Cosmic microwave background search for fine-structure constant evolution

Hurum Maksora Tohfa[1,2,3,*] Jack Crump,[4] Ethan Baker[4] Luke Hart,[5]
Daniel Grin,[4,†] Madeline Brosius[2] and Jens Chluba[5]

[1]*Department of Astronomy, University of Washington, Seattle, Washington 98195, United States*
[2]*Department of Physics, Bryn Mawr College, 101N Merion Avenue, Bryn Mawr, Pennsylvania 19010, United States*
[3]*Department of Physics and Astronomy, University of California,*
Riverside, Riverside, California 92521, United States
[4]*Department of Physics and Astronomy, Haverford College,*
370 Lancaster Avenue, Haverford, Pennsylvania 19041, United States
[5]*Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester,*
Manchester M13 9PL, United Kingdom

In some extensions of the standard model of particle physics, the values of the fundamental coupling constants vary in space and time. Some observations of quasars hint at time and spatial variation of the fine structure constant α. Here, the Bekenstein-Sandvik-Barrow-Magueijo (BSBM) model (which posits the existence of a scalar field driving evolution in the fundamental electric charge e) is tested against quasar and *Planck* satellite cosmic microwave background (CMB) data. In this model, variations in e are coupled to the matter density through a factor $\zeta_{\mathrm{m}} M_{\mathrm{pl}}^2 / \bar{\omega}$, where ζ_m is the fractional electromagnetic contribution to nucleon rest masses, $\bar{\omega}$ is the energy-scale squared of new (BSBM) physics, and M_{pl} is the Planck energy scale. Simulations conducted here do not support claims that the electrostatic contribution to ζ_m is completely shielded. Other common approximations used in BSBM modeling are found to be adequate. Principal components of the CMB data with respect to variations in α are used to obtain constraints of $\zeta_{\mathrm{m}} M_{\mathrm{pl}}^2 / \bar{\omega} \lesssim 9.3 \times 10^{-9}$ for a massless field. A forecast anticipating the promise of the Simons Observatory CMB experiment shows that SO will be sensitive to values of $\zeta_{\mathrm{m}} M_{\mathrm{pl}}^2 / \omega \geq 2.2 \times 10^{-9}$, significantly improving on the uncertainty in e-variation from quasar spectra alone.

I. INTRODUCTION

Thanks to measurements of cosmic microwave background (CMB) anisotropies by the *Planck* satellite and ground-based efforts like the South-Pole Telescope (SPT) and Atacama Cosmology Telescope (ACT), cosmic acceleration as probed by type Ia supernovae, and clustering/lensing of galaxies, there is a standard cosmological model. This concordance model has relic-density parameters of $\Omega_b h^2 = 0.0224 \pm 0.0001$, $\Omega_c h^2 = 0.1200 \pm 0.0012$, and $\Omega_{\mathrm{DE}} = 0.68.547 \pm 0.0073$ (where b denotes baryons, c denotes cold dark matter, or CDM, and DE represents dark energy) [1–4].

Using *Planck* data and future results from nearly cosmic-variance-limited CMB polarization experiments (e.g. CMB-S4 and the Simons Observatory or SO), the CMB can also be used to test models of dark-sector contents and interactions. Existing data could even provide evidence for an early dark-energy component [5–10], which could reconcile tension between late-time supernovae and CMB inferences of the Hubble constant [11,12]. The evolution of cosmological perturbations probed by the CMB is influenced by the photon diffusion length (see Refs. [13,14] and citations therein), the redshift of the last scattering surface, and the detailed physics of recombination [15–18]. The CMB is thus also sensitive to the possibility that the fundamental "constants" are in fact dynamical rather than constant.

The possibility that fundamental parameters like e, m_e, m_p, G, $\hbar$, and even the speed-of-light c vary in time (or space) was raised by Dirac and others, who noticed that certain combinations of these parameters with units of time were numerically comparable to the age of the Universe [19–22]. They posited that these ratios were equal to the age of the Universe at *all* times, implying a specific time evolution for the fundamental parameters. Although the simplest realizations of this idea are readily ruled out on anthropic grounds, more general scalar-tensor theories of gravitation [23] and electromagnetism [24] were then developed, as well as some UV completions of

*htohfa@uw.edu
†dgrin@haverford.edu

103529-1

the standard model that predicted evolution of fundamental constants [25–28]

The simplest extension of Maxwell electromagnetism that supports variation in e [and thus the fine-structure constant $\alpha = e^2/(\hbar c)$] while recovering the predictions of standard electromagnetism was put forward by Bekenstein [24]. This theory relies on a new scalar ψ, which modulates the Maxwell Lagrangian for electrodynamics and has a Brans-Dicke kinetic term with coupling ω. This model was extended to include the gravitational interactions of ψ in Refs. [29,30]. Subsequent work (see Ref. [31] for a review) generalized the scalar field to a massive one with non-trivial field dependence in ω [32,33], included spatial perturbations [34,35], and developed more extensive modeling of the theory's dynamics [30,36–38].

Broadly, varying fundamental parameters occur in standard-model extensions with extra dimensions (beginning with Kaluza-Klein theory [39] and continuing with string-inspired ideas like the runaway dilaton [40–48]). Other cases include disformal theories, in which radiation and matter geodesics are given by different metrics [49,50]. These scalars could constitute DE [51–54] or DM [55–58] with novel astronomical and lab phenomenology resulting, such as evolution in the cosmic equation of state, laser interferometry signals, and violations of the weak equivalence principle [31,59,60].

There have been hints of variation in α from high-resolution spectra of redshift $z \sim 3$–7 quasistellar objects (QSOs) [61–66], with disputes about the interpretation of these results [67–72]. Current/future observations promise unprecedented sensitivity to α variations with the potential to resolve these disputes [73,74] (e.g. results from the ESPRESSO spectrograph, which impose the constraint $\Delta\alpha/\alpha \sim 10^{-6}$ [75]).

Any theory of varying α will also make a prediction at $z \simeq 1100$, the epoch of CMB decoupling. The Thomson scattering rate scales as α^2, while the hydrogen $2s \to 1s$ transition rate (the bottleneck for recombination) scales as $\propto \alpha^8$. The redshift and width of the last-scattering surface are influenced by model parameters [13,76–78]. CMB temperature/polarization measurements can be used to probe variations in α, as shown by the BOOMERanG/WMAP/*Planck*/SPT/ACT upper limits of Refs. [79–87] and forecasts of Refs. [88–90]. Correlations with other fundamental constants (e.g., G) were considered in Refs. [83–87]. All these analyses assumed a single non-standard α value at early times. Evolution of G was constrained in Refs. [91–95].

Simply parametrized time evolution in α and connections to the Hubble tension were explored in Refs. [96,97]. Given the range of theoretical possibilities and observational controversy, we use principal component analysis (PCA), a model-independent technique in which the eigenvectors of the information matrix are found. PCA may be used to test data for novel physics, even without a compelling

model, and has been applied to explore dark matter annihilation, nonstandard recombination, and late-time cosmic acceleration [98–105].

Models with large projections onto these eigenvectors will be best constrained by the data. Any model can be constrained through projection onto principal components (PCs), without the need to rerun a full Monte Carlo Markov chain (MCMC). PCs can elucidate the epochs driving the constraint. Techniques for probing cosmic recombination with PCA were developed and applied in Refs. [106,107]. In Ref. [108], PCA was applied to to obtain the constraint of $\Delta\alpha/\alpha = 0.0010 \pm 0.0024$.

Causality requires that α variation occur in space if it occurs in time. The resulting non-Gaussian statistics [109] were used to impose $\sim 10^{-2}$ level constraints to spatial α variation [110,111]. Secondary CMB sight lines to galaxy clusters constrain variation in α [112] because of the Sunyaez-Zel'dovich (SZ) effect [113,114], although these inferences remain challenging due to relativistic corrections [115]. Large volumes will be mapped by upcoming observations of neutral hydrogen at high redshift [116], probing varying-α theories [117,118].

Here, we model the evolution of ψ and α dynamically in the Bekenstein-Sandvik-Barrow-Magueijo (BSBM) model. We allow for ψ to have a mass m, motivated by recent work finding that light scalars and pseudoscalars are numerous in string-theory realizations and possibly cosmologically important [119–124].

We use the PCA decomposition for α variation and constraints from *Planck* data in Ref. [108] to test the BSBM model, determining the allowed values for the coupling ζ_m/ω (where ζ_m is the ratio of nuclear electromagnetic to total rest-mass energy) and m. Our constraint (at 95% C.L.) is $\zeta_m/\omega \leq 9.3 \times 10^{-9}$ for $m = 0$, with constraints relaxing for $mc^2 \gtrsim 1.4 \times 10^{-32}$ eV.

We thus obtain some of the first constraints to the BSBM model that include the full time dependence of the model, going back to the recombination era.[1] The next decade of CMB measurements will bring an order of magnitude improvement in precision, and the possibility of testing many novel physical scenarios for the dark sector, neutrinos [127,128], and varying α. We conduct a forecast of the sensitivity of the ground-based SO [128] to test the BSBM model. We obtain a $\zeta_m/\omega \leq 2.2 \times 10^{-9}$ sensitivity forecast. Results for $m \neq 0$ are presented in the body of the paper.

In the BSBM model, ψ couples to the trace of the radiation stress-energy tensor ($\propto E^2 - B^2$), where E and B

---

[1]In the preparation of this work, we became aware of other recent work that applies CMB data to the BSBM model [125,126]. These results are complementary, as we allow for the possibility that ψ is massive, use PCA methods, and obtain forecasts for future CMB experiments. We also obtain QSO prior-free results in our CMB-only analysis.

are electric and magnetic field amplitudes. This trace is nontrivial to calculate [129–135]. Observational probes of BSBM typically assume $E^2 - B^2 \propto \rho_m$, the total cosmological matter density, with a proportionality constant ζ_m. Estimates of ζ_m vary in sign and in magnitude by 2–3 decades [129–135], depending on whether it is dominated by electric or magnetic contributions [135], classical vs quantum approaches, and our ignorance of the dark-sector contribution [129–135].

We ran a simulation to test recent claims [135] that magnetic contributions to ζ_m dominate over electrostatic terms. Our results indicate that they do not. We then put constraints on ζ_m/ω. Most comparisons of the BSBM model to data rely on a non-energy-conserving approximation in the equations of motion. We find that our constraints are insensitive to this approximation.

We begin with a discussion of scalar field dynamics and numerical methods for their evolution in Sec. II. We review principal component methods and explain our use of them in Sec. III. We summarize our MCMC modeling techniques and results in Sec. IV, and conclude in Sec. V. In Appendix A, we test the cancellation of electrostatic contributions to ζ_m. In Appendix B, we explore the impact of an energy nonconserving approximation made throughout the literature to solve the BSBM equations of motion, and find it to be negligible.

II. BSBM THEORY

In Bekenstein's original formulation of this α variation theory [24], the scalar field ψ was coupled to the standard Maxwell Lagrangian, but the coupling to gravity was not included. Later work by Sandvik, Barrow, and Magueijo included this inevitable coupling [29]. In the resulting BSBM theory [29,135], the electric charge e evolves, while Planck's constant h and the speed of light c are constant. A real scalar field modulates the Maxwell Lagrangian, causing variations in α: $e_0 \to e = e_0 \epsilon(x^\mu)$. Here ϵ is a dimensionless scalar field. Then ϵ couples to the electromagnetic gauge field A_μ in the Lagrangian. Under the usual local gauge transformation $U(x) = e^{i\theta(x)}$, the electromagnetic action is still invariant with a modified gauge-field transformation, $\epsilon A_\mu \to \epsilon A_\mu + \partial_\mu \theta(x)$.

Recasting the equations in terms of a more standard scalar field ψ defined by $\psi \equiv \ln(\epsilon)$, the equation of motion for the scalar field is [135]

$$\Box \psi = \frac{2}{\bar{\omega}} e^{-2\psi} \mathcal{L}_{em}, \qquad (1)$$

where $\bar{\omega} = \hbar c/l^2$ is a coupling constant and l is some new length scale below which Coulomb's law breaks down. In Planck units, $\bar{\omega}$ has units [energy]^2, and so we employ the reparametrization $\bar{\omega} = M_{pl}^2 \omega$ in terms of the reduced Planck

mass $M_{pl} = 1/\sqrt{8\pi G}$ and a dimensionless parameter ω that quantifies the amplitude of evolution in α.

The standard Maxwell Lagrangian, $\mathcal{L}_{em} = F_{\mu\nu} F^{\mu\nu}/4$, vanishes for pure radiation (as $\mathcal{L}_{em} \propto E^2 - B^2 = 0$, where E and B are the electric and magnetic-field vector amplitudes, respectively). It would, however, be excited by plasma-sourced electrostatic and magnetic fields **E** and **B** in the early Universe. The approximation $\mathcal{L} \simeq \zeta_m \rho_m$ is used throughout the literature, where ζ_m is a dimensionless constant quantifying the electromagnetic self-energy of nonrelativistic matter. This replacement warrants justification and requires a numerical estimate of ζ_m.

A. Values of ζ_m

We summarize past results and quantitatively assess claims [135] that prior calculations of ζ_m were in severe error, using a numerical plasma simulation in Appendix A. The simplest possibility is that the dark sector does not source scalar-field evolution. Straightforward estimates of the baryonic contribution to ζ_m may be obtained by extrapolating a semi-empirical mass formula for nuclear electromagnetic self-energy to the case of a primordial (hydrogen + helium) plasma [136,137]:

$$E_{Coulomb} = -a_C \frac{Z^2}{A^{1/3}}. \qquad (2)$$

In this equation, a_C is an empirically determined constant, Z is the atomic number of the nucleus, and A is the mass number of the nucleus. This term represents the decrease in nuclear binding energy that is caused by the electrostatic repulsion between the positively charged protons in the nucleus.

Using the fact that $a_C \approx 0.7$ MeV [24], the estimate $\zeta_m \sim 1.3 \times 10^{-2}$ was obtained through classical approximations to the nucleus. Late-time nucleosynthesis could lead to additional time dependence [24,130]. These estimates are dominated by E^2 (rather than B^2). Subsequent estimates modeled quark electromagnetic fields within the nucleus [29], applying the Born term in the Cottingham formula and following the methods of Ref. [129] to obtain $\zeta_m \approx 10^{-4}$.

As ρ_m is dominated by dark matter (whose nature is unknown), ζ_m could be dominated by nonbaryonic contributions [29]. One can then have $-1 < \zeta_m < 1$ [with the bounds saturated for dark matter composed of superconducting strings]. A negative value or ζ_m is appealing because it can explain QSO results hinting at a decrease of α with time.

Subsequently [135], the author of Ref. [24] solved the modified Maxwell equations of BSBM theory and the static limit of Eq. (1), claiming that (under some approximations) the ψ field configuration shields out the electrostatic contribution to Eq. (1), leaving only the B^2 term.

This cancellation is also claimed to allow the BSBM theory to evade terrestrial, Solar System constraints to weak-equivalence principle (WEP) violation [135], although the magnetic term might also lead to detectable WEP violation [131]. More generally, WEP bounds may be satisfied for $|\zeta_m| > 10^{-3}$ for the ζ_m/ω values that saturate the CMB constraints of Sec. IV, motivating us to continue testing this theory empirically.

If this cancellation occurs, then the magnetic contributions from baryons dominate, and one can have negative ζ_m without exotic dark sector contributions of the type discussed in Refs. [29,55]. Reference [135] provides estimates of ζ_m, integrating the classical solution for $B^2/8\pi$ outside of the Compton radius of a proton, weighting quantities by the abundances of hydrogen and helium in the Universe, and obtaining the estimate $\zeta_m \approx -1.98 \times 10^{-5}$. This estimate does not apply the effective field theory methods of Ref. [129].

Accounting more accurately for the composite and quantum mechanical nature of the nucleus, Ref. [131] applies methods introduced in Refs. [134,138] to compute ζ_m in terms of expectation values of nuclear electromagnetic density and current operators. Quantum mechanical identities (e.g. the Thomas-Reich-Kuhn rule) can be used to rewrite these quantities in terms of the photoabsorption cross section and other empirically measurable properties of nuclei. An estimate of

$$\zeta_m(A) \approx -\frac{8.60465 \times 10^{-6}}{A^{1/3}} \qquad (3)$$

is obtained [132], although this calculation does not yet apply the fully relativistic field-theory techniques of Ref. [129].

Seeking to test the analytic approximations of Ref. [135], we conducted a simulation of the early Universe plasma to assess if the electrostatic contribution to ζ_m is shielded. Our methods and results are presented in Appendix A. We find that the term of opposite sign potentially driving cancellation in ζ_m is orders of magnitude smaller than other terms, and thus that electrostatic cancellation does not occur. Further investigation is needed, but we recommend the use of $\zeta_m \approx 10^{-4}$ as a "standard" value for the BSBM variant in which ψ is not coupled to the dark sector.

Broadly, there are two logical possibilities. In one case, α variation is sourced by a coupling of ψ to baryonic matter. In the other case, α variation is sourced by a coupling of ψ to the dark sector. In the case that ζ_m is sourced only by standard electromagnetism, it is important to compute ζ_m accurately, so that measurements of or limits to α variation can be turned into allowed ranges for $\sqrt{\bar{\omega}}$, the energy scale of new physics. Put another way, an accurate value of ζ_m is essential to go from the limits obtained in Sec. IV to an empirically allowed range for $l = \hbar c/\sqrt{\bar{\omega}}$, the length scale

at which Coulomb's law breaks down due to interactions of standard model fields with the BSBM scalar field.

In the case that α variation is sourced by dark-sector interactions of ψ, $\zeta_m M_{\text{pl}}^2/\bar{\omega}$ is a single effective (unknown) dimensionless coupling constant (e.g. g in a dilaton warp factor of the form $e^{-g\psi}$ in a nonminimal coupling term of a Lagrangian) to be determined empirically or predicted in an effective dark sector theory, as discussed extensively in Ref. [55]. There, values of $\zeta_m M_{\text{pl}}^2/\bar{\omega} = -\sqrt{2}/4$, 1/2, 0.05, and $-1/\sqrt{16\Omega + 24}$ are obtained, for the string dilaton, supersymmetric Bekenstein-Magueijo model, gaugino-driven modulus, and Brans-Dicke electromagnetism model variants, respectively. Here Ω is the usual Brans-Dicke coupling parameter. For the remainder of this paper, we treat $\zeta_m M_{\text{pl}}^2/\bar{\omega}$ as a parameter to be empirically determined from cosmological data.

B. Scalar field dynamics

The evolution of the scalar field from Eq. (1) can then be written

$$\ddot{\psi} + 3H\dot{\psi} = -\frac{2}{\bar{\omega}} e^{-2\psi} \zeta_m \rho_m. \qquad (4)$$

Assuming a spatially flat, homogeneous, and isotropic universe, we write Eq. (4) as a function of redshift z using the standard Hubble parameter definition $H \equiv \dot{a}/a$, where $H = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda}$ and standard relation between scale factor a and z, $a = 1/(1+z)$:

$$\frac{d^2\psi}{dz^2} - \frac{d\psi}{dz}\left\{\frac{2}{(1+z)} - \frac{d\ln E(z)}{dz}\right\}$$
$$+ \frac{\tilde{m}^2\psi}{(1+z)^2[\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda]}$$
$$= -\frac{6\zeta_m}{\omega}\frac{\Omega_m(1+z)}{\sqrt{\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda}}e^{-2\psi}, \quad (5)$$

where $E(z) \equiv H(z)/H_0$.

Here we also allow a mass m for the scalar field to allow a more general set of dynamical variation models for α. In the Klein-Gordon equation with meters-kilograms-seconds (MKS) units, one has a term $= (mc^2/\hbar)^2\psi$. After rewriting the Klein-Gordon equation through the preceding series of transformations, it is straightforward to see that $\tilde{m} = mc^2/(\hbar H_0) = mc^2/(2.13 \times 10^{-33} \, h\,\text{eV})$, where h here is the usual dimensionless Hubble constant. For the analysis in the main body of this paper, we have assumed the standard $\rho_m \propto a^{-3}$ scaling. As discussed in Appendix B, this does not self-consistently allow for energy conservation under the $\mathcal{L}_{\text{em}} \to \zeta_m\rho_m$ approximation. We assess this issue quantitatively in Appendix B, and find that the resulting correction to our field evolution, observables, and constraints is negligible.
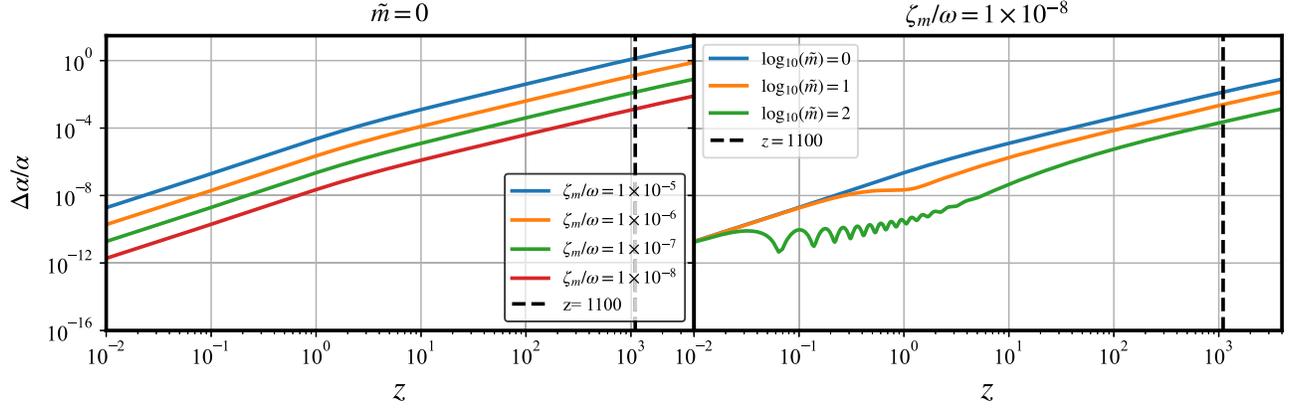
FIG. 1.   Evolution of α variation as a function of z for a range of BSBM model parameters.

The coefficient on the rhs of Eq. (5) is defined relative to the reduced Planck mass and thus appears different from that shown in some other work (e.g. Ref. [125]) on this topic. We treat the scalar-field coupling ζ_m/ω and $\tilde{m}$ as free parameters to be empirically constrained. The scalar field, ψ is related to α by $\alpha = e^{2\psi}e_0^2/hc$, implying that $\Delta\alpha/\alpha \simeq 2\psi$ for small ψ.

We solved Eq. (5) going from $z = 0$ to higher values, using an eighth-order Runge-Kutta (RK) method [139] with 80000 linear steps in z. The number of steps was increased until the relative fractional numerical error in $\Delta\alpha/\alpha$ was smaller than $\sim 10^{-7}$, the fractional α variation implied ($\sim 10^{-5}$) by some QSO measurements [29]. Note that this calculation itself does not depend on the QSO measurement. All the high-z behavior is determined by Eq. (5), along with the chosen values of ζ_m/ω, $\psi(z = 0)$ and $\psi'(z = 0)$. Here the $'$ denotes a derivative with respect to z.

To verify the stability and accuracy of this method, we ran a test in which we evolved our numerical solver backward in time to high z and used $\psi(z_{\text{high}})$ and $\psi'(z_{\text{high}})$ as the initial conditions for a forward integration in time (decreasing z), comparing $z = 0$ and intermediate values of ψ and ψ' with the results obtained from the high-to-low z evolution. We recovered the desired fractional $\sim 10^{-7}$ precision in $\Delta\alpha/\alpha$. Assuming that $\psi = 0$ and $\psi' = 0$ at $z = 0$ (needed for consistency with present-day lab constraints), we show the resulting cosmological α variation as a function of z in Fig. 1 for a variety of ζ_m and $\tilde{m}$ values.

Reference [31] summarizes Keck HIRES QSO observations hinting at α variation. There, a slightly different convention for BSBM constants is used. Effectively, they constrain $\tilde{\zeta} = 8\pi\zeta_m/\omega$, obtaining $\tilde{\zeta} \leq 3.7 \times 10^{-6}$ at 95% C.L. and $\zeta_m/\omega \sim 10^{-7}$. In Fig. 1, we see that if these constraints are saturated, one expects $\Delta\alpha/\alpha \sim 10^{-2}$ at the recombination epoch, well within reach of the CMB's sensitivity to α variation [79–87,97]. For values $\log_{10}\tilde{m} \gtrsim 1.5$, the oscillations in ψ are fast enough that many (tens or more) have occurred in a period comparable to the Hubble timescale today.

In this regime, we use an adaptive version of the eighth-order Runge-Kutta solver (step size is set by comparing the value for the nonadaptive regime with $\sim 1/10$ of a Hubble time, and choosing the minimum). We verify that field evolution agrees between fixed and adaptive step size solvers at the $\sim 5\%$ level for $\tilde{m} = 1.0$. We have verified that constraints to ζ_m/ω are insensitive to an increase of time resolution (by a factor of ~ 2) at the ~ 1–5% level depending on the precise value of $\tilde{m}$.

III. PRINCIPAL COMPONENT ANALYSIS

One powerful approach for probing nonstandard physics is PCA, which eschews a specific model, and instead determines a nonparametric family of template functions. These PCs are data-driven models that capture the variance between observations and a fiducial (or best-fit) model.

PCA has already been used to great effect to test models of dark energy [as parametrized by its equation-of-state parameter $w(z)$] [99,104,105], the cosmic reionization history [100,101], nonstandard cosmic recombination [106], as well as more exotic physics like dark-matter decay and annihilation [103]. In this work, we use the α variation in PCs already determined in prior work by some of us [97,107] to probe time-varying α, setting the stage for the work described here, as well as future analysis of a broad family of theoretical models. Full details of the techniques used to apply and constrain the PCs can be found in Refs. [97,107], but we briefly review the technique below.

PCs are obtained by diagonalizing the Fisher information matrix [140,141]:

$$F_{ij} = \left\langle \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j} \right\rangle, \qquad (6)$$

where $\mathcal{L}$ is the log-likelihood function of a data set or simulation given values of all model parameters θ_i evaluated at their fiducial values. Schematically, the parameter vector $\boldsymbol{\theta} = \{\mathbf{p}, \mathbf{q}\}$, where $\mathbf{p}$ denotes fiducial model parameters and $\mathbf{q}$ denotes the expansion coefficients of non-standard

deviations from the fiducial model for quantities (e.g. α) usually treated as constant.

For our case, this means that

$$\frac{\Delta\alpha(z)}{\alpha_0} = \sum_i q_i f_i(z), \tag{7}$$

where the basis functions f_i are some (complete, but not necessarily orthogonal) set of smooth basis functions centered at some set of redshifts z_i and q_i are expansion coefficients. The vector $\mathbf{p} = \{A_s, n_s, \Omega_b h^2, \Omega_c h^2, \tau_{\text{reion}}, H_0\}$ contains the standard cosmological parameters of the dimensionless amplitude of the primordial perturbation power spectrum, its spectral index, the relic baryon density, CDM density, optical depth to reionization, and Hubble constant, respectively.

For a Gaussian likelihood, the CMB Fisher matrix is given by

$$F_{ij} = \sum_\ell f_{\text{sky}} \left(\frac{2\ell+1}{2}\right) \frac{\partial \mathbf{C}_\ell}{\partial \theta_i} \mathbf{\Sigma}_l^{-1} \frac{\partial \mathbf{C}_\ell}{\partial \theta_j}, \tag{8}$$

where $\mathbf{C}_\ell = \{C_\ell^{\text{TT}}, C_\ell^{\text{EE}}, C_\ell^{\text{TE}}\}$ is the theoretically predicted set of CMB (temperature/polarization auto and cross) power spectra, $\mathbf{\Sigma}_l$ is the covariance matrix of observationally estimated CMB power spectra at multipole index ℓ, including the effect of instrumental noise and cosmic variance, and f_{sky} is the fraction of sky covered by an experiment.

Using a set of Gaussian basis functions, a modified version of the CAMB [142] code interfaced with the COSMOREC recombination code [143], the *Planck* 2018 likelihood function, and the usual analytic approximations for the instrumental properties of the SO (an ongoing ground-based CMB experiment), a set of PCs for time-varying α was obtained in Ref. [108]. CAMB was run including the impact of gravitational lensing [144], which smooths the high-ℓ anisotropies. We apply these PCs here to test the BSBM model.

Several PCs are shown for the SO case in Fig. 2. These PCs peak in the range $800 \lesssim z \lesssim 1200$. As described in Refs. [96,97,107,108] (and references therein), evolution in α affects the CMB by modulating the Lyman-α absorption cross section, the hydrogen two-photon $2s \to 1s$ transition rage (both of which impact the free electron fraction), and the Thomson scattering cross section. All these rates in turn affect the median redshift and width of the last-scattering surface, the rate of diffusion damping, and the efficiency with which CMB polarization is generated. All these processes occur primarily during the peak z range of the PCs.

We see that at low $\tilde{m}$, the BSBM model is relatively featureless compared to the PCs in the z range of interest. Models with higher $\tilde{m}$ values show an overlap of large amplitude α variation and oscillation, increasing the importance of higher-index PCs. Perhaps more complicated
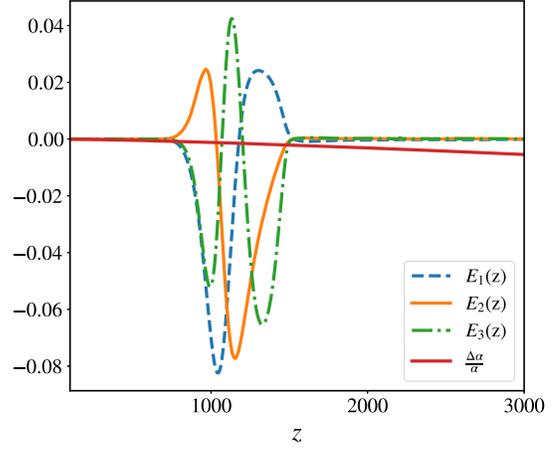

FIG. 2. Evolution of α variation as a function of z assuming $\tilde{m} = 0$, $\zeta_m/\omega = 1 \times 10^{-8}$ and first three SO eigenmodes.

potential energy functions (which likely begin coherent oscillation earlier) would show an overlap of large amplitude α variation and oscillation, leading to more interesting interaction with the PCs. We will explore this possibility more in further work.

In terms of the PCs $E_i(z)$, any model may be expressed

$$\frac{\Delta\alpha(z)}{\alpha_0} = \sum_i \rho_{\text{M},i} E_i(z), \tag{9}$$

where the PCs can be expressed as linear combinations of the original basis functions

$$E_i(z) = \sum_j e_{ij} f_j(z), \tag{10}$$

where e_{ij} denotes the jth basis component of the ith Fisher-matrix eigenvector $\mathbf{e}_i$. For sufficiently dense basis sets, the PCs should themselves be numerically convergent (checked for the α-variation case in Ref. [107]) and basis-independent (checked for variations in the cosmic recombination history in Ref. [106]).

The projection amplitudes for any specific model realization are given by

$$\rho_{\text{M},i} = \int \frac{\Delta\alpha(z)}{\alpha_0} \times E_i(z) dz \tag{11}$$

and may be used to accurately reexpress the variation around the fiducial model as long as it is small.

These can be used to construct a χ^2 diagnostic with the best-fit (from the data) values for the model-expansion coefficients in the PC basis, $\rho_{\text{D},i}$:

$$\chi^2 = \sum_i (\rho_{\text{M},i} - \rho_{\text{D},i})(\mathcal{F})_{ij}^{-1}(\rho_{\text{M},i} - \rho_{\text{D},i}). \tag{12}$$

Here the sum is over PC indices.

One advantage of this method is that even if the likelihood with respect to model (e.g. BSBM) parameters is non-Gaussian, the likelihood with respect to PC amplitudes is still very close to Gaussian. Here the covariance is given by

$$\sigma_{\rho_{M,i}} = \sqrt{(\mathcal{F}^{-1})_{ii}} \simeq \frac{1}{\sqrt{\lambda_i}}, \tag{13}$$

where $\mathcal{F}$ is the (nearly diagonal) α-variation Fisher matrix. The quantity λ_i is the ith eigenvalue of the α-variation Fisher matrix. As discussed in Ref. [106], two types of PCs are possible for new physics. In one, standard cosmological parameters are held fixed, and $\mathcal{F}$ only includes variations in α. In the other, the Fisher matrix includes cosmological parameter variations and α variation (and thus their full degeneracy structure). The former are marginalized over to obtain a Fisher matrix for α variation, which was then diagonalized to obtain the PCs.

Applying the results of Ref. [108], we use these post-marginalization PCs, whose best-fit values and allowed ranges were originally found through a full MCMC in which cosmological parameters were simultaneously varied with PC amplitudes [108]. At Fisher level, expanding around the fiducial model, parameter changes (from $\rho \to \zeta_m/\omega$) commute with cosmological parameter marginalization, and so we do not expect our constraints to get less stringent due to any neglected covariances. Nonetheless, in future work, we will more fully account for these covariances by using the original samples from the MCMC to construct a kernel-density likelihood in which PC amplitudes and cosmological parameters can be simultaneously varied, as in e.g. Ref. [145].

To construct our *Planck* posterior probability for BSBM model coefficients, we use Eqs. (11) and (12) and the usual $\mathcal{L} \propto e^{-\chi^2/2} \times \pi$. To efficiently explore a broad parameter space, we use log-flat and flat priors π for $\tilde{m}$ and ζ_m/ω, respectively, and the parameter ranges $-8 \le \log_{10}(\tilde{m}) \le 1$ and $-10^{-2} \le \zeta_m/\omega \le 10^{-2}$. In Sec. IV, we will explain a number of tests done to make sure our constraints are robust and not driven by our choice of prior.

We use the best-fit PC amplitudes and errors (obtained from MCMC runs) from Ref. [97]. We computed $\Delta\alpha(z)/\alpha$ using the method discussed in Sec. II. To conduct forecasts for SO, we built a mock likelihood, assuming that the fiducial model is true (e.g. that $\rho_{D,i} = 0$), and used Fisher-forecast errors on the PC amplitudes. For our *Planck* analysis, three PCs were used, while 10 PCs were included for SO, to allow for its higher information content. It can be helpful to assess the relative information content and utility of different PCs, and to determine how many are truly needed to properly capture a data set. One useful tool is the signal-to-noise (SNR) contribution from each mode, which is

$$(S/N)_i = \sqrt{\lambda_i \rho_i^2}. \tag{14}$$

Another interesting quantity in PCA is the risk factor [98,99,102], defined via

$$\sigma^2[\alpha(z_j)] = \left(\sum_i^{N_{PC}} \frac{e_i^2(z_j)}{\lambda_i} \right), \tag{15}$$

$$b(z_j) = \frac{\Delta\alpha(z_j)}{\alpha} - \sum_i^{N_{PC}} [\rho_i E_i(z_j)], \tag{16}$$

$$\text{Risk}[N_{PC}] = \sum_j \{b^2(z_j) + \sigma^2[\alpha(z_j)]\}, \tag{17}$$

where N_{PC} is the number of PCs used to test a model. Here $b(z_j)$ is defined to be the bias in $\Delta\alpha(z)/\alpha$ induced by using an incomplete set of PCs, which competes with the variance (which decreases when PCs are filtered out to reduce the error in the data if the fiducial model is actually true). This quantity is model dependent and minimized when an optimal number of PCs is chosen to test a specific model. In Sec. IV, we compute this quantity to assess the impact of different PCs on our BSBM model constraints.

IV. DATA ANALYSIS AND RESULTS

Using the *Planck* 2018 α-variation likelihood described in Sec. III and our scalar-field integrator, we found best fit values of ζ_m/ω for the $m = 0$ case. We then set a broad parameter range of $-0.01 \le \zeta_m/\omega \le 0.01$ and $-8 \le \log_{10}(\tilde{m}) \le 1$ and ran an MCMC simulation to determine the allowed ζ_m/ω and $\log_{10}(\tilde{m})$ parameter space.

We used the EMCEE package, which applies Goodman and Weare's affine invariant sampler to run MCMCs and checked for convergence using the autocorrelation time [146]. For the CMB analysis, we established two-dimensional MCMC convergence as follows: For the *Planck* analysis, we ran 32 chains for 12000 samples. We obtained a correlation length of $\sim$150 for ζ_m/ω and $\sim$60 samples for $\log_{10}(\tilde{m})$, less than $12000/50 = 240$. For the SO analysis, we ran 32 chains for 10000 samples. We obtained an autocorrelation length of $\sim$10 samples ζ_m/ω and 58 in $\log_{10}(\tilde{m})$. Both are less than $10000/50 = 200$, and so all these CMB chains are converged. For our analysis, we used a burn-in fraction of 0.3 throughout. Visualizations and confidence intervals were generated with the GetDist package [147].

The results are shown in Fig. 3. As we can see, the constraint becomes less stringent near $\log_{10}(\tilde{m}) \simeq 0$. The predicted decoupling-era α variation of parameter sets beyond this threshold is smaller than for lower values of $\log_{10}(\tilde{m})$. To be sure that our allowed parameter space [which is highly non-Gaussian in $\log_{10}(\tilde{m})$] is robust, we did a series of one-dimensional MCMC runs at a discrete grid of m values near and beyond the transition point, as also
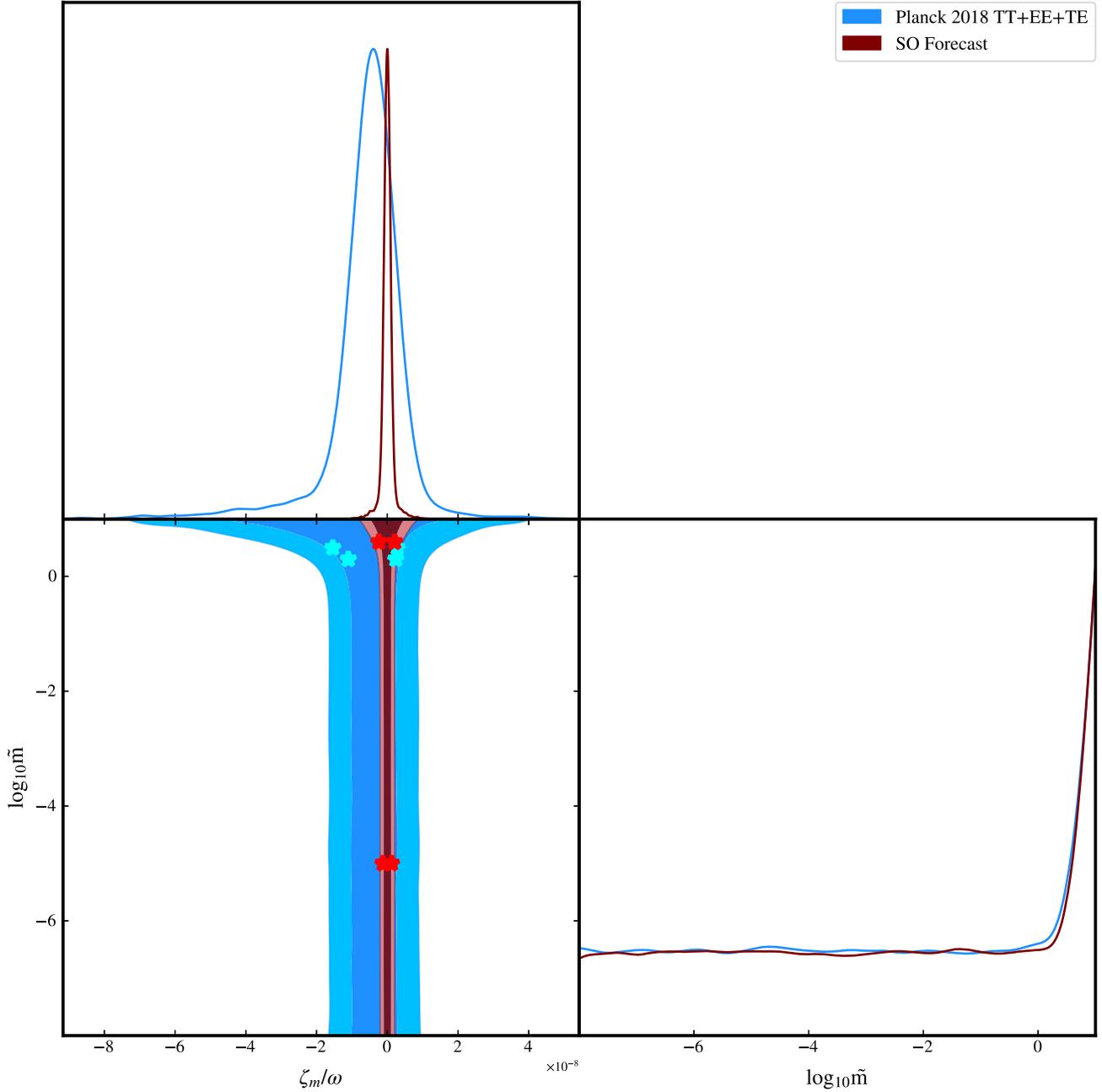
FIG. 3.   68.5% (dark) and 95% C.L. (light) contours for ζ_m/ω and $\tilde{m}$ using *Planck* 2018 data. Overlaid are forecasts for upcoming SO data, assuming forecast SO error bars on principal component amplitudes and a fiducial value $\zeta_m/\omega = 0$. Five-sided red (blue) stars show *Planck* (SO) results from the one-dimensional MCMCs described in Sec. IV.

done to obtain constraints to ultra-light axions in Ref. [122]. These are presented in Table I.

We made sure that the range used for $\log_{10}(\tilde{m})$ overlapped with that used in the two-dimensional MCMCs, finding agreement that validates both sets of constraints (the constraints from overlapping one-dimensional simulations are superimposed with stars in Fig. 3). This confirms that our two-dimensional constraints are not dominated by our choice of prior for $\log_{10}(\tilde{m})$. Additionally, we computed the data-theory χ^2 on a coarse grid and obtained preliminary

limits that validate our ultimate results (from both one and two-dimensional MCMCs).

For $\log_{10}(\tilde{m}) \geq 1$, we used the adaptive solver described above. Chains (24 for every mass) for $\log_{10}(\tilde{m}) \leq 2$ were 3000 samples long, while those with $\log_{10}(\tilde{m}) > 2$ were 4500 samples long. The autocorrelation length was ≤ 77 samples for the high $\log_{10}(\tilde{m})$ cases, and ≤ 30 for the low $\log_{10}(\tilde{m})$ cases, and are thus all converged. The consolidated results (constraints and forecasts) at high values of $\log_{10}(\tilde{m})$ are shown in Figs. 4 and 5.
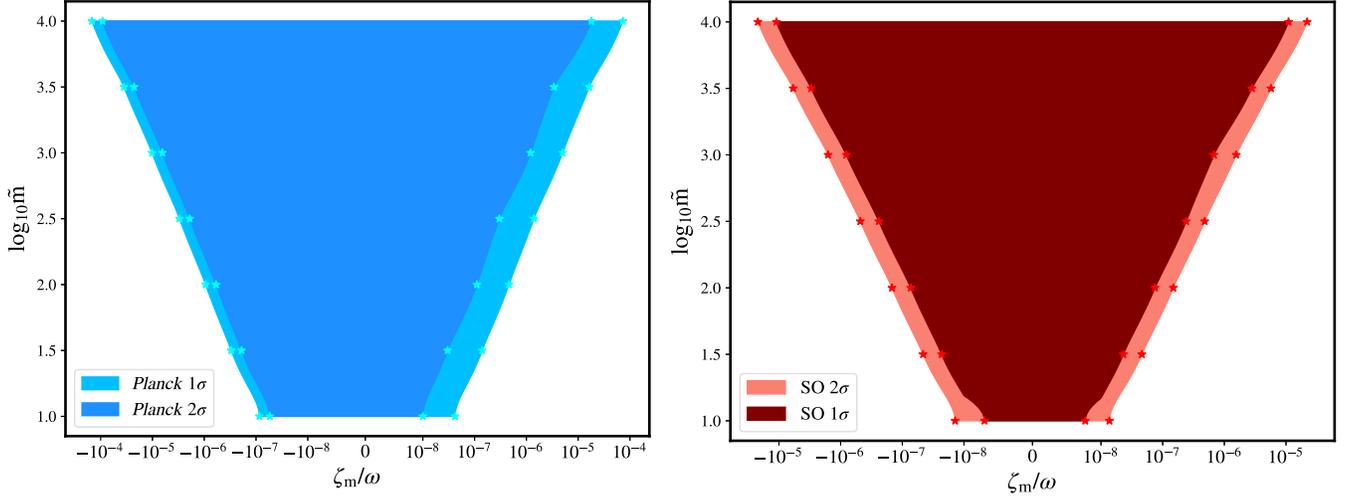
FIG. 4. 68.5% (dark) and 95% C.L. (light) contours for ζ_m/ω and $\tilde{m}$ using our SO forecast and Planck data, obtained with the adaptive solver and one-dimensional EMCEE runs.

Summarizing our two-dimensional results, we see that at 95% C.L., *Planck* data impose the constraint $\zeta_m/\omega \leq 9.3 \times 10^{-9}$, with future SO data offering an order-magnitude improvement in sensitivity. This constraint relaxes nearly completely for $\log_{10}(\tilde{m}) \geq 1$. For SO, we find that at the lowest $\tilde{m}$ values, SO would be sensitive to values of $\zeta_m/\omega \simeq 2.2 \times 10^{-9}$ and higher, as can be seen separately in Fig. 6. These results are also shown in Table II.

It is interesting to examine the change induced in CMB observables for parameter values saturating our constraints in order to understand what features drive our sensitivity to the BSBM model. We have $\Delta C_{\ell,j}^{XY}$ for each of the PCs obtained via

$$
\frac{dC_{\ell,j}^{XY}}{d\rho_j} = \sum_i \frac{dC_\ell^{XY}}{dq_i} \frac{dq_i}{d\rho_j},
$$
$$
= \sum_i \frac{dC_\ell^{XY}}{dq_i} \int dz E_j(z) f_i(z), \quad (18)
$$

where $X, Y \in \{TT, EE, TE\}$. We then use these expressions to calculate the BSBM-induced change to observables, applying the fact that

$$
\Delta C_{\ell,\text{BSBM}}^{XY} = \sum_{j=1}^{N_{pc}} \rho_{\text{BSBM},j} \frac{dC_{\ell,j}^{XY}}{d\rho_j}. \quad (19)
$$

The changes to C_ℓ for *Planck* best-fit values of ζ_m/ω, as well as 68.5% and 95% C.L. constraint-saturating values, are shown in Fig. 7, along with the same quantities for SO. For this plot we used $\log_{10}(\tilde{m}) = -3.0$. The changes are normalized to the cosmic variance per multipole

$$
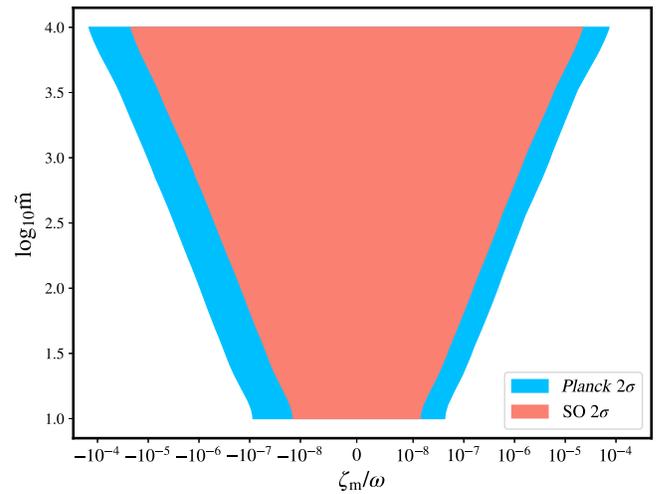\sigma_{C_\ell^{XX}} = \sqrt{\frac{2}{2l+1}} C_\ell^{XX} \quad (20)
$$

TABLE I. Constraints on ζ_m/ω at different masses from *Planck* 2018 analysis using one-dimensional EMCEE runs and a fixed set of $\tilde{m}$ values.

$\log_{10} \tilde{m}$	(ζ_m/ω) (Planck 2018)
0.3	$(-4.3 \pm 6.7) \times 10^{-9}$
0.5	$(-6.3 \pm 9.1) \times 10^{-9}$
0.6	$(-0.8 \pm 1.1) \times 10^{-8}$
0.8	$(-1.4 \pm 2.3) \times 10^{-8}$
1	$(-2.2 \pm 3.2) \times 10^{-8}$
1.5	$(-0.8 \pm 1.1) \times 10^{-7}$
2	$(-2.4 \pm 3.5) \times 10^{-7}$
2.5	$(-0.8 \pm 1.1) \times 10^{-6}$
3	$(-2.6 \pm 3.8) \times 10^{-6}$
3.5	$(-0.96 \pm 1.3) \times 10^{-5}$
4	$(-3.7 \pm 5.5) \times 10^{-5}$

FIG. 5. 95% C.L. (light) contours for ζ_m/ω and $\tilde{m}$ using our SO forecast and Planck data, obtained with the adaptive solver and one-dimensional EMCEE runs.
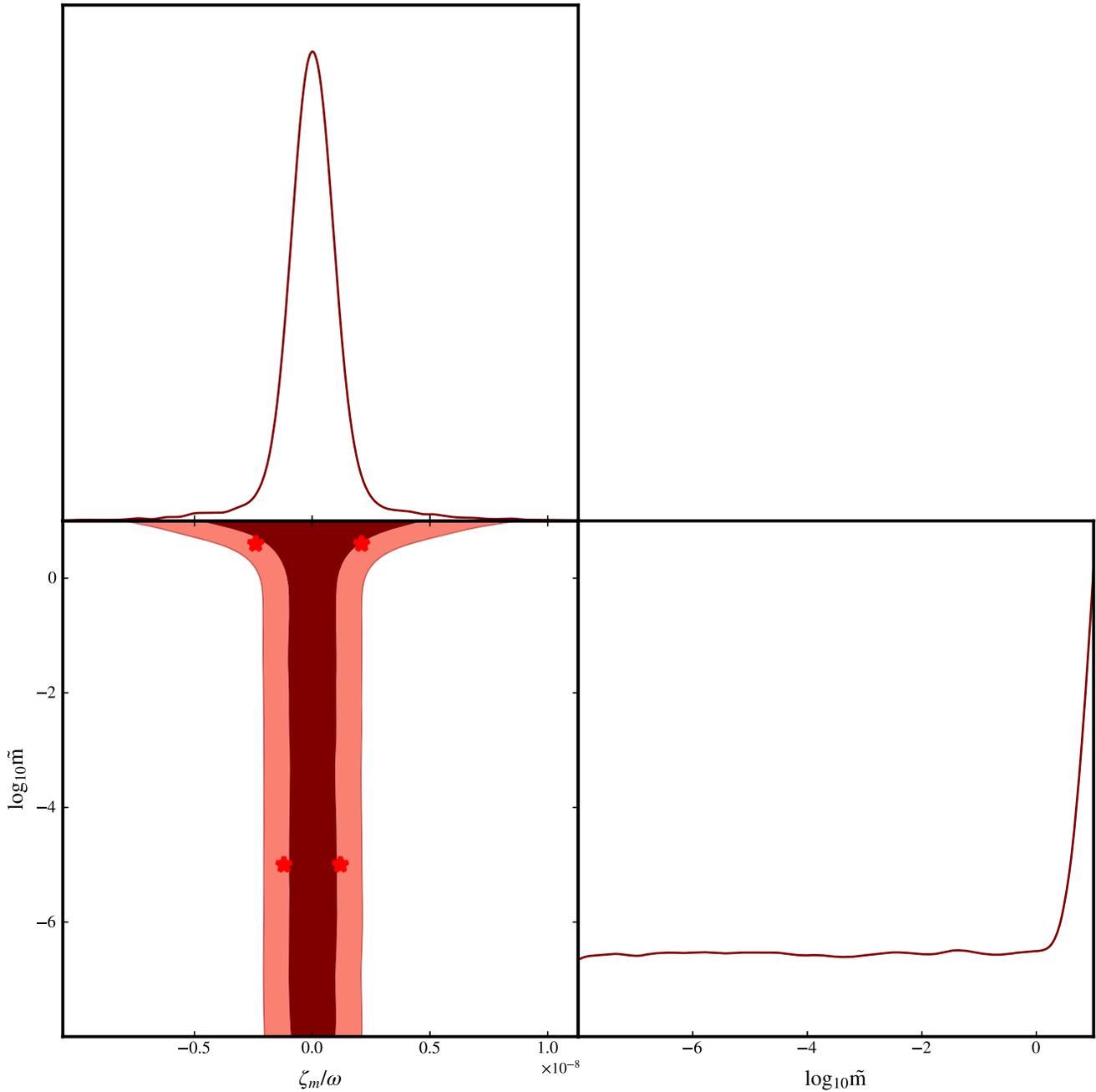
FIG. 6. 68.5% (dark) and 95% C.L. (light) contours for ζ_m/ω and $\tilde{m}$ using our SO forecast. Overlaid are forecasts for upcoming SO data, assuming forecast SO error bars on principal component amplitudes and a fiducial value $\zeta_m/\omega = 0$. Five-sided stars show SO forecasts from the one-dimensional MCMCs described in Sec. IV.

of the fiducial model. To avoid spikes near zero crossing of TE, we use the usual convention that

$$\sigma_{C_\ell^{\mathrm{TE}}} = \sqrt{\frac{2}{2l+1}}\sqrt{(C_\ell^{\mathrm{TE}})^2 + C_\ell^{\mathrm{TT}}C_\ell^{\mathrm{EE}}}. \qquad (21)$$

We see that in temperature, the dominant effect is a decrease in high-ℓ anisotropies. This corresponds to a shift of the diffusion damping tail to lower-ℓ (larger angular scales). Larger positive values of ζ_m/ω correspond to larger

values of α in the past, more efficient scattering, later decoupling, and a surface of last-scattering closer to the observer, driving features to lower ℓ. In temperature, this geometric effect dominates over higher scattering rates yielding lower diffusion-damping lengths (which would enhance rather than depress low-ℓ anisotropies).

Higher values of ζ_m/ω mean larger values of α in the past. This means that Thomson scattering rates were higher, and E-mode polarization anisotropies were generated more efficiently at scales at ℓ below the Silk damping scale.

TABLE II. Forecast sensitivity levels on ζ_m/ω at different masses from the SO forecast, using one-dimensional EMCEE runs and a fixed set of $\tilde{m}$ values.

$\log_{10}\tilde{m}$	(ζ_m/ω) (SO)
-5	$(0.0 \pm 1.2) \times 10^{-9}$
0.6	$(0.2 \pm 2.1) \times 10^{-9}$
0.8	$(0.0 \pm 4.9) \times 10^{-9}$
1	$(0.1 \pm 6.9) \times 10^{-9}$
1.5	$(0.0 \pm 2.3) \times 10^{-8}$
2	$(0.1 \pm 7.4) \times 10^{-8}$
2.5	$(0.0 \pm 2.4) \times 10^{-7}$
3	$(-0.2 \pm 7.9) \times 10^{-7}$
3.5	$(-0.1 \pm 2.9) \times 10^{-6}$
4	$(-0.0 \pm 1.1) \times 10^{-5}$

FIG. 8. Compilation of possible α variation as a function of time, inferred from analysis of QSO spectra, from Ref. [31].

As a check, we also used EMCEE with a QSO data set and error bars shown in Fig. 8 to probe the BSBM parameter space. We used the χ^2 between the model BSBM variation $\Delta\alpha_M(z)$ (for any given values for m and ζ_m/ω) and the reported variation inferred from QSO data $\Delta\alpha_D(z)$:

$$\chi^2 = \sum_i \frac{[\Delta\alpha_M(z_i) - \Delta\alpha_D(z_i)]^2}{\sigma_i^2}. \quad (22)$$

The QSO redshifts are denoted z_i and the measurement errors in $\Delta\alpha_D(z)$ are denoted σ_i. For the QSO analysis, we used 32 one-dimensional MCMC chains of 2000 samples
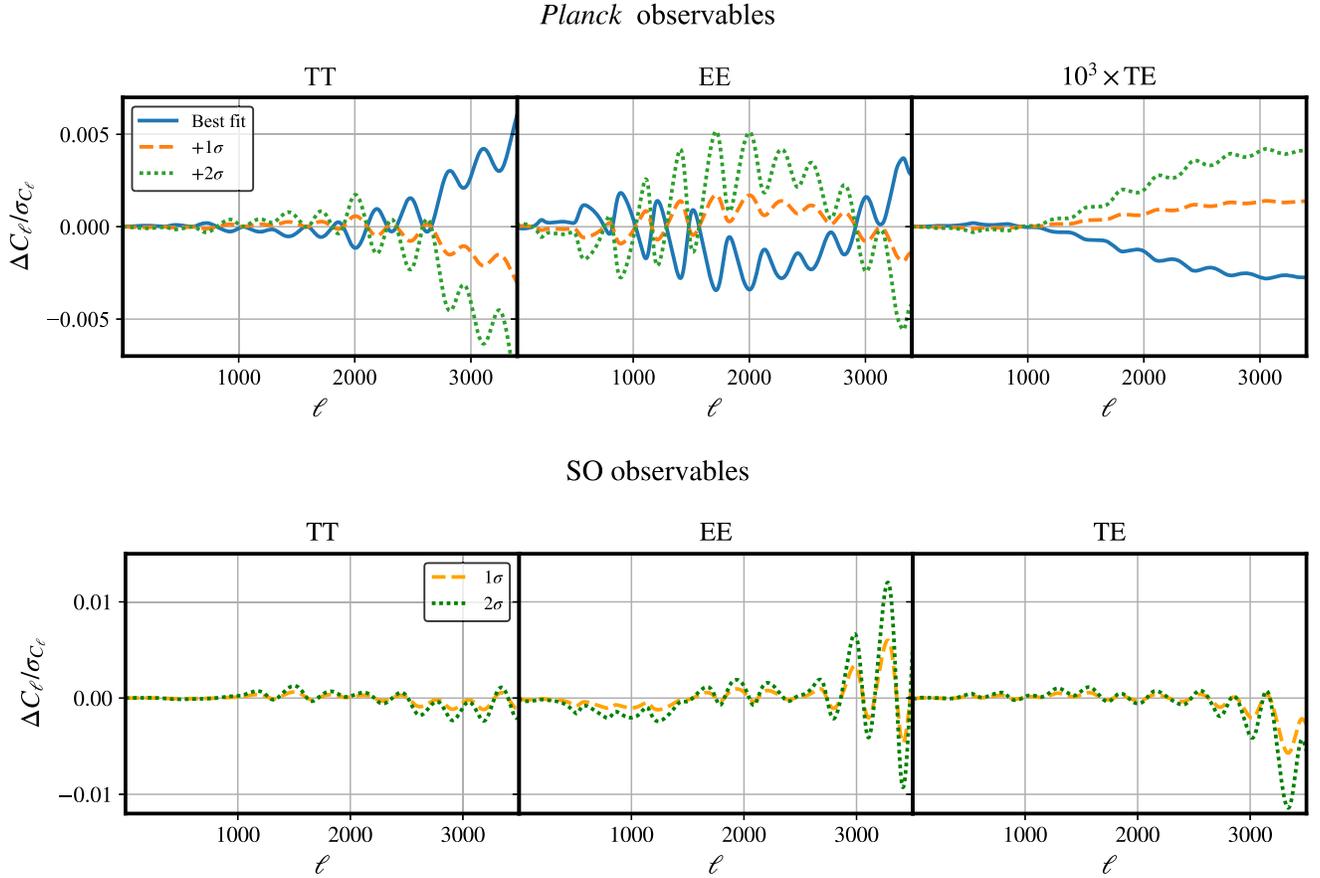
FIG. 7. Top row: fractional change $\Delta C_\ell/\sigma_{C_\ell}$ for *Planck* 2018 best fit values, and for values saturating 68.5% C.L. and 95% C.L. constraints to ζ_m/ω, all for $\log_{10}(\tilde{m}) = -3.0$. Here σ_{C_ℓ} is the cosmic variance per multipole. Bottom row: same quantity, now assuming assuming models that saturate SO error bars (themselves determined using the fiducial model).

TABLE III.   Constraints on ζ_m/ω at different masses from QSO data, using one-dimensional EMCEE runs and a fixed set of $\tilde{m}$ values.

$\log_{10} \tilde{m}$	(ζ_m/ω) (QSO)
-5	$(-1.1 \pm 1.0) \times 10^{-7}$
0	$(-1.4 \pm 1.1) \times 10^{-7}$
1	$(-1.4 \pm 1.2) \times 10^{-6}$
1.5	$(-1.24 \pm 0.89) \times 10^{-5}$
2	$(-7.4 \pm 0.83) \times 10^{-5}$
2.5	$(7.8 \pm 0.82) \times 10^{-4}$
3	(-0.1 ± 0.0084)
3.5	(-0.085 ± 0.08)
4	(-0.01 ± 0.008)

each for each $\tilde{m}$ value. The autocorrelation length was ~ 17 samples, far less than $2000/50 = 40$, and so the chains are converged.

Our reanalysis of the QSO data summarized in Ref. [31] (mostly from the Keck HIRES spectrograph) yields a 95% C.L. limit $\zeta_m/\omega \leq 2.8 \times 10^{-7}$ when $\log_{10}(\tilde{m}) = -5$. Translating into the different normalization of BSBM couplings used there ($\tilde{\zeta} = 8\pi\zeta_m/\omega$), this implies $\tilde{\zeta} \leq 7 \times 10^{-6}$, consistent with the limits given in Ref. [31]. QSO results for higher $\tilde{m}$ are given in Table III. The *Planck* limits presented in this work (alone, without a QSO prior) to BSBM parameters are thus tighter than those imposed by the HIRES data.

Recently, extremely precise constraints to α variation have been obtained with the ESPRESSO spectrograph at the European Southern Observatory and applied to test the BSBM and related models [125,126]. That work includes an external constraint imposed from the CMB, but features QSO (and other more local) results. A 95% C.L. limit of $\tilde{\zeta} \leq 1.5 \times 10^{-6}$ is given there (obtained from QSO data), equivalent to $\zeta_m/\omega \leq 6.0 \times 10^{-8}$. This limit folds in a significant nonzero centroid for the PDF of ζ_m/ω in their analysis, however, and the sensitivity of that work (with QSO data set) has a 95% error bar on ζ_m/ω of $\sim 3.3 \times 10^{-7}$, less stringent than our *Planck* analysis.

Our SO forecasts predict a sensitivity level better than this, and so it stands to reason that future CMB experimental efforts (e.g. S4, HD) will outpace some lower-z probes of α variation in testing the BSBM model and related ideas. More broadly, this is an independent technique, depending on different physics, with different systematics, and evidence for the model at high-redshift could be reconciled with low-redshift upper limits in the context of a theory with different time evolution than initially assumed (the same could be said for a high-redshift upper limit in conflict with potential low-redshift evidence).

In future work, we will investigate the power of combining QSO and CMB data sets. We have developed

a version of the Boltzmann code CLASS which evolves the spatial fluctuations in the scalar field of the Bekenstein model. Our code includes both gravitational effects and nonminimal coupling of ψ to DM.

While we are exploring the possible impact of this *spatial* fluctuation on CMB observables, we have already found that the predicted angular spectrum of α variations at decoupling is blue. The observable impact of these fluctuations is currently only computable within a separate universe limit (see Ref. [148] for a discussion) and produces a spatial variation (on scales that are super-horizon at recombination) in α with a root mean-squared value at the $\sim 10^{-5}$ fractional level for values $\zeta_m/\omega = 1$. This signal is small, but the correlation with the underlying dark-matter density field (and possible early Universe quantum contributions to ψ fluctuations), could still induce an observable signal through non-Gaussian signatures in the CMB along the lines described/sought in Refs. [109,110,149].

A more detailed discussion of the results and methods of those efforts is beyond the scope of this work and will be presented in a future manuscript [150]. The small amplitude of the signal (due to the causal growth of ψ fluctuations and the smallness of primordial fluctuations) is consistent with predictions of Ref. [37].

To understand which PCs are driving constraints, it is interesting to examine the SNR of each mode for parameter values saturating the 95% C.L. BSBM model constraints for $\tilde{m} = 0$ with *Planck* data, applying Eq. (14). These quantities are shown in Fig. 9. We see that the first two PCs have significant constraining power for *Planck*, while three PCs of SO data will be constraining. Unsurprisingly, SO will generally have a much higher SNR for the BSBM model.
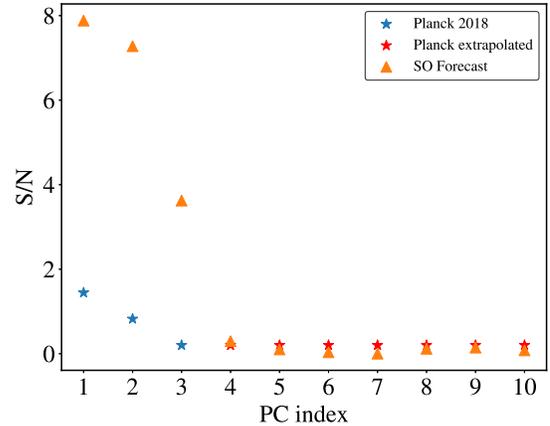

FIG. 9.   Signal-to-noise ratio as a function of PC index, assuming true ζ_m/ω that saturate *Planck* 95% C.L. constraints with $\tilde{m} = 0$. PC eigenvalues for higher indices in *Planck* are extrapolated to estimate SNR (as the eigenvalues appear to follow a clear power law) and are shown in red for clarity (red).
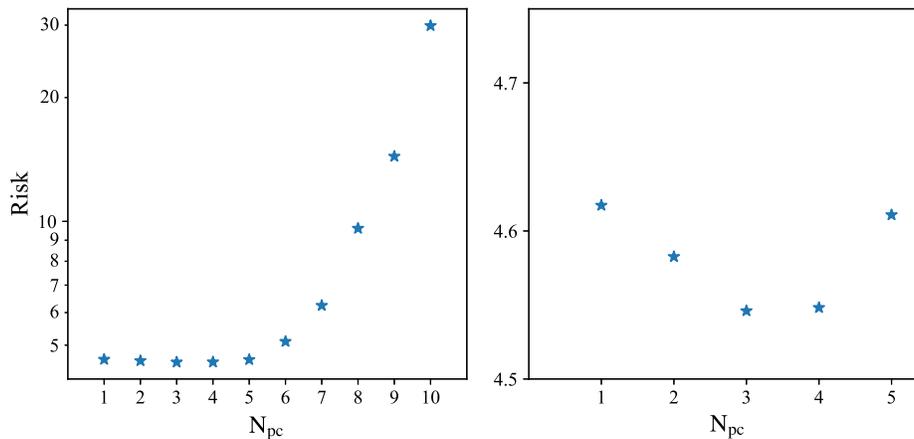
FIG. 10. Risk factor for the BSBM model, as a function of the number of PCs used for SO sensitivity levels, assuming ζ_m/ω values that saturate *Planck* 95% C.L. constraints with $\tilde{m} = 0$. Left panel: full range. Right panel: zoom-in to resolve presence of a minimum in the risk.

We also compute the risk [Eq. (17)] as a function of N_{PC}, the number of PCs used in the analysis, assuming SO noise levels and a signal saturating the *Planck* 95% constraint level for $\tilde{m} = 0$. The results are shown in Fig. 10 for SO, indicating that three PCs should suffice to test the BSBM model adequately while minimizing risk.

V. CONCLUSIONS

We have used *Planck* 2018 data and its principal components for variations in α to constrain the BSBM theory of varying α, obtaining a 95% C.L. constraint of $\zeta_m/\omega \leq 9.3 \times 10^{-9}$. Assuming the null hypothesis holds, we have found that the Simons Observatory will have sensitivity to values as low as $\zeta_m/\omega = 2.2 \times 10^{-9}$. These results apply not only to the BSBM theory, but also to a related family of theoretical ideas, such as the string dilaton, the supersymmetric BSBM, the gaugino-driven modulus, and Brans-Dicke electromagnetism [55].

Looking forward, it will be interesting to extend our results to further theoretical models for α variation [including, for example, a scalar-field potential and coupling function $\omega(\psi)$ [32,33]], or models where other coupling constants (e.g. m_e [151], G [91–95], or even c [152–154]) are dynamical.

Using kernel-density estimates of the full-PCA $+\Lambda$CDM likelihood, we will more fully probe the covariance of BSBM model parameters with standard cosmological parameters. Causality dictates that variations of α (or other fundamental parameters) in time require variations of α. While these variations are likely to be very small, they would induce non-Gaussian signatures in the CMB [109–111], and it would be useful to explore if this signal could be better extracted by applying CMB delensing, and harnessing the cross-correlation of α with the underlying density field [155]. The induced CMB bi- and trispectra by these models have, in principle, a distinct shape from more standard effects like CMB lensing [148,155]. In future work, we will assess if this can be used to better extract spatial variations in α from the CMB.

Thinking towards future measurements of CMB anisotropies, it will be interesting to explore the power of the planned CMB-S4 experiment [127,156], as well as more futuristic concepts like CMB-HD [157], which could probe the CMB at much smaller angular scales than ever before, promising much improved leverage on varying fundamental constants.

In coming years, new cosmological frontiers will open, with likely measurements of the neutral 21-cm signature of cosmic reionization and the cosmic dark ages [116]. The global 21-cm signal and anisotropies should be strongly sensitive to α [117,118]. Furthermore, CMB spectral distortions from the Silk damping and recombination eras would have spatial and frequency dependence that is strongly dependent on α [158]. In future work, we will assess the sensitivity of these powerful measurements to the full family of theories enumerated here. The era of precision cosmology is here, and we should look forward to harnessing its data products, not only to characterize the energy budget of the Universe but also to test the constancy of the fundamental parameters.

APPENDIX A: ELECTROMAGNETIC SOURCING OF α VARIATION

1. Argument for shielding of electrostatic energy

Here we recapitulate Ref. [135]'s argument that electrostatic contributions to ζ_m are shielded. We follow the discussion there but provide additional details as needed. We begin with the static contributions to the Poisson equation:

$$\nabla^2\psi = 4\pi\kappa^2\left[\sum_i\frac{\partial m_i c^2}{\partial\psi}\delta^3(\mathbf{x}-\mathbf{z}_i)+\frac{1}{4\pi}e^{-2\psi}E^2\right]. \quad (A1)$$

The sum here is over all source particles and the constant $\kappa = l/\sqrt{4\pi\hbar c}$ is a renormalized Brans-Dicke (kinetic) coupling for ψ.

Bekenstein integrates the total contributions from each particle in a volume $\mathcal{V}$ in the first term of the rhs of Eq. (A1) and then replaces $(\partial m_i c^2)/\partial\psi$ with $\kappa^{-1}e_{0i}\tan[\kappa\Phi(\mathbf{z}_i)]$ [Eq (43) of Ref. [35]].

The first term [on the rhs of Eq. (A1)] then becomes

$$\sum_i\frac{\partial m_i c^2}{\partial\psi}\delta^3(\mathbf{x}-\mathbf{z}_i)$$

$$\to -\frac{1}{\mathcal{V}}\int_{\mathcal{V}}d^3x\sum_{i\in\mathcal{V}}\kappa^{-1}e_{0i}\tan[\kappa\Phi(\mathbf{z}_i)]\delta^3(\mathbf{x}-\mathbf{z}_i)$$

$$= -\frac{1}{\mathcal{V}}\sum_{i\in\mathcal{V}}e_{0i}\tan[\kappa\Phi(\mathbf{z}_i)].$$

Taylor expanding $\tan[\kappa\Phi(\mathbf{z}_i)]$ about $\Phi(\mathbf{z_i})$ and discarding terms of $\mathcal{O}(\kappa^3\Phi^3)$, the expression

$$-\frac{1}{\mathcal{V}}\sum_{i\in\mathcal{V}}e_{0i}\Phi(\mathbf{z}_i). \quad (A2)$$

is then obtained for the first term.

The second term is somewhat more complicated. Modified Maxwell equations of the BSBM theory imply that $\mathbf{E}$ with $\mathbf{E} = -e^{2\psi}\boldsymbol{\nabla}\Phi$ [Eq. (41) of Ref. [135]], and since $e^{2\psi} = \sec^2\kappa\Phi$ [Eq. (45) of Ref. [35]], linearization yields

$$e^{2\psi} = \sec^2(\kappa\Phi) \approx 1 + \kappa^2\Phi^2 + \mathcal{O}(\kappa^4\Phi^4). \quad (A3)$$

Discarding terms of $\mathcal{O}(\kappa^4\Phi^4)$, the spatial average of the second term is

$$\frac{1}{4\pi}e^{-2\psi}\mathbf{E}^2 \to \frac{1}{4\pi\mathcal{V}}\int_{\mathcal{V}}d^3x[(\boldsymbol{\nabla}\Phi)^2+\kappa^2\Phi^2\mathbf{E}^2]. \quad (A4)$$

Integrating the first term of this integral using the divergence theorem and applying Eq. (A5), one finds

$$\nabla^2\Phi = -4\pi\sum_i e_{0i}\delta^3(\mathbf{x}-\mathbf{z}_i). \quad (A5)$$

The final approximation (applying the immediately preceding result) for the second term on the rhs of Eq. (A1) is

$$\frac{1}{\mathcal{V}}\sum_{i\in\mathcal{V}}e_{0i}\Phi(\mathbf{z}_i)-\frac{1}{4\pi\mathcal{V}}\left[\oint_{\partial\mathcal{V}}\Phi\mathbf{E}\cdot d\mathbf{s}-\kappa^2\int_{\mathcal{V}}d^3x\Phi^2\mathbf{E}^2\right]. \quad (A6)$$

Summing Eqs. (A2) and (A6), it is immediately clear that the only terms remaining are the two integrals in Eq. (A6), and Ref. [135] argues they both are negligible—the argument goes as follows. One can approximate the surface integral as $\mathcal{V}^{-1}\langle\Phi\rangle\sum_{i\in\mathcal{V}}e_{0i}$, using Gauss's law where $\langle\Phi\rangle$ is the surface average of Φ. This quantity is much less than $\mathcal{V}^{-1}\sum_{i\in\mathcal{V}}e_{0i}\Phi(\mathbf{z}_i)$, so it is negligible.

The only remaining quantity is the second integral in Eq. (A6). An upper bound on $|\Phi|$ for a unit charge is obtained divided by the smallest length scale Φ varies over, which is at most 10^{-17} cm, if quarks are the smallest particles of interest. Then, $\kappa^2\Phi^2\mathbf{E}^2/4\pi$ has an upper bound of roughly $10^{-34}(l/l_p)^2\mathbf{E}^2/4\pi$, where l is the characteristic length of Bekenstein's theory. This implies that using $\mathbf{E}^2/4\pi$ vastly overestimates electrostatic energy as a source for α variation in the BSBM model.

2. Testing electrostatic cancellation

In order to the claim that ψ arranges itself to self-shield the electrostatic contribution to its equation of motion, we conduct a particle-in-cell (PIC) plasma simulation of the early Universe (proton and neutron) plasma in PYTHON.

These particles interact and cluster in ways that produce electric and magnetic fields.

We compute these and use them to estimate $\zeta_m = \langle \mathbf{E}^2 - \mathbf{B}^2 \rangle / \rho_m$. We determined field quantities by solving the modified Maxwell's equations using a Fourier spectral method. The equations we solved were the following, given in both real and k space:

$$\mathbf{E} = \boldsymbol{\nabla}\Phi \Rightarrow \tilde{\mathbf{E}} = i\tilde{\Phi}\mathbf{k}, \tag{A7}$$

$$\nabla^2\Phi - 4\pi\rho \Rightarrow \tilde{\Phi} = -4\pi\tilde{\rho}K^2, \tag{A8}$$

$$\nabla^2\mathbf{A} = \frac{4\pi}{c}\mathbf{J} \Rightarrow \tilde{\mathbf{A}} = \frac{4\pi}{c}\tilde{\mathbf{J}}K^2, \tag{A9}$$

where $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$, $\mathbf{k} = \{dx^{-1}\sin(k_x dx), dy^{-1}\sin(k_y dy), dz^{-1}\sin(k_z dz)\}$, and $K = K_x^2 + K_y^2 + K_z^2$ with $K_x = 2dx^{-1}\sin(k_x dx/2)$ and analogous definitions for K_y and K_z [161]. These definitions for k and K_x, K_y, K_z come from taking the Fourier transform of the finite difference forms of ∇ and ∇^2, respectively.

We use these algebraic equations to solve for the fields of interest in the simulation. At each time step, we interpolate the particles onto a spatially uniform grid using a linear spline and periodic boundary conditions. Then, we use the node locations and appropriate weighting to determine the particle number density and current density at each node. We then take the discrete Fourier transform of these quantities and solve for $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{A}}$ using Eq. (A7) through (A9). Next, we take the inverse Fourier transform of $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{A}}$ to get $\mathbf{E}$ and $\mathbf{A}$. Finally, we calculate $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$. $\mathbf{E}$ and $\mathbf{B}$ are then reinterpolated onto particle positions. These are the quantities used to push the particles using the relativistic Boris method [162]. In order to calculate the discrete Fourier transforms, we used a built-in Fourier transform package from the NUMPY module for PYTHON [159].

The simulation uses 4096 particles on a $128 \times 128 \times 128$ grid at $z = 1100$. The particles are arranged uniformly on the grid initially and their initial velocities are drawn from a Maxwell-Boltzmann distribution. The domain of the simulation was a box where each side was $Nn_0^{-1/3}$ long, where N is the number of particles along one side and n_0 is the number density of protons in the Universe at $z = 1100$. Also, $dt = 0.1dx/v_0$, where v_0 is the mode of the initial velocity distribution for electrons. This was set to ensure that no particle traverses more than one grid cell in a time step.

We tested the code using several methods. First, we tested the particle-pushing method by placing a single electron under the influence of a uniform electric and uniform magnetic field. In both cases, the error of our particle pusher method is of order machine precision (e.g. Fig. 11).

Next, we tested our Fourier equation solving method by attempting to reproduce the analytical results for known functions. For example, instead of using the actual current
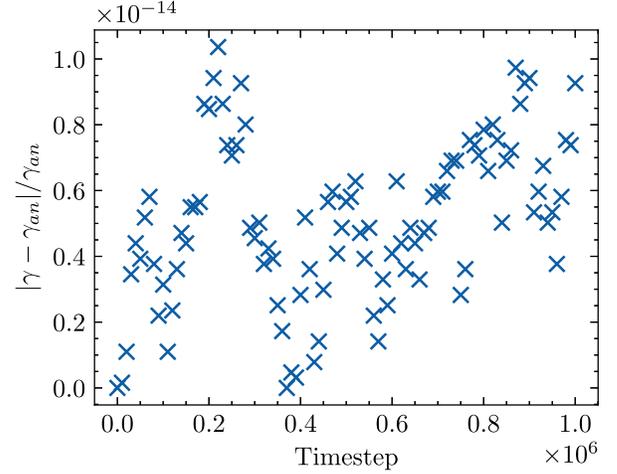

FIG. 11. The fractional error in the difference between the calculated Lorentz factor γ and the analytic Lorentz factor $\gamma_{an} = \sqrt{1 + v_{0y}^2}$ of a particle with $m = 1$, $q = 1$, and $c = 1$ in a constant $\mathbf{B}$ field. The time step was $dt = 2\pi/(10^8)$ and proceeded for 10^6 time steps. Initial $v_0 = (0, 1 - 5 \times 10^{-13}, 0)$ and $\mathbf{B} = (0, 0, 10^6)$. These quantities follow the recommendations in Ref. [162].

density from the simulation, we used a known function [$f(x, y) = \sin x + \sin y$, for example] and applied the same field solver to it. This test confirmed that the Fourier solving method accurately calculates ∇f and $\nabla^2 f$. For an additional test of the Fourier solving method, we tested it alongside a more traditional grid-based solver in a 1D plasma simulation and another 1D simulation using the same initial conditions as our final 3D simulation [163]. We did the calculations separately from start to finish in both MKS and Gaussian units. The simulation produced identical results for all quantities in each context. Additionally, we used a classic 1D two-stream instability test with two streams of particles each with $q/m = -1$ with opposite velocities. This test showed the characteristic swirls in phase space [161]. As a final test, we confined two identical particles to a line and verified that they exhibited simple harmonic motion, as is expected given the periodic boundary conditions of the simulation.

We tested for numerical convergence. For a fixed number of particles along one axis N, $\mathbf{E}$, and $\mathbf{B}$ converge as the number of cells along one axis N_x increases. Increasing N_x increases the grid resolution, and so we would expect both $\mathbf{E}$ and $\mathbf{B}$ to converge to some value. This is indeed what we observe. As an important note, one would not expect $\mathbf{E}$ and $\mathbf{B}$ to converge as N increases. This is because n_0 is physically set by the physics of the Universe at a given z. Therefore, increasing N simply increases the domain size of the simulation without altering the electrodynamics of the plasma. This means that for a fixed ratio N/N_x, the simulation should produce identical results, which we also observed. Of course, N should be set high enough to allow

for sufficient interactions between particles, but once $N \approx 16$, any additional increase is not necessary.

Increasing N higher than necessary also means that a higher N_x is needed to increase the accuracy of the results. Finally, it takes about 100 time steps for the simulation to properly equilibrate, so results to this point should be discarded. After this point however, the values of **E** and **B** converge as N_x increases.

We used the uninterpolated **E** and **B** to calculate $\zeta_m = \langle E^2 - B^2 \rangle / \rho_m$. The averaging was calculated by taking the arithmetic average of the values of E^2 and B^2 at each point on the grid. Using this simulation, we estimate $\zeta_m \approx 10^{-13}$ which is eight orders of magnitude lower in absolute than other estimates—this is not surprising, our calculation is obtained for a diffuse plasma on scales well above nuclear length scales—our simulation is not appropriate to test the absolute scale of ζ_m from nuclei in the early Universe.

It does give us the tools to test the claim that the electrostatic contribution to ζ_m cancels out in a specific environment that we can directly simulate. Our conclusions may need to be revised to properly account for nuclear scales, but offer an interesting first test of the (analytic) claims of electrostatic shielding made in Ref. [135].

In contrast to that approach, we calculated all the electric and magnetic fields generated through inhomogeneities in a neutral plasma like the early universe. We did not consider nucleons or macroscopic objects as sources for fields. We calculated **E** and **B** directly in the simulation using Eqs. (A7)–(A9). However, this simulation differed from that used to estimate ζ_m because we introduced ψ as a scalar field responsible for α variation. This required us to alter our equation for **E**:

$$\mathbf{E} = -e^{2\psi} \nabla \Phi. \qquad (A10)$$

Because $e^{\psi} = \sec(\kappa \Phi)$, we were also able to calculate the new **E** directly in the simulation. **B** is unchanged according to Ref. [135]. Additionally, because κ is a free parameter, we estimate it as 8.11×10^{-26} cm$^{1/2}$ erg$^{-1/2}$ [135]. While this has an effect on the precise values of the terms in Eq. (A1), it does not affect cancellation.

We thus calculated all field quantities in Eq. (A1) directly. The only approximation made was in the delta function in Eq. (A1):

$$\delta^3(\mathbf{x} - \mathbf{z}_i) \approx \frac{1}{a\sqrt{\pi}} \exp\left(-\frac{(\mathbf{x} - \mathbf{z}_i)^2}{a^2}\right) \qquad (A11)$$

in order to properly account for finite grid resolution. The other parameters of the simulation were the same as in our simulation to estimate ζ_m.

Fig. 12 shows our results for our test of Bekenstein's Cancellation Theorem. The combined rhs is only marginally different from the E^2 term. On average, the E^2 term is
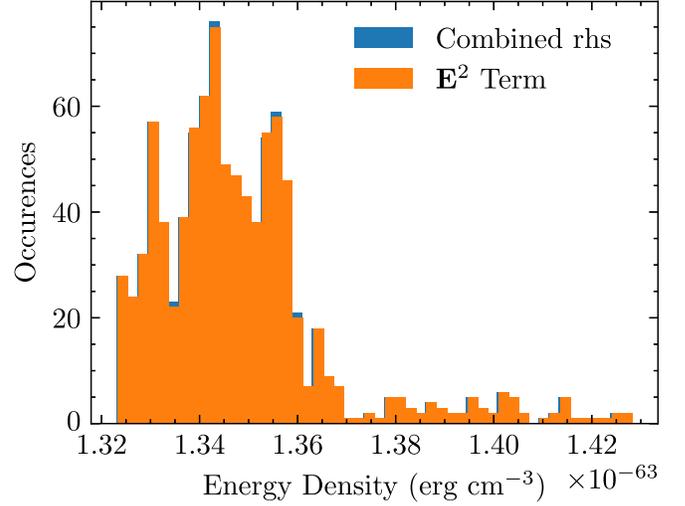

FIG. 12. The values of the E^2 term and combined rhs in the equation of motion for ψ for a recombination density neutral plasma. The $\partial m_i / \partial \psi$ term is not plotted because it is four orders of magnitude less than the other terms. Simulation used 4096 particles on a $128 \times 128 \times 128$ grid over 1000 time steps. The first 100 time steps are not plotted.

about 1×10^{-63} erg cm^{-3}, as is the combined rhs. However the $\partial m_i / \partial \psi$ term is about -2×10^{-67} erg cm^{-3}, about four orders of magnitude less than the E^2 term. Because this term is so much smaller, it is not plotted in Fig. 12. Because of the large difference between the E^2 term and the $\partial m_i / \partial \psi$ term, our results do not support cancellation.

In order to further test the cancellation theorem, we further simulated a neutral plasma at a nuclear density $n_0 = 1.2 \times 10^{38}$ protons/cm^3. Here, our simulation still does not support cancellation. In this case, the two terms in Eq. (A1) are 24 orders of magnitude apart. The E^2 term is about 10^{-16} erg cm^{-3} and the $dm/d\psi$ term is about 10^{-40} erg cm^{-3}.

Finally, we repeated these calculations for a proton-only plasma at the same density as the recombination-density neutral plasma. Bekenstein's arguments center on the properties of charged macroscopic objects, so it is possible that for a plasma with a net charge, the cancellation theorem holds. First, we recalculate ζ_m. For the positive plasma with a $128 \times 128 \times 128$ grid, $\zeta_m \approx 10^{-14}$, which is slightly lower than for a neutral plasma. The results for cancellation are in Fig. 13. Here, the two terms from Eq. (A1) still differ significantly. The $dm/d\psi$ term is about -10^{-66} erg cm^{-3} and the E^2 term is about 10^{-63} erg cm^{-3}. While the two terms still differ greatly, Fig. 13 demonstrates that cancellation is more plausible in this case, as there is some difference between each histogram, as we would expect if cancellation is occurring.

One possible reason for our observed lack of cancellation is Bekenstein's reliance on approximations to convert between microscopic and macroscopic scales. We considered
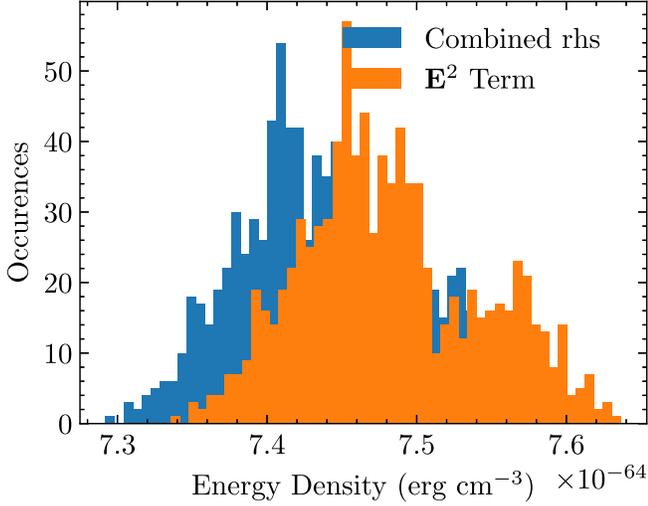
FIG. 13. The values of the E^2 term and combined rhs in the equation of motion for ψ in a recombination density positive plasma. The $\partial m_i/\partial\psi$ term is not plotted because it is two orders of magnitude less than the other terms. Simulation used 4,096 particles on a $128 \times 128 \times 128$ grid over 1000 time steps. The first 100 time steps are not plotted.

instead fields calculated directly in the simulation. Some of Bekenstein's approximations to perform this conversion are robust. We tested the approximation of the second term of Eq. (A1), which is given by Eq. (A6). In the recombination-density and nuclear density simulations, the approximation is off by at most an order of magnitude, indicating the approximation is reasonably robust. In turn, this suggests that the problem lies with the $dm/d\psi$ term in Eq. (A1). This approximation relies more on the analytic specifics of Bekenstein's theory and is therefore more difficult to test computationally. Further work should be aimed at testing the robustness of the analytic derivations of this term.

APPENDIX B: ENERGY CONSERVATION

The equations of motion used earlier (and throughout the BSBM literature) do not conserve energy. In particular, using the fact that the field energy density is $\bar{\omega}\dot{\psi}^2/2$ and the field equation of motion [Eq. (4)], we see that $\dot{\rho}_\psi = -3\bar{\omega}\dot{\psi}^2 H - 2e^{-2\psi}\zeta_m\dot{\psi}$. Given that a field with no potential has equation-of-state parameter $w_\psi = 1$, this can be rewritten as

$$\dot{\rho}_\psi + 3(1+w_\psi)\rho_\psi = -2\sqrt{\frac{2}{\bar{\omega}}}e^{-2\psi}\zeta_m\rho_m\sqrt{\rho_\psi}. \quad \text{(B1)}$$

The left-hand side of Eq. (B1) is of course the standard term for energy flow out a fixed physical volume in an expanding universe. The right-hand side is an energy flow term from the scalar field into matter.

The usual ΛCDM continuity equation for matter density,

$$\dot{\rho}_m + 3H\rho_m = 0, \quad \text{(B2)}$$

must acquire an additional term if energy conservation is to hold, and so we have that

$$\dot{\rho}_m + 3H\rho_m - 2\sqrt{\frac{2}{\bar{\omega}}}e^{-2\psi}\zeta_m\rho_m\sqrt{\rho_\psi} = 0. \quad \text{(B3)}$$

There is nothing *ad hoc* about using Eq. (B3). The additional term on the right-hand side of Eq. (B1) comes from the substitution $\mathcal{L}_{em} \to \zeta\rho_m$—this amounts to saying that the electrostatic influence of charged particles contributing to $\mathcal{L}_{em}$ leads to an additional interaction between ψ and matter, in some sense, integrating out the relevant electromagnetic fields. These fields should backreact on the plasma that sources them, and if we are to take this Lagrangian substitution at face value, the Bianchi identity implies the $\nabla_\mu(e^{-2\psi}T^{\mu\nu,\text{matter}}) = 0$. This can be readily applied in a Freedman-Robertson-Walker (FRW) cosmology to obtain Eq. (B3).

This modified equation for matter density and the second-order differential equation for the scalar field can then be written as a system of three first-order ordinary differential equations as follows:

$$\frac{df}{dz} = \frac{3}{1+z}f + \frac{2}{g(1+z)}e^{-2\psi}\zeta_m f u, \quad \text{(B4)}$$

$$\frac{d\psi}{dz} = -\frac{u}{g(1+z)}, \quad \text{(B5)}$$

$$\frac{du}{dz} = \frac{3}{1+z}u + \frac{6\Omega_{m,0}}{g(1+z)}\frac{1}{\omega}e^{-2\psi}\zeta_m f, \quad \text{(B6)}$$

where $f(z) = \Omega_m(z)/\Omega_{m,0}$, $u(z) = -\dot{\psi}/[H_0 g(1+z)]$ is a dimensionless velocity for the scalar field and

$$g^2 = \Omega_{m,0}f(1+|\zeta_m|e^{-2\psi}) + \Omega_{r,0}(1+z)^4 e^{-2\psi} + \frac{\omega}{6}u^2 + \Omega_{\Lambda,0} \quad \text{(B7)}$$

represents the modified dimensionless Hubble parameter.

This system of differential equations was then solved numerically over a range of redshifts and compared to the non energy-conserving implementation with $\zeta_m = 10^{-8}$ and $\omega = 1$. The left panel of Fig. 14 shows the fractional difference in the evolution of $\Delta\alpha/\alpha$ over z for these two models. Agreement is sufficient for our purposes (a $\sim$10% correction to our constraints). We also compared $\Delta\alpha/\alpha$ values at recombination for a range of ζ_m values from -10^{-8} to 10^{-8} (roughly the constraint from *Planck*), with ω still set to 1. The results are displayed in the right panel of Fig. 14, which shows that the two calculations agree well
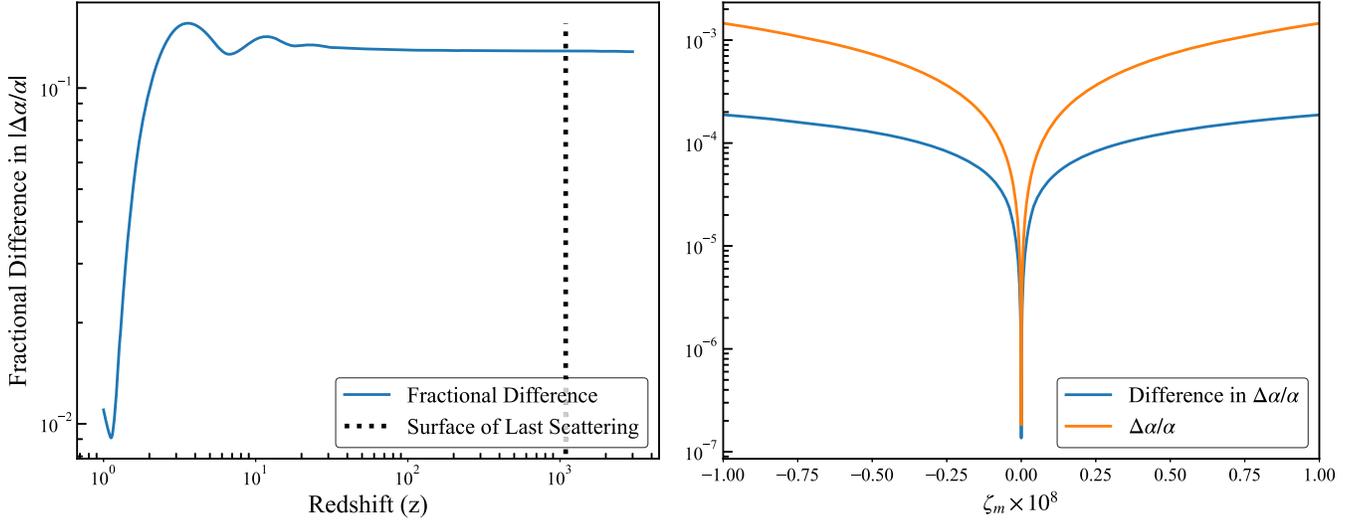
FIG. 14.   Left panel: fractional error as a function of z between energy-conserving and nonconserving BSBM model implementations with overall level of fine-structure constant variation, $\Delta\alpha/\alpha$. The parameter $\zeta_m = 10^{-8}$, while $\omega = 1$, consistent with the *Planck* constraint. Right panel: comparison of error at recombination between energy-conserving and nonconserving BSBM model implementations with overall level of fine-structure constant variation, $\Delta\alpha/\alpha$. The parameter ζ_m is varied in the range -10^{-8} to 10^{-8}, while $\omega = 1$.

over this range of ζ_m, and vanish when $\zeta_m \to 0$. The agreement improves with even higher $|\zeta_m|$.

It is worth pausing to consider the interpretation of these corrections. In particular, we can state the matter continuity equation as

$$\dot{\rho}_m + 3H\rho_m = -2e^{-2\psi}\zeta_m\dot{\psi}\rho_m. \tag{B8}$$

With the ansatz $\rho \equiv a^{-3}f(a)$, we find (applying separation of variables) that

$$\rho_m = \frac{C}{a^3}e^{\zeta_m e^{-2\psi}}. \tag{B9}$$

This scaling suggests the definition $\tilde{\rho}_m \equiv \rho_m e^{-\zeta_m e^{-2\psi}}$ as a physical matter density with the right redshift dependence. The matter-dependent term in the Friedmann equation (in the BSBM model) is [29,30]

$$H_m^2 = \frac{8\pi G}{3}\{\rho_m + \zeta_m e^{2\psi}\rho_m\}, \tag{B10}$$

$$= \frac{8\pi G}{3}e^{\zeta_m e^{-2\psi}}[1 + \zeta_m e^{2\psi}]\tilde{\rho}_m, \tag{B11}$$

$$= \frac{8\pi G\tilde{\rho}_m}{3}\frac{m_m(\psi)}{m_m(\psi = 0)}, \tag{B12}$$

where $\tilde{m}_m$ can be interpreted as the ψ-modulated mass of nonrelativistic particles (e.g. baryons, dark matter). If ρ_m is treated (erroneously) as a substance that redshifts as a^{-3}, then this amounts to the approximation $m_m(\psi) = m_m(\psi = 0)$. The analysis above shows that the error in α

evolution induced by this approximation in the BSBM is negligible.

This effective modulation of the mass of nonrelativistic particles is a well-known feature of theories with non-minimally coupled scalar fields, in which the relevant matter fields travel on geodesics of a different metric than the one satisfying Einstein's equations (see e.g. Refs. [164,165] for recent applications to early dark energy phenomenology and Refs. [166–169] for earlier applications).

In Sec. II B, the contribution of ψ to the Hubble expansion itself was not included. The dynamics modeled
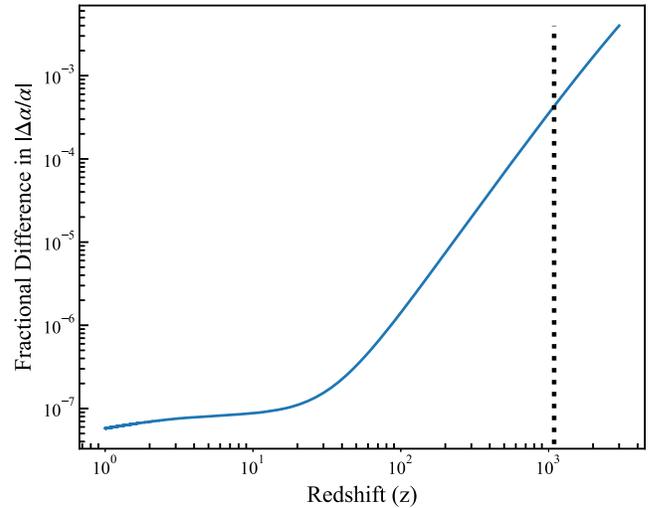

FIG. 15.   Fractional difference in $|\Delta\alpha/\alpha|$ between BSBM model implementation with and without contribution of ψ to the Friedmann equation, as a function of redshift z. Here $\zeta_m = 10^{-8}$ and $\omega = 1$.

above do include this contribution, and we have verified that the overall correction to $\Delta\alpha/\alpha$ does not affect the remaining results of this paper. It is instructive to consider the contribution of ψ to the Hubble expansion (separate and apart from the energy nonconservation or modulated scalar mass effect discussed above). The results are shown in Fig. 15. We see that this effect (taken alone) is even smaller than the energy conservation correction.

Additionally, Eq. (B7) is only valid for massless neutrinos, whereas the fiducial cosmology used in *Planck* data analysis assumes a single neutrino with $m_\nu = 0.06$ eV. Using the equations in Ref. [170], which correct the Hubble factor analytically for the contribution of massive neutrinos (through their full relativistic to nonrelativistic transition), we assess if neutrinos induce any change to the values of α determined in the preceding portions of this section. We find a fractional error in $\Delta\alpha/\alpha$ at $z = 1100$ of $\sim 10^{-3}$, meaning that this effect does not alter any of the conclusions of this paper.

[1] B. A. Benson *et al.* (SPT-3G Collaboration), Proc. SPIE Int. Soc. Opt. Eng. **9153**, 91531P (2014).

[2] J. Frieman, M. Turner, and D. Huterer, Annu. Rev. Astron. Astrophys. **46**, 385 (2008).

[3] M. Kowalski *et al.* (Supernova Cosmology Project Collaboration), Astrophys. J. **686**, 749 (2008).

[4] N. Aghanim *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A6 (2020); **652**, C4(E) (2021).

[5] J. C. Hill and E. J. Baxter, J. Cosmol. Astropart. Phys. 08 (2018) 037.

[6] V. Poulin, T. L. Smith, D. Grin, T. Karwal, and M. Kamionkowski, Phys. Rev. D **98**, 083525 (2018).

[7] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, Phys. Rev. Lett. **122**, 221301 (2019).

[8] T. L. Smith, V. Poulin, and M. A. Amin, Phys. Rev. D **101**, 063523 (2020).

[9] T. L. Smith, V. Poulin, J. L. Bernal, K. K. Boddy, M. Kamionkowski, and R. Murgia, Phys. Rev. D **103**, 123542 (2021).

[10] T. L. Smith, M. Lucca, V. Poulin, G. F. Abellan, L. Balkenhol, K. Benabed, S. Galli, and R. Murgia, Phys. Rev. D **106**, 043526 (2022).

[11] A. G. Riess *et al.*, Astrophys. J. **826**, 56 (2016).

[12] F. Renzi, E. Di Valentino, and A. Melchiorri, Phys. Rev. D **97**, 123534 (2018).

[13] J. Silk, Astrophys. J. **151**, 459 (1968).

[14] W. T. Hu, Wandering in the background: A CMB explorer, Other thesis, University of California Berkeley, 1995.

[15] D. Grin and C. M. Hirata, Phys. Rev. D **81**, 083005 (2010).

[16] J. Chluba and R. M. Thomas, Mon. Not. R. Astron. Soc. **412**, 748 (2011).

[17] Y. Ali-Haïmoud and C. M. Hirata, Phys. Rev. D **83**, 043513 (2011).

[18] J. Chluba and R. A. Sunyaev, Astron. Astrophys. **446**, 39 (2006).

[19] P. A. M. Dirac, Nature (London) **139**, 323 (1937).

[20] E. Teller, Phys. Rev. **73**, 801 (1948).

[21] G. Gamow, Phys. Rev. Lett. **19**, 759 (1967).

[22] G. Gamow, Phys. Rev. Lett. **19**, 913 (1967).

[23] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).

[24] J. D. Bekenstein, Phys. Rev. D **25**, 1527 (1982).

[25] C. J. Isham, A. Salam, and J. A. Strathdee, Phys. Rev. D **3**, 1805 (1971).

[26] H. Terazawa, Y. Chikashige, K. Akama, and T. Matsuki, Phys. Rev. D **15**, 1181 (1977).

[27] H. Terazawa, Phys. Lett. **101B**, 43 (1981).

[28] H. Terazawa, Nonlinear Phenom. Complex Syst. **20**, 241 (2014).

[29] H. B. Sandvik, J. D. Barrow, and J. Magueijo, Phys. Rev. Lett. **88**, 031302 (2002).

[30] J. D. Barrow and D. F. Mota, Classical Quantum Gravity **20**, 2045 (2003).

[31] C. J. A. P. Martins, Rep. Prog. Phys. **80**, 126902 (2017).

[32] J. D. Barrow and B. Li, Phys. Rev. D **78**, 083536 (2008).

[33] J. D. Barrow and S. Z. W. Lip, Phys. Rev. D **85**, 023514 (2012).

[34] J. D. Barrow, J. Magueijo, and H. B. Sandvik, Phys. Rev. D **66**, 043515 (2002).

[35] D. F. Mota, Variations of the fine structure constant in space and time, Other thesis, University of Cambridge, 2004.

[36] D. F. Mota and J. D. Barrow, Mon. Not. R. Astron. Soc. **349**, 291 (2004).

[37] J. D. Barrow, Phys. Rev. D **71**, 083520 (2005).

[38] J. D. Barrow and A. A. H. Graham, Phys. Rev. D **88**, 103513 (2013).

[39] O. Klein, Z. Phys. **37**, 895 (1926).

[40] A. Chodos and S. L. Detweiler, Phys. Rev. D **21**, 2167 (1980).

[41] W. J. Marciano, Phys. Rev. Lett. **52**, 489 (1984).

[42] E. W. Kolb, M. J. Perry, and T. P. Walker, Phys. Rev. D **33**, 869 (1986).

[43] T. Damour and A. M. Polyakov, Nucl. Phys. **B423**, 532 (1994).

[44] T. Damour, F. Piazza, and G. Veneziano, Phys. Rev. Lett. **89**, 081601 (2002).

[45] T. Damour, F. Piazza, and G. Veneziano, Phys. Rev. D **66**, 046007 (2002).

[46] J. M. Overduin and P. S. Wesson, Phys. Rep. **283**, 303 (1997).

[47] P. Brax, C. van de Bruck, A. C. Davis, and C. S. Rhodes, Astrophys. Space Sci. **283**, 627 (2003).

[48] G. A. Palma, P. Brax, A. C. Davis, and C. van de Bruck, Phys. Rev. D **68**, 123519 (2003).

[49] C. van de Bruck and J. Morrice, J. Cosmol. Astropart. Phys. 04 (2015) 036.

[50] C. van de Bruck, J. Mifsud, and N. J. Nunes, J. Cosmol. Astropart. Phys. 12 (2015) 018.

[51] C. Wetterich, Nucl. Phys. **B302,** 668 (1988).

[52] E. J. Copeland, N. J. Nunes, and M. Pospelov, Phys. Rev. D **69,** 023501 (2004).

[53] D. Parkinson, B. A. Bassett, and J. D. Barrow, Phys. Lett. B **578,** 235 (2004).

[54] A. A. H. Graham, Classical Quantum Gravity **32,** 015019 (2015).

[55] K. A. Olive and M. Pospelov, Phys. Rev. D **65,** 085044 (2002).

[56] K. A. Olive and M. Pospelov, Phys. Rev. D **77,** 043524 (2008).

[57] Y. V. Stadnik and V. V. Flambaum, Phys. Rev. Lett. **114,** 161301 (2015).

[58] Y. V. Stadnik and V. V. Flambaum, Phys. Rev. Lett. **115,** 201301 (2015).

[59] J.-P. Uzan, Rev. Mod. Phys. **75,** 403 (2003).

[60] C. J. A. P. Martins and A. M. M. Pinho, Phys. Rev. D **95,** 023008 (2017).

[61] J. K. Webb, V. V. Flambaum, C. W. Churchill, M. J. Drinkwater, and J. D. Barrow, Phys. Rev. Lett. **82,** 884 (1999).

[62] M. T. Murphy, J. K. Webb, V. V. Flambaum, C. W. Churchill, and J. X. Prochaska, Mon. Not. R. Astron. Soc. **327,** 1223 (2001).

[63] T. Chiba, T. Kobayashi, M. Yamaguchi, and J. Yokoyama, Phys. Rev. D **75,** 043516 (2007).

[64] J. K. Webb, J. A. King, M. T. Murphy, V. V. Flambaum, R. F. Carswell, and M. B. Bainbridge, Phys. Rev. Lett. **107,** 191101 (2011).

[65] J. A. King, Searching for variations in the fine-structure constant and the proton-to-electron mass ratio using quasar absorption lines, Ph.D. thesis, New South Wales U., 2010.

[66] J. A. King, J. K. Webb, M. T. Murphy, V. V. Flambaum, R. F. Carswell, M. B. Bainbridge, M. R. Wilczynska, and F. E. Koch, Mon. Not. R. Astron. Soc. **422,** 3370 (2012).

[67] P. Molaro, D. Reimers, I. I. Agafonova, and S. A. Levshakov, Eur. Phys. J. Special Topics **163,** 173 (2008).

[68] P. Molaro *et al.*, Astron. Astrophys. **555,** A68 (2013).

[69] A. Songaila and L. L. Cowie, Astrophys. J. **793,** 103 (2014).

[70] S. M. Kotuš, M. T. Murphy, and R. F. Carswell, Mon. Not. R. Astron. Soc. **464,** 3679 (2017).

[71] M. T. Murphy, A. L. Malec, and J. X. Prochaska, Mon. Not. R. Astron. Soc. **461,** 2461 (2016).

[72] M. T. Murphy and K. L. Cooksey, Mon. Not. R. Astron. Soc. **471,** 4930 (2017).

[73] T. M. Schmidt *et al.*, Astron. Astrophys. **646,** A144 (2021).

[74] M. R. Wilczynska *et al.*, Sci. Adv. **6,** eaay9672 (2020).

[75] M. T. Murphy *et al.*, Astron. Astrophys. **658,** A123 (2022).

[76] W. Hu and N. Sugiyama, Astrophys. J. **444,** 489 (1995).

[77] W. Hu and S. Dodelson, Annu. Rev. Astron. Astrophys. **40,** 171 (2002).

[78] K. T. Story *et al.*, Astrophys. J. **779,** 86 (2013).

[79] R. A. Battye, R. Crittenden, and J. Weller, Phys. Rev. D **63,** 043505 (2001).

[80] 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, Phys. Rev. D **64,** 103505 (2001).

[81] C. J. A. P. Martins, A. Melchiorri, R. Trotta, R. Bean, G. Rocha, P. P. Avelino, and P. T. P. Viana, Phys. Rev. D **66,** 023505 (2002).

[82] G. Rocha, R. Trotta, C. J. A. P. Martins, A. Melchiorri, P. P. Avelino, and P. T. P. Viana, New Astron. Rev. **47,** 863 (2003).

[83] C. J. A. P. Martins, A. Melchiorri, G. Rocha, R. Trotta, P. P. Avelino, and P. T. P. Viana, Phys. Lett. B **585,** 29 (2004).

[84] K. Ichikawa, T. Kanzaki, and M. Kawasaki, Phys. Rev. D **74,** 023515 (2006).

[85] E. Menegoni, S. Galli, J. G. Bartlett, C. J. A. P. Martins, and A. Melchiorri, Phys. Rev. D **80,** 087302 (2009).

[86] E. Menegoni, M. Archidiacono, E. Calabrese, S. Galli, C. J. A. P. Martins, and A. Melchiorri, Phys. Rev. D **85,** 107301 (2012).

[87] C. J. A. P. Martins, E. Menegoni, S. Galli, G. Mangano, and A. Melchiorri, Phys. Rev. D **82,** 023532 (2010).

[88] M. Kaplinghat, R. J. Scherrer, and M. S. Turner, Phys. Rev. D **60,** 023516 (1999).

[89] S. Hannestad, Phys. Rev. D **60,** 023515 (1999).

[90] G. Rocha, R. Trotta, C. J. A. P. Martins, A. Melchiorri, P. P. Avelino, R. Bean, and P. T. P. Viana, Mon. Not. R. Astron. Soc. **352,** 20 (2004).

[91] A. Hojjati, G.-B. Zhao, L. Pogosian, A. Silvestri, R. Crittenden, and K. Koyama, Phys. Rev. D **85,** 043508 (2012).

[92] A. Hojjati, L. Pogosian, A. Silvestri, and G.-B. Zhao, Phys. Rev. D **89,** 083505 (2014).

[93] K. Koyama, Rep. Prog. Phys. **79,** 046902 (2016).

[94] L. Pogosian and A. Silvestri, Phys. Rev. D **94,** 104014 (2016).

[95] J. Espejo, S. Peirone, M. Raveri, K. Koyama, L. Pogosian, and A. Silvestri, Phys. Rev. D **99,** 023512 (2019).

[96] L. Hart and J. Chluba, Mon. Not. R. Astron. Soc. **474,** 1850 (2018).

[97] L. Hart and J. Chluba, Mon. Not. R. Astron. Soc. **493,** 3255 (2020).

[98] L. Wasserman, C. J. Miller, R. C. Nichol, C. Genovese, W. Jang, A. J. Connolly, A. W. Moore, and J. Schneider (PICA Group), arXiv:astro-ph/0112050.

[99] D. Huterer and G. Starkman, Phys. Rev. Lett. **90,** 031301 (2003).

[100] W. Hu and G. P. Holder, Phys. Rev. D **68,** 023001 (2003).

[101] M. J. Mortonson and W. Hu, Astrophys. J. **672,** 737 (2008).

[102] J. Samsing, E. V. Linder, and T. L. Smith, Phys. Rev. D **86,** 123504 (2012).

[103] D. P. Finkbeiner, S. Galli, T. Lin, and T. R. Slatyer, Phys. Rev. D **85,** 043522 (2012).

[104] V. Miranda and C. Dvorkin, Phys. Rev. D **98,** 043537 (2018).

[105] A. Diaz Rivero, V. Miranda, and C. Dvorkin, Phys. Rev. D **100,** 063504 (2019).

[106] M. Farhang, J. R. Bond, and J. Chluba, Astrophys. J. **752,** 88 (2012).

[107] L. Hart and J. Chluba, Mon. Not. R. Astron. Soc. **495,** 4210 (2020).

[108] L. Hart and J. Chluba, Mon. Not. R. Astron. Soc. **510,** 2206 (2022).

[109] K. Sigurdson, A. Kurylov, and M. Kamionkowski, Phys. Rev. D **68**, 103509 (2003).

[110] J. O'Bryan, J. Smidt, F. De Bernardis, and A. Cooray, Astrophys. J. **798**, 18 (2015).

[111] T. L. Smith, D. Grin, D. Robinson, and D. Qi, Phys. Rev. D **99**, 043531 (2019).

[112] I. de Martino, C. J. A. P. Martins, H. Ebeling, and D. Kocevski, Phys. Rev. D **94**, 083008 (2016).

[113] R. A. Sunyaev and Y. B. Zeldovich, Annu. Rev. Astron. Astrophys. **18**, 537 (1980).

[114] R. A. Sunyaev and Y. B. Zeldovich, Mon. Not. R. Astron. Soc. **190**, 413 (1980).

[115] E. Lee, D. Anbajagane, P. Singh, J. Chluba, D. Nagai, S. T. Kay, W. Cui, K. Dolag, and G. Yepes, Mon. Not. R. Astron. Soc. **517**, 5303 (2022).

[116] J. R. Pritchard and A. Loeb, Rep. Prog. Phys. **75**, 086901 (2012).

[117] R. Khatri and B. D. Wandelt, Phys. Rev. Lett. **98**, 111301 (2007).

[118] L. Lopez-Honorez, O. Mena, S. Palomares-Ruiz, P. Villanueva-Domingo, and S. J. Witte, J. Cosmol. Astropart. Phys. 06 (2020) 026.

[119] P. Svrcek and E. Witten, J. High Energy Phys. 06 (2006) 051.

[120] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, Phys. Rev. D **81**, 123530 (2010).

[121] M. Cicoli, M. Goodsell, and A. Ringwald, J. High Energy Phys. 10 (2012) 146.

[122] R. Hložek, D. Grin, D. J. E. Marsh, and P. G. Ferreira, Phys. Rev. D **91**, 103512 (2015).

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

[124] A. Drlica-Wagner et al., arXiv:2209.08215.

[125] C. J. A. P. Martins et al., Phys. Rev. D **105**, 123507 (2022).

[126] L. Vacher, N. Schöneberg, J. D. F. Dias, C. J. A. P. Martins, and F. Pimenta, Phys. Rev. D **107**, 104002 (2023).

[127] K. N. Abazajian et al. (CMB-S4 Collaboration), arXiv:1610.02743.

[128] P. Ade et al. (Simons Observatory Collaboration), J. Cosmol. Astropart. Phys. 02 (2019) 056.

[129] J. Gasser and H. Leutwyler, Phys. Rep. **87**, 77 (1982).

[130] M. Livio and M. Stiavelli, Astrophys. J. Lett. **507**, L13 (1998).

[131] L. Kraiselburd and H. Vucetich, Int. J. Mod. Phys. E **20**, 101 (2011).

[132] L. Kraiselburd, M. Miller Bertolami, P. Sisterna, and H. Vucetich, Astron. Astrophys. **529**, A125 (2011).

[133] L. Kraiselburd, F. L. Castillo, M. E. Mosquera, and H. Vucetich, Phys. Rev. D **97**, 043526 (2018).

[134] C. M. Will, Theory and Experiment in Gravitational Physics (Cambridge University Press, Cambridge, England, 2018).

[135] J. D. Bekenstein, Phys. Rev. D **66**, 123514 (2002).

[136] E. Segrè, Nuclei and Particles. an Introduction to Nuclear and Subnuclear Physics, 2nd ed. (W. A. Benjamin, Berlin, Heidelberg, and New York, 1977).

[137] K. R. Lang, Astrophysical Formulae (Springer, New York, 1998), Vol. 1.

[138] M. P. Haugan and C. M. Will, Phys. Rev. D **15**, 2711 (1977).

[139] G. J. Cooper and J. H. Verner, SIAM J. Numer. Anal. **9**, 389 (1972).

[140] M. Tegmark, A. Taylor, and A. Heavens, Astrophys. J. **480**, 22 (1997).

[141] W. Hu, D. J. Eisenstein, and M. Tegmark, Phys. Rev. Lett. **80**, 5255 (1998).

[142] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).

[143] J. Chluba and R. M. Thomas, CosmoRec: Cosmological Recombination code, Astrophysics Source Code Library, record ascl:1304.017 (2013).

[144] D. Hanson, A. Challinor, and A. Lewis, Gen. Relativ. Gravit. **42**, 2197 (2010).

[145] C. H. Heinrich, V. Miranda, and W. Hu, Phys. Rev. D **95**, 023513 (2017).

[146] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Publ. Astron. Soc. Pac. **125**, 306 (2013).

[147] A. Lewis, arXiv:1910.13970.

[148] C. H. Heinrich, D. Grin, and W. Hu, Phys. Rev. D **94**, 043534 (2016).

[149] T. L. Smith, J. B. Muñoz, R. Smith, K. Yee, and D. Grin, Phys. Rev. D **96**, 083508 (2017).

[150] Z. Yuan, D. Grin, and T. L. Smith, Testing matter-correlated variation in the fundamental parameters using the CMB (to be published).

[151] J. D. Barrow and J. Magueijo, Phys. Rev. D **72**, 043521 (2005).

[152] A. Albrecht and J. Magueijo, Phys. Rev. D **59**, 043516 (1999).

[153] J. Magueijo, Phys. Rev. D **62**, 103521 (2000).

[154] J. Magueijo, Rep. Prog. Phys. **66**, 2025 (2003).

[155] C. He, D. Grin, and W. Hu, Phys. Rev. D **92**, 063018 (2015).

[156] K. Abazajian et al., arXiv:1907.04473.

[157] N. Sehgal et al., Bull. Am. Astron. Soc. **51**, 6 (2019).

[158] L. Hart and J. Chluba, Mon. Not. R. Astron. Soc. **519**, 3664 (2023).

[159] C. R. Harris et al., Nature (London) **585**, 357 (2020).

[160] P. Virtanen (SciPy 1.0 Contributors), Nat. Methods **17**, 261 (2020).

[161] C. K. Birdsall, Plasma Physics via Computer Simulation, The Adam Hilger Series on Plasma Physics (Adam Hilger, Bristol, Eng., 1991).

[162] B. Ripperda, F. Bacchini, J. Teunissen, C. Xia, O. Porth, L. Sironi, G. Lapenta, and R. Keppens, Astrophys. J. Suppl. Ser. **235**, 21 (2018).

[163] P. Mocz, Create Your Own Plasma PIC Simulation (With Python) (medium.com, 2011).

[164] T. Karwal, M. Raveri, B. Jain, J. Khoury, and M. Trodden, Phys. Rev. D **105**, 063535 (2022).

[165] M.-X. Lin, E. McDonough, J. C. Hill, and W. Hu, Phys. Rev. D **107**, 103523 (2023).

[166] R. Bean and J. Magueijo, Phys. Lett. B **517**, 177 (2001).

[167] R. Bean, E. E. Flanagan, and M. Trodden, Phys. Rev. D **78**, 023009 (2008).

[168] R. Bean, E. E. Flanagan, I. Laszlo, and M. Trodden, Phys. Rev. D **78**, 123514 (2008).

[169] V. Miranda, M. Carrillo González, E. Krause, and M. Trodden, Phys. Rev. D **97**, 063511 (2018).

[170] E. Komatsu et al. (WMAP Collaboration), Astrophys. J. Suppl. Ser. **192**, 18 (2011).