $n_e = 2\eta\zeta(3)T_p^3/\pi^2$,

$$\frac{\delta m_e(T_p)}{m_e} = \alpha \left(\frac{4\pi}{12} - \frac{\eta\zeta(3)}{\pi} \right) \frac{T_p^2}{m_e^2}, \quad (2.42)$$

which is order 10^{-11} when $T_p \sim \alpha^2 m_e$ (i.e., the binding energy of hydrogen) and even smaller by the time of photon decoupling, $T_p \approx 0.26$ eV.

Similarly, electric charges are screened in plasmas. In a nonrelativistic and dilute plasma, the electrostatic Coulomb potential between two charges, which in vacuum is α/r , is suppressed by a factor $\exp(-r/\lambda_D) \approx 1 - r/\lambda_D$, where

$$\lambda_D = \sqrt{\frac{1}{4\pi\alpha} \frac{T_p}{\sum_X q_X^2 n_X}} \quad (2.43)$$

is the Debye length of the plasma [111]. This screening effect may be interpreted as a shift $|\delta\alpha(T_p)/\alpha| = r/\lambda_D$ of the fine-structure constant for processes taking place over a characteristic length scale r . In other words, plasma charges appear as neutral on scales greater than the Debye length.

There are at least two relevant length scales around recombination. The first is the Bohr radius $r_B \sim 1/\alpha m_e$: $\alpha(1 - r_B/\lambda_D)$ is essentially the fine-structure constant “seen” by atomic charges. For example, in the extreme case $\lambda_D \ll r_B$, proton charges would be screened on

scales smaller than atomic sizes and atoms could not form. Parametrically, $r_B/\lambda_D \sim \alpha^{-1/2}(n_p/T_p^3)^{1/2}(T_p/m_e) \sim \alpha^{-1/2}\sqrt{\eta}(T_p/m_e) \sim 10^{-8}$ at the time of recombination. The second important length scale is the wavelength of Thomson-scattered photons. Electrons appear neutral to photons with wavelengths longer than λ_D , leading to a suppression of scattering. In principle the bath contains photons of all wavelengths, most conveniently parametrized in units of the plasma temperature T_p for which $1/\lambda_D T_p \sim \alpha^{1/2}(n_e/T^3)^{1/2} \sim \alpha^{1/2}\sqrt{\eta} \sim 10^{-6}$. Photons with longer wavelengths experience a larger suppression but represent an exponentially suppressed fraction of the bath. SM plasma effects are thus subdominant to the level of precision with which variations in m_e and α can be tested at present.

E. Cosmological dynamics

We conclude our overview of the theory of models of varying fundamental constants by applying the preceding results to the cosmological dynamics of new, coupled scalars. We consider scenarios in which φ has a nonzero, spatially homogeneous, “misaligned” initial condition for which an expansion into a homogeneous component and a small spatial perturbation, $\varphi(t, \mathbf{x}) = \bar{\varphi}(t) + \delta\varphi(t, \mathbf{x})$, is appropriate. For the purposes of this work (except those results relying on Ref. [60]), only the homogeneous component has important cosmological implications. Per Eq. (2.2), $\bar{\varphi}(t)$ evolves according to the homogeneous Klein-Gordon equation,

$$0 = \ddot{\bar{\varphi}} + 3H\dot{\bar{\varphi}} + m_{\text{eff}}(a)^2 \bar{\varphi} \quad (2.44a)$$

with

$$m_{\text{eff}}(a)^2 \equiv m_\varphi^2 + \delta m_T(a)^2, \quad (2.44b)$$

where the “thermal mass” δm_T is given by Eqs. (2.37) and (2.40). Note that $\delta m_T(a)^2$ is proportional to the quadratic couplings $d_\lambda^{(2)}$ and therefore is not necessarily positive. The field φ itself contributes to the Friedmann equation for the Hubble parameter H but is subdominant so long as $m_{\text{eff}}^2 \varphi^2 \ll H^2$, which we assume to be the case. To first approximation, expansion thus proceeds as in standard Λ cold dark matter (Λ CDM) cosmology.

1. Free evolution

In the bare-mass-dominated limit ($m_\varphi \gg |\delta m_T|$), Eq. (2.44) reduces to the standard equation of motion for a massive scalar, a regime that has been extensively studied in the context where φ is, e.g., a preinflationary axion, some other light scalar dark matter candidate, or a dynamical subcomponent of dark energy. In such a context, solutions to Eq. (2.44) are effectively frozen at the initial condition $\bar{\varphi}_i$ until H drops below m_φ , at which point

damped, harmonic oscillations begin. If the field is sufficiently light at recombination ($m_\phi < H_\star \approx 3 \times 10^{-29}$ eV) but is also heavy today ($m_\phi > H_0$) oscillations begin between recombination and the present day. The fundamental constants are then spacetime independent over the Universe's history up to (and including) recombination but different from their current values. Hence, the free evolution of a hyperlight (10^{-33} eV $< m_\phi < 10^{-28}$ eV) field realizes the scenario considered in previous phenomenological studies of shifted fundamental constants at big bang nucleosynthesis (BBN) and recombination. Here and in Ref. [60] we focus on this “hyperlight” regime.

Since recombination occurs after matter-radiation equality and matter–dark-energy equality was relatively recent, the scalar begins oscillating in the matter-dominated era in the parameter space we consider. In a matter-dominated Universe (for which $H = 2/3t$), the solution to Eq. (2.44) with $\delta m_T = 0$ and initial conditions $\phi \rightarrow \bar{\phi}_i$ and $\dot{\phi} \rightarrow 0$ as $m_\phi t \rightarrow 0$ is $\bar{\phi}(t) = \bar{\phi}_i \sin(m_\phi t)/m_\phi t$. As anticipated, $\phi(t)$ is nearly constant until $t \sim 1/m_\phi$ and oscillates with frequency m_ϕ afterward. We therefore define the time t_{osc} when oscillations begin by $H(t_{\text{osc}}) = 2/3t_{\text{osc}} = 2m_\phi/3$. The associated scale factor a_{osc} and redshift $z_{\text{osc}} \equiv 1/a_{\text{osc}} - 1$ are determined via the first Friedmann equation by

$$m_\phi^2 = \frac{9H(t_{\text{osc}})^2}{4} \approx \frac{3\bar{\rho}_c(t_{\text{osc}}) + \bar{\rho}_b(t_{\text{osc}})}{4M_{\text{pl}}^2} \approx \frac{3\bar{\rho}_{c,0} + \bar{\rho}_{b,0}}{4a_{\text{osc}}^3 M_{\text{pl}}^2}. \quad (2.45)$$

Solutions in the oscillatory regime may be extended until after matter- Λ equality via

$$\bar{\phi}(t) = \mathcal{A}(t) \sin(m_\phi t), \quad (2.46)$$

where the slowly varying oscillation amplitude is

$$\mathcal{A}(t) \approx \bar{\phi}_i \left(\frac{a(t)}{a_{\text{osc}}} \right)^{-3/2} \quad (2.47)$$

at times $t > t_{\text{osc}}$.

2. Impact of matter potentials

We now study modifications to the scalar dynamics incurred at larger couplings for which the effective in-medium potential is non-negligible, specializing to interactions quadratic in ϕ . Section II C shows that electron-positron annihilation separates two qualitatively distinct regimes. All probes we consider only depend on dynamics after electron-positron annihilation, except for BBN; the impact of matter potentials before annihilation on BBN was studied in detail in Refs. [37,59]. The dynamics before annihilation do impact the relationship between the post-annihilation initial condition $\bar{\phi}_i$ and the initial conditions at much earlier times. We discuss analytic solutions for the

preannihilation dynamics and their implications in Appendix A 2 and here consider only dynamics after annihilation, for which the thermal mass is given by Eqs. (2.37) and (2.38).

Since the thermal mass evolves with the energy density of SM matter, m_{eff} is generically proportional to Hm_{eff} (except at electron-positron annihilation); when $m_{\text{eff}} \gtrsim H$, Eq. (2.44) is then amenable to a Wentzel-Kramers-Brillouin (WKB) approximation [37,59,98]. The solutions are therefore well described by an oscillatory function times a slowly varying amplitude,

$$\bar{\phi}(t) = \mathcal{A}(t) \sin \left(\int_{t_{\text{ref}}}^t m_{\text{eff}}(t') dt' + \vartheta(t_{\text{ref}}) \right), \quad (2.48a)$$

where t_{ref} is some reference time for which $m_{\text{eff}}^2 > H^2$, $\vartheta(t_{\text{ref}})$ the phase at that reference time, and

$$\mathcal{A}(t) \propto \frac{1}{a(t)^{3/2}} \frac{1}{m_{\text{eff}}(t)^{1/2}}. \quad (2.48b)$$

In the limit that $m_\phi \gg |\delta m_T|$ and $m_\phi \gg H$, we recover $\mathcal{A} \propto a^{-3/2}$ as for free evolution [Eq. (2.47)]. On the other hand, if δm_T is much larger than both H and m_ϕ , the envelope evolves as $\mathcal{A} \propto a^{-3/2} \bar{\rho}_b^{-1/4} \propto a^{-3/4}$ (after electron-positron annihilation).⁴

When $m_{\text{eff}}(a)^2 \lesssim H^2$, the WKB method does not apply. Because δm_T redshifts, generally $m_\phi \ll |\delta m_T|$ during some period of time before t_{osc} unless couplings are very small or the bare mass is relatively large. Because Eq. (2.44) then describes an oscillator whose mass and damping coefficient both evolve with time, the scalar is not necessarily frozen prior to t_{osc} . In Appendix A, we obtain analytic solutions to the equation of motion at times $t \lesssim t_{\text{osc}}$ when the bare mass m_ϕ is irrelevant to the dynamics. These solutions are uniquely parametrized by the combination (see Appendix A 1)

$$\mathcal{D} \equiv \frac{3}{2} \frac{(Q_b)_\lambda d_\lambda^{(2)}}{1 + \omega_c/\omega_b}, \quad (2.49)$$

where the dilatonic charges of baryonic matter $(Q_b)_\lambda$ are given by Eq. (2.35) and are of order 10^{-4} . Here $\omega_X = \bar{\rho}_{X,0}/3H_{100}^2 M_{\text{pl}}^2$ denotes the rescaled, present-day energy density of species X ; in Λ CDM cosmologies, $\omega_b/\omega_c = 0.1855 \pm 0.0028$ [108].

⁴We do not treat the regime of negative δm_T much larger in magnitude than m_ϕ ; the resulting tachyonic instability makes an analysis unwieldy and the phenomenology sensitive to initial conditions and dynamics at earlier times. An understanding thereof is also not required to describe the parameter regimes our analysis ultimately constrains.

Note that $|\delta m_T^2|/H^2$ increases during radiation domination and is constant and equal to $|\mathcal{D}|$ during matter domination, behavior reflected in Fig. 1. Appendix A 1 shows that the field begins evolving before matter-radiation equality if $|\mathcal{D}| \gtrsim 1$, eventually reaching the WKB regime at larger couplings. On the other hand, when $|\mathcal{D}| \lesssim 1$, $|\delta m_T^2|$ is always smaller than H^2 and the thermal mass never causes the field to oscillate. Instead, the solutions presented in Appendix A 1 are well approximated by

$$\bar{\varphi}(a) \simeq \bar{\varphi}_i (1 + 2a/3a_{\text{eq}})^{-2\mathcal{D}/3} \quad (2.50)$$

(determined empirically; see Appendix A 1 for genuine but more cumbersome analytic solutions). Moreover, in this regime the thermal mass drops below the bare mass in magnitude before $t = t_{\text{osc}} = m_\varphi^{-1}$, so the field only ever oscillates due to its bare mass alone.

3. Evolution of energy density and fundamental constants

We conclude our treatment of the theoretical description of coupled scalar models by delineating useful parametrizations of their phenomenology—both their impact on fundamental constants and gravitation. The energy density in the homogeneous field $\bar{\varphi}$ with effective mass m_{eff} is $\bar{\rho}_\varphi = M_{\text{pl}}^2 \dot{\bar{\varphi}}^2 + m_{\text{eff}}^2 M_{\text{pl}}^2 \bar{\varphi}^2$. We quantify the scalar's present-day energy density relative to that in cold dark matter via

$$F_\varphi \equiv \frac{\bar{\rho}_{\varphi,0}}{\bar{\rho}_{c,0}}, \quad (2.51)$$

with $\bar{\rho}_{c,0} \equiv 3H_{100}^2 M_{\text{pl}}^2 \omega_c \approx (1.76 \text{ meV})^4$ [108]. When matter effects are negligible, the scalar and CDM redshift in tandem once oscillations begin and

$$F_\varphi = \frac{\bar{\rho}_\varphi(a)}{\bar{\rho}_c(a)} = \frac{m_\varphi^2 M_{\text{pl}}^2 \bar{\varphi}_i^2 / (a/a_{\text{osc}})^3}{\bar{\rho}_c(a_0)/a^3} = \frac{3(1 + \omega_b/\omega_c) \bar{\varphi}_i^2}{4}, \quad (2.52)$$

which notably is independent of m_φ . More generally, the late-time evolution of the scalar is well described by the WKB approximation in Eq. (2.48) unless the field is extremely light ($m_\varphi < H_0$) or has a negative thermal mass squared that is larger in magnitude than m_φ . The scalar's energy density is then $\bar{\rho}_\varphi \approx 2m_{\text{eff}}^2 M_{\text{pl}}^2 \mathcal{A}^2$ and evolves according to

$$\frac{\bar{\rho}_\varphi(a)}{\bar{\rho}_c(a)} = F_\varphi \frac{m_{\text{eff}}(a)}{m_{\text{eff}}(a_0)} \quad (2.53)$$

in the WKB regime. Extracting the amplitude from the energy density,

$$\mathcal{A}(a) \approx \sqrt{\frac{\rho_\varphi(a)}{m_{\text{eff}}(a)^2 M_{\text{pl}}^2}} = 1.28 \times 10^{-3} \left(\frac{m_{\text{eff}}(a)}{10^{-30} \text{ eV}} \right)^{-1} \left(\frac{\rho_\varphi(a)}{\bar{\rho}_{c,0}} \right)^{1/2}. \quad (2.54)$$

At the present, the scalar's dynamics may therefore be specified in terms of its frequency $m_{\text{eff}}(a_0)$, its abundance relative to cold dark matter F_φ , and a phase ϑ_0 . For quadratic couplings [Eq. (2.8)], the shift in the cosmological value of fundamental constants with time is

$$\frac{\lambda(t) - \lambda_0}{\lambda_0} \approx \frac{\left(d_\lambda^{(2)}/2 \right) \left[\mathcal{A}(t)^2 \sin^2 \left(\int_{t_0}^t m_{\text{eff}}(t') dt' + \vartheta_0 \right) - \mathcal{A}_0^2 \sin^2(\vartheta_0) \right]}{1 + d_\lambda^{(2)} \mathcal{A}_0^2 \sin^2(\vartheta_0)/2}, \quad (2.55)$$

where λ_0 is the cosmological value of the fundamental constant today.

How present-day conditions map into the initial condition $\bar{\varphi}_i$ (and therefore the early-time fundamental constants λ_i) depends on whether $|\mathcal{D}|$ is large or small compared to unity. In the latter case, Eq. (2.52) holds and

$$\bar{\varphi}_i \approx \sqrt{\frac{4}{3} \frac{F_\varphi}{1 + \omega_b/\omega_c}} \quad (2.56)$$

$$\frac{\lambda_i - \lambda_0}{\lambda_0} \approx \frac{d_\lambda^{(2)} \bar{\varphi}_i^2}{2} \approx \frac{2}{3} \frac{d_\lambda^{(2)} F_\varphi}{1 + \omega_b/\omega_c}, \quad (2.57)$$

taking $\bar{\varphi}_i \gg \mathcal{A}_0$, i.e., $a_{\text{osc}} \ll 1$. The small corrections due to matter couplings with $|\mathcal{D}| \lesssim 1$ may be accounted for via Eqs. (2.48) and (2.50). Perhaps surprisingly, a scalar with Planck-suppressed couplings (i.e., $d_\lambda^{(2)} \sim 1$) can shift the early-time fundamental constants substantially—that is, to the same extent as the scalar contributes to the total dark matter density.

On the other hand, when $|\mathcal{D}| \gtrsim 1$, the thermal mass induces substantial evolution before t_{osc} , as described in Sec. II E 2. In particular, $\bar{\rho}_\varphi(a)/\bar{\rho}_c(a) \propto a^{-3/2}$ until at least a_{eq} , the scale factor of matter-radiation equality. Thus, the abundance of the scalar relative to cold dark

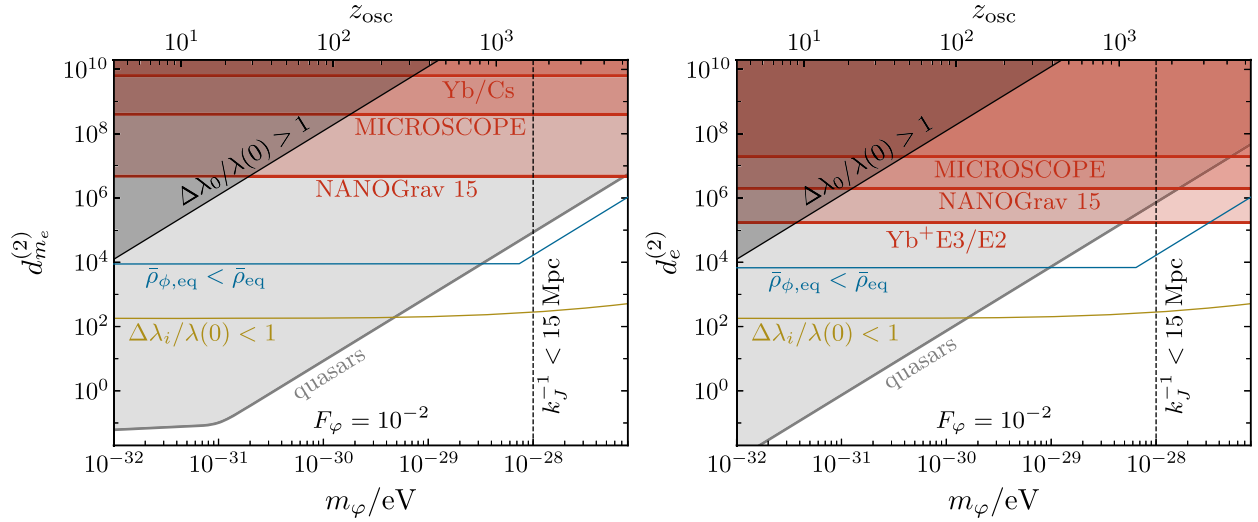


FIG. 2. Constraints on a hyperlight scalar's quadratic couplings to the electron (left) and photon (right) from late-time astrophysical (gray) and laboratory (red) measurements. Bounds assume the scalar comprises a present-day cosmological energy density $F_\varphi \equiv \bar{\rho}_{\varphi,0}/\bar{\rho}_{c,0} = 10^{-2}$ relative to that of cold dark matter. Also shown are theoretical consistency curves delineating the minimum couplings for which the perturbative expansion would break down at the present time (black) and where the scalar energy density extrapolated from the present F_φ back to matter-radiation equality would dominate the CDM [Eq. (2.59), blue]. Finally, the gold line displays the coupling values for which the scalar coupling would violate perturbativity at its initial misalignment (which is inversely proportional to F_φ). The limits from atomic clocks, quasars, and perturbativity of the coupling expansion scale inversely with F_φ , while bounds from tests of the universality of free fall scale with $F_\varphi^{-1/2}$. The matter-radiation equality constraint is proportional to F_φ^{-2} above the break in scaling. The dependence of these bounds on F_φ is further detailed in Sec. III. For simplicity we only report bounds on positive couplings; bounds from quasars on negative couplings differ by an order-unity factor. The vertical dashed lines indicate the heaviest mass for which the scalar's present-day Jeans length is larger than the size of the local supercluster; at higher masses, the scalar would cluster with the rest of the dark matter to some degree, making the reported bounds conservative in this regime. Section IV shows that early-time data from the CMB and BBN supersede these late-time constraints for all but the lightest masses.

matter can be much larger than F_φ at earlier times. Requiring $\bar{\rho}_\varphi(a) < \bar{\rho}_c(a)$ yields

$$(Q_b)_\lambda d_\lambda^{(2)} \leq \frac{m_\varphi^2}{3\omega_b H_{100}^2/2} \begin{cases} \frac{F_\varphi^2+1}{F_\varphi^2/a^3-1} & F_\varphi > a^{3/2} \\ \infty & F_\varphi < a^{3/2} \end{cases}. \quad (2.58)$$

No bound applies when the scalar cannot reach $F_\varphi(a) = 1$ even if its density redshifts like $a^{-3/2}$ (due to its in-medium mass) from a all the way until the present. Evaluated at matter-radiation equality and for $a_{\text{eq}}^3 \ll F_\varphi^2 \ll 1$, Eq. (2.58) restricts

$$(Q_b)_\lambda d_\lambda^{(2)} \lesssim \max \left[\left(\frac{m_\varphi/F_\varphi}{7.8 \times 10^{-29} \text{ eV}} \right)^2, \frac{2}{3}(1 + \omega_c/\omega_b) \right], \quad (2.59)$$

including the truncation of the bound when $\mathcal{D}$ drops below order unity (i.e., when the scalar only ever redshifts under the influence of its bare mass).

The impact of in-medium masses on dynamics affects the interpretation of a certain density in φ today in terms of its initial conditions and energy density at earlier

times. Conversely, a limit (i.e., from CMB data) on the abundance of an interacting scalar around matter-radiation equality translates to a more stringent limit on its present-day abundance relative to CDM for larger couplings. The hyperlight scalar parameter space, including the constraint from Eq. (2.59), is displayed in Fig. 2. Fig. 2 also includes a conservative limit that requires perturbativity of the coupling expansion at the present day, i.e., $g_\lambda(\varphi_0) < 1$ with φ_0 estimated by Eq. (2.54).

III. LATE-TIME PROBES

In this section, we assess the constraining power of numerous late-time ($z \lesssim 10$) probes of variations in fundamental constants as applied to hyperlight dark matter. We discuss the impact of a hyperlight scalar's slow oscillations in Sec. III A and the spatial profile of a scalar sourced by a central body in Sec. III B. We then evaluate the constraining power of quasar absorption spectra (Sec. III C), tests of the universality of free fall (Sec. III D), the Oklo phenomenon (Sec. III E), atomic clocks and pulsar timing arrays (Sec. III F), and stellar emission (Sec. III G). Fig. 2 summarizes the results of this section.

A. Coherent oscillations

Local signatures of scalar dark matter depend on its amplitude, and therefore its energy density, near Earth. The density in *ultralight* scalars is often measured relative to the energy density of virialized galactic dark matter at Earth, $\rho_{c,\oplus} \approx 0.4 \text{ GeV/cm}^3 \approx (41.9 \text{ meV})^4$ [112], which is about $\sim 10^6$ times the cosmological background density. A hyperlight scalar, however, is inherently cosmological: it does not cluster on galactic scales and its local amplitude is specified strictly as a fraction of the average cosmological density. Specifically, density perturbations of a scalar do not grow below the Jeans length $1/k_J \gtrsim 15 \text{ Mpc} \sqrt{m_{\text{eff}}/10^{-28} \text{ eV}}$ [113].

Interpolating between the regimes in which the scalar does and does not track the local dark matter density would require a more quantitative understanding of the dynamics of hyperlight subcomponents of dark matter. Here, we restrict our results to the mass range $m_\phi \ll 10^{-28} \text{ eV}$ and furthermore to couplings small enough that $|\delta m_T^2| \lesssim m_\phi^2$ holds before $z = 10$. The scalar's Jeans length at and after the time of galaxy formation is then larger than the local supercluster; this happens to be the same mass range relevant to the cosmological analysis of Ref. [60]. In this regime, the field is entirely specified by the dimensionless amplitude of the cosmologically homogeneous field, Eq. (2.54), which we rewrite as

$$\mathcal{A} = 0.720 \left(\frac{m_{\text{eff}}}{10^{-30} \text{ eV}} \right)^{-1} \left(\frac{\rho_\phi}{\rho_{c,\oplus}} \right)^{1/2}, \quad (3.1)$$

which, for ease of comparison with probes at heavier masses, we have normalized to the local CDM energy density. The scalar oscillates over a period $T_\phi = 2\pi/m_\phi \approx 1.3 \text{ Myr} (10^{-28} \text{ eV}/m_\phi)$, which is much longer than the duration of any experiment.

Most experimental measurements also take place over timescales very short compared to the coherence time of ultralight, galactic dark matter candidates. The local amplitude of virialized bosonic dark matter is then effectively stochastic—i.e., there is some probability that Earth resides in a region that is underdense compared to the expectation for the Milky Way [112]. Current bounds on ultralight scalar dark matter account for the distribution of possible scalar amplitudes by reducing the expectation from Eq. (2.54) by a factor $f_{\text{stoch}} \approx 3$ [114]. Because hyperlight scalars do not cluster on galactic scales, their local amplitude is instead set deterministically by Eq. (3.1).

Consider a measurement taken over a period of time T at some cosmological time t . When $t \gg T_\phi \gg T$, it is intractable—if not impossible—to know where the scalar is in its oscillation at t , due to the limited precision with which we know both cosmological parameters and the scalar's dynamics. For example, the scalar's present-day phase scales as $m_\phi t \propto 2m_\phi/3hH_{100}$, with t the age of the

Universe and $H_{100} \approx 2.133 \times 10^{-33} \text{ eV}$. Thus, even fixing all other cosmological parameters, percent-level variations in h result in $\sim 2\pi(m_\phi/10^{-30} \text{ eV})$ uncertainties in the scalar's phase. Since observables depend on, e.g., $d_\lambda^{(n)} \varphi(t)^n \approx d_\lambda^{(n)} \mathcal{A}(t)^n \sin(m_\phi t)^n$, direct constraints on the combination $d_\lambda^{(n)} \mathcal{A}(t)^n$ are intrinsically limited by the unknown phase—namely, one cannot exclude the possibility that the signal vanishes during the experiment due to the scalar being presently near an antinode.

To concretely reason about the effect of phase marginalization, we ground our analysis with a likelihood. To capture the variety of signals we consider, we begin with a general parametrization of a Gaussian likelihood with measurement mean $\hat{x}$ and variance σ_x . We decompose the model prediction at parameters θ into amplitude and phase factors $A(t; \theta)P(t; \theta)$, where P is a periodic function bounded by unity in magnitude and A encodes the signal magnitude if $P(t; \theta)$ were at an antinode. In the relevant limit where the phase ϑ of P is effectively unknown and treated as a uniform variate, the phase factor simply reduces to $P(t; \theta) \rightarrow P(\vartheta)$. The phase-marginalized posterior is then

$$p(\theta|\hat{x}, \sigma_x) = \pi(\theta) \int_0^{2\pi} \frac{d\vartheta}{2\pi} \frac{1}{\sqrt{2\pi\sigma_x^2}} \times \exp \left[-\frac{1}{2} \left(\frac{A(t; \theta)P(\vartheta) - \hat{x}}{\sigma_x} \right)^2 \right], \quad (3.2)$$

with $\pi(\theta)$ the prior over all other parameters. The 95% highest density interval, used to place upper limits, depends on the behavior of the phase-marginalized posterior in its tails, i.e., where $|A(t; \theta)| \gg |\hat{x} \pm \sigma_x|$. In this regime, the integrand of the marginal posterior is dominated where $|P(\vartheta)| \ll 1$. The only important property of $P(\vartheta)$ is therefore how quickly it approaches zero at its steepest node. If $P(\vartheta) \propto \vartheta^m$ (taking the relevant node to be at $\vartheta = 0$ for simplicity), the integral's support is then dominated where $|\vartheta|^m < |\hat{x} \pm \sigma_x|/A(t; \theta)$. Approximating the Gaussian as constant over this interval reveals that the marginal posterior scales with $1/\sqrt[m]{A(t; \theta)}$ when $A(t; \theta)$ is large, a finding reproduced by numerical quadrature. Given that $A(t; \theta)$ is often linear in an unknown coupling $d_\lambda^{(n)}$, the marginal posterior over the coupling is improper (since the integral over $d_\lambda^{(n)}$ does not converge at infinity).

Meaningful posterior constraints thus require regularization by some means. If one simply takes $P(\vartheta) > P_{\min}$ *a priori* (i.e., excluding the fraction of the scalar's period where its oscillation amplitude is below some threshold), then the integrand of Eq. (3.2) falls off exponentially when $|A(t; \theta)| \gtrsim |\hat{x} \pm \sigma_x|/P_{\min}$. Such a prior, however, is entirely arbitrary without concrete justification. On the other hand, experimental measurements with duration T are guaranteed to observe the scalar when it is at least $T/2T_\phi$ away from a

node in phase for some portion of the experiment's duration, effectively imposing a prior that $P(\vartheta) \gtrsim (T/2T_\varphi)^m$ for some interval. Therefore, marginalizing over the unknown phase of a scalar with period $T_\varphi \gg T$ penalizes bounds by a factor $\propto m_\varphi^{-m}$. Another simple solution is to combine more than m independent measurements with spacing in time greater than T_φ , each marginalized over their own independent phase; the posterior's tails then fall off faster than $1/d_\lambda^{(n)}$, making it proper. Given that in our parameter space T_φ is much longer than the history of human civilization, this option is only practical for distant astrophysical signals such as quasar absorption (Sec. III C).

Finally, we reiterate that phase marginalization does not so severely penalize the actual discovery potential of a given search: while a null signal can be consistent with a large coupling if the scalar were near a node, a measurement of a nonzero signal cannot be explained by an arbitrarily small coupling, no matter the scalar's phase. The above arguments concern measures of the tail of posterior distributions and still apply in that regime, but the posterior would still localize about finite couplings if the model could match the detected signal. When observing only a small fraction of an oscillation, however, it may be challenging to disentangle the phase factor $P(\vartheta)$ from the actual amplitude $A(t; \theta)$ and measure the latter at better than the order-unity level.

B. Spatial variations about a central body

Just as a cosmological abundance of matter sources an effective potential for the scalar (Sec. II C), localized bulk matter distributions—such as Earth or any central mass—source a potential that affects the scalar's spatial distribution. The field profile generated by a central mass depends on the specific form of the φ -matter interaction and is obtained by solving the classical equation of motion for φ [Eq. (2.2)] with the effective potential $V_{\text{matter}}(\varphi)$ generated by the localized matter distribution. A central mass thus generally introduces spatial variations in fundamental constants which, among other phenomenology, leads to an apparent “fifth force” toward the central mass (Sec. III D).

For couplings that are linear in φ , the space-dependent profile sourced by a localized matter volume simply adds to any ambient field constituting dark matter. For a spherical, time-independent, uniform distribution of nonrelativistic matter [28]:

$$\varphi(t, \mathbf{x}) = \bar{\varphi}(t) - d_{M_C}^{(1)} \frac{GM_C}{|\mathbf{x}|} e^{-m_\varphi |\mathbf{x}|}, \quad (3.3)$$

for a central body C with total mass M_C and radius R_C , for $|\mathbf{x}| > R_C$ and in the limit $m_\varphi R_C \ll 1$. Here the effective dilatonic coupling of C is defined as $d_{M_C}^{(n)} = \sum_\lambda (Q_{M_C})_\lambda d_\lambda^{(n)}$, where λ is summed over SM parameters. The boundary

condition at $|\mathbf{x}| \rightarrow \infty$ corresponds to the unvirialized cosmological abundance of φ , as is appropriate for a hyperlight scalar. Notably, the sourced contribution to Eq. (3.3) is independent of the background field, yielding signatures regardless of whether the scalar makes up any dark matter at all; searches for equivalence principle violation stringently constrain linear couplings for this reason.

For the quadratic couplings we focus on, however, the matter distribution scatters the incident ambient dark matter; experiments then probe the combination of the incident and scattered profiles. For $|\mathbf{x}| > R_C$ [26],

$$\varphi(t, \mathbf{x}) \approx \bar{\varphi}(t) \left[1 - d_{M_C}^{(2)} \frac{GM_C}{|\mathbf{x}|} s_C^{(2)} \left(d_{M_C}^{(2)} \frac{GM_C}{R_C} \right) \right], \quad (3.4)$$

which satisfies the same boundary condition as Eq. (3.3). The screening factor $s_C^{(2)}(y)$ is a function of the quadratic coupling times the compactness (or surface gravity) GM_C/R_C of the source mass, which can have an important effect for, e.g., terrestrial probes of variations in fundamental constants. Because $s_C^{(2)}(y \gg 1) \approx y^{-1}$, strongly coupled compact objects yield a field profile

$$\varphi(t, \mathbf{x}) \approx \bar{\varphi}(t) \left[1 - \frac{R_C}{|\mathbf{x}|} \right] \quad (3.5)$$

for $|\mathbf{x}| \geq R_C$ and assuming $d_{M_C}^{(2)} > 0$. In this limit, the amplitude of temporal variations in fundamental constants is strongly suppressed near the surface of the central mass and independent of the coupling strength of the central mass at any $|\mathbf{x}|$, while the size of spatial gradients $\nabla\varphi$ is also independent of the coupling. However, if the matter couplings are small compared to the inverse compactness,

$$d_{M_C}^{(2)} \ll R_C/G_N M_C, \quad (3.6)$$

then the relevant limit is $s_C(y \rightarrow 0) \approx 1$.

Recall from Eq. (2.17) that the couplings of composite bulk matter to φ are suppressed relative to the fundamental Lagrangian couplings by the dilatonic charges Eq. (2.20), reflecting that only a fraction of the total energy couples to φ . The dilatonic charge of an atomic mixture is the average of the dilatonic charges of the atomic constituents, weighted by mass. Earth is approximately 32% iron and 67% silicate SiO_2 by mass [115–117]. Using Eq. (2.20) and accounting for isotopic abundances [26], one has $(Q_{m_{\text{Fe}}})_{m_e} \approx 2.5 \times 10^{-4}$ and $(Q_{m_{\text{SiO}_2}})_{m_e} \approx 2.7 \times 10^{-4}$, as well as $(Q_{m_{\text{Fe}}})_e \approx 2.6 \times 10^{-3}$ and $(Q_{m_{\text{SiO}_2}})_e \approx 1.6 \times 10^{-3}$. All in all, $(Q_{M_\oplus})_{m_e} \approx 2.6 \times 10^{-4}$ and $(Q_{M_\oplus})_e \approx 1.9 \times 10^{-3}$. For Earth, the screening condition Eq. (3.6) translates to $d_{M_\oplus}^{(2)} \ll 10^9$; because the dilatonic charges of Earth are small, screening is only relevant at fundamental couplings much larger than those considered here. In conclusion, the

scalar amplitude and time dependence are well approximated by the background density in measurements in the vicinity of Earth, as relevant for terrestrial probes (Secs. III E and III F), while the spatial gradients remain proportional to $d_{M_\oplus}^{(2)}$.

Finally, very light scalars can be thermally produced in stars and compact objects (Sec. III G). Possible processes include pair production of scalars from Standard Model particles or single scalar production in the presence of the cosmological background field. In the latter case, central-body effects can suppress (or enhance) the background field inside the star. We discuss this further in Sec. III G.

C. Quasar absorption lines

The precision measurement of quasar absorption spectra has played a key role in the search for variations of fundamental constants due to their cosmologically non-trivial distances [63, 118–122]. Comparisons of atomic transition lines allow for a search for variations of α and the rotational and vibrational transitions of molecules for variations of the proton-to-electron mass ratio $\mu \equiv m_p/m_e$ by comparing the line frequencies to those measured today in the laboratory. The redshift z_{abs} at which distant quasar light passes through the absorptive medium of interest (like a molecular cloud) is obtained by measuring multiple lines or even a broad spectrum, which enables distinguishing underlying differences in fundamental constants from the determination of the redshift and additional systematics such as Doppler shifts and instrumental effects [63].

In Table I we gather measurements of $\mu(z_{\text{abs}})/\mu_0 - 1$ and $\alpha(z_{\text{abs}})/\alpha_0 - 1$ from an ensemble of quasar results that mitigate the systematic distortions of wavelength calibration [123]. As described in Sec. III A, the uncertainty in the scalar’s oscillation phase at z_{abs} and today should be taken into account. A complication beyond the discussion of Sec. III A is that the comparison of values of λ means each quasar measurement is also sensitive to the scalar’s unknown *present-day* phase:

$$\begin{aligned} \frac{\lambda[\varphi(t)]}{\lambda[\varphi(t_0)]} - 1 &\approx g_\lambda[\varphi(t)] - g_\lambda[\varphi(t_0)] \\ &= \frac{d_\lambda^{(n)}}{n!} [\mathcal{A}(t)^n \sin^n \vartheta(t) - \mathcal{A}(t_0)^n \sin^n \vartheta(t_0)]. \end{aligned} \quad (3.7)$$

Given that we consider quadratic couplings and that most of the absorbers in Table I reside at redshifts $z_{\text{abs}} > 1$, usually $\mathcal{A}(t_{\text{abs}})^n > 10\mathcal{A}(t_0)^n$, suppressing the impact of marginalization over $\vartheta(t_0)$. Moreover, combining an ensemble of measurements with independent phases makes the posterior steeper about its peak and less sensitive to such effects deeper in the tails. Ignoring the present-day contribution greatly simplifies the analysis, as

TABLE I. Quasars absorption measurements of variations in the proton-to-electron mass ratio μ and the fine-structure constant α , tabulated from references listed in the table. The reported uncertainties add systematic and statistical errors in quadrature where needed.

Quasar	z_{abs}	$10^6 \Delta\mu(z_{\text{abs}})/\mu$
HE0027-1836 [118]	2.4	-7.6 ± 10.2
Q0347-383 [118]	3.02	5.1 ± 4.5
Q0405-443 [118]	2.59	7.5 ± 5.3
Q0528-250 [118]	2.81	-0.5 ± 2.7
B0642-5038 [118]	2.66	10.3 ± 4.6
Q1232 + 082 [118]	2.34	140 ± 60
J1237 + 064 [118]	2.69	-5.4 ± 7.2
J1443 + 2724 [118]	4.22	-9.5 ± 7.5
J2123-005 [118]	2.05	7.6 ± 3.5
Q2348-011 [118]	2.43	-6.8 ± 27.8
B0218 + 357 [124]	0.685	0 ± 0.9
PKS1830-211 [125]	0.886	-0.1 ± 0.13
PMNJ0134-0931 [126]	0.765	-3.31 ± 2.74
PKS1413 + 135 [127]	0.247	-1 ± 1.3

Quasar	z_{abs}	$10^6 \Delta\alpha(z_{\text{abs}})/\alpha$
J0058 + 0041 [128]	1.072	-1.35 ± 7.16
PHL957 [128]	2.309	-0.2 ± 12.9
J0108-0037 [128]	1.371	-8.45 ± 7.34
J0226-2857 [128]	1.023	3.54 ± 8.87
J0841 + 0312 [128]	1.342	5.67 ± 4.71
J1029 + 1039 [128]	1.622	-1.7 ± 10.1
J1237 + 0106 [128]	1.305	-4.54 ± 8.66
Q1755 + 57 [128]	1.971	4.72 ± 4.71
Q2206 – 1958 [128]	1.921	-4.65 ± 6.41
J0120 + 2133 [129]	0.576	-9.12 ± 40
J0120 + 2133 [129]	0.729	0.73 ± 6.42
J0120 + 2133 [129]	1.048	5.47 ± 18.7
J0120 + 2133 [129]	1.325	2.6 ± 4.19
J0120 + 2133 [129]	1.343	8.36 ± 12.2
J1944 + 7705 [129]	1.738	12.7 ± 16.2
HS1549 + 1919 [130]	1.143	-7.49 ± 5.53
HS1549 + 1919 [130]	1.342	-0.7 ± 6.61
HS1549 + 1919 [130]	1.802	-6.42 ± 7.25
J1120 + 0641 [131]	7.05852	128 ± 684
J1120 + 0641 [131]	6.17097	-102 ± 152
J1120 + 0641 [131]	5.95074	-228 ± 174
J1120 + 0641 [131]	5.50726	74.2 ± 111

one may marginalize the phase for N measurements independently—otherwise, marginalization requires an $N + 1$ -dimensional integral that does not factorize.

For a collection of observations at multiple redshifts, their spacing in time is also important, since the scalar’s phases for two nearly coincident (relative to T_φ) observations are deterministically related. To accurately account for these correlations, and to provide realistic bounds for scalars so light that their phases at the relevant times are not effectively stochastic, an analysis must fully model the scalar’s dynamics and marginalize over the background

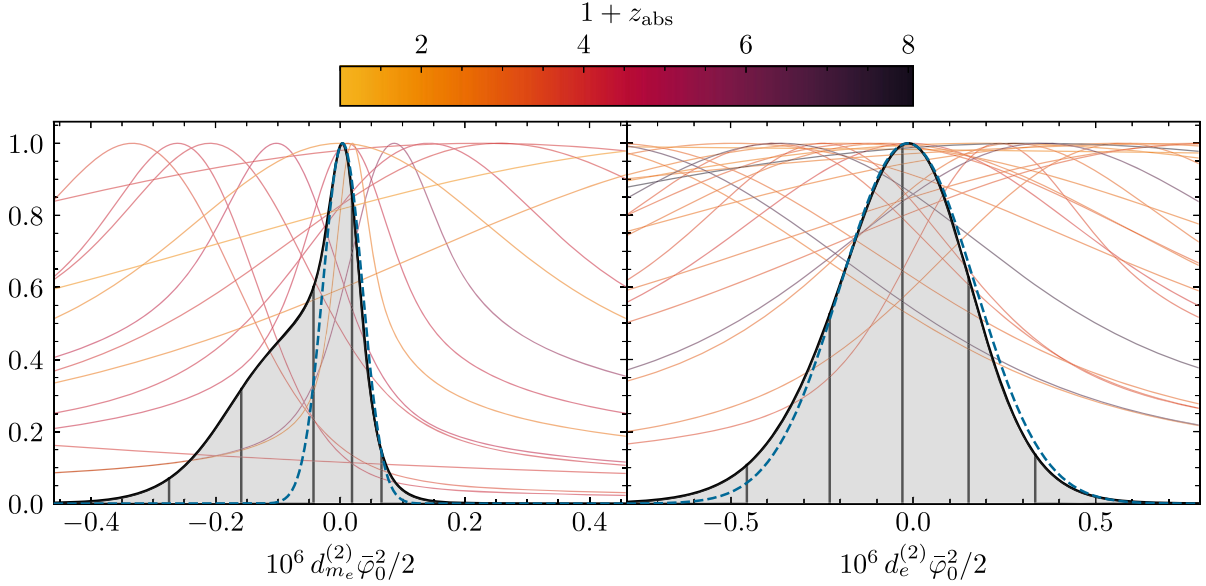


FIG. 3. Individual and combined posteriors from various observations of quasar absorption lines for a scalar coupled quadratically to the electron (left) or photon (right). Posteriors are plotted relative to their maximum value and against the present-day shift in a parameter relative to its vacuum value (evaluated at the antinode), $\Delta\lambda_0/\lambda(0) = d_\lambda^{(2)} \mathcal{A}_0^2/2$; the shift at the absorption redshift z_{abs} is given by $(1 + z_{\text{abs}})^3 \Delta\lambda_0 \sin^2 \theta(z_{\text{abs}})$. The transparent curves depict posteriors from individual observations (despite the fact that they are improper), colored by $1 + z_{\text{abs}}$ per the color bar. The black curves depict the joint posterior for the quasar observations in Table I, independently marginalized over each unknown phase $\theta(z_{\text{abs}})$. Vertical lines denote the $\pm 1\sigma$ and $\pm 2\sigma$ quantiles and the median for the joint posterior. The dashed blue curves depict the would-be joint posteriors were one to instead assume $\sin^2 \theta(z_{\text{abs}}) = 1/2$ at each z_{abs} .

cosmology. Here we perform a simpler analysis where we treat the scalar's phase at each z_{abs} in the sample as independent and stochastic and perform a few cross checks on subsets of these samples. Fig. 3 demonstrates the impact of phase marginalization on combined quasar constraints. We take each measurement to specify a simple Gaussian posterior over $\Delta\lambda_0/\lambda_0$ and independently marginalize over the phase of the scalar's oscillation for each measurement [as in Eq. (3.2)]. Fig. 3 assumes quadratic couplings to the electron or photon, insofar as the measured bounds on $\Delta\lambda(z_{\text{abs}})/\lambda_0$ are converted to present-day shifts via $\Delta\lambda(z_{\text{abs}}) = (1 + z_{\text{abs}})^3 \Delta\lambda_0$ and the phase factor is taken to be $\sin^2 \theta(z_{\text{abs}})$.

Fig. 3 includes curves that combine all measurements by either marginalizing over each phase $\theta(z_{\text{abs}})$ independently or by fixing $\sin^2 \theta(z_{\text{abs}}) = 1/2$ for each absorption redshift z_{abs} . For the varying- α results, the two results are quite similar near the peak, reflecting that the larger collection of $\Delta\alpha$ constraints are effectively randomly scattered about zero. (As explained in Sec. III A, the tails of the distribution do fall off as a power law in the phase-marginalized case rather than exponentially.) The difference is much more striking for constraints on the electron mass: while the phase-marginalized result resembles the phase-fixed one near the peaks of the distributions, it exhibits greatly enhanced support at $\Delta m_{e,0}/m_{e,0} \sim -10^{-7}$. This skew arises due to a cluster of higher-redshift observations that themselves skew to negative $\Delta m_{e,0}/m_{e,0}$, which are evident in

Fig. 3. When marginalizing over each phase, other measurements centered near zero or positive values of $\Delta m_{e,0}/m_{e,0}$ each only contribute a square-root suppression at negative $\Delta m_{e,0}$ as opposed to an exponential one (when the phases are instead fixed).

Fig. 3 displays 1σ intervals of order 10^{-7} for $d_\lambda^{(2)} \mathcal{A}_0^2/2$. As a cross check, we compare these results to those with subsets of the quasar samples partitioned by redshift. For both couplings, selecting only quasars with $z_{\text{abs}} > 1.5$ hardly affects the width of the posteriors. However, for this subsample the electron coupling constraints shift further toward negative values (the median lying 1.4σ rather than 0.5σ below zero); these observations source the skew evident in Fig. 3. (All other tested subsamples exhibit an expected degree of scatter in central values.) Only four of the 14 quasar measurements of $\Delta\mu/\mu$ reside at $z_{\text{abs}} < 1.5$, and the bound deriving from them alone degrades by a factor of ~ 10 . Twelve of the 22 $\Delta\alpha/\alpha$ measurements reside at $z_{\text{abs}} < 1.5$, broadening the posterior by a factor of ~ 2 . Many of the $\Delta\alpha/\alpha$ measurements have z_{abs} near 1, with numerous pairs at relatively similar redshifts. Taking subsets that are separated in scale factor by at least 1%–10% again only degrades constraints by a factor $\lesssim 2$. Moreover, unlike for the electron coupling, the photon coupling bounds are hardly affected by phase marginalization for any subsample.

We reiterate that marginalizing over each phase independently is conservative for deriving upper limits—in the

absence of a systematic or physical effect, combining multiple measurements at nearby times with random scatter about zero would improve constraining power. In principle, however, doing so could obfuscate correlations that in fact favor nonzero $\Delta\lambda_0$ (if multiple measurements at nearby redshifts favor similar $\Delta\lambda$). Given that the joint, phase-marginalized posterior for $\Delta m_{e,0}$ in Fig. 3 displays non-negligible skew, this possibility merits a more dedicated analysis which we do not pursue here.

In conclusion, we take the result from Fig. 3, $-0.3 \lesssim 10^6 d_{m_e}^{(2)} \mathcal{A}_0^2/2 \lesssim 0.07$, as our most conservative estimate of the 95% credible interval. Because this bound is strongly influenced by higher- z_{abs} observations, it is inappropriate for $m_\phi \lesssim 10^{-31}$ eV; we therefore take the $z_{\text{abs}} < 1$ result, $10^6 |d_{m_e}^{(2)}| \mathcal{A}_0^2/2 \lesssim 5$, to indicate the degree to which bounds degrade at lower masses. (Scalars with masses below 10^{-32} eV are not captured by our treatment here, as they would not have redshifted sufficiently by the present day for their effect on the present-day values of the constants to be negligible.) Using Eq. (2.54), the bounds on the quadratic electron coupling are

$$-0.36 \lesssim d_{m_e}^{(2)} \left(\frac{m_\phi}{10^{-31} \text{ eV}} \right)^{-2} \frac{F_\phi}{10^{-2}} \frac{\bar{\rho}_{c,0}}{(1.76 \text{ MeV})^4} \lesssim 8.6 \times 10^{-2} \quad (3.8)$$

at $m_\phi \gtrsim 10^{-31}$ eV, with the bounds for negative and positive couplings respectively degrading by relative factors of ~ 70 and 17 at $m_\phi \approx 10^{-32}$ eV. For the photon coupling we take $10^6 |d_e^{(2)}| \mathcal{A}_0^2/2 \lesssim 0.6$ as a representative bound. Given that this result is relatively insensitive to the choice of redshift cut (or even whether phases are marginalized over), we take it to apply for all $m_\phi \gtrsim 10^{-32}$ eV. The resulting bound on the quadratic photon coupling is

$$|d_e^{(2)}| \lesssim 0.73 \left(\frac{m_\phi}{10^{-31} \text{ eV}} \right)^2 \left(\frac{F_\phi}{10^{-2}} \right)^{-1} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ MeV})^4} \right)^{-1}. \quad (3.9)$$

The bounds from quasars established in Eqs. (3.8) and (3.9) provide the most stringent constraints on the lightest end of hyperlight scalar parameter space, as seen in Fig. 2 and Fig. 5 in Sec. IV C.

D. Universality of free fall

The effects considered so far derive from temporal variations in fundamental constants. *Spatial* variations in the value of fundamental constants, such as those induced by localized matter distributions, exert forces on matter. Because the effects of coupled scalars on matter can be composition dependent, they source apparent violations

of the universality of free fall (UFF) toward a central mass [14,21,22,26,62,63].

A species X whose effective mass [i.e., $m_X(\phi)$ from Sec. II C 3] depends on space and time through ϕ experiences an acceleration

$$\mathbf{a}_X(t, \mathbf{x}) \approx -\frac{\nabla m_X(\phi)}{m_X(\phi)} = -g'_X(\phi(t, \mathbf{x})) \nabla \phi(t, \mathbf{x}), \quad (3.10)$$

with g'_X defined by Eq. (2.3). Additional terms proportional to temporal gradients or suppressed by $\dot{\mathbf{x}}$ are omitted. For the power-series expansion of $g_{m_X}(\phi)$ [Eq. (2.6)], the scalar sources an acceleration

$$\begin{aligned} \mathbf{a}_X(t, \mathbf{x}) &\approx -\nabla \phi(t, \mathbf{x}) \sum_{n=1} \frac{d_{m_X}^{(n)}}{(n-1)!} \phi(t, \mathbf{x})^{n-1} \\ &= -\nabla \phi(t, \mathbf{x}) \left[d_{m_X}^{(1)} + d_{m_X}^{(2)} \phi(t, \mathbf{x}) + \dots \right], \end{aligned} \quad (3.11)$$

as one might have inferred directly from the Lagrangian. In the vicinity of a localized matter distribution (Sec. III B) the scalar's gradient $\nabla \phi$ generally has a component toward the central mass that may be interpreted as a new, long-range “fifth” force. Species with different compositions (different dilatonic charges) at the same location experience different accelerations toward the central mass.

The degree of UFF violation for test masses A and B attracted toward a central mass C (often taken to be Earth) is quantified by the Eötvös parameter $\eta \equiv |\mathbf{a}_A - \mathbf{a}_B|/|\mathbf{g}|$, where $\mathbf{g} = -GM_C \mathbf{x}/r^3$ is the acceleration due to the gravitational attraction between the test and central masses. Linearly coupled scalars are directly sourced by a central body independent of their cosmological or galactic abundance (if any), with local profile given by Eq. (3.3) when the scalar only otherwise has a homogeneous, cosmological component (and when $m_\phi R_C \ll 1$). The resulting Eötvös parameter is [26,132]

$$\eta^{(1)} \approx \left| d_{M_C}^{(1)} \left(d_{m_A}^{(1)} - d_{m_B}^{(1)} \right) \right| e^{-m_\phi r} (1 + m_\phi r). \quad (3.12)$$

On the other hand, the UFF violation due to a quadratically coupled scalar is suppressed by its local field value (which, for hyperlight scalars, corresponds to its cosmological abundance). From Eq. (3.4) [26,133],

$$\eta^{(2)} \approx \left| d_{M_C}^{(2)} \left(d_{m_A}^{(2)} - d_{m_B}^{(2)} \right) \right| \bar{\phi}(t)^2. \quad (3.13)$$

The preceding calculations are classical, arising at tree level in field theory. The value of η is independent of the location $\mathbf{x}$ because the tree-level force Eq. (3.11) with the boundary condition Eq. (3.4) goes as $1/|\mathbf{x}|^2$, like gravity. Contributions to the force from loops are negligible as long as the cosmological abundance satisfies $\phi(t)^2 \gg G/|\mathbf{x}|^2$, or

equivalently as long as the momentum exchange $\sim |\mathbf{x}|^{-1}$ is much less than $\varphi(t)$ [134–136].

Terrestrial [137] and space-based [138] experiments have obtained stringent constraints on η , requiring models with new scalars to suppress their violation of the UFF by some means. The MICROSCOPE experiment in particular measured $10^{15}\eta = -1.5 \pm 2.7$ between test masses made of a platinum-rhodium alloy and of a titanium-aluminium-vanadium alloy, in a satellite orbiting Earth at an altitude of 710 km [138]. Since the dilatonic charges of these alloys and of Earth are of order 10^{-3} and 10^{-5} (for photon and electron couplings, respectively) [26], MICROSCOPE limits the linear couplings of light scalars (with $m_\varphi r \ll 1$) as $d_e^{(1)} \lesssim 3 \times 10^{-5}$ and $d_{m_e}^{(1)} \lesssim 6 \times 10^{-4}$. The dependence of UFF violation [Eq. (3.13)] for quadratically coupled scalars on their background values weakens bounds on the couplings because the scalar is sub-Planckian ($\varphi \ll 1$) and, for hyperlight scalars, due to the penalty from marginalizing over the scalar’s unknown phase. MICROSCOPE took data over a two year period; per Sec. III A, the Eötvös parameter predicted at fixed couplings $d_\lambda^{(2)}$ and abundance F_φ is suppressed by a factor of $(m_\varphi/6.6 \times 10^{-23} \text{ eV})^2$ (for $m_\varphi < 6.6 \times 10^{-23} \text{ eV}$). Simply placing an upper limit of $|\eta| \lesssim 2.7 \times 10^{-15}$ yields constraints on quadratic couplings of

$$|d_{m_e}^{(2)}| \lesssim 4.1 \times 10^8 \left(\frac{F_\varphi}{10^{-2}} \right)^{-1/2} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1/2} \quad (3.14a)$$

and

$$|d_e^{(2)}| \lesssim 1.9 \times 10^7 \left(\frac{F_\varphi}{10^{-2}} \right)^{-1/2} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1/2}. \quad (3.14b)$$

As evident in Fig. 2, these bounds are extremely weak in comparison to, e.g., those from quasar absorption spectra [Eqs. (3.8) and (3.9)]. They are, however, mass independent in the hyperlight regime.

E. The Oklo phenomenon

The Oklo phenomenon refers to the only known, natural occurrence of a sustained nuclear fission reaction, which occurred 1.8 Gyr ago and lasted $\mathcal{O}(10^5)$ yr in what is now an open-pit uranium mine in Oklo, Gabon, Africa [62,63,139,140]. Constraints on the value of α at redshift $z \simeq 0.14$ may therefore be obtained from the relative abundance of nuclear isotopes in the Oklo mines with an analysis somewhat similar to that for BBN (Sec. IV A). The key process is neutron capture on certain isotopes under exposure to a flux of neutrons from the sustained nuclear fission process. The conversion process is described by a

coupled system of nuclear equations and depends on the cross section σ_n for neutron capture. The cross section for such a “heavy nucleus plus slow neutron” [141] process exhibits a resonant structure $\sigma_n \propto (E - E_r)^{-1}$, where E is the energy of the incoming neutron and E_r the resonant energy. Constraints on $E_r(z \simeq 0.14) - E_{r,0}$ are thus obtained by measuring relative isotopic abundances in the Oklo mines; by modeling the dependence of E_r on fundamental constants, one can infer constraints thereof. Because no leptons are involved, the Oklo phenomenon is insensitive to variations in m_e .

As emphasized by Ref. [62], constraints on α from the Oklo rocks involve a number of assumptions, not least of which is the dependence of E_r on fundamental constants. Other sources of uncertainties are the temperature of the reaction, variability between the rock samples used, the specific shape of the neutron spectrum assumed for the analysis, and the geometry of the natural reactor. Bounds on α vary by $\mathcal{O}(1)$ ([62] and references therein), but

$$\left| \frac{\Delta\alpha(z \simeq 0.14)}{\alpha_0} \right| \lesssim 10^{-8} \quad (3.15)$$

is a conservative estimate. As discussed in Sec. III C, converting this bound into constraints on $d_e^{(2)}$ requires comparing the range of possible values taken on by the scalar at $z \simeq 0.14$ to those today [via Eq. (3.7)]. Even over the $\sim 10^5$ yr that the reaction lasted, any scalar with mass $m_\varphi \ll 10^{-27} \text{ eV}$ would have been effectively constant. On the other hand, the phase shift between $z = 0.14$ and today is only $\lesssim 2\pi$ for $m_\varphi < 10^{-31} \text{ eV}$; otherwise, the scalar’s phases at both epochs cannot be precisely related to one another. This interval in mass is precisely that which we seek to constrain. Given that the squared amplitude of the scalar is only reduced by a factor of $1.14^3 \approx 1.5$ in this interval, for much of parameter space the shift in α could be negligible sheerly by a coincidental cancellation in Eq. (3.7). Marginalizing over such possibilities (as described in Sec. III A but for both phases) impedes drawing any robust constraint from Eq. (3.15).

F. Atomic clocks and pulsar timing arrays

A dynamical scalar field coupled to the SM imparts spacetime dependence to atomic energy levels that can be probed by laboratory experiments. Searches for this effect compare two levels with different dilatonic charges or “sensitivities” in order to identify a relative difference that can be attributed to a scalar field. Atomic clocks are currently the most precise laboratory systems, and comparisons of different transitions within the same system or different systems yield stringent constraints on oscillatory signals of scalar dark matter [142–148] and also linearly drifting fundamental constants more generally [149–155] (see Ref. [156] for a review). In addition to atomic

transitions, nature provides another precise clock: pulsars with a steady rotation rate. The residuals in pulse arrival times at Earth may be compared to a local laboratory clock, providing another means to search for scalar dark matter [25,157,158]. Due to the effects discussed in Sec. III A, both yield weaker constraints than cosmological probes in the hyperlight regime we consider; as such, we only briefly summarize the constraints on quadratically coupled, hyperlight scalars as inferred from bounds on linearly coupled scalars that do follow the galactic dark matter profile.

A signal that is quadratic in an oscillating scalar has time dependence $\sim \sin 2m_\phi t$; a bound on the linear coupling of a scalar with mass m_ϕ therefore translates to one at $m_\phi/2$ for a quadratic coupling. Following the considerations of Sec. III A, we may recast limits on a locally stochastic, linearly coupled scalar with local density $\rho_{\phi,\oplus}$ and mass $2m_\phi$ into those for a cosmologically homogeneous, quadratically coupled scalar with mass m_ϕ that comprises a fraction F_ϕ of the dark matter density by equating their values for $\Delta\lambda_0/\lambda(0) \approx g_\lambda(\phi)$. Namely, Eq. (2.54) sets

$$\frac{d_\lambda^{(2)} \text{ cosmo.}}{d_\lambda^{(1)} \text{ galactic}} \approx \frac{4.45 \times 10^{13}}{f_{\text{stoch}}} \sqrt{\frac{\rho_{\phi,\oplus}}{0.4 \text{ GeV/cm}^3}} \frac{m_\phi}{10^{-22} \text{ eV}} \times \frac{(1.76 \text{ meV})^4}{F_\phi \bar{\rho}_{c,0}} \quad (3.16)$$

where f_{stoch} is the stochasticity factor discussed below Eq. (3.1).

In the regime of hyperlight masses, atomic clocks effectively probe the rate of change of fundamental constants—e.g., a linear drift rate, $d \ln \lambda / dt$. Marginalization over the scalar's unknown phase still penalizes bounds in this case, but with different parametric dependence on m_ϕ . In general, the theoretical drift rate is $dg_\lambda(\phi)/dt = m_\phi d_\lambda^{(n)} \sin^{n-1} \vartheta \cos \vartheta / (n-1)!$; per Sec. III A, the presence of nodes in phase that scale with ϑ^{n-1} penalizes bounds by a factor of m_ϕ^{-1} for $n = 1$ and 2 and by $m_\phi^{-(n-1)}$ for larger n . As such, converting bounds from linear couplings to quadratic couplings requires no additional suppression with m_ϕ in the regime where the experimental runtime is much shorter than the scalar's oscillation period.

The strongest constraint on the photon coupling at low masses to date is that from an optical clock frequency comparison of the electric octupole (E3) and electric quadrupole (E2) transitions of the $^{171}\text{Yb}^+$ ion with observations spanning over ~ 6 years [148], giving

$$|d_e^{(2)}| \lesssim 1.8 \times 10^5 \left(\frac{F_\phi}{10^{-2}} \right)^{-1} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1}. \quad (3.17)$$

The strongest clock constraint on the electron coupling at low masses is placed by a comparison between a ^{171}Yb

optical lattice clock and a ^{133}Cs fountain microwave clock which yields, over a ~ 298 -day duration [147],

$$|d_{m_e}^{(2)}| \lesssim 7.7 \times 10^9 \left(\frac{F_\phi}{10^{-2}} \right)^{-1} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1}. \quad (3.18)$$

Electron coupling limits are generally weaker as transition energies typically vary much less in their sensitivity to the electron mass than to the fine-structure constant; that is, the dilatonic charges of the transition energies are small.

Bounds from pulsar timing are driven by the ϕ dependence of the cesium clocks used to measure the pulsar residuals, which may therefore be interpreted as an “Earth term” in the analysis [157,158]. The pulsar rotation rate itself does depend on fundamental parameters due to angular momentum conservation; however, its sensitivity to the electron mass is five orders of magnitude weaker than that of the clock transition because the fraction of electrons in pulsars by mass is small. The dependence of the pulsar rotation rate on α is assumed to be negligible. As such, the pulsars effectively provide a steady reference clock that is independent of ϕ [157,158]. The recent NANOGrav 15-yr results constrain the linear couplings of hyperlight scalars [158] which translate to

$$|d_e^{(2)}| \lesssim 2.0 \times 10^6 \left(\frac{F_\phi}{10^{-2}} \right)^{-1} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1} \quad (3.19a)$$

$$|d_{m_e}^{(2)}| \lesssim 4.9 \times 10^6 \left(\frac{F_\phi}{10^{-2}} \right)^{-1} \left(\frac{\bar{\rho}_{c,0}}{(1.76 \text{ meV})^4} \right)^{-1}. \quad (3.19b)$$

The bounds listed in this section all only apply for non-clustering scalars with periods well below the total time span covered by the experimental measurements. As discussed in Sec. III A, both conditions are satisfied below masses $m_\phi \sim 10^{-28}$ eV. The ytterbium ion bound is the leading clock-based bound on the quadratic photon coupling, and the NANOGrav bound is that for the quadratic electron coupling.

G. Stars

New light particle degrees of freedom weakly coupled to the SM can be produced efficiently in hot stellar plasma and their subsequent escape provides additional cooling pathways. We find that stellar evolution is a relatively negligible constraint on hyperlight, quadratically coupled scalars considered here, but we note interesting features that arise due to the effect of the cosmological scalar background.

Reference [23] considers the production of two scalar quanta in stellar plasmas. Higher-dimensional operators like the ones considered here— $\phi^2 F_{\mu\nu} F^{\mu\nu}$, $\phi^2 \bar{e}e$, and $\phi^2 \bar{N}N$ (for N some nucleon)—are suppressed by higher powers of the Planck mass. Thermal production rates therefore depend steeply on temperature, making core-collapse supernovae ($T \sim 30$ MeV) the most constraining systems.

Reference [23] limits $d_e^{(2)} \lesssim 10^{15}$ and $d_{m_N}^{(2)} \lesssim 10^{14}$ from photon-pair annihilations ($\gamma + \gamma \rightarrow \varphi + \varphi$) and nucleon-nucleon bremsstrahlung ($N + N \rightarrow N + N + \varphi + \varphi$). Limits on $d_{m_e}^{(2)}$ from electron-pair annihilations ($e^+ + e^- \rightarrow \varphi + \varphi$) and Compton-like scattering ($e + \gamma \rightarrow e + \varphi + \varphi$) are further suppressed by $\sim(m_e/T)^2$ and $\sim\alpha(m_e/T)^2$, respectively, making the constraints irrelevant both compared to cosmological and laboratory bounds (Fig. 2).

In the presence of a cosmological background $\bar{\varphi}(t)$, one of the “legs” of the emission process can be replaced with the background field. The emission process in the stellar environment is then proportional to that of a single energetic φ particle, which can give stronger constraints as single-particle production arises from relatively lower-dimension operators and suffers less phase space suppression. For example, energy loss through the Primakoff process ($\gamma + Ze \rightarrow \varphi + Ze$) in globular clusters constrain $d_e^{(1)} \lesssim 10^8$ [159]. Similarly, single-particle emission via electron-ion bremsstrahlung ($e^- + Ze \rightarrow e^- + Ze + \varphi$) in red giants, horizontal branch stars, and galactic white dwarves constrain both $d_{m_e}^{(1)}$ and $d_{m_N}^{(1)} \lesssim 10^6$ [160,161]. Given that the dilatonic charges of nuclei for the photon coupling are order 10^{-3} , the latter translates to $d_e^{(1)} \lesssim 10^6/(Q_{m_N})_e \sim 10^9$.

Rates for background-assisted single φ emission may be obtained by substituting a slowly varying coupling $d^{(1)}(t) \rightarrow d^{(2)}\varphi_{\text{core}}(t)$, where $\varphi_{\text{core}}(t)$ is the scalar’s value inside the core of the star. For sufficiently small couplings to the star $d_{M_\star}^{(2)}$, the scalar at the core is well approximated by its cosmological value, $\varphi_{\text{core}}(t) \approx \bar{\varphi}(t)$. However, above the critical coupling Eq. (3.6), the field inside the star is exponentially suppressed [26]:

$$\varphi_{\text{core}}(t) \approx \frac{\bar{\varphi}(t)}{2} \exp\left(-\sqrt{d_{M_\star}^{(2)} \frac{GM_\star}{R_\star}}\right), \quad (3.20)$$

for $d_{M_\star}^{(2)} > 0$ and a nonrelativistic, nondegenerate star; we expect a suppression to also occur in relativistic and/or degenerate systems. Consequently, background-assisted emission in stellar cores drops off sharply above some (positive) critical coupling, in contrast to signals at the surface of Earth which still monotonically increase with coupling even beyond the critical value. For a star like the Sun, the condition that the scalar amplitude is well approximated by the cosmological background, Eq. (3.6), translates to $d_{m_\odot}^{(2)} \ll 5 \times 10^5$. For a neutron star (with radius ~ 10 km and 1.4 solar masses), $d_{m_{\text{NS}}}^{(2)} \ll 5$; while this can serve as a first estimate, Eq. (3.6) cannot be applied rigorously to a very degenerate system like a neutron star. For $d_{M_\star}^{(2)} < 0$, the solution of Ref. [26] in principle predicts that the scalar diverges at the core for a discrete set of

couplings: $\varphi_{\text{core}}(t) \approx \bar{\varphi}(t) \sec \sqrt{|d_{M_\star}^{(2)}| GM_\star/R_\star}$, but the theory becomes nonperturbative in this regime.

The nonmonotonic dependence on coupling of background-assisted emission in stars raises the possibility of a gap in what couplings are constrained, with a given star constraining a finite range of couplings, extending to larger couplings for less compact stars. A rigorous bound would also depend on the integrated effect of $\bar{\varphi}(t)$ as it redshifts and oscillates over stellar lifetimes (including cosmological in-medium effects on its evolution, relevant at large couplings). Moreover, even without taking screening into account, the above bounds on linear couplings and the fact that $\bar{\varphi}(t) \ll 1$ at all times translate to constraints on $d_e^{(2)}$ and $d_{m_e}^{(2)}$ much above 10^9 and 10^6 , respectively, significantly weaker than many depicted in Fig. 2.

IV. EARLY-TIME PROBES

We proceed by considering cosmological probes of the fundamental constants in the early Universe. In Sec. IV A we review the dependence of nucleosynthesis on fundamental constants and derive constraints on the couplings of hyperlight scalar fields. Section IV B then places bounds from the impact of varying fundamental constants on CMB anisotropies, building off the results of Ref. [60].

A. Big bang nucleosynthesis

A deviation in the value of fundamental constants from their present-day values during BBN would change the predicted yields of primordial elements, which are currently in excellent agreement with observations [37,59,64,162,163]. The vast majority of nuclei of primordial origin are either hydrogen or helium-4. Their partitioning is typically quantified by the helium yield by nucleon number, $Y_p \equiv 4n_{\text{He}}/(n_{\text{H}} + 4n_{\text{He}})$, which differs slightly from the yield by mass Y_{He} used in Sec. II C and CMB literature. We now review the basic elements of the nucleosynthesis calculations that predict Y_p in order to establish its expected scaling with various fundamental parameters, focusing on α and m_e ; we then derive approximate bounds on their early-time values from astrophysical measurements of Y_p . We assume the parameters are simply fixed to some time-independent values α_i and $m_{e,i}$ that differ from their present-day values, α_0 and $m_{e,0}$, as would be the case for the models we consider in Sec. II when the scalar field is frozen at early times. Our discussion follows that of Refs. [37,164].

The primordial yield of ${}^4\text{He}$ is largely determined by the abundance of free neutrons when fusion to helium becomes efficient, as nearly all free neutrons become bound in primordial ${}^4\text{He}$ nuclei at that time. The neutron abundance at helium fusion is in turn determined in two stages [164]. First, neutron-proton conversions $n + e^+ \leftrightarrow p + \bar{\nu}_e$ and $n + \nu_e \leftrightarrow p + e^-$ become inefficient when the weak

interactions decouple (at a temperature $T_W \approx \text{MeV}$), at which point their relative abundance freezes out. Solving the Boltzmann equation for evolution of the relative neutron abundance sets

$$X_{n,W} \equiv \frac{\bar{n}_n(T_W)}{\bar{n}_n(T_W) + \bar{n}_p(T_W)} \approx \int_0^\infty \frac{dT}{T} \frac{m_{np} e^{-I(T)}}{2T[1 + \cosh(m_{np}/T)]} \quad (4.1)$$

to a good approximation. Here $m_{np} = m_n - m_p$ is the neutron-proton mass difference and

$$I(T) = \int_0^T \frac{dT'}{T'} \frac{1 + 3g_A^2}{\pi^3} \frac{G_F^2 T'^5}{H} J(m_{np}/T') (1 + e^{-m_{np}/T'}), \quad (4.2)$$

where the weak axial-vector coupling constant of the nucleon $g_A \approx 1.2755$ [165], $G_F \approx 1.166379 \times 10^{-5} \text{ GeV}^{-2}$ is the Fermi constant, and

$$J(x) = \frac{45\zeta(5)}{2} + \frac{21\zeta(4)}{2}x + \frac{3\zeta(3)}{2} \left(1 - \frac{m_e^2}{2m_{np}^2}\right) x^2 \quad (4.3)$$

is the phase space form factor associated with neutron-to-proton conversion. During this epoch one may approximate $H^2 = \pi^2 g_\star(T) T^4 / 90 M_{\text{pl}}^2$ using $g_\star \approx 10.75$.⁵

After weak decoupling, protons no longer convert efficiently to neutrons but the neutron abundance decreases steadily via free neutron decay, $n \rightarrow p + e^- + \bar{\nu}_e$. Neutron decay continues until the so-called deuterium bottleneck is cleared, at which point efficient fusion begins (at a time t_{BBN}) and

$$X_{n,\text{BBN}} \approx X_{n,W} e^{-t_{\text{BBN}}/\tau_n}, \quad (4.4)$$

where τ_n is the lifetime of the neutron. Virtually all neutrons present at t_{BBN} then ultimately end up in primordial ^4He nuclei, so

$$Y_p^{\text{th}} \approx 2X_{n,\text{BBN}}. \quad (4.5)$$

In the Standard Model, the neutron lifetime is given by [165,166]

$$\frac{1}{\tau_n} = \frac{1 + 3g_A^2}{2\pi^3} |V_{ud}|^2 G_F^2 m_e^5 (1 + \text{RC}) P(m_{np}/m_e), \quad (4.6)$$

where the quark mixing matrix element $V_{ud} \approx 0.9742$ and the contribution from electroweak radiative corrections

⁵In trading integrals over time for ones over temperature, we approximate aT and therefore also $g_\star$ as constant, neglecting their actual time dependence near, e.g., electron-positron annihilation.

$\text{RC} \approx 0.03886$ [165,166]. The phase space factor $P(m_{np}/m_e)$ is approximately given by [164]

$$P(x) = \frac{1}{60} \left[\sqrt{x^2 - 1} (2x^4 - 9x^2 - 8) + 15x \ln(x + \sqrt{x^2 - 1}) \right], \quad (4.7)$$

but more detailed calculations yield $P(m_{np}/m_e) \approx 1.6887$ [165–167]. Experiments measure the present-day value of the neutron lifetime to be $\tau_{n,0} = 878.6 \pm 0.5 \text{ s}$ [72].

Eqs. (4.1), (4.4), and (4.6) show that the electron mass affects the helium yield by setting the phase space available to all weak reactions. Namely, heavier decay products reduce the phase space available to free neutron decay, extending the lifetime of the neutron and increasing Y_p . (For example, in the extreme case that $m_e > m_{np}$, decay via weak interactions is forbidden.) One can see that, evaluated near $(m_e/m_{np})^2 \simeq 0.15$, a change in the electron mass introduces only a fractional difference to the form factor $J(x)$, at least when the argument x is small ($\partial \ln P / \partial \ln m_e \ll 1$). Heuristically, this is because $J(x)$ comes about as an integral over the broad distribution of relativistic thermal leptons. In contrast, there are no temperature scales involved in the decay of free neutrons. This fact alone would not be sufficient to make the neutron lifetime sensitive to the electron mass scale were it not also the case that $m_e \sim 0.4 m_{np}$, such that electron products are only semirelativistic. As a result, $\tau_n^{-1} \sim m_e^5 (m_{np}/m_e - 1)^{7/2} + \dots$, resulting in $\partial \ln \tau_n^{-1} / \partial \ln m_e \sim \mathcal{O}(1)$.

The fine-structure constant affects Eqs. (4.1) and (4.4) via the electrodynamic contribution to the neutron-proton mass splitting,

$$m_{np} \approx m_{np}^{(\text{QED})} + m_{np}^{(\text{QCD})} \approx b\alpha\Lambda_{\text{QCD}} + (m_d - m_u), \quad (4.8)$$

where m_u and m_d are the masses of the up and down quarks and b is a proportionality constant. Lattice calculations of the QED contribution to m_{np} yield $m_{np}^{(\text{QED})} \approx -1 \text{ MeV}$ [168,169], with order 10% precision.⁶ Assuming no variation in either b or Λ_{QCD} , the neutron-proton mass splitting then varies with α via

$$\frac{\Delta m_{np}}{m_{np}} = \frac{m_{np}^{(\text{QED})}}{m_{np}} \frac{\Delta\alpha}{\alpha} \approx -\frac{1 \text{ MeV}}{m_{np}} \frac{\Delta\alpha}{\alpha}. \quad (4.9)$$

Like m_e , m_{np} affects the phase space available to all weak reactions, but m_{np} also controls the relative abundance of

⁶Calculations by alternative means yield a value $m_{np}^{(\text{QED})} \approx -0.76 \text{ MeV}$ [170] that has often been used in past literature; other recent approaches yield $m_{np}^{(\text{QED})} \approx -0.58 \text{ MeV}$ [171,172]. See, e.g., Ref. [173] for comparative discussion.

protons and neutrons in thermal equilibrium. Consequently, electroweak freeze-out is more sensitive to variations in m_{np} than to variations in m_e . An increase in α decreases the magnitude of the neutron-proton mass splitting, therefore increasing Y_p . (In the extreme case that $m_{np} \rightarrow 0$, their abundances are equal and $Y_p \rightarrow 1$.) The neutron lifetime is also slightly more sensitive to variations in m_{np} than m_e because m_{np} is the dominant dimensionful scale that enters relative to the Fermi coupling scale $G_F^{-1/2}$, as evidenced by considering neutron decay to massless leptons.

The net amount of neutron decay depends on the age of the Universe when the deuterium bottleneck clears, which itself depends upon the expansion history up until that point. It so happens that the bottleneck clears during electron-positron annihilation (which completes at temperatures around 20 keV in the SM [109,110]); the resulting entropy transfer causes the SM temperature to redshift more slowly than a^{-1} and the Universe to expand slightly more slowly. A heavier electron annihilates earlier, so the Universe is older when nucleosynthesis occurs and neutrons are less abundant. This effect slightly offsets the increase in neutron lifetime for a heavier electron, reducing the sensitivity of the helium yield to variations in m_e . While there may be other such effects on the dynamics of the plasma at play in an analysis more detailed than ours, we account for this change in the thermal history since it only makes the bounds more conservative. We employ the fitting functions from Ref. [174] to calculate the effective degrees of freedom $g_*(T)$.

To assess the dependence of Y_p on the fundamental constants λ , we numerically evaluate Eq. (4.1) and (4.4) at various values thereof. Though the level of calculation described here is insufficient to accurately predict the actual value of Y_p , we take it to provide a reasonable estimate of its scaling with the early-time values λ_i . We compute

$$\frac{\Delta Y_p}{Y_p} \approx 0.34 \frac{\Delta m_{e,i}}{m_{e,0}} + 2.6 \frac{\Delta \alpha_i}{\alpha_0}. \quad (4.10)$$

The coefficient of $\Delta m_{e,i}$ increases to 0.42 if the aforementioned effects of shifting $e^+ - e^-$ annihilation are neglected, which is consistent with the findings of, e.g., Ref. [162]. For all fundamental constants at their present-day value during BBN, detailed numerical calculations marginalized over *Planck*'s determination of the baryon abundance [108] yield

$$Y_p^{\text{th}} = 0.24672 \pm 0.00061, \quad (4.11)$$

while the measured value is currently [72]

$$Y_p^{\text{obs}} = 0.245 \pm 0.003. \quad (4.12)$$

Converting the $\pm 2\sigma$ interval into constraints on α and m_e varied independently yields

$$-1.2 \times 10^{-2} \lesssim \frac{\Delta \alpha_i}{\alpha_0} \lesssim 6.9 \times 10^{-3}, \quad (4.13a)$$

$$-9.2 \times 10^{-2} \lesssim \frac{\Delta m_{e,i}}{m_{e,0}} \lesssim 5.1 \times 10^{-2}. \quad (4.13b)$$

Observe that the intervals are slightly skewed toward early-time decreases in α or m_e because the standard theoretical prediction is slightly larger than the median of the observational results.

To translate Eq. (4.13) into constraints on quadratic scalar couplings, note first that the above analysis only applies to time-independent shifts in fundamental constants and therefore to scalars that are frozen during BBN. The pertinent mass range extends far beyond that for which the scalar remains frozen during recombination as well, which is our main focus. Constraints in this hyperlight regime ($m_\phi \lesssim 10^{-28}$ eV) are mass independent at fixed F_ϕ (see Sec. II E) and are well described by Eq. (2.57), which yields

$$-2.1 \lesssim d_e^{(2)} \frac{F_\phi}{10^{-2}} \lesssim 1.2 \quad (4.14a)$$

$$-16 \lesssim d_m^{(2)} \frac{F_\phi}{10^{-2}} \lesssim 9.1. \quad (4.14b)$$

These bounds mostly exclude the regime in which in-medium contributions to the scalar's potential are important (see Appendix A 2), but upper limits for scalar abundances F_ϕ smaller than the fiducial one chosen in Eq. (4.14) would require a dedicated analysis including such effects.

The abundance of deuterium is also sensitive to α and m_e [162] and measured at the percent level or better [175,176]; uncertainties in nuclear reaction rates important to theoretical predictions of the deuterium abundance have recently been improved substantially [177,178]. The deuterium abundance, unlike that of helium, is also sensitive to the baryon abundance (via the baryon-to-photon number ratio η) and drives BBN-derived constraints on ω_b . In principle, the correlation of ω_b with m_e imposed by CMB constraints would allow deuterium measurements to constrain the early-time electron mass, but obtaining precise bounds is nontrivial due to the intrinsic dependence on α and m_e (via their effect on the neutron lifetime and, for α , on m_{np} and nuclear binding energies) [162]. We defer such analyses to future work.

B. Cosmic microwave background anisotropies

The CMB anisotropies precisely measure the spatial inhomogeneities of the SM plasma at the time of hydrogen recombination as viewed on the sky today. The rates governing recombination dynamics and the diffusion of

acoustic waves in the plasma depend on the fine-structure constant and electron mass, providing a probe of their values at redshifts $\gtrsim 1000$. The hyperlight scalar responsible for early-time shifts in α or m_e can also reduce the distance to photon last scattering and the rate of structure growth; the former impacts the angular extent of the acoustic peaks on the sky and the latter suppresses the extent to which they are smeared by gravitational lensing. We study these dynamics in detail and perform a quantitative analysis of current cosmological data—measurements of the CMB anisotropies from *Planck* and low-redshift distances via baryon acoustic oscillations (BAO) and supernova (SN) brightnesses—in Ref. [60].

The CMB alone strongly constrains the early-time value of the fine-structure constant via its modulation of the diffusion rate, which alters the degree to which small-scale anisotropies are suppressed. High-resolution observations like *Planck* measure the so-called damping tail with sufficient precision to constrain α_i at the permille level. The fine-structure constant also affects the rate at which photons and free protons recombine into hydrogen, which impacts the generation of CMB polarization in a manner that scales the overall height of the polarization power spectrum. *Planck* also measured the *E*-mode component of the polarization spectrum (and its correlation with temperature anisotropies), though less precisely than it did the temperature power spectrum. On its own, *Planck* yields a 68% posterior interval $\alpha_i/\alpha_0 - 1 \approx (0.7 \pm 2.4) \times 10^{-3}$, which tightens to $(1.3 \pm 2.1) \times 10^{-3}$ when combined with baryon acoustic oscillation data [60].

Cosmological constraints on electron-mass variations are more complex because of the degeneracy with Λ CDM parameters that arise in cosmological observables. The electron mass determines the temperature (and so redshift) of recombination by scaling the hydrogen energy levels but, unlike the fine-structure constant, does not alter the relevant rate of photon diffusion. In fact, all of the physical effects apparent in the primary CMB anisotropies that are modulated by the redshift of recombination may be compensated for by adjusting the energy densities in baryonic matter, dark matter, and dark energy—all of which are free parameters in the Bayesian analysis. These degeneracies are of contemporary interest because early recombination shrinks the sound horizon of the photon-baryon plasma, for which CMB data then infer a larger Hubble constant; the model is thus a modification of early-time physics that could explain the present tension of Hubble constant measurements derived from the CMB and the distance ladder [179,180]. We discuss this possibility (with updated constraints using the most recent BAO and SNe datasets) in detail in Ref. [60].

The CMB by itself provides weak bounds on electron-mass variations (at the 10% level), but BAO and SNe data break degeneracies when combined with CMB data by constraining the late-time expansion history (irrespective of

the Hubble tension, i.e., not including the SNe calibrations that lead to the tension). In particular, these low-redshift datasets measure the redshift at which dark energy overtook matter in the Universe’s energy budget; along the CMB’s degeneracy between the electron mass and Λ CDM parameters, the redshift of matter–dark-energy equality changes substantially. However, measurements of this redshift (or, equivalently, the fraction of Universe’s energy in matter) are mildly discrepant between different probes—in particular, BAO data prefer earlier transitions to dark-energy domination (lower present-day matter fractions) than recent SNe datasets [60,181,182]. This discrepancy is puzzling even for standard Λ CDM cosmology, given that both probes measure cosmological distances over a common interval of the Universe’s history.

Reference [60] examines the manner in which these differences in low-redshift datasets propagate to constraints on the early-time electron mass (and other Λ CDM parameters, including the Hubble constant); at the extreme ends, *Planck* data combined with either SNe data from the Dark Energy Survey [183] or BAO data from the Dark Energy Spectroscopic Instrument [181] yield $10^2(m_{e,i}/m_{e,0} - 1) \approx -1.6 \pm 1$ and 1.7 ± 0.7 , respectively. Using BAO data from the completed Sloan Digital Sky Survey (SDSS) [184,185] and Six-degree Field Galaxy Survey (6dFGS) [186] instead yields 0.8 ± 0.7 , while using alternative SNe samples from Pantheon+ [187,188] and Union3 [189] yield -0.6 ± 1.1 and -2 ± 1.4 , respectively.⁷ This lack of concordance impedes a consensus cosmological constraint, as all dataset combinations yield percent-level measurements that are offset by 1 to 3σ at worst. Rather than arbitrarily selecting one measurement to highlight, we present results deriving from each dataset combination in order to be forthright about the discordance.

The preceding discussion made no reference to the gravitational impact of a hyperlight scalar field because current CMB data turn out to disfavor its suppression of structure. While Ref. [60] shows that the late-time increase to the matter density from such a scalar by itself opens up a new parameter degeneracy (that extends not just to the primary CMB anisotropies but BAO and SNe data simultaneously), *Planck* data do not tolerate the degree of structure suppression induced by a scalar field whose abundance can substantially realize this degeneracy. While this predicted parameter relationship is mildly evident in posterior constraints, its impact on marginal measurements of $m_{e,i}$ or α_i is only moderate. Similar conclusions apply to SM neutrinos [60,182], whose masses are truly free parameters and which have similar phenomenology to hyperlight scalars. The constraints on the abundance of coupled, hyperlight scalars from Ref. [60]

⁷Curiously, the measurement using an older version of the Pantheon sample [190] is $1^{+1.5}_{-1.4}$, a result more consistent with current BAO than current SNe data.

thus do not differ substantially from prior bounds on uncoupled scalars (at the same mass) [97]: over the mass range $10^{-31} \text{ eV} \lesssim m_\phi \lesssim 10^{-29} \text{ eV}$, the 95th posterior percentile is $F_\phi \approx 1\%-2\%$. On the other hand, Ref. [182] showed that cosmological neutrino mass measurements also vary substantially among more recent low-redshift datasets; constraints on the abundance of uncoupled, hyperlight scalars are likely also subject to a similar degree of variability.

While the detailed relationships between the fundamental constants, standard Λ CDM parameters, and current datasets are interesting in their own right, for the purposes of this work we need only understand how to extract bounds on the scalar's fundamental couplings $d_\lambda^{(n)}$ and abundance F_ϕ . Doing so requires understanding the regime in which the analysis of Ref. [60] is valid (Sec. IV B 1) and the relationship between constraints on the fundamental constants and the fundamental couplings (Sec. IV B 2). In Appendix B we discuss the impact of analysis choices in cosmological parameter inference, including the impact of priors (Appendix B 1) and of consistently accounting for the fundamental-constant dependence of the helium yield Y_p per Sec. IV A (Appendix B 2).

1. Effects of the matter potential

In Ref. [60] we assume that the scalar remains, to a good approximation, constant in time at least until last scattering. However, as discussed in Sec. II C, the effective in-medium mass is not necessarily negligible compared to H and m_ϕ [see Eqs. (2.37) and (2.40)]. Fig. 1 depicts the time evolution of the squared in-medium mass relative to $H^2 \propto \bar{\rho}$, up to an overall factor of the dimensionless coupling $d_\lambda^{(2)}$, and shows that electron-positron annihilation separates two qualitatively distinct regimes. While early-time dynamics inform the conditions of the field after inflation, our analysis of the CMB is only strictly affected by dynamics after annihilation; here we focus on the latter regime, deferring a detailed study of both to Appendix A. In particular, we quantitatively assess the impact of the matter potential as a function of coupling on the simplifying assumptions made in the analysis of Ref. [60].

In the late Universe, matter-induced dynamics are dictated by the in-medium potential Eq. (2.37) and parametrized by the dimensionless combination $\mathcal{D}$ from Eq. (2.49). Appendix A 1 shows that the field begins evolving before matter-radiation equality if $|\mathcal{D}| \gtrsim 1$ —that is, if the in-medium mass is comparable in size to H . Because the CMB visibility function peaks around $a_\star/a_{\text{eq}} \approx 3$ (where a_{eq} is the scale factor of matter-radiation equality), the fundamental constants would evolve non-negligibly during recombination for such values of the couplings. These dynamics would be inconsistent with the treatment of Ref. [60] (and prior studies of the CMB with varying constants); they are not precluded *a priori* but are

simply beyond the scope of the present work. In the regime $|\mathcal{D}| \lesssim 1$, the approximate solution Eq. (2.50) sets an estimated shift in fundamental constants λ between early times (post annihilation) and recombination of

$$|\lambda_\star/\Delta\lambda_i - 1| = |\bar{\phi}(a_\star)^2/\bar{\phi}_i^2 - 1| \approx \left| (Q_b)_\lambda d_\lambda^{(2)} \right| / 3. \quad (4.15)$$

Requiring the early-time shift in α or m_e to remain constant at the 10% level then requires $|d_\lambda^{(2)}| \lesssim 500$.

As discussed in Sec. II E, matter-induced dynamics can also modify the relationship between the scalar's early-time value $\bar{\phi}_i$ and its present-day abundance F_ϕ , which we take to be given exactly by the free-evolution result, Eq. (2.52), in this section. The total deviation is the sum of that accumulated through Eq. (2.50) evaluated at $a(t_{\text{osc}})$ and the small deviation from an $a^{-3/2}$ oscillation envelope at $t > t_{\text{osc}}$ (see Appendix A 1). The former effect is more pronounced for lighter fields that begin oscillating later. Requiring the matter-free result Eq. (2.52) to be accurate at the 10% level constricts the regime of validity by a further factor of two to five for scalars with masses $10^{-29} - 10^{-31} \text{ eV}$.

2. Cosmological constraints

The late-Universe probes discussed in Sec. III each only constrain a particular combination of model parameters—that is, they technically cannot disentangle the size of a dimensionless coupling $d_\lambda^{(n)}$ from the present-day scalar amplitude $\bar{\phi}_0$ on their own. While $\bar{\phi}_0$ is related to the scalar's mass m_ϕ and abundance F_ϕ relative to CDM, in the hyperlight regime the probes of Sec. III have no information on the scalar's oscillation frequency. Cosmological probes are unique in their sensitivity not just to the scalar's effect on the SM but also to its gravitational effects. The CMB, for instance, measures scale-dependent information through the integrated Sachs-Wolfe effect that depends on both the scalar's abundance and its mass (see Ref. [60]). A hypothetical cosmological detection of a coupled, hyperlight scalar could therefore measure not just the magnitude of the scalar's effect on SM parameters but also its mass and abundance, allowing for a direct constraint on its dimensionless coupling $d_\lambda^{(n)}$.

The results of Ref. [60] do not in fact demonstrate a preference for a nonzero abundance of a new hyperlight scalar, thereby limiting the data's potential to independently probe the couplings $d_\lambda^{(2)}$. Nevertheless, to better assess any information on $d_\lambda^{(2)}$ contained in the cosmological datasets, we display in Fig. 4 two-dimensional joint posteriors on the abundance F_ϕ and the coupling strengths $d_\lambda^{(2)}$, as well as the corresponding marginalized one-dimensional posteriors. The most salient feature in the two-dimensional posteriors is that, at lower F_ϕ and away from $d_\lambda^{(2)} \approx 0$, the mass levels

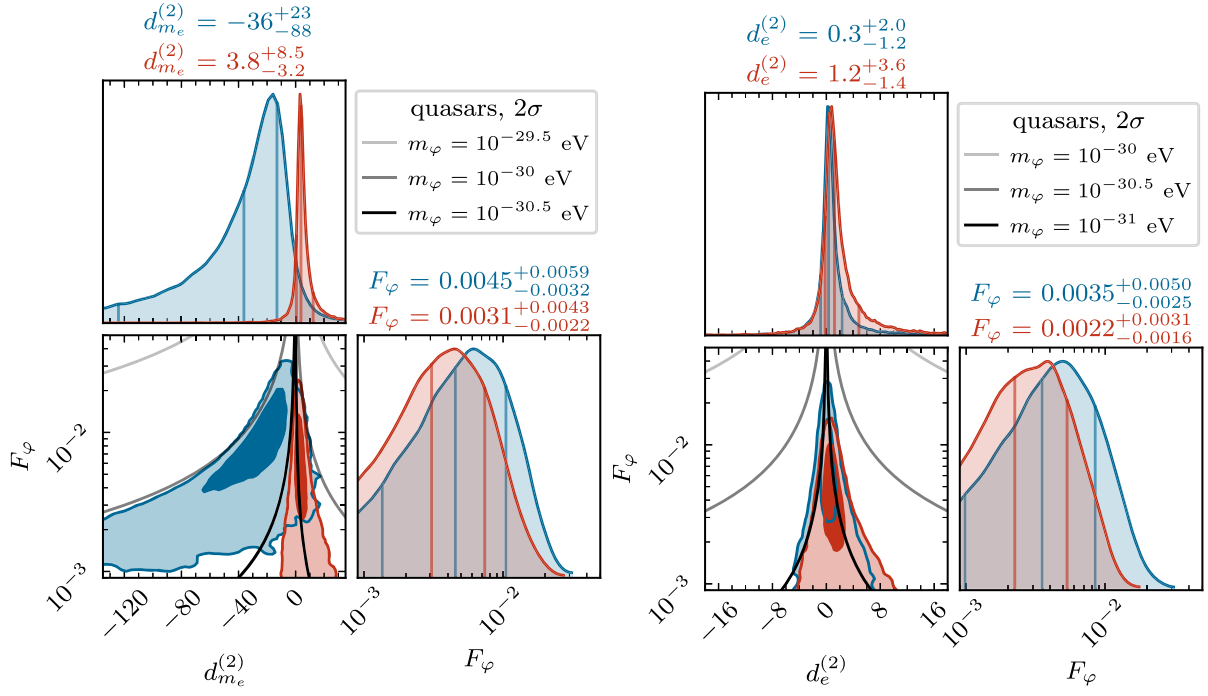


FIG. 4. Posterior distributions over a hyperlight scalar field's present energy density relative to that of cold dark matter, $F_\varphi = \omega_\varphi/\omega_c$, and its quadratic couplings to the electron (left) and the photon (right). The diagonal panels depict one-dimensional marginal distributions, and the lower left panel depicts 1 and 2 σ mass levels of the joint, two-dimensional posterior from parameter inference using *Planck* 2018 data alone (blue) and combined with BAO data (red). Gray lines depict 2 σ limits from quasar absorption lines (derived in Sec. III C) for a set of scalar masses m_φ as labeled. For the early-time constraints, the scalar mass is fixed to $m_\varphi = 10^{-30}$ eV, but these results do not change qualitatively within the range $10^{-32} \lesssim m_\varphi/\text{eV} \lesssim 10^{-29}$. Note that parameter inference for the early-time constraints place flat priors on F_φ and the early-time parameter value λ_i ; the increase in posterior density with F_φ at low values is the result of plotting the distribution over $\log_{10} F_\varphi$.

of the joint posteriors follow lines of constant $d_\lambda^{(2)} F_\varphi$, i.e., constant early-time shift in α or m_e as per Eq. (2.57). The joint constraints in this regime are simply driven by the posteriors over α_i and $m_{e,i}$ themselves. As such, the constraints on the electron coupling are substantially narrower when BAO data are combined with *Planck*, reflecting the strong degeneracy of the early-time electron mass with Λ CDM parameters [60,179]. Since *Planck* data on their own generally prefer the electron to be lighter at recombination than today, the posterior over $d_{m_e}^{(2)}$ skew toward negative values, but only at the 1 σ level; the posterior shifts to positive values when combined with BAO, but again only by $\sim 1\sigma$.

Since the CMB-derived bounds on coupled scalars are relatively insensitive to the scalar's mass [60], Fig. 4 takes $m_\varphi = 10^{-30}$ eV as a benchmark. However, the bounds from quasars do depend on mass; we therefore overlay the 2 σ upper limits from the analysis of Sec. III C for a variety of masses. Comparing these lines to the posteriors from *Planck* and BAO data indicates the mass at which the two probes have comparable constraining power.

Because Fig. 4 takes uniform priors over F_φ , the marginalized posteriors over $\log_{10} F_\varphi$ grow exponentially

at small F_φ ; they are then sharply cut off at values near half a percent, a feature shared with results for scalars with purely gravitational interactions [97] that derives solely from the disfavored effects of hyperlight scalar fields on structure growth. Namely, the small- F_φ tail of the marginalized posterior is driven mostly by the prior, while the sharp decline of posterior probability beyond $F_\varphi = \mathcal{O}(10^{-2})$ is driven by the data. Further, because uniform priors disfavor extremely small values for F_φ , the bulk of each posterior always lies within the regime of couplings where matter effects on the scalar's dynamics (Sec. IV B 1) are negligible. This holds even for the broad, *Planck*-only constraints on the electron coupling, for which 95% of posterior samples lie at values of $d_{m_e}^{(2)}$ above -500 . We discuss the impact of alternative priors in Appendix B 1.

C. Joint constraints

We now explore the joint constraining power of the early- and late-time datasets discussed in Secs. III and IV. The constraints from nucleosynthesis, the cosmic microwave background, quasar absorption spectra, and laboratory experiments span a broad range of cosmic history and are therefore sensitive to distinct regions of parameter space.

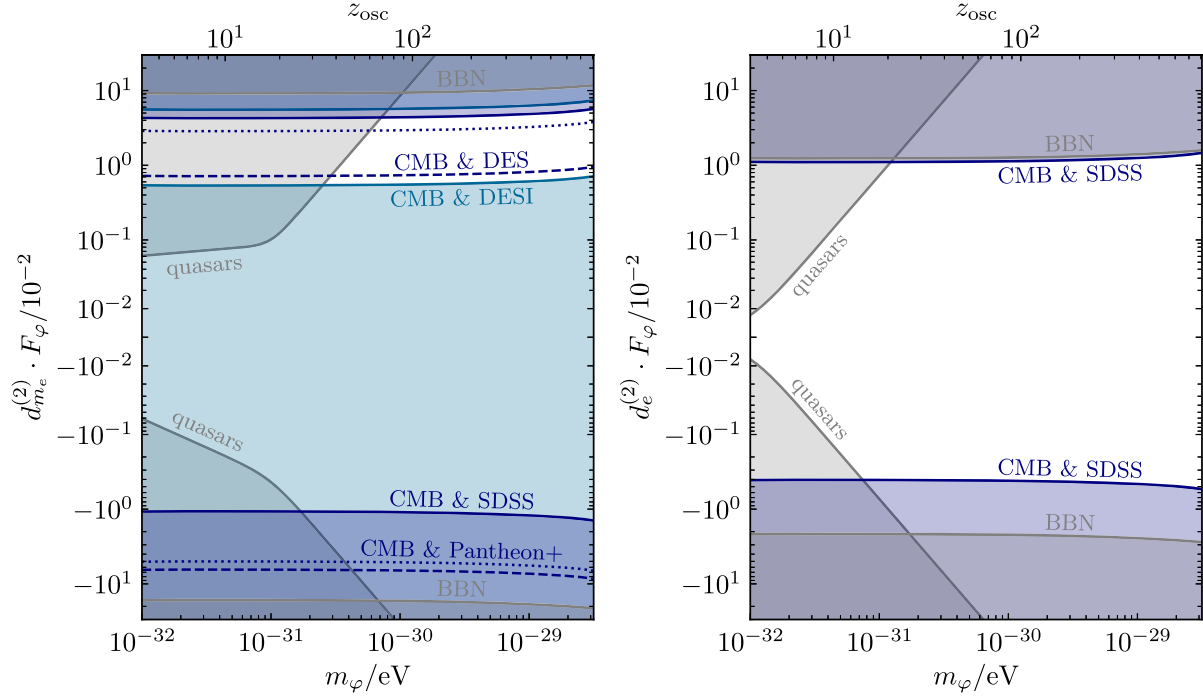


FIG. 5. Exclusion regions for the quadratic couplings of a new scalar to the electron (left column) and photon (right column) that are positive (top row) or negative (bottom row). Results are shown for a fiducial abundance of the scalar (relative to that of CDM) of $F_\varphi \equiv \bar{\rho}_{\varphi,0}/\bar{\rho}_{c,0} = 10^{-2}$, which is roughly the largest allowed in our analysis for the mass range depicted here [60]. The bounds on quadratic couplings, however, are mostly driven by the effect on the early-time values of α and m_e , as are the constraints from quasars (Sec. III C) and BBN (Sec. IV A). The CMB data are from the 2018 *Planck* release, which are presented in combination with BAO data from SDSS [184,185] and 6dFGS [186] (dark blue), BAO data from DESI [181] (light blue), and SNe data from either DES [183] (dashed) or Pantheon+ [187,188] (dotted). Observe that the combination with DESI yields an apparent preference for nonzero, positive electron-mass couplings at 2σ that is not particularly consistent with the other dataset combinations, all of which are consistent with zero; see Sec. IV B and Ref. [60] for exposition on the cosmological sensitivity of these measurements. At masses higher than those depicted, the scalar deviates from its initial condition by more than 10% at recombination, a regime which requires a dedicated analysis. Results approximately account for the impact on the relationship between mass and initial misalignment for scalars that begin oscillating around or before matter-radiation equality. Results are plotted on a symmetric logarithmic scale that interpolates linearly about zero between values of $\pm 10^{-2}$ via a transformation using an inverse hyperbolic function (arcsinh).

As explained in Sec. III, present-day probes (whether from equivalence principle violation or atomic clocks) are ultimately not competitive in their sensitivity to hyperlight scalars that are too light to make up all of dark matter. Fig. 5 displays the remaining constraints from quasars, BBN, and cosmological observations. The impressive precision and redshift coverage of quasar absorption spectra enables strong constraints at the lightest masses. Sensitivity falls off with increasing mass simply because heavier scalars have a smaller amplitude at fixed present-day abundance F_φ . In the mass range for which our treatment is valid, constraints from *Planck* (combined with BAO observations to break degeneracies with Λ CDM parameters, as discussed in Sec. IV B) improve upon those from BBN by upward of an order of magnitude, depending on the coupling and its sign, and dominate that from quasars at masses $m_\varphi \gtrsim 10^{-31}$ eV.

Measurements of the helium yield constrain the scalar's value at very high redshifts $z \gtrsim 10^9$, at which point it is frozen unless $m_\varphi \gtrsim 10^{-17}$ eV or the quadratic couplings are

large enough that the effects of matter on the scalar's dynamics are important (Sec. IV B 1; see also Ref. [37]). Our analysis of *Planck* data combined with BAO or SNe distance measurements (to break degeneracies with Λ CDM parameters) in Ref. [60] also applies only when the scalar is frozen through the end of recombination, which is the case for the entire mass range depicted in Fig. 5. At a fiducial $F_\varphi = 10^{-2}$, both the CMB and BBN constraints on $d_e^{(2)}$ and $d_{m_e}^{(2)}$ are strong enough that the effective potentials due to matter are always irrelevant. Larger F_φ are ruled out by a scalar's independent gravitational effects on the CMB. The constraints on couplings weaken at smaller F_φ ; for F_φ smaller than 10^{-3} or so, the bounds on couplings are large enough that the matter effects discussed in Sec. IV B 1 can be relevant around recombination. However, a hyperlight scalar with so small an abundance is gravitationally irrelevant. Our analysis is therefore only inapplicable for F_φ small enough such that (at fixed α_i or $m_{e,i}$) the scalar's

quadratic couplings are large enough that the fundamental constants are dynamical during recombination. Per Sec. IV B 1, the limits reported in Fig. 5 only enter this regime for $F_\phi \lesssim 10^{-4}$ or 10^{-5} , depending on the coupling under consideration (and its sign).

V. DISCUSSION AND CONCLUSIONS

New scalar fields can couple to the Standard Model as effective spacetime modulations of its fundamental parameters. In this work, we investigate scenarios in which a new field—light enough to be cosmologically frozen in amplitude until after recombination, but heavy enough to constitute a subcomponent of the dark matter today—modulates the mass of the electron and the electromagnetic fine-structure constant on cosmological timescales. We demonstrate that a scalar in this hyperlight mass range, even when comprising a fraction of the energy density of dark matter and with couplings to the SM as weak as gravity, may still have an observable impact on the values of fundamental constants throughout cosmological history. Furthermore, our companion work shows that such scalars, which inevitably contribute to expansion and the growth of structure, open up new degeneracies in cosmological parameters by adding to the matter density late in cosmological history [60]. We now summarize the theoretical considerations and phenomenological constraints presented in this work and discuss a number of avenues for future study they motivate. The leading constraints on the hyperlight parameter space are presented in Fig. 5.

A. Models of cosmologically varying constants

A substantial body of work has studied the cosmological implications of varying constants on a purely phenomenological basis; our work grounds this literature in concrete microphysical models and enables an assessment of their consistency. Section II reviews the field-theoretic framework of dynamical variations of fundamental constants induced by a new, light scalar field. We first outline a general description that promotes SM parameters to spacetime-dependent quantities via a scalar in Sec. II A. In Sec. II B we then review the low-energy effective theory of the scalar’s interactions with composite matter, providing a mapping from microphysical couplings to phenomenological consequences for both cosmological dynamics of SM matter and signatures in laboratory experiments [21,22].

Any evolution in fundamental constants is determined by the product of the cosmologically evolving amplitude of a scalar field (to some power n) and its interaction strength with the relevant SM operator, $d_\lambda^{(n)}$. The interactions responsible for the variation of fundamental constants, in turn, inevitably modify the scalar’s own dynamics. Using background-field methods in thermal field theory, Sec. II C derives the effective scalar potential sourced by

the cosmological abundance of SM matter in the relativistic and nonrelativistic regimes, presenting results relevant before and after electron-positron annihilation (see Fig. 1). We then apply these results in Sec. II E to delineate the regimes of the scalar’s cosmological dynamics, where we show that the approximations of prior work—time-independent shifts in fundamental constants leading up to recombination—are valid only in the small-coupling (and thus non-negligibly gravitating) regime. After electron-positron annihilation and for quadratic couplings to both photon and electron, the magnitude of the effective potential is suppressed relative to the Hubble scale by the small fractional contribution to the net mass in SM matter from electrons and QED effects; specifically, matter effects may be neglected for dimensionless couplings $d_\lambda^{(2)} \ll 10^4$. We study the dynamics of the regime of stronger coupling in more detail in Appendix A. Section II D also comments on the (currently undetectably small) extent to which in-medium effects within the SM itself effectively modify the fundamental constants.

B. Local and late-time searches

Numerous astrophysical and local probes—including quasar absorption spectra, atomic clocks, the Oklo phenomenon, stars, and tests of the universality of free fall—constrain variations of fundamental constants at times well after recombination. The constraining power of these probes relative to those from early-Universe cosmology depends on the form of the scalar’s potential and interactions, insofar as they determine both its cosmological redshifting and its response to matter. In particular, strong bounds from equivalence principle tests (Sec. III D) on light, linearly coupled scalars are unique in being independent of the scalar’s local or cosmological abundance, leaving no parameter space with observable cosmological signatures. The signatures of quadratically coupled scalars, on the other hand, do depend on the scalar’s abundance, leaving open the possibility of unconstrained parameter space with observable effects in the early Universe.

In Sec. III we extend the analysis of these late-time probes to handle several qualitatively distinct features of hyperlight scalars, with results summarized in Fig. 2. Measurements long after a hyperlight scalar begins oscillating are plagued by its extremely long oscillation period: the statistical possibility that the measurement occurs near a node of the signal steeply penalizes the reach of single, finite-duration probes. As a result, atomic clocks and tests of the equivalence principle are limited in constraining power.⁸ Section III A argues that the incurred penalty depends on the signal’s scaling with phase about its nodes. Equivalence principle tests (Sec. III D) for quadratic couplings are therefore penalized by two powers of mass

⁸The penalty does not apply to the discovery *reach* of such searches; see Sec. III A for further discussion.

times inverse experimental duration, while atomic clock and pulsar timing arrays (Sec. III F), which essentially probe the drift rate of fundamental constants, incur a penalty of one power for both linear and quadratic couplings. Combinations of measurements can mitigate this penalty only if spaced in time by a substantial fraction of the scalar's oscillation period because of the greatly suppressed chance that the scalar's amplitudes sampled at random points in its oscillation all occur at nodes. Constraints on varying constants from quasar absorption lines (Sec. III C) span redshifts as high as 7 and therefore remain competitive when analyzed as an ensemble (Fig. 3), even if during epochs when the field oscillates, setting the leading constraints at the lowest masses.

C. Cosmological signatures and constraints

Searches for hyperlight, coupled scalars in the early Universe have the crucial advantage of observing a hyperlight scalar at earlier times when it is not only more abundant but also frozen rather than oscillating. Namely, the amplitude of misaligned scalars (and therefore also that of their signatures) redshifts with expansion and is therefore greater at earlier times. Moreover, regimes for which the scalar is frozen during the pertinent epoch, like nucleosynthesis (for $m_\phi \lesssim 10^{-17}$ eV) or recombination ($m_\phi \lesssim 10^{-28}$ eV), are not plagued by any uncertainty in the scalar's instantaneous amplitude as in the oscillatory regime.

Section IV shows that the quadratic couplings of hyperlight scalars are best constrained by probes at higher redshift, including measurements of the primordial abundances of light elements (Sec. IV A) and the cosmic microwave background anisotropies (Sec. IV B). An ensemble of quasar absorption line measurements from various redshifts exceeds cosmological probes in reach for masses $m_\phi \lesssim 10^{-31}$ eV (see Figs. 4 and 5) since such scalars remain frozen longer and therefore redshift less by the time of the absorption that produces the lines. We studied the complementary power of these probes in Sec. IV C, showing that results deriving from CMB data in Ref. [60] provide leading bounds over nearly two decades in mass. The photon coupling is robustly constrained via its effects on the polarization and small-scale anisotropies of the CMB, limited to $-0.4 < d_e^{(2)} < 1.1$ (as quantified relative to gravity) for scalars that make up a percent of the CDM density (roughly the largest allowed fraction [60]). The electron coupling, on the other hand, is subject to strong degeneracies with cosmological parameters [60]; low-redshift data from baryon acoustic oscillations or supernovae can effectively break them, but current datasets do so inconsistently. Combined with CMB data, DESI has a preference for a positive, non-zero quadratic electron coupling whereas DES, SDSS, and Pantheon datasets rule out $|d_{m_e}^{(2)}|$ smaller than a few. Despite the lack of concordance in these joint constraints at

a quantitative level, cosmological data still provide leading bounds for $10^{-31} \lesssim m_\phi/\text{eV} \lesssim 10^{-28}$, as shown in Fig. 5.

Furthermore, this work provides theoretical grounding for the substantial body of literature constraining cosmological variations in fundamental constants on a phenomenological basis. The framework we develop, in the style of effective field theory, enables a relatively general comparison of the constraining power from cosmological data and independent laboratory and astrophysical searches. Prior work noted that linearly coupled scalars in principle do not redshift sufficiently after recombination to satisfy atomic clock constraints and simultaneously yield observably large variations in the CMB [191]. Our work shows that such arguments apply only to linearly coupled scalars that do not oscillate at late times: the fundamental-constant variation (or oscillation amplitude thereof) due to higher-order couplings redshifts faster, and marginalizing over the scalar's unknown oscillation phase parametrically penalizes limits from terrestrial probes (see Sec. III A). On the other hand, equivalence principle limits already entirely exclude any observable cosmological signatures from linearly coupled scalars (Sec. III D), largely independent of the scalar's abundance and potential [in the regime of interest with $V(\phi) \ll \text{eV}^4$].

To assess the implications of laboratory and astrophysical searches for cosmological signatures, we may straightforwardly extrapolate the noncosmological limits presented in Fig. 5 to limits on variations at recombination. In the mass range for which the scalar remains frozen at recombination, the early-time variation in fundamental constants is $\sim d_\lambda^{(2)} F_\phi$ [Eq. (2.57)]. The limits reported in Fig. 5, taking fiducial $F_\phi = 10^{-2}$, thus correspond to the percent variation allowable at recombination. Limits from quasars preclude subpermille early-time variations only for $m_\phi \lesssim 10^{-31}$ eV and quickly fall off at heavier masses. The limits from nucleosynthesis [Eq. (4.13)] directly apply at recombination, so long as the scalar is static in the intervening epoch—i.e., if neither the bare nor in-medium potentials are important. The latter holds for scalar abundances near our fiducial choice, but for much less abundant scalars (at fixed early-time variation in fundamental constants) the dimensionless coupling, and therefore in-medium effects, are enhanced and nontrivially modify these conclusions. In summary, while this work establishes minimal models for which early-time variations in fundamental constants (including those motivated as early-recombination solutions to the Hubble tension) are viable in light of both theory and independent searches, these considerations essentially require it makes up a non-negligible fraction of the total matter density today (the regime studied in Ref. [60]).

D. Future directions

In order to realize the phenomenological scenarios of prior study in which the change to the fundamental

constants is time independent through recombination, we restricted our analysis to regimes in which the effect of matter on the scalar's dynamics is negligible. The analysis of Sec. IV B 1 and Appendix A shows that, though such effects are only relevant for relatively large quadratic couplings, they lead the scalar to grow or decay during the matter era, depending on the sign of the coupling. That the field dynamically reacts to large matter couplings in particular complicates the “phenomenological” limit of varying-constant scenarios where time-independent shifts in fundamental constants during recombination are in effect realized by a strongly coupled yet negligibly abundant new scalar. Our analysis of concrete models enables a quantitative assessment of the consistency of such treatments, revealing that scalars which remain frozen during recombination must contribute non-negligibly to gravity. A consistent treatment of the opposite regime, with strongly time-dependent fields leading up to and during recombination (beyond existing studies of phenomenological toy models [192,193]), is necessary to place genuine constraints on couplings $|d_\lambda^{(2)}| \gtrsim 500$ [Eq. (4.15)] and scalar abundances $F_\varphi \lesssim 10^{-4}$. By the same token, scalars heavier than those we considered here would begin oscillating before last scattering. Such nontrivial dynamics before recombination are beyond the scope of this work and the cosmological analysis of Ref. [60] and are important targets for future work.

The possibility that the coupling to Standard Model matter invoked to change fundamental constants itself explains the scalar's dynamics is also theoretically attractive. The cosmological dynamics induced by SM matter were central to early studies of variations in fundamental constants driven by massless scalar fields (“quintessence”) [14,20]. As discussed in Sec. IV B 1 and Appendix A, such couplings to matter induce nontrivial scalar dynamics starting only around matter-radiation equality, which could potentially explain why the scalar begins diluting after recombination without invoking an arbitrarily light mass, albeit without resolving the issue of large radiative corrections induced by the same matter coupling. Models of early dark energy that attempt to address to Hubble tension are subject to a similar (but more severe) coincidence problem in the timing of recombination and the start of the scalar's dynamical evolution. Indeed, proposed solutions invoke couplings to the dark matter [194,195] or neutrinos [196–198] that induce the field to start rolling around matter-radiation equality, in exact analogy to the dynamics discussed here.

We treat only electron and photon interactions, but in a UV complete theory a scalar likely possesses further couplings with the SM and specific relationships between them. Couplings to quarks, gluons, and the Higgs all alter the masses of nucleons, whose cosmological signatures (beyond BBN) have not been studied in detail. We also consider only the simplest models of scalars with bare

masses, which nominally must be extremely tuned against radiative corrections; our study of such prototype models may provide inspiration for the construction of more complete theories.

To sum up, the phenomenology of scalar fields coupled to the SM is rich, and cosmological observations offer a powerful probe of unique regimes of parameter space, both via precise sensitivity to altered microphysics in the early Universe and through the gravitational impact of the new fields. Moreover, future opportunities to search for coupled scalars (hyperlight or otherwise) are myriad [199], including nuclear clocks [200–203], high-resolution CMB observations [204,205], CMB spectral distortions [206], the Lyman- α forest [207], and absorption spectra measurements from quasars [208] and emission line galaxies [209]. In the near term, further data from DESI and other future galaxy surveys [210–214] and supernova surveys [215–217] should improve limits on electron couplings and clarify the discordant results from current datasets. Studying the interesting dynamical regimes of strong couplings and heavier masses, as well as couplings to the strong sector, would paint a complete picture of new hyperlight, interacting scalars, which may well be first revealed through their cosmological signatures in the wealth of upcoming data.

ACKNOWLEDGMENTS

We thank Asimina Arvanitaki, Kimberly Boddy, Thomas Bouley, David Cyncynates, Thibault Damour, John Donoghue, Michael Fedderke, Subhjit Ghosh, Junwu Huang, Mikhail Ivanov, Marilena Loverde, Maxim Pospelov, Sergey Sibiryakov, Ken Van Tilburg, Neal Weiner, and Tien-Tien Yu for many useful discussions and correspondence and Marios Galanis for collaboration on initial stages of this work. M. B. is supported by the U.S. Department of Energy Office of Science under Award No. DE-SC0024375. M. B. and Z. J. W. are supported by the Department of Physics and College of Arts and Science at the University of Washington; Z. J. W. is also supported in part by the Dr. Ann Nelson Endowed Professorship. O. S. is supported by a Princeton Center for Theoretical Science Fellowship at Princeton University, and the Department of Physics and was supported by a DARE Fellowship from the Office of the Vice Provost for Graduate Education at Stanford University at the time this research was initiated. M. B. is grateful for the hospitality of Kavli Institute for Theoretical Physics (KITP) and Perimeter Institute where part of this work was carried out. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. This work was also supported by a grant from the Simons Foundation (1034867, Dittrich). This research was supported in part by Grant No. NSF

PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP). This work made use of the software packages EMCEE [218–220], CORNER.PY [221], NUMPY [222], SCIPY [223], MATPLOTLIB [224], XARRAY [225], ARVIZ [226], SYMPY [227], and CMASHER [228].

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: SCALAR DYNAMICS WITH EFFECTIVE POTENTIALS SOURCED BY MATTER

In this appendix we describe solutions to the Klein-Gordon equation with an effective potential sourced by matter, as discussed in Sec. II C. The matter potentials are qualitatively different before and after electron-positron annihilation, so we treat each regime separately. In both cases we discuss what restrictions on parameter space ensure our treatment of the theory remains self-consistent.

1. After annihilation

We first solve Eq. (2.2) for a homogeneous scalar with Eq. (2.37) contributing to its dynamics. Written in terms of

$N \equiv \ln a/a_{\text{eq}}$ where a_{eq} is the scale factor at matter-radiation equality, the equation of motion is

$$\partial_N^2 \varphi + \left(3 + \frac{\partial_N H}{H}\right) \partial_N \varphi = -\left(\frac{m_\varphi^2}{H^2} + \frac{\mathcal{D}}{1 + e^{-N}}\right) \varphi \quad (\text{A1})$$

where

$$\mathcal{D} \equiv \frac{3(Q_b)_e d_e^{(2)} + (Q_b)_{m_e} d_{m_e}^{(2)}}{2(1 + \omega_c/\omega_b)} \quad (\text{A2})$$

and $\partial_N H/H = -3(1+w)/2$ in a Universe with a constant equation of state w . Taking the scalar's bare mass to be negligible, Eq. (A1) admits analytic solutions in the limits of pure matter and radiation domination, $N \gg 0$ and $N \ll 0$. The radiation-era solution satisfying $\varphi \rightarrow \varphi_i$ and $\dot{\varphi} \rightarrow 0$ at early times $a \rightarrow 0^9$ is

$$\varphi_{\text{RD}}(a) = \varphi_i \frac{J_1(2\sqrt{\mathcal{D}a/a_{\text{eq}}})}{\sqrt{\mathcal{D}a/a_{\text{eq}}}} \quad (\text{A3})$$

where J_1 is the order-one Bessel function of the first kind. Since the scalar begins rolling (to some degree, at least) during the radiation era, we take general initial conditions φ_{eq} and $\dot{\varphi}_{\text{eq}}$ for the matter-era solution, which is

$$\varphi_{\text{MD}}(a) = \frac{(-\beta_- \varphi_{\text{eq}} + \partial_N \varphi_{\text{eq}})(a/a_{\text{eq}})^{\beta_+} + (\beta_+ \varphi_{\text{eq}} - \partial_N \varphi_{\text{eq}})(a/a_{\text{eq}})^{\beta_-}}{\sqrt{9/4 - 4\mathcal{D}}} \quad (\text{A4})$$

where

$$\beta_{\pm} = -3/4 \pm \sqrt{9/16 - \mathcal{D}}. \quad (\text{A5})$$

The matter-era solution can be matched onto the radiation-era one by setting φ_{eq} and $\partial_N \varphi_{\text{eq}}$ according to Eq. (A3); however, the transition between the two epochs is gradual enough that matching only qualitatively reproduces full solutions to Eq. (A1), which are depicted in Fig. 6 for positive couplings $d_\lambda^{(2)}$. The dynamics well before and well after equality match the analytic solutions (A3) and (A4) well, and clearly when $(Q_b)_\lambda d_\lambda^{(2)} \gtrsim 1$ the scalar rolls substantially before recombination. For negative couplings, the squared in-medium mass is negative, yielding growing rather than decaying solutions; for $|(Q_b)_\lambda d_\lambda^{(2)}| \gtrsim 1$, the growth is exponential [as captured by Eq. (A3) when evaluating the Bessel function with complex argument] and clearly not viable cosmologically. For $|(Q_b)_\lambda d_\lambda^{(2)}| \lesssim 1$, the growth remains a power law described by Eq. (A4).

We now determine what maximum $|d_\lambda^{(2)}|$ permits the scalar to be sufficiently stable that we may treat the variations in fundamental constants as themselves time

independent before recombination. For $|(Q_b)_\lambda d_\lambda^{(2)}| \lesssim 1$, most of the scalar's evolution takes place in the short period between matter-radiation equality and recombination and can be understood with the matter-era solution; we empirically find $\varphi/\varphi_i \approx (1 + 2a/3a_{\text{eq}})^{\beta_+}$ provides an excellent approximation, yielding Eq. (2.50) at small $\mathcal{D}$. For *Planck*'s preferred Λ CDM parameters [108], the scale factor of recombination relative to that at equality is $a_\star/a_{\text{eq}} \approx 3$ and $\omega_c/\omega_b \approx 5.4$. In the small- $\mathcal{D}$ limit, $\beta_+ \approx -2\mathcal{D}/3$ and the *shift* in λ varies by recombination by an amount given by Eq. (4.15).

The scalar departs from its initial condition even more by the time it begins oscillating due to its bare mass, i.e., as in Eq. (2.50) with a_{osc} substituted for $a_\star$. Between t_{osc} and today, the field amplitude oscillates approximately at the

⁹To be completely clear, this initial condition is specified long before matter-radiation equality but after electron-positron annihilation, since the scalar is also in general dynamical during and before annihilation (see Appendix A 2). The specified initial conditions remain appropriate because the scalar refreezes after annihilation unless its coupling is so large ($|d_\lambda^{(2)}| \gtrsim 10^9$) that the two regimes cannot be considered separately.

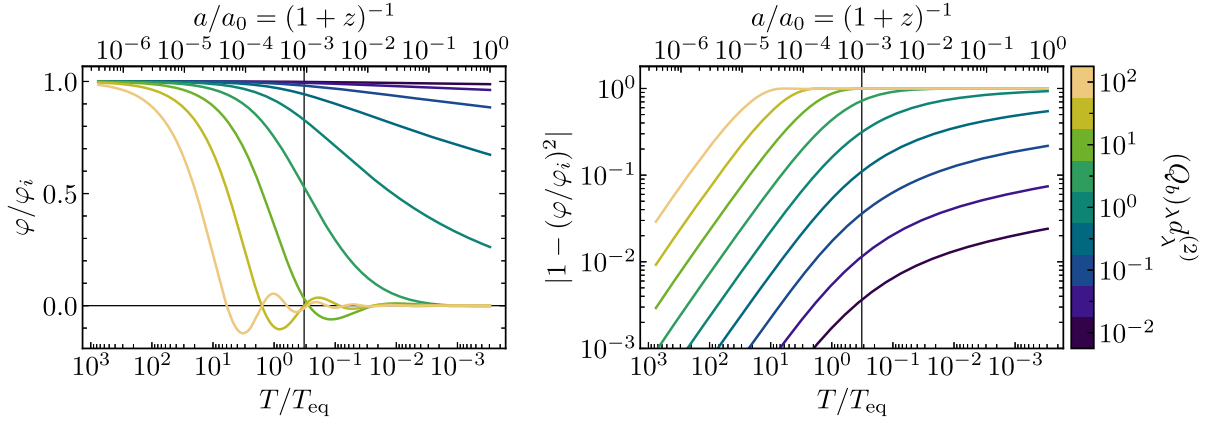


FIG. 6. Solutions to the Klein-Gordon equation (A1) for a quadratically coupled scalar with a nonrelativistic matter potential, with values of $(Q_b)_\lambda d_\lambda^{(2)}$ varying by color. Vertical black lines indicate the moment of recombination (i.e., for the fiducial case that the fundamental constants do not change with time). The right panel depicts the fractional difference in the shift in λ from the initial condition, i.e., $|\lambda/\lambda_i - 1|$.

bare mass frequency m_φ with an envelope well described by a WKB approximation [37,59,98],

$$\varphi(a) \propto \frac{1}{a^{3/2}} \frac{1}{\sqrt{m_\varphi^2 + \delta m_T(a)^2}}, \quad (\text{A6})$$

with $\delta m_T(a)^2 = \mathcal{D}H^2$ in the matter era. The scalar's energy density relative to that of CDM is therefore not constant after t_{osc} when the in-medium mass is non-negligible. Combined with Eq. (2.50), at small $|\mathcal{D}|$ and $a > a_{\text{osc}}$ we have

$$\begin{aligned} \frac{\bar{\rho}_\varphi(t)}{\bar{\rho}_c(t)} &\approx F_\varphi \frac{1}{(1 + 2a_{\text{osc}}/3a_{\text{eq}})^{4\mathcal{D}/3}} \frac{m_\varphi^2 + \delta m_T(a_{\text{osc}})^2}{m_\varphi^2 + \delta m_T(a(t))^2} \\ &\approx F_\varphi \frac{1 + \mathcal{D}(1 - [a(t)/a_{\text{osc}}]^{-3})}{(1 + 2a_{\text{osc}}/3a_{\text{eq}})^{4\mathcal{D}/3}} \end{aligned} \quad (\text{A7})$$

where F_φ is the scalar-CDM ratio given by Eq. (2.52), i.e., for an uncoupled scalar. Eq. (A7) encodes the correction to the relationship between the early-time parameter shift $\Delta\lambda_i/\lambda(0)$, $d_\lambda^{(2)}$, and the late-time abundance F_φ for an uncoupled scalar, Eq. (2.57). As explained in Sec. IV C, because these effects are relevant only at large $d_\lambda^{(2)}$ but cosmological data place relatively strong constraints on $\Delta\lambda_i/\lambda(0)$, they are only relevant in the regime in which the scalar's abundance is negligibly small. Any changes to the scalar's dynamics are then themselves negligible. Eq. (A7) still encodes the correction between F_φ [taken to be defined as the combination in Eq. (2.57)] and the present-day scalar abundance relevant to, e.g., laboratory probes. But these effects are again small in the parameter space we consider.

2. Before annihilation

In the early Universe, the Klein-Gordon equation with the in-medium mass Eq. (2.40) (in the relativistic limit) may be written as

$$\partial_N^2 \varphi + \partial_N \varphi = - \left(\frac{m_\varphi^2}{H^2} + \frac{15d_{m_e}^{(2)}}{2\pi^2} e^{2N} + \frac{25\alpha(0)d_e^{(2)}}{8\pi} \right) \varphi \quad (\text{A8})$$

where $N = \ln a/a_{\text{ann}} \equiv \ln m_e/T$. Eq. (A8) may be solved analytically for each coupling individually; we also compare to numerical solutions using the full potential Eq. (A8) which is valid through the end of annihilation.

Because the scalar's in-medium mass becomes exponentially suppressed at annihilation (for both couplings), the scalar refreezes except for very large $d_\lambda^{(2)}$. Though non-trivial dynamics before annihilation therefore have no direct effect on phenomenology at later times, they do affect how the scalar's value at, e.g., recombination ($\varphi_\star$) is related to its value at earlier times. For large enough couplings and $\varphi_\star$, the scalar at earlier times could probe regimes that are beyond the scope of the effective theory Eq. (2.7). For instance, interactions at higher order in φ can in principle be present in the full theory, but such higher-dimension operators should be suppressed in the regime that the effective field theory remains valid. In the sense that the power series in Eq. (2.6) remains perturbative, the leading-order contribution should remain smaller than unity to avoid sensitivity to unknown details of the theory's UV completion. Moreover, our calculation of the effective potential in Sec. II C relied on an expansion in small $g_\lambda(\varphi)$. In our case, $g_\lambda(\varphi) = d_\lambda^{(2)}\varphi^2/2$ should then remain small at all times. One could also require the scalar to remain below the Planck scale ($\varphi \lesssim 1$), but this condition is weaker than the former unless $d_\lambda^{(2)} < 1$.

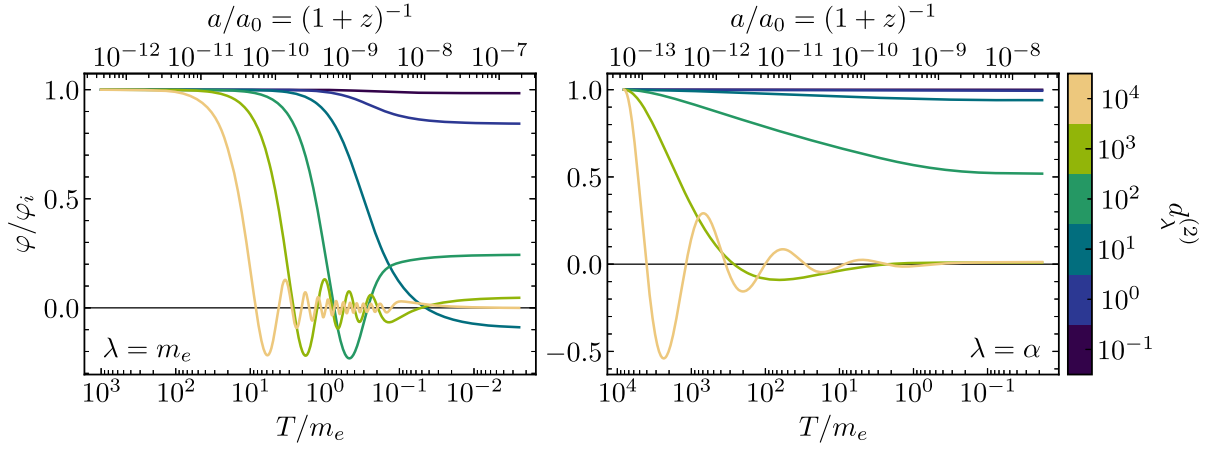


FIG. 7. Solutions to the Klein-Gordon equation (A8) for a scalar coupled quadratically to the electron (left) and photon (right), with couplings $d_\lambda^{(2)}$ varying by color.

Second, our treatment does not capture possible backreaction of the scalar on the SM plasma (nor the expansion of the Universe). Setting aside the current lack of a formalism to treat it, such backreaction would likely have a significant effect on nucleosynthesis (which is sensitive to plasma dynamics during this epoch). As a proxy, one might require that the effective potential Eq. (2.39) be negligible compared to the plasma energy density; backreaction would otherwise be important on the grounds of energy conservation. Since $\rho_{\text{SM}}(T) = \pi^2 g_\star(T) T^4/30$, this condition amounts to requiring that the quantity in brackets in Eq. (2.39) remain smaller than unity.

To determine the limits imposed by these considerations, we solve Eq. (A8). We again compare approximate, analytic results to numerical results, the latter of which are presented in Fig. 7. Under the electron coupling, the solution to Eq. (A8) with an initially frozen value (say, whatever value φ_{RH} it took on at reheating) is

$$\varphi_{m_e}(a) = \varphi_{\text{RH}} \frac{\sin(\sqrt{\mathcal{D}_{m_e}} a / a_{\text{ann}})}{\sqrt{\mathcal{D}_{m_e}} a / a_{\text{ann}}}. \quad (\text{A9})$$

where $\mathcal{D}_{m_e} \equiv 15d_{m_e}^{(2)}/2\pi^2 g_\star(T)$. Eq. (A9), however, is only valid while electrons and positrons are relativistic. When they annihilate, the in-medium mass is exponentially suppressed and the scalar refreezes.

From the form of Eq. (A9), one would expect the scalar to refreeze at a value $\sqrt{\mathcal{D}_{m_e}}$ times smaller than its initial value φ_{RH} . While numerical results match this scaling when comparing solutions evaluated precisely when $T = m_e$, the subsequent dynamics are not captured by matching onto solutions to the massless Klein-Gordon equation because annihilation is not instantaneous. The dynamics during annihilation—in part that the in-medium mass becomes negligible later the larger $d_{m_e}^{(2)}$ is—are nontrivial, and the scalar ultimately refreezes at values that are instead $\sim \sqrt[4]{\mathcal{D}_{m_e}}$ times smaller than its initial value, as displayed

in Fig. 8. Moreover, the precise timing of the scalar's oscillations as its in-medium mass shuts off allows the scalar to ultimately refreeze near zero. As a function of the refrozen shift in the electron mass, the constraint that $g_{m_e}(\varphi_{\text{RH}}) \lesssim 1$ requires

$$d_{m_e}^{(2)} \lesssim 2 \left(\frac{\Delta m_{e,i}}{m_e(0)} \right)^{-2}, \quad (\text{A10})$$

where $\Delta m_{e,i}$ is the change in the electron mass postannihilation—i.e., that which would be the initial condition for the dynamics during recombination.

Oscillatory solutions for positive couplings again switch to exponentially growing ones for negative couplings. For large negative couplings $-d_{m_e}^{(2)} \gtrsim 100$, numerical solutions determine that the scalar refreezes at a value $\sim e^3 \sqrt{|\mathcal{D}_{m_e}|}$ times larger than φ_{RH} ; the dependence on $\sqrt{|\mathcal{D}_{m_e}|}$ is as

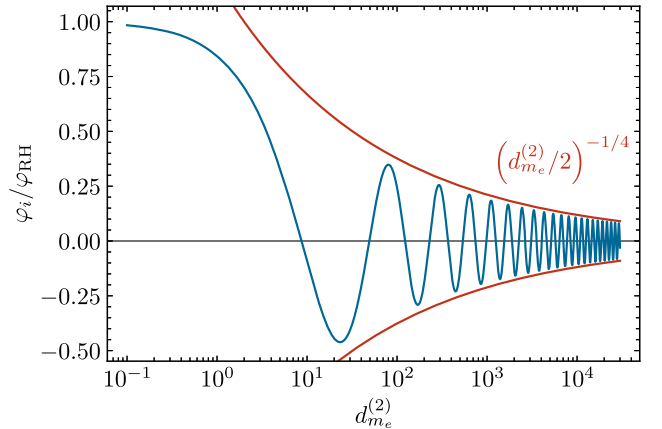


FIG. 8. Value φ_i at which a scalar refreezes after electron-positron annihilation as a function of its quadratic coupling to the electron, obtained by numerically solving the Klein-Gordon equation with effective potential Eq. (2.39), itself evaluated by numerical quadrature of Eq. (2.26).

expected from Eq. (A9), while the numerical coefficient reflects the particular dynamics of refreezing as the in-medium mass drops exponentially during annihilation. At the classical level, there is no theoretical issue with the scalar having an arbitrarily small value before reheating that is substantially amplified to its postannihilation value. However, during inflation, a quantized scalar field with $m_\phi < H_{\text{inf}}$ acquires an average amplitude of order $\sim H_{\text{inf}}^2/m_\phi > H_{\text{inf}}$ [229]. Heuristically, this is because an inflationary Universe has a IR cutoff at the inflationary scale H_{inf} , meaning that the smallest physically meaningful scale is $\mathcal{O}(H_{\text{inf}})$. At the very minimum, inflation had to occur in time for reheating to complete at a temperature $T_{\text{RH}} \approx 4 \text{ MeV}$ [85–91], limiting $H_{\text{inf}} \gtrsim 10^{-14} \text{ eV}$. Requiring that $|\varphi_{\text{RH}}| > H_{\text{inf}}/\sqrt{2}M_{\text{pl}}$ implies

$$\sqrt{\frac{2|\Delta m_{e,i}/m_e(0)|}{|d_e^{(2)}|}} e^{-3\sqrt{|\mathcal{D}_e|}} \gtrsim \frac{H_{\text{inf}}}{\sqrt{2}M_{\text{pl}}}, \quad (\text{A11})$$

so

$$-d_e^{(2)} \lesssim \frac{2\pi^2 g_\star(T_{\text{RH}})}{135} W^2 \left(\sqrt{\frac{270}{\pi^2 g_\star(T_{\text{RH}})}} \sqrt{\frac{\Delta m_{e,i}}{m_e(0)}} \frac{M_{\text{pl}}}{H_{\text{inf}}} \right) \quad (\text{A12})$$

where $W(z)$ is the Lambert W function, which at large argument is $W(z) \approx \ln z - \ln \ln z$.

The scalar's effective mass induced by its coupling to photons is proportional to H at all times after reheating.¹⁰ Now defining $\mathcal{D}_e = 25\alpha(0)d_e^{(2)}/8\pi g_\star$, the solution under the photon coupling is

$$\varphi_e(a) = \frac{\beta_+ \varphi_{\text{RH}}(a/a_{\text{RH}})^{\beta_-} - \beta_- \varphi_{\text{RH}}(a/a_{\text{RH}})^{\beta_+}}{\sqrt{1 - 4\mathcal{D}_e}} \quad (\text{A13})$$

where $\beta_\pm = -1/2 \mp \sqrt{1/4 - \mathcal{D}_e}$, where we have taken $\dot{\varphi}_e(a_{\text{RH}}) = 0$. The solutions are oscillatory if $\mathcal{D}_e > 1/4$, i.e., when $d_e^{(2)} \gtrsim 370$. To leading order in small $\mathcal{D}_e$, $\beta_- \approx -\mathcal{D}_e$ and $\beta_+ \approx -1$. Eq. (A13) precisely reproduces the numerical solutions displayed in Fig. 7 well before annihilation, at which point the scalar again refreezes.

For positive couplings the dominant mode is $\propto a^{-\mathcal{D}_e}$, meaning the scalar would have been roughly $(T_{\text{RH}}/m_e)^{\mathcal{D}_e}$ times larger at reheating than at annihilation. Requiring $|g_e(\varphi_{\text{RH}})| \lesssim 1$ then amounts to

¹⁰Technically, the rest of the charged SM content would contribute to the interacting free energy at higher temperatures. The SM presumably also generates an effective potential for the scalar before it fully thermalizes, but it would not be captured by the thermal field theory calculations Sec. II C 2 relies on. We set aside these complex and uncertain details for the sake of simplicity.

$$d_e^{(2)} \lesssim -\frac{8\pi g_\star}{50\alpha(0)} \frac{\ln[\Delta\alpha_i/\alpha(0)]}{\ln(T_{\text{RH}}/m_e)} \quad (\text{A14})$$

where $\Delta\alpha_i$ is the change in α postannihilation that is the initial condition for the dynamics leading up to recombination. Again taking the minimum reheat temperature $T_{\text{RH}} \approx 4 \text{ MeV}$, this bound amounts to $d_e^{(2)} \lesssim 360$; for instantaneous reheating after high-scale inflation, the bound tightens by a factor ~ 20 .

For negative photon couplings, the solutions are always growing power laws; for large and negative $\mathcal{D}_e$, the growing mode is $\beta_- \approx \sqrt{-\mathcal{D}_e}$. Again assuming an initial field value generated stochastically during inflation $\sim H_{\text{inf}}^2/m_\phi$ and assuming reheating was instantaneous ($H_{\text{RH}} = H_{\text{inf}}$), $|\varphi_{\text{RH}}| > H_{\text{inf}}/\sqrt{2}M_{\text{pl}}$ if

$$\sqrt{\frac{2|\Delta\alpha_i/\alpha(0)|}{|d_e^{(2)}|}} \left(\frac{m_e}{T_{\text{RH}}} \right)^{\sqrt{-\mathcal{D}_e}} \gtrsim \frac{H_{\text{inf}}}{\sqrt{2}M_{\text{pl}}}, \quad (\text{A15})$$

so

$$-d_e^{(2)} \lesssim \frac{8\pi g_\star(T_{\text{RH}})}{25\alpha(0)} \frac{1}{\ln^2(T_{\text{RH}}/m_e)} W^2 \times \left(\ln[T_{\text{RH}}/m_e] \sqrt{\frac{25\Delta\alpha_i}{2\pi g_\star(T_{\text{RH}})}} \frac{M_{\text{pl}}}{H_{\text{inf}}} \right). \quad (\text{A16})$$

APPENDIX B: ANALYSIS VARIANTS FOR COSMOLOGICAL PARAMETER INFERENCE

1. Impact of priors on marginalized constraints

As the discussion of Sec. IV B 2 suggests, the marginal posteriors over both F_ϕ and $d_\lambda^{(2)}$ are strongly sensitive to the choice of prior, which impedes drawing robust constraints on the parameters of the underlying model. We now consider the effect of priors in more detail. In Sec. IV B 2 we selected uniform priors over F_ϕ and the early-time parameter values α_i and $m_{e,i}$ (as taken in all previous phenomenological studies). Such priors lack any theoretical input but allow the data to drive the shape of the posteriors, at least in determining where the tails of the posterior are located. The effective prior on the fundamental parameters $d_\lambda^{(2)}$ and $\bar{\varphi}_i$ induced by taking uniform priors over F_ϕ and λ_i/λ_0 is

$$d(\lambda_i/\lambda_0) dF_\phi = \frac{3\bar{\varphi}_i^3(1 + \omega_b/\omega_c)}{4} dd_\lambda^{(2)} d\bar{\varphi}_i. \quad (\text{B1})$$

Notably, this prior rather steeply favors $\bar{\varphi}_i$ as large as possible, akin to the implicit preference in models of axionlike early dark energy for Planckian initial misalignments [230].

While any choice of prior on phenomenological parameters ultimately lacks direct theoretical motivation, even that for fundamental parameters is ambiguous. For instance, from the perspective of effective field theory one might consider the dimensionless parameter $d_\lambda^{(2)}$ an order-unitary Wilson coefficient suitable for a uniform prior. On the other hand, $M_{\text{pl}}/d_\lambda^{(2)}$ may be better described as a new-physics scale for a UV completion of the model, and one might choose for it a log-uniform prior in order to avoid arbitrarily selecting a preferred energy scale. [That is, a uniform prior between a and b identifies $(b-a)/2$ as a “special” value, whereas a log-uniform prior allocates equal weight to each decade within the allowed limits.] A choice for $\bar{\varphi}_i$ could also be informed by the choice for $d_\lambda^{(2)}$ —perhaps $\bar{\varphi}_i$ is itself a characteristic energy scale in the theory, or derives from some other physics that generates the initial conditions, or so on. Posteriors, of course, are most useful when they are least sensitive to prior choices.

Given that the analysis of Ref. [60] does not demonstratively “detect” an early-time shift in α or m_e nor the presence of a hyperlight scalar, such issues are of less practical concern at the present. To illustrate what effect an prior defined instead over the fundamental parameters can have on the analysis in Ref. [60], we consider a uniform prior over $d_\lambda^{(2)}$ and a log-uniform prior over $\bar{\varphi}_i$. The resulting prior over phenomenological parameters is

$$dd_\lambda^{(2)} d \ln \bar{\varphi}_i = \frac{3(1 + \omega_b/\omega_c)}{4F_\varphi^2} d(\lambda_i/\lambda_0) dF_\varphi. \quad (\text{B2})$$

Compared to a uniform prior over F_φ and λ_i/λ_0 , this prior penalizes large F_φ quadratically. If the posteriors over $d_\lambda^{(2)}$ are largely driven by the likelihood’s dependence on λ_i/λ_0 , we would expect such a prior to drive the marginal posteriors over the quadratic couplings $d_\lambda^{(2)}$ to values even larger in magnitude. Priors must be proper, and the chosen limits on, e.g., $\bar{\varphi}_i$ determine the relative prior weight allocated to percent-level values of F_φ —i.e., those that were allowed in the results taking uniform phenomenological priors. Specifically, we impose the effective prior Eq. (B2) with boundaries corresponding to $d_\lambda^{(2)} \sim \mathcal{U}(-10^3, 10^3)$ and $\log_{10} \bar{\varphi}_i \sim \mathcal{U}(-2, 0)$, where $\mathcal{U}$ denotes a uniform distribution. We additionally restrict $10^{-4} \leq F_\varphi \leq 0.3$ in order to roughly match the range of values covered by the posteriors using uniform, phenomenological priors.

Fig. 9 compares marginalized posteriors for this prior with those using the same uniform prior over phenomenological parameters applied in Ref. [60]. The F_φ dependence of the prior Eq. (B2) clearly translates directly to the marginalized posterior; the prior penalizes larger F_φ so severely that the data’s disfavor for $F_\varphi \gtrsim 10^{-2}$ is not even apparent. As a result, the posterior over $d_{m_e}^{(2)}$ broadens by about an order of magnitude. Since the data only directly constrain F_φ and $m_{e,i}$, we therefore cannot derive any meaningful marginalized constraints on $d_{m_e}^{(2)}$ that are driven by the likelihood rather than the prior.

However, the posterior over $m_{e,i}$ is much more robust, only shifting marginally toward lower values. The degeneracy discussed in Ref. [60] allows for larger $m_{e,i}$ correlated

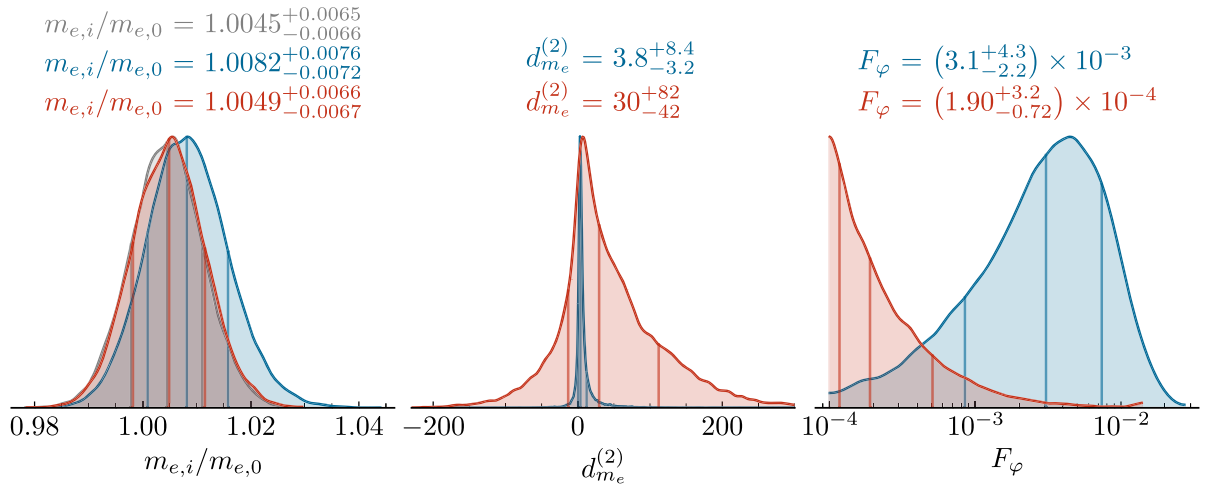


FIG. 9. Comparison of marginalized posteriors over $m_{e,i}$, $d_{m_e}^{(2)}$, and F_φ with uniform priors over both $m_{e,i}$ and F_φ (blue) and a uniform prior over $d_{m_e}^{(2)}$ and a log-uniform prior over $\bar{\varphi}_i$ (red). The marginal posterior over $m_{e,i}$ for purely a phenomenological model (with no scalar field) appears in gray. All cases use *Planck* 2018 data and BAO measurements and fix the scalar’s mass to 10^{-30} eV, where applicable. Vertical lines depict the median and $\pm 1\sigma$ quantiles of the marginal posteriors, which are also reported above each panel. Results are obtained as described in Ref. [60]. The evident prior sensitivity of the marginalized posteriors for $d_{m_e}^{(2)}$ and F_φ only slightly affects that of $m_{e,i}$, as interpreted in the text.

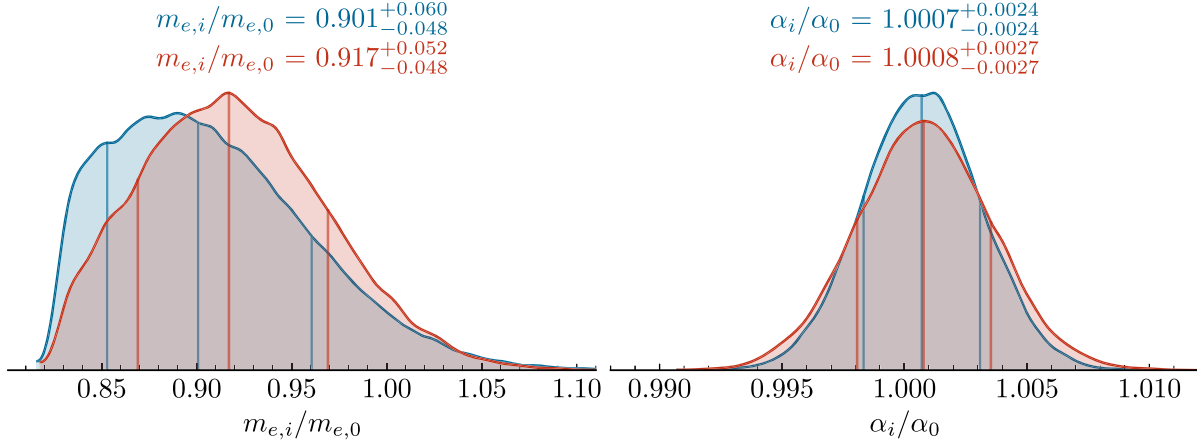


FIG. 10. Marginalized posterior distributions over the early-time electron mass (left) and fine-structure constant (right), each varied independently, for the corresponding scalar field models with (red) and without (blue) the effect of the variation in fundamental constants on BBN accounted for. The posteriors only use the *Planck* 2018 likelihoods, since the difference is negligible when including data that better constrain the models. Vertical lines mark the median and 16th and 84th percentiles; the median and corresponding $\pm 1\sigma$ uncertainties for each parameter are reported above each panel.

to a non-negligible scalar abundance, $F_\phi > 0$, but Eq. (B2) effectively precludes any gravitational impact of the scalar. Given that *Planck* data ultimately disfavor a substantial scalar abundance [60], the impact of this effect on the marginalized posteriors is minor. We therefore conclude that constraints derived directly from the early-time values of fundamental constants (and therefore the combination $d_\lambda^{(2)} F_\phi$) are robust to the choice of prior. Bounds on $d_\lambda^{(2)}$ at particular values of F_ϕ from cosmological data are thus on par with quasar- and BBN-derived bounds, which are insensitive to gravitational effects of the scalar in the first place.

2. Impact of BBN consistency

In the class of models we consider (Sec. II), shifts in the fundamental constants during BBN are generally the same as during recombination (unless the matter effects of Sec. II C are important, in which case for positive couplings the constants differ from their present-day values even more at earlier times). Following the discussion in Ref. [60], the primary effects of the helium yield on the CMB anisotropies are on diffusion damping rate and the width of the

visibility function. The change to the helium yield predicted by BBN in particular introduces *implicit* dependence on $m_{e,i}$ per Eq. (4.10) that (in principle) breaks the degeneracy of $m_{e,i}$ with Λ CDM parameters.

Here we use the approximate calculation of Sec. IV A to assess the impact of varying constants at BBN on CMB-derived constraints thereof. Fig. 10 demonstrates that the most extreme values of $m_{e,i}$ allowed by *Planck* data are more disfavored when this implicit dependence is accounted for. Namely, the marginalized posterior over $m_{e,i}$ skews less toward values smaller than one when treating BBN consistently. The degeneracy is therefore only meaningfully broken for $\mathcal{O}(10\%)$ early-time shifts in the electron mass, far outside of the strongest constraints obtained in Ref. [60] (that come from combining *Planck* data with BAO or supernova datasets). The CMB independently constrains the fine-structure constant to so narrow a region that the effect on Y_{He} is even more minor. The diffusion damping rate already explicitly depends on α_i^3 , while the dependence of the helium yield on α_i [using Eq. (4.10)] slightly decreases this scaling to $\sim \alpha_i^{2.6}$ [60], leading to the marginally broader posterior evident in Fig. 10.

- [1] F. Englert and R. Brout, Broken symmetry and the mass of gauge vector mesons, *Phys. Rev. Lett.* **13**, 321 (1964).
- [2] P. W. Higgs, Broken symmetries and the masses of gauge bosons, *Phys. Rev. Lett.* **13**, 508 (1964).
- [3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Global conservation laws and massless particles, *Phys. Rev. Lett.* **13**, 585 (1964).

- [4] R. D. Peccei and H. R. Quinn, *CP* conservation in the presence of instantons, *Phys. Rev. Lett.* **38**, 1440 (1977).
- [5] F. Wilczek, Problem of strong *P* and *T* invariance in the presence of instantons, *Phys. Rev. Lett.* **40**, 279 (1978).
- [6] S. Weinberg, A new light boson?, *Phys. Rev. Lett.* **40**, 223 (1978).

- [7] C. Brans and R.H. Dicke, Mach's principle and a relativistic theory of gravitation, *Phys. Rev.* **124**, 925 (1961).
- [8] T. Damour and G. Esposito-Farese, Testing gravity to second post-Newtonian order: A field theory approach, *Phys. Rev. D* **53**, 5541 (1996).
- [9] Y. Fujii and K. Maeda, *The Scalar-Tensor Theory of Gravitation*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2007).
- [10] J. Scherk and J. H. Schwarz, Dual models for nonhadrons, *Nucl. Phys.* **B81**, 118 (1974).
- [11] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory Vol. 2: 25th Anniversary Edition*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2012).
- [12] T. R. Taylor and G. Veneziano, Dilaton couplings at large distances, *Phys. Lett. B* **213**, 450 (1988).
- [13] Y. M. Cho, Unified cosmology, *Phys. Rev. D* **41**, 2462 (1990).
- [14] T. Damour and A. M. Polyakov, String theory and gravity, *Gen. Relativ. Gravit.* **26**, 1171 (1994).
- [15] D. B. Kaplan and M. B. Wise, Couplings of a light dilaton and violations of the equivalence principle, *J. High Energy Phys.* **08** (2000) 037.
- [16] M. Gasperini, F. Piazza, and G. Veneziano, Quintessence as a runaway dilaton, *Phys. Rev. D* **65**, 023508 (2002).
- [17] T. Damour, F. Piazza, and G. Veneziano, Runaway dilaton and equivalence principle violations, *Phys. Rev. Lett.* **89**, 081601 (2002).
- [18] S. Dimopoulos and G. F. Giudice, Macroscopic forces from supersymmetry, *Phys. Lett. B* **379**, 105 (1996).
- [19] H. Terazawa, Cosmological origin of mass scales, *Phys. Lett.* **101B**, 43 (1981).
- [20] J. D. Bekenstein, Fine structure constant: Is it really a constant?, *Phys. Rev. D* **25**, 1527 (1982).
- [21] T. Damour and J. F. Donoghue, Phenomenology of the equivalence principle with light scalars, *Classical Quantum Gravity* **27**, 202001 (2010).
- [22] T. Damour and J. F. Donoghue, Equivalence principle violations and couplings of a light dilaton, *Phys. Rev. D* **82**, 084033 (2010).
- [23] K. A. Olive and M. Pospelov, Environmental dependence of masses and coupling constants, *Phys. Rev. D* **77**, 043524 (2008).
- [24] A. Arvanitaki, J. Huang, and K. Van Tilburg, Searching for dilaton dark matter with atomic clocks, *Phys. Rev. D* **91**, 015015 (2015).
- [25] P. W. Graham, D. E. Kaplan, J. Mardon, S. Rajendran, and W. A. Terrano, Dark matter direct detection with accelerometers, *Phys. Rev. D* **93**, 075029 (2016).
- [26] A. Hees, O. Minazzoli, E. Savalle, Y. V. Stadnik, and P. Wolf, Violation of the equivalence principle from light scalar dark matter, *Phys. Rev. D* **98**, 064051 (2018).
- [27] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, The hierarchy problem and new dimensions at a millimeter, *Phys. Lett. B* **429**, 263 (1998).
- [28] E. G. Adelberger, B. R. Heckel, and A. E. Nelson, Tests of the gravitational inverse square law, *Annu. Rev. Nucl. Part. Sci.* **53**, 77 (2003).
- [29] J. Preskill, M. B. Wise, and F. Wilczek, Cosmology of the invisible axion, *Phys. Lett.* **120B**, 127 (1983).
- [30] L. F. Abbott and P. Sikivie, A cosmological bound on the invisible axion, *Phys. Lett.* **120B**, 133 (1983).
- [31] M. Dine and W. Fischler, The not so harmless axion, *Phys. Lett.* **120B**, 137 (1983).
- [32] M. S. Turner, Cosmic and local mass density of invisible axions, *Phys. Rev. D* **33**, 889 (1986).
- [33] A. D. Linde, Inflation and axion cosmology, *Phys. Lett. B* **201**, 437 (1988).
- [34] W. Hu, R. Barkana, and A. Gruzinov, Cold and fuzzy dark matter, *Phys. Rev. Lett.* **85**, 1158 (2000).
- [35] F. Piazza and M. Pospelov, Sub-eV scalar dark matter through the super-renormalizable Higgs portal, *Phys. Rev. D* **82**, 043533 (2010).
- [36] L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, Ultralight scalars as cosmological dark matter, *Phys. Rev. D* **95**, 043541 (2017).
- [37] T. Bouley, P. Sørensen, and T.-T. Yu, Constraints on ultralight scalar dark matter with quadratic couplings, *J. High Energy Phys.* **03** (2023) 104.
- [38] A. Banerjee, G. Perez, M. Safronova, I. Savoray, and A. Shalit, The phenomenology of quadratically coupled ultralight dark matter, *J. High Energy Phys.* **10** (2023) 042.
- [39] S. Hannestad, Possible constraints on the time variation of the fine structure constant from cosmic microwave background data, *Phys. Rev. D* **60**, 023515 (1999).
- [40] M. Kaplinghat, R. J. Scherrer, and M. S. Turner, Constraining variations in the fine structure constant with the cosmic microwave background, *Phys. Rev. D* **60**, 023516 (1999).
- [41] P. P. Avelino, C. J. A. P. Martins, G. Rocha, and P. T. P. Viana, Looking for a varying α in the cosmic microwave background, *Phys. Rev. D* **62**, 123508 (2000).
- [42] R. A. Battye, R. Crittenden, and J. Weller, Cosmic concordance and the fine structure constant, *Phys. Rev. D* **63**, 043505 (2001).
- [43] P. P. Avelino, S. Esposito, G. Mangano, C. J. A. P. Martins, A. Melchiorri, G. Miele, O. Pisanti, G. Rocha, and P. T. P. Viana, Early universe constraints on a time varying fine structure constant, *Phys. Rev. D* **64**, 103505 (2001).
- [44] S. J. Landau, D. D. Harari, and M. Zaldarriaga, Constraining nonstandard recombination: A worked example, *Phys. Rev. D* **63**, 083505 (2001).
- [45] C. J. A. P. Martins, A. Melchiorri, G. Rocha, R. Trotta, P. P. Avelino, and P. T. P. Viana, WMAP constraints on varying α and the promise of reionization, *Phys. Lett. B* **585**, 29 (2004).
- [46] G. Rocha, R. Trotta, C. J. A. P. Martins, A. Melchiorri, P. P. Avelino, R. Bean, and P. T. P. Viana, Measuring α in the early universe: CMB polarization, reionization and the fisher matrix analysis, *Mon. Not. R. Astron. Soc.* **352**, 20 (2004).
- [47] P. Stefanescu, Constraints on time variation of fine structure constant from WMAP-3 yr data, *New Astron.* **12**, 635 (2007).
- [48] M. Nakashima, R. Nagata, and J. Yokoyama, Constraints on the time variation of the fine structure constant by the 5-year WMAP data, *Prog. Theor. Phys.* **120**, 1207 (2008).

- [49] E. Menegoni, S. Galli, J. G. Bartlett, C. J. A. P. Martins, and A. Melchiorri, New constraints on variations of the fine structure constant from CMB anisotropies, *Phys. Rev. D* **80**, 087302 (2009).
- [50] E. Menegoni, M. Archidiacono, E. Calabrese, S. Galli, C. J. A. P. Martins, and A. Melchiorri, The fine structure constant and the CMB damping scale, *Phys. Rev. D* **85**, 107301 (2012).
- [51] J. Kujat and R. J. Scherrer, The effect of time variation in the Higgs vacuum expectation value on the cosmic microwave background, *Phys. Rev. D* **62**, 023510 (2000).
- [52] K. Ichikawa, T. Kanzaki, and M. Kawasaki, CMB constraints on the simultaneous variation of the fine structure constant and electron mass, *Phys. Rev. D* **74**, 023515 (2006).
- [53] S. J. Landau, M. E. Mosquera, C. G. Scoccola, and H. Vucetich, Early universe constraints on time variation of fundamental constants, *Phys. Rev. D* **78**, 083527 (2008).
- [54] C. G. Scoccola, S. J. Landau, and H. Vucetich, WMAP 5-year constraints on time variation of α and m_e in a detailed recombination scenario, *Phys. Lett. B* **669**, 212 (2008).
- [55] M. Nakashima, K. Ichikawa, R. Nagata, and J. Yokoyama, Constraining the time variation of the coupling constants from cosmic microwave background: Effect of Λ CDM, *J. Cosmol. Astropart. Phys.* **01** (2010) 030.
- [56] S. J. Landau and C. G. Scoccola, Constraints on variation in α and m_e from WMAP 7-year data, *Astron. Astrophys.* **517**, A62 (2010).
- [57] C. G. Scoccola *et al.*, The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic survey: Constraints on the time variation of fundamental constants from the large-scale two-point correlation function, *Mon. Not. R. Astron. Soc.* **434**, 1792 (2013).
- [58] N. Schöneberg and L. Vacher, The mass effect—variations of masses and their impact on cosmology, *J. Cosmol. Astropart. Phys.* **03** (2025) 004.
- [59] S. Sibiryakov, P. Sørensen, and T.-T. Yu, BBN constraints on universally-coupled ultralight scalar dark matter, *J. High Energy Phys.* **12** (2020) 075.
- [60] M. Baryakhtar, O. Simon, and Z. J. Weiner, Cosmology with varying fundamental constants from hyperlight, coupled scalars, *Phys. Rev. D* **110**, 083505 (2024).
- [61] K. A. Olive, M. Pospelov, Y.-Z. Qian, A. Coc, M. Casse, and E. Vangioni-Flam, Constraints on the variations of the fundamental couplings, *Phys. Rev. D* **66**, 045022 (2002).
- [62] J.-P. Uzan, Varying constants, gravitation and cosmology, *Living Rev. Relativity* **14**, 2 (2011).
- [63] J.-P. Uzan, Fundamental constants: From measurement to the universe, a window on gravitation and cosmology, *arXiv:2410.07281*.
- [64] A. Coc, N. J. Nunes, K. A. Olive, J.-P. Uzan, and E. Vangioni, Coupled variations of fundamental couplings and primordial nucleosynthesis, *Phys. Rev. D* **76**, 023511 (2007).
- [65] L. W. H. Fung, L. Li, T. Liu, H. N. Luu, Y.-C. Qiu, and S. H. H. Tye, Axi-Higgs cosmology, *J. Cosmol. Astropart. Phys.* **08** (2021) 057.
- [66] L. W. H. Fung, L. Li, T. Liu, H. N. Luu, Y.-C. Qiu, and S. H. H. Tye, Hubble constant in the axi-Higgs universe, *Phys. Rev. Res.* **5**, L022059 (2023).
- [67] H. N. Luu, Axion-Higgs cosmology: Cosmic microwave background and cosmological tensions, *Phys. Rev. D* **107**, 023513 (2023).
- [68] J. E. Kim, Weak interaction singlet and strong CP invariance, *Phys. Rev. Lett.* **43**, 103 (1979).
- [69] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Can confinement ensure natural CP invariance of strong interactions?, *Nucl. Phys.* **B166**, 493 (1980).
- [70] M. Dine, W. Fischler, and M. Srednicki, A simple solution to the strong CP problem with a harmless axion, *Phys. Lett.* **104B**, 199 (1981).
- [71] A. R. Zhitnitsky, On possible suppression of the axion hadron interactions (In russian), *Sov. J. Nucl. Phys.* **31**, 260 (1980).
- [72] R. L. Workman *et al.* (Particle Data Group), Review of particle physics, *Prog. Theor. Exp. Phys.* **2022**, 083C01 (2022).
- [73] S. M. Carroll, Quintessence and the rest of the world, *Phys. Rev. Lett.* **81**, 3067 (1998).
- [74] A. Lue, L.-M. Wang, and M. Kamionkowski, Cosmological signature of new parity violating interactions, *Phys. Rev. Lett.* **83**, 1506 (1999).
- [75] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, String axiverse, *Phys. Rev. D* **81**, 123530 (2010).
- [76] P. Agrawal, A. Hook, and J. Huang, A CMB millikan experiment with cosmic axiverse strings, *J. High Energy Phys.* **07** (2020) 138.
- [77] P. Diego-Palazuelos *et al.*, Cosmic birefringence from the Planck data release 4, *Phys. Rev. Lett.* **128**, 091302 (2022).
- [78] W. W. Yin, L. Dai, J. Huang, L. Ji, and S. Ferraro, New probe of cosmic birefringence using galaxy polarization and shapes, *Phys. Rev. Lett.* **134**, 161001 (2025).
- [79] R. M. Sullivan, A. Abghari, P. Diego-Palazuelos, L. T. Hergt, and D. Scott, Planck PR4 (NPIPE) map-space cosmic birefringence, *arXiv:2502.07654*.
- [80] V. Flambaum, S. Lambert, and M. Pospelov, Scalar-tensor theories with pseudoscalar couplings, *Phys. Rev. D* **80**, 105021 (2009).
- [81] A. Smith, M. Mylova, P. Brax, C. van de Bruck, C. P. Burgess, and A.-C. Davis, CMB implications of multi-field axio-dilaton cosmology, *J. Cosmol. Astropart. Phys.* **12** (2024) 058.
- [82] A. Smith, M. Mylova, P. Brax, C. van de Bruck, C. P. Burgess, and A.-C. Davis, A minimal axio-dilaton dark sector, *arXiv:2410.11099*.
- [83] D. Cyncynates and Z. J. Weiner, Experimental targets for dark photon dark matter, *arXiv:2410.14774*.
- [84] D. Brzeminski, Z. Chacko, A. Dev, and A. Hook, Time-varying fine structure constant from naturally ultralight dark matter, *Phys. Rev. D* **104**, 075019 (2021).
- [85] G. F. Giudice, E. W. Kolb, and A. Riotto, Largest temperature of the radiation era and its cosmological implications, *Phys. Rev. D* **64**, 023508 (2001).
- [86] M. Kawasaki, K. Kohri, and N. Sugiyama, MeV scale reheating temperature and thermalization of neutrino background, *Phys. Rev. D* **62**, 023506 (2000).
- [87] M. Kawasaki, K. Kohri, and T. Moroi, Big-Bang nucleosynthesis and hadronic decay of long-lived massive particles, *Phys. Rev. D* **71**, 083502 (2005).

- [88] S. Hannestad, What is the lowest possible reheating temperature?, *Phys. Rev. D* **70**, 043506 (2004).
- [89] K. Ichikawa, M. Kawasaki, and F. Takahashi, The oscillation effects on thermalization of the neutrinos in the universe with low reheating temperature, *Phys. Rev. D* **72**, 043522 (2005).
- [90] P. F. de Salas, M. Lattanzi, G. Mangano, G. Miele, S. Pastor, and O. Pisanti, Bounds on very low reheating scenarios after Planck, *Phys. Rev. D* **92**, 123534 (2015).
- [91] T. Hasegawa, N. Hiroshima, K. Kohri, R. S. L. Hansen, T. Tram, and S. Hannestad, MeV-scale reheating temperature and thermalization of oscillating neutrinos by radiative and hadronic decays of massive particles, *J. Cosmol. Astropart. Phys.* **12** (2019) 012.
- [92] A. Arvanitaki, S. Dimopoulos, V. Gorbenko, J. Huang, and K. Van Tilburg, A small weak scale from a small cosmological constant, *J. High Energy Phys.* **05** (2017) 071.
- [93] R. Hlozek, D. Grin, D. J. E. Marsh, and P. G. Ferreira, A search for ultralight axions using precision cosmological data, *Phys. Rev. D* **91**, 103512 (2015).
- [94] R. Hlozek, D. J. E. Marsh, D. Grin, R. Allison, J. Dunkley, and E. Calabrese, Future CMB tests of dark matter: Ultralight axions and massive neutrinos, *Phys. Rev. D* **95**, 123511 (2017).
- [95] R. Hlozek, D. J. E. Marsh, and D. Grin, Using the full power of the cosmic microwave background to probe axion dark matter, *Mon. Not. R. Astron. Soc.* **476**, 3063 (2018).
- [96] A. Laguë, J. R. Bond, R. Hlozek, K. K. Rogers, D. J. E. Marsh, and D. Grin, Constraining ultralight axions with galaxy surveys, *J. Cosmol. Astropart. Phys.* **01** (2022) 049.
- [97] K. K. Rogers, R. Hlozek, A. Laguë, M. M. Ivanov, O. H. E. Philcox, G. Cabass, K. Akitsu, and D. J. E. Marsh, Ultralight axions and the S_8 tension: Joint constraints from the cosmic microwave background and galaxy clustering, *J. Cosmol. Astropart. Phys.* **06** (2023) 023.
- [98] 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).
- [99] H. M. Tohfa, J. Crump, E. Baker, L. Hart, D. Grin, M. Brosius, and J. Chluba, Cosmic microwave background search for fine-structure constant evolution, *Phys. Rev. D* **109**, 103529 (2024).
- [100] J. Gasser and H. Leutwyler, Quark masses, *Phys. Rep.* **87**, 77 (1982).
- [101] L. Dolan and R. Jackiw, Symmetry behavior at finite temperature, *Phys. Rev. D* **9**, 3320 (1974).
- [102] S. Weinberg, Gauge and global symmetries at high temperature, *Phys. Rev. D* **9**, 3357 (1974).
- [103] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, England, 2014).
- [104] J. I. Kapusta and C. Gale, *Finite-Temperature Field Theory: Principles and Applications*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2011).
- [105] B. Batell and A. Ghalsasi, Thermal misalignment of scalar dark matter, *Phys. Rev. D* **107**, L091701 (2023).
- [106] A. L. Erickcek, N. Barnaby, C. Burrage, and Z. Huang, Chameleons in the early universe: Kicks, rebounds, and particle production, *Phys. Rev. D* **89**, 084074 (2014).
- [107] P. B. Arnold and C.-x. Zhai, The three loop free energy for high temperature QED and QCD with fermions, *Phys. Rev. D* **51**, 1906 (1995).
- [108] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
- [109] C. Grayson, C. T. Yang, M. Formanek, and J. Rafelski, Electron-positron plasma in BBN: Damped-dynamic screening, *Ann. Phys. (Amsterdam)* **458**, 169453 (2023).
- [110] L. C. Thomas, T. Dezen, E. B. Grohs, and C. T. Kishimoto, Electron-positron annihilation freeze-out in the early universe, *Phys. Rev. D* **101**, 063507 (2020).
- [111] K. Thorne and R. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics* (Princeton University Press, Princeton, NJ, 2017).
- [112] P. J. McMillan, Mass models of the milky way, *Mon. Not. R. Astron. Soc.* **414**, 2446 (2011).
- [113] D. J. E. Marsh, Axion cosmology, *Phys. Rep.* **643**, 1 (2016).
- [114] G. P. Centers *et al.*, Stochastic fluctuations of bosonic dark matter, *Nat. Commun.* **12**, 7321 (2021).
- [115] Y. Su, B. R. Heckel, E. G. Adelberger, J. H. Gundlach, M. Harris, G. L. Smith, and H. E. Swanson, New tests of the universality of free fall, *Phys. Rev. D* **50**, 3614 (1994).
- [116] T. Damour and D. Vokrouhlicky, The equivalence principle and the moon, *Phys. Rev. D* **53**, 4177 (1996).
- [117] J. W. Morgan and E. Anders, Chemical composition of earth, venus, and mercury, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6973 (1980).
- [118] W. Ubachs, J. Bagdonaite, E. J. Salumbides, M. T. Murphy, and L. Kaper, Search for a drifting proton-electron mass ratio from H_2 , *Rev. Mod. Phys.* **88**, 021003 (2016).
- [119] M. Savedoff, Physical constants in extra-galactic nebulae, *Nature (London)* **178**, 688 (1956).
- [120] P. Petitjean, R. Srianand, H. Chand, A. Ivanchik, P. Noterdaeme, and N. Gupta, Constraining fundamental constants of physics with quasar absorption line systems, *Space Sci. Rev.* **148**, 289 (2009).
- [121] R. Srianand, P. Petitjean, H. Chand, P. Noterdaeme, and N. Gupta, Probing the variation of fundamental constants using QSO absorption lines, *Mem. Soc. Astron. Ital.* **80**, 842 (2009).
- [122] N. Kanekar, Probing fundamental constant evolution with redshifted radio lines, *Mem. Soc. Astron. Ital.* **80**, 895 (2009).
- [123] J. B. Whitmore and M. T. Murphy, Impact of instrumental systematic errors on fine-structure constant measurements with quasar spectra, *Mon. Not. R. Astron. Soc.* **447**, 446 (2015).
- [124] M. T. Murphy, V. V. Flambaum, S. Muller, and C. Henkel, Strong limit on a variable Proton-to-electron mass ratio from molecules in the distant universe, *Science* **320**, 1611 (2008).
- [125] J. Bagdonaite, M. Daprà, P. Jansen, H. L. Bethlem, W. Ubachs, S. Muller, C. Henkel, and K. M. Menten, Robust constraint on a drifting proton-to-electron mass ratio at

- $z = 0.89$ from methanol observation at three radio telescopes, *Phys. Rev. Lett.* **111**, 231101 (2013).
- [126] N. Kanekar, G. I. Langston, J. T. Stocke, C. L. Carilli, and K. M. Menten, Constraining fundamental constant evolution with HI and OH lines, *Astrophys. J. Lett.* **746**, L16 (2012).
- [127] N. Kanekar, T. Ghosh, and J. N. Chengalur, Stringent constraints on fundamental constant evolution using conjugate 18 cm satellite OH lines, *Phys. Rev. Lett.* **120**, 061302 (2018).
- [128] M. T. Murphy, A. L. Malec, and J. X. Prochaska, Precise limits on cosmological variability of the fine-structure constant with zinc and chromium quasar absorption lines, *Mon. Not. R. Astron. Soc.* **461**, 2461 (2016).
- [129] 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).
- [130] T. M. Evans *et al.*, The UVES large program for testing fundamental physics—III. Constraints on the fine-structure constant from three telescopes, *Mon. Not. R. Astron. Soc.* **445**, 128 (2014).
- [131] M. R. Wilczynska *et al.*, Four direct measurements of the fine-structure constant 13 billion years ago, *Sci. Adv.* **6**, eaay9672 (2020).
- [132] T. A. Wagner, S. Schlamminger, J. H. Gundlach, and E. G. Adelberger, Torsion-balance tests of the weak equivalence principle, *Classical Quantum Gravity* **29**, 184002 (2012).
- [133] R. Oswald *et al.*, Search for dark-matter-induced oscillations of fundamental constants using molecular spectroscopy, *Phys. Rev. Lett.* **129**, 031302 (2022).
- [134] F. Ferrer and J. A. Grifols, Long range forces from pseudoscalar exchange, *Phys. Rev. D* **58**, 096006 (1998).
- [135] F. Ferrer and J. A. Grifols, Effects of Bose-Einstein condensation on forces among bodies sitting in a boson heat bath, *Phys. Rev. D* **63**, 025020 (2001).
- [136] M. Bauer and G. Rostagni, Fifth forces from QCD axions scale differently, *Phys. Rev. Lett.* **132**, 101802 (2024).
- [137] S. Schlamminger, K. Y. Choi, T. A. Wagner, J. H. Gundlach, and E. G. Adelberger, Test of the equivalence principle using a rotating torsion balance, *Phys. Rev. Lett.* **100**, 041101 (2008).
- [138] P. Touboul *et al.*, Result of the MICROSCOPE weak equivalence principle test, *Classical Quantum Gravity* **39**, 204009 (2022).
- [139] R. Naudet, *Oklo, des réacteurs nucléaires fossiles: étude physique* (Eyrolles, Paris, France, 1991).
- [140] A. Shlyakhter, Direct test of the constancy of the fundamental nuclear constants using the Oklo natural reactor, Preprint of LNPI-260 (1976).
- [141] A. I. Shlyakhter, Direct test of the constancy of fundamental nuclear constants, *Nature (London)* **264**, 340 (1976).
- [142] K. Van Tilburg, N. Leefer, L. Bougas, and D. Budker, Search for ultralight scalar dark matter with atomic spectroscopy, *Phys. Rev. Lett.* **115**, 011802 (2015).
- [143] A. Hees, J. Guéna, M. Abgrall, S. Bize, and P. Wolf, Searching for an oscillating massive scalar field as a dark matter candidate using atomic hyperfine frequency comparisons, *Phys. Rev. Lett.* **117**, 061301 (2016).
- [144] C. J. Kennedy, E. Oelker, J. M. Robinson, T. Bothwell, D. Kedar, W. R. Milner, G. E. Marti, A. Derevianko, and J. Ye, Precision metrology meets cosmology: Improved constraints on ultralight dark matter from atom-cavity frequency comparisons, *Phys. Rev. Lett.* **125**, 201302 (2020).
- [145] K. Beloy *et al.* (BACON Collaboration), Frequency ratio measurements at 18-digit accuracy using an optical clock network, *Nature (London)* **591**, 564 (2021).
- [146] I. Kozyryev, Z. Lasner, and J. M. Doyle, Enhanced sensitivity to ultralight bosonic dark matter in the spectra of the linear radical SrOH, *Phys. Rev. A* **103**, 043313 (2021).
- [147] T. Kobayashi *et al.*, Search for ultralight dark matter from long-term frequency comparisons of optical and microwave atomic clocks, *Phys. Rev. Lett.* **129**, 241301 (2022).
- [148] 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).
- [149] T. M. Fortier *et al.*, Precision atomic spectroscopy for improved limits on variation of the fine structure constant and local position invariance, *Phys. Rev. Lett.* **98**, 070801 (2007).
- [150] T. Rosenband *et al.*, Frequency ratio of Al⁺ and Hg⁺ single-ion optical clocks; Metrology at the 17th decimal place, *Science* **319**, 1154622 (2008).
- [151] J. Guena, M. Abgrall, D. Rovera, P. Rosenbusch, M. E. Tobar, P. Laurent, A. Clairon, and S. Bize, Improved tests of local position invariance using Rb-87 and Cs-133 fountains, *Phys. Rev. Lett.* **109**, 080801 (2012).
- [152] N. Leefer, C. T. M. Weber, A. Cingöz, J. R. Torgerson, and D. Budker, New limits on variation of the fine-structure constant using atomic dysprosium, *Phys. Rev. Lett.* **111**, 060801 (2013).
- [153] M. E. Tobar *et al.*, Testing local position and fundamental constant invariance due to periodic gravitational and boost using long-term comparison of the SYRTE atomic fountains and H-masers, *Phys. Rev. D* **87**, 122004 (2013).
- [154] N. Huntemann, B. Lipphardt, C. Tamm, V. Gerginov, S. Weyers, and E. Peik, Improved limit on a temporal variation of m_p/m_e from comparisons of Yb⁺ and Cs atomic clocks, *Phys. Rev. Lett.* **113**, 210802 (2014).
- [155] R. M. Godun, P. B. R. Nisbet-Jones, J. M. Jones, S. A. King, L. A. M. Johnson, H. S. Margolis, K. Szymaniec, S. N. Lea, K. Bongs, and P. Gill, Frequency ratio of two optical clock transitions in Yb + 171 and constraints on the time variation of fundamental constants, *Phys. Rev. Lett.* **113**, 210801 (2014).
- [156] M. S. Safronova, D. Budker, D. DeMille, D. F. J. Kimball, A. Derevianko, and C. W. Clark, Search for new physics with atoms and molecules, *Rev. Mod. Phys.* **90**, 025008 (2018).
- [157] D. E. Kaplan, A. Mitridate, and T. Trickle, Constraining fundamental constant variations from ultralight dark matter with pulsar timing arrays, *Phys. Rev. D* **106**, 035032 (2022).

- [158] A. Afzal *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Search for signals from new physics, *Astrophys. J. Lett.* **951**, L11 (2023); **971**, L27(E) (2024).
- [159] M. J. Dolan, F. J. Hiskens, and R. R. Volkas, Advancing globular cluster constraints on the axion-photon coupling, *J. Cosmol. Astropart. Phys.* **10** (2022) 096.
- [160] E. Hardy and R. Lasenby, Stellar cooling bounds on new light particles: Plasma mixing effects, *J. High Energy Phys.* **02** (2017) 033.
- [161] S. Bottaro, A. Caputo, G. Raffelt, and E. Vitagliano, Stellar limits on scalars from electron-nucleus bremsstrahlung, *J. Cosmol. Astropart. Phys.* **07** (2023) 071.
- [162] T. Dent, S. Stern, and C. Wetterich, Primordial nucleosynthesis as a probe of fundamental physics parameters, *Phys. Rev. D* **76**, 063513 (2007).
- [163] 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).
- [164] V. F. Mukhanov, Nucleosynthesis without a computer, *Int. J. Theor. Phys.* **43**, 669 (2004).
- [165] A. Czarnecki, W. J. Marciano, and A. Sirlin, Neutron lifetime and axial coupling connection, *Phys. Rev. Lett.* **120**, 202002 (2018).
- [166] A. Czarnecki, W. J. Marciano, and A. Sirlin, Precision measurements and CKM unitarity, *Phys. Rev. D* **70**, 093006 (2004).
- [167] D. H. Wilkinson, Phase space for neutron beta-decay: An update, *Nucl. Instrum. Methods Phys. Res., Sect. A* **404**, 305 (1998).
- [168] S. Borsanyi *et al.* (BMW Collaboration), *Ab initio* calculation of the neutron-proton mass difference, *Science* **347**, 1452 (2015).
- [169] A. W. Thomas, X. G. Wang, and R. D. Young, Electromagnetic contribution to the proton-neutron mass splitting, *Phys. Rev. C* **91**, 015209 (2015).
- [170] J. Gasser and H. Leutwyler, Implications of scaling for the proton—neutron mass—difference, *Nucl. Phys.* **B94**, 269 (1975).
- [171] J. Gasser, H. Leutwyler, and A. Rusetsky, Sum rule for the Compton amplitude and implications for the proton—neutron mass difference, *Eur. Phys. J. C* **80**, 1121 (2020).
- [172] J. Gasser, H. Leutwyler, and A. Rusetsky, On the mass difference between proton and neutron, *Phys. Lett. B* **814**, 136087 (2021).
- [173] A. Walker-Loud, On the cottingham formula and the electromagnetic contribution to the proton-neutron mass splitting, *Proc. Sci.*, CD2018 (2019) 045 [arXiv:1907.05459].
- [174] K. Saikawa and S. Shirai, Primordial gravitational waves, precisely: The role of thermodynamics in the standard model, *J. Cosmol. Astropart. Phys.* **05** (2018) 035.
- [175] R. J. Cooke, M. Pettini, and C. C. Steidel, One percent determination of the primordial deuterium abundance, *Astrophys. J.* **855**, 102 (2018).
- [176] P. A. Kislitsyn, S. A. Balashev, M. T. Murphy, C. Ledoux, P. Noterdaeme, and A. V. Ivanchik, A new precise determination of the primordial abundance of deuterium: Measurement in the metal-poor sub-DLA system at $z = 3.42$ towards quasar J1332 + 0052, *Mon. Not. R. Astron. Soc.* **528**, 4068 (2024).
- [177] V. Mossa *et al.*, The baryon density of the universe from an improved rate of deuterium burning, *Nature (London)* **587**, 210 (2020).
- [178] T.-H. Yeh, K. A. Olive, and B. D. Fields, The impact of new $d(p, \gamma)^3$ rates on big bang nucleosynthesis, *J. Cosmol. Astropart. Phys.* **03** (2021) 046.
- [179] T. Sekiguchi and T. Takahashi, Early recombination as a solution to the H_0 tension, *Phys. Rev. D* **103**, 083507 (2021).
- [180] L. Hart and J. Chluba, Updated fundamental constant constraints from Planck 2018 data and possible relations to the Hubble tension, *Mon. Not. R. Astron. Soc.* **493**, 3255 (2020).
- [181] A. G. Adame *et al.* (DESI Collaboration), DESI 2024 VI: Cosmological constraints from the measurements of baryon acoustic oscillations, *J. Cosmol. Astropart. Phys.* **02** (2025) 021.
- [182] M. Loverde and Z. J. Weiner, Massive neutrinos and cosmic composition, *J. Cosmol. Astropart. Phys.* **12** (2024) 048.
- [183] T. M. C. Abbott *et al.* (DES Collaboration), The dark energy survey: Cosmology results with ~ 1500 new high-redshift type Ia supernovae using the full 5 yr data set, *Astrophys. J. Lett.* **973**, L14 (2024).
- [184] J. E. Bautista *et al.* (eBOSS Collaboration), The completed SDSS-IV extended Baryon Oscillation Spectroscopic survey: Measurement of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic correlation function between redshifts 0.6 and 1, *Mon. Not. R. Astron. Soc.* **500**, 736 (2020).
- [185] H. Gil-Marín *et al.* (eBOSS Collaboration), The completed SDSS-IV extended Baryon Oscillation Spectroscopic survey: Measurement of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic power spectrum between redshifts 0.6 and 1.0, *Mon. Not. R. Astron. Soc.* **498**, 2492 (2020).
- [186] 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).
- [187] D. Brout *et al.*, The Pantheon+ Analysis: Cosmological constraints, *Astrophys. J.* **938**, 110 (2022).
- [188] D. Scolnic *et al.*, The Pantheon+ Analysis: The full data set and light-curve release, *Astrophys. J.* **938**, 113 (2022).
- [189] D. Rubin *et al.*, Union through UNITY: Cosmology with 2,000 SNe using a unified Bayesian framework, *arXiv*: 2311.12098.
- [190] D. M. Scolnic *et al.* (Pan-STARRS1 Collaboration), 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).
- [191] L. Vacher and N. Schöneberg, Incompatibility of fine-structure constant variations at recombination with local observations, *Phys. Rev. D* **109**, 103520 (2024).
- [192] 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).

- [193] 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).
- [194] T. Karwal, M. Raveri, B. Jain, J. Khoury, and M. Trodden, Chameleon early dark energy and the Hubble tension, *Phys. Rev. D* **105**, 063535 (2022).
- [195] M.-X. Lin, E. McDonough, J. C. Hill, and W. Hu, Dark matter trigger for early dark energy coincidence, *Phys. Rev. D* **107**, 103523 (2023).
- [196] J. Sakstein and M. Trodden, Early dark energy from massive neutrinos as a natural resolution of the Hubble tension, *Phys. Rev. Lett.* **124**, 161301 (2020).
- [197] M. Carrillo González, Q. Liang, J. Sakstein, and M. Trodden, Neutrino-assisted early dark energy: theory and cosmology, *J. Cosmol. Astropart. Phys.* **04** (2021) 063.
- [198] M. Kamionkowski and A. Mathur, Thermo-coupled early dark energy, *Phys. Rev. D* **111**, 063551 (2025).
- [199] D. Antypas *et al.*, New horizons: Scalar and vector ultralight dark matter, [arXiv:2203.14915](https://arxiv.org/abs/2203.14915).
- [200] V. V. Flambaum, Enhanced effect of temporal variation of the fine structure constant and the strong interaction in Th-229, *Phys. Rev. Lett.* **97**, 092502 (2006).
- [201] C. J. Campbell, A. G. Radnaev, A. Kuzmich, V. A. Dzuba, V. V. Flambaum, and A. Derevianko, A Single-Ion nuclear clock for metrology at the 19th decimal place, *Phys. Rev. Lett.* **108**, 120802 (2012).
- [202] E. Peik, T. Schumm, M. S. Safronova, A. Pálffy, J. Weitenberg, and P. G. Thirolf, Nuclear clocks for testing fundamental physics, *Quantum Sci. Technol.* **6**, 034002 (2021).
- [203] D. Brzemiński, Z. Chacko, A. Dev, I. Flood, and A. Hook, Searching for a fifth force with atomic and nuclear clocks, *Phys. Rev. D* **106**, 095031 (2022).
- [204] K. N. Abazajian *et al.* (CMB-S4 Collaboration), CMB-S4 science book, first edition, [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [205] P. Ade *et al.* (Simons Observatory Collaboration), The Simons Observatory: Science goals and forecasts, *J. Cosmol. Astropart. Phys.* **02** (2019) 056.
- [206] 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).
- [207] L. Hamaide, H. Müller, and D. J. E. Marsh, Searching for dilaton fields in the Lyman- α forest, *Phys. Rev. D* **106**, 123509 (2022).
- [208] C. J. A. P. Martins *et al.* (ANDES Collaboration), Cosmology and fundamental physics with the ELT-ANDES spectrograph, *Exper. Astron.* **57**, 5 (2024).
- [209] L. Jiang *et al.* (DESI Collaboration), Constraints on the spacetime variation of the fine-structure constant using DESI emission-line galaxies, *Astrophys. J.* **968**, 120 (2024).
- [210] A. Aghamousa *et al.* (DESI Collaboration), The DESI experiment part I: Science, targeting, and survey design, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036).
- [211] D. J. Schlegel *et al.* (DESI Collaboration), A spectroscopic road map for cosmic frontier: DESI, DESI-II, stage-5, [arXiv:2209.03585](https://arxiv.org/abs/2209.03585).
- [212] Y. Mellier *et al.* (Euclid Collaboration), Euclid. I. Overview of the Euclid mission, *Astron. Astrophys.* **697**, A1 (2025).
- [213] M. Archidiacono *et al.* (Euclid Collaboration), Euclid preparation. LIV. Sensitivity to neutrino parameters, *Astron. Astrophys.* **693**, A58 (2025).
- [214] T. Eifler *et al.*, Cosmology with the Roman Space Telescope—multiprobe strategies, *Mon. Not. R. Astron. Soc.* **507**, 1746 (2021).
- [215] M. Rigault *et al.*, ZTF SN Ia DR2: Overview, *Astron. Astrophys.* **694**, A1 (2025).
- [216] R. Mandelbaum *et al.* (LSST Dark Energy Science Collaboration), The LSST Dark Energy Science Collaboration (DESC) science requirements document, [arXiv:1809.01669](https://arxiv.org/abs/1809.01669).
- [217] D. O. Jones *et al.* (Young Supernova Experiment Collaboration), The Young Supernova Experiment: Survey goals, overview, and operations, *Astrophys. J.* **908**, 143 (2021).
- [218] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [219] D. W. Hogg and D. Foreman-Mackey, Data analysis recipes: Using Markov chain Monte Carlo, *Astrophys. J. Suppl. Ser.* **236**, 11 (2018).
- [220] D. Foreman-Mackey, W. Farr, M. Sinha, A. Archibald, D. Hogg, J. Sanders, J. Zuntz, P. Williams, A. Nelson, M. de Val-Borro, T. Erhardt, I. Pashchenko, and O. Pla, Emcee v3: A PYTHON ensemble sampling toolkit for affine-invariant MCMC, *J. Open Source Software* **4**, 1864 (2019).
- [221] D. Foreman-Mackey, corner.py: Scatterplot matrices in PYTHON, *J. Open Source Software* **1**, 24 (2016).
- [222] C. R. Harris *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [223] P. Virtanen *et al.*, SCIPY 1.0—Fundamental algorithms for scientific computing in PYTHON, *Nat. Methods* **17**, 261 (2020).
- [224] J. D. Hunter, MATPLOTLIB: A 2D graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [225] S. Hoyer and J. Hamman, xarray: N-D labeled arrays and datasets in PYTHON, *J. Open Res. Software* **5**, 10 (2017).
- [226] R. Kumar, C. Carroll, A. Hartikainen, and O. Martin, Arviz a unified library for exploratory analysis of Bayesian models in PYTHON, *J. Open Source Software* **4**, 1143 (2019).
- [227] A. Meurer *et al.*, SYMPY: Symbolic computing in PYTHON, *PeerJ Comput. Sci.* **3**, e103 (2017).
- [228] E. van der Velden, CMasher: Scientific colormaps for making accessible, informative and ‘cmashing’ plots, *J. Open Source Software* **5**, 2004 (2020).
- [229] A. A. Starobinsky and J. Yokoyama, Equilibrium state of a selfinteracting scalar field in the de Sitter background, *Phys. Rev. D* **50**, 6357 (1994).
- [230] 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).