Priors on the effective dark energy equation of state in scalar-tensor theories

Marco Raveri,[1,2] Philip Bull,[3,4] Alessandra Silvestri,[2] and Levon Pogosian[5]

[1]*Kavli Institute for Cosmological Physics, Enrico Fermi Institute,*
University of Chicago, Chicago, Illinois 60637, USA
[2]*Institute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, Netherlands*
[3]*California Institute of Technology, Pasadena, California 91125, USA*
[4]*Jet Propulsion Laboratory, California Institute of Technology,*
4800 Oak Grove Drive, Pasadena, California, USA
[5]*Department of Physics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada*
(Received 25 March 2017; published 11 October 2017)

Constraining the dark energy (DE) equation of state, w_{DE}, is one of the primary science goals of ongoing and future cosmological surveys. In practice, with imperfect data and incomplete redshift coverage, this requires making assumptions about the evolution of w_{DE} with redshift z. These assumptions can be manifested in a choice of a specific parametric form, which can potentially bias the outcome, or else one can reconstruct $w_{DE}(z)$ nonparametrically, by specifying a prior covariance matrix that correlates values of w_{DE} at different redshifts. In this work, we derive the theoretical prior covariance for the effective DE equation of state predicted by general scalar-tensor theories with second order equations of motion (Horndeski theories). This is achieved by generating a large ensemble of possible scalar-tensor theories using a Monte Carlo methodology, including the application of physical viability conditions. We also separately consider the special subcase of the minimally coupled scalar field, or quintessence. The prior shows a preference for tracking behaviors in the most general case. Given the covariance matrix, theoretical priors on parameters of any specific parametrization of $w_{DE}(z)$ can also be readily derived by projection.

I. INTRODUCTION

The cosmological constant Λ is the simplest form of dark energy (DE) capable of driving the observed accelerating cosmic expansion [1,2]. In general relativity (GR), the bare (geometric) cosmological constant is a free parameter. However, Λ also includes a constant vacuum energy density that receives contributions from the zero-point fluctuations of known particles [3]. To match the observed value, the geometric contribution to Λ needs to be separately fine-tuned to compensate for the contribution of each different particle type individually, making the fine-tuning technically unnatural [4].

The extraordinary degree of tuning required to explain the observed value of Λ within GR has spurred a broad search for mechanisms that can naturally set Λ to a small value. A common proposal involves finding ways to prevent the vacuum from gravitating altogether, as in the so-called "degravitation" [5], "self-tuning" [6], or "sequestering" [7] scenarios, which would solve the naturalness problem at the cost of invoking some new mechanism—generally a new field or force—to cause the cosmic acceleration. These closely related issues have motivated studies of alternative models of DE and modifications of gravity (MG). While not necessarily successful in resolving the fine-tuning problem themselves, many of these models can account for the acceleration, often predicting an effective DE that is dynamical. As the quest for a complete understanding of cosmic acceleration continues, it is of significant interest to know the extent to which DE can be dynamical and still remain consistent with observations.

Given the values of the Hubble parameter today, H_0, and the current matter density fraction, Ω_M, an arbitrary expansion history can be reproduced by assuming a flat Friedmann-Robertson-Walker (FRW) universe with a DE component that has an equation of state (EoS) w_{DE} with an appropriately chosen dependence on redshift [8]. Constraining $w_{DE}(z)$, which we define here as the effective EoS of all nondust contributions to the Friedmann equation at late times, is one of the primary science goals of ongoing and future surveys of large scale structure. A detection of $w_{DE}(z) \neq -1$ would be strong evidence of new gravitational physics and would constitute a potentially vital clue in understanding the source of cosmic acceleration.

Fitting a constant w_{DE} to current data results in good agreement with -1 [9], but such fits would have missed subtle variations in $w_{DE}(z)$, especially if the average happened to be close to -1. By the same token, using the Chevallier-Polarski-Linder (CPL) parametrization [10,11], or any particular parametric form of $w_{DE}(z)$, is prone to biasing the outcome [12]. Forecasts using principle component analysis (PCA) [13,14] have shown that future surveys like Euclid and LSST will be able to constrain several eigenmodes of $w_{DE}(z)$, justifying the use of more flexible parametrizations.

A relatively general way to proceed is to use a piecewise representation, defining $w_{DE}(z)$ in terms of its values in

discrete bins in z. In the limit of many bins, fitting such a parametrization to data would avoid any bias, as any arbitrary $w_{DE}(z)$ could be reproduced. In practice, however, if more than a few bins are used, many of them will be left unconstrained by the data, with values in neighboring bins effectively becoming degenerate. A possible way to remedy this problem is to introduce correlations between the bins by supplementing the likelihood with a prior covariance [15]. In this approach, the bias in the reconstructed $w_{DE}(z)$ can be controlled for a given survey by using a PCA forecast. Specifically, as demonstrated in [15,16], one can tune the strength of the prior to avoid the bias in the reconstructed evolution of $w_{DE}(z)$ on timescales larger than a given correlation scale. A similar effect can be achieved using Gaussian processes [17,18], which model correlations in the value of w_{DE} as a smooth function of redshift. Another approach, taken in [17,19,20], attempts to derive $w_{DE}(z)$ directly from data, by taking derivatives of the luminosity distances $d_L(z)$. While seemingly model independent, such methods require smoothing of the (noisy) data before derivatives can be taken, which effectively amounts to adopting a correlation prior.

Since any attempt to constrain $w_{DE}(z)$ will require a prior in one form or the other, one would like to develop reconstruction tools based on theoretically motivated priors that are explicit and easy to interpret. Scalar fields (fundamental or effective) are ubiquitous in theoretical cosmology and particle physics, and are well suited for representing a dynamical DE. A natural way to construct a theoretically consistent, yet not-too-model-specific prior, is by generating a representative ensemble of scalar field models. Such approach was taken, for example, in [21,22], where priors on the CPL parameters w_0 and w_a were derived based on an ensemble of very general quintessence DE models, i.e., minimally coupled scalar fields [8,23]. In this paper, we go beyond minimally coupled scalars and consider the most general scalar-tensor models that are "stable", in the sense that they lack ghosts, gradient instabilities, and other pathologies. To generate the ensemble of models, we use the so-called "unifying" or "effective field theory" (EFT) approach to DE and MG [24–30]. Rather than working with w_0 and w_a, we derive the prior covariance for a binned $w_{DE}(z)$, which makes it possible to project onto other parametrizations, and allows for reconstruction of $w_{DE}(z)$ using the methods introduced in [15,16].

We generate the prior covariance of $w_{DE}(z)$ for Horndeski theories, i.e., the general class of scalar-tensor theories with up to second order equations of motion [31,32]. We also separately consider the widely studied subclasses of the minimally couple scalar field, or quintessence [8,23], and models with the canonical form of the scalar field kinetic energy, i.e., the generalized Brans-Dicke (GBD) models [33,34]. A nonminimally coupled scalar field mediates a fifth force which is tightly constrained in the Solar System. We will not concern ourselves with

satisfying such constraints, as the models may include a screening mechanism that suppresses the fifth force in dense environments. For our purposes, a scalar-tensor theory is viable if it has stable cosmological perturbations.

The paper is organized as follows. In Sec. II, we introduce the EFT description of the cosmological background and linear perturbations in scalar-tensor theories and discuss a couple of representative forms of $w_{DE}(z)$. We detail our method for sampling the space of viable Horndeski models in Sec. III. In Sec. IV, we present our results for the prior probability distribution for $w_{DE}(z)$ and its covariance. We also provide an analytical form that accurately represents the correlation of $w_{DE}(z)$ at any two points in z. We conclude with a summary in Sec. V.

II. MODELING VIABLE SCALAR FIELD DARK ENERGY

We focus on the broad class of scalar-tensor models of gravity with second order equations of motion. The corresponding action in $(3+1)$ dimensions was derived by Horndeski [31] and later rediscovered in the context of generalized Galileons [32]. For the purpose of our analysis, rather than working with the full covariant action of Horndeski gravity, it is convenient to employ the unifying framework, or effective theory approach (EFT), formulated in [24,25] and further developed in [26–30]. This allows us to model the background evolution and linear perturbations in a model independent way, in terms of a handful of free functions of time. The relevant EFT action reads

$$
\begin{aligned}
S = \int d^4x \sqrt{-g} \bigg\{ & \frac{m_0^2}{2}[1+\Omega(\tau)]R + \Lambda(\tau) - c(\tau)a^2\delta g^{00} \\
& + \frac{M_2^4(\tau)}{2}(a^2\delta g^{00})^2 - \frac{\bar{M}_1^3(\tau)}{2}a^2\delta g^{00}\delta K_\mu^\mu \\
& + \frac{\bar{M}_3^2(\tau)}{2}\left[(\delta K_\mu^\mu)^2 - \delta K_\nu^\mu \delta K_\mu^\nu - \frac{a^2}{2}\delta g^{00}\delta\mathcal{R}\right] + \ldots \bigg\} \\
& + S_m[g_{\mu\nu}, \chi_m],
\end{aligned}
\tag{1}
$$

where R is the four-dimensional Ricci scalar, δg^{00}, δK_ν^μ, δK_μ^μ, and $\delta\mathcal{R}$ are, respectively, the perturbations of the upper time-time component of the metric, the extrinsic curvature, its trace, and the three-dimensional spatial Ricci scalar of the constant-time hypersurfaces. The six functions, $\{\Omega, \Lambda, c, M_2^4, \bar{M}_1^3, \bar{M}_3^2\}$, are arbitrary functions of time allowed by the breaking of the time-diffeomorphism invariance, to which we refer to as the "EFT functions". Finally, S_m is the action for all matter fields χ_m minimally coupled to the metric $g_{\mu\nu}$. Action (1) is built in the unitary gauge, where the additional scalar field is absorbed into the metric and represents an expansion to quadratic order in perturbations around the flat Friedmann-Robertson-Walker (FRW) universe. A few extra terms were present in the original formulation of EFT of dark energy [24,25], and different generalizations of (1) have recently been studied

in the literature [35–37] to include models outside the Horndeski class.

Perturbations in any specific Horndeski model can be mapped onto the action (1), with the correspondence between the EFT functions and the functions that appear in the full Lagrangian of the model given in [27]. In the absence of a preferred scalar-tensor theory, one can adopt the agnostic point of view and treat the EFT functions as free functions of time. We note that the GBD theories are represented by terms in the first line of (1), while quintessence only requires specifying $\Lambda(\tau)$, with all other functions set to zero.

One of the advantages of using the EFT formulation is the ability to separate the terms that affect the background evolution from those that only concern the perturbations. In particular, only three of the EFT functions, Ω, c, and Λ, play a role in determining the dynamics of the background; hence, we will refer to them as the *background* EFT functions. Furthermore, as detailed in the Appendix, only two out of these three functions are sufficient to fix the background dynamics. In our approach, we specify $\{\Omega(a), \Lambda(a)\}$ and find the Hubble parameter, $\mathcal{H}(a)$, by solving a differential equation,

$$\left(1 + \Omega + \frac{1}{2} a\Omega'\right) \frac{dy}{d\ln a} + (1 + \Omega + 2a\Omega' + a^2\Omega'')y$$
$$+ \left(\frac{P_m a^2}{m_0^2} + \frac{\Lambda a^2}{m_0^2}\right) = 0, \qquad (2)$$

where $y \equiv \mathcal{H}^2$, and the prime indicates differentiation with respect to the scale factor. Given the solution for $\mathcal{H}(a)$, the effective DE EoS is defined via

$$w_{\mathrm{DE}} \equiv \frac{P_{\mathrm{DE}}}{\rho_{\mathrm{DE}}} = \frac{-2\dot{\mathcal{H}} - \mathcal{H}^2 - P_m a^2/m_0^2}{3\mathcal{H}^2 - \rho_m a^2/m_0^2}, \qquad (3)$$

where ρ_m and P_m are the combined energy density and the pressure of all particle species, and the overdot denotes a conformal time derivative. The full details of this procedure are given in the Appendix. By sampling from a broad range of possible $\{\Omega(a), \Lambda(a)\}$ functions, as described in Sec. III, we are able to generate a representative ensemble of possible expansion histories in single field DE models. We note that this procedure is different from the so-called designer approach used in, e.g., [38], in which one provides $\mathcal{H}$ (and Ω), and solves for $\{c, \Lambda\}$.

Since we are interested in reproducing the expansion histories of theoretically viable Horndeski models, we sample the functions $\{M_2^4, \bar{M}_1^3, \bar{M}_3^2\}$ along with $\{\Omega(a), \Lambda(a)\}$. While the former affect only the perturbations and, hence, would appear not to matter for the prior on $w_{\mathrm{DE}}(z)$, they do in fact play a role in determining whether a given background solution corresponds to a stable Horndeski theory. By perturbing the action (1) around a given background, one can derive constraints on combinations of EFT

functions and their derivatives that exclude instabilities. Specifically, in our analysis, we impose the no-ghost and no-gradient-type instability conditions for scalar and tensor modes [37]. We implement the procedure for solving for the background discussed above in EFTCAMB[1] [39,40] and use its stability module to filter out models with ghost and gradient type instabilities. We find that imposing these conditions leads to the model acceptance rate ∼50% for quintessence, ∼10% for GBD, and ∼1% for Horndeski, thus removing a noticeable fraction of the parameter space. The details on technical implementation of the stability conditions can be found in [41].

A. Two simple case studies

To gain intuition into the expected behavior of $w_{\mathrm{DE}}(z)$ in scalar-tensor theories, let us examine the effect of varying Λ and Ω.

First, we consider a model in which all EFT functions except Λ are set to their ΛCDM values. This choice corresponds to the minimally coupled quintessence models, with Λ representing the Lagrangian of the scalar field, or the difference between its kinetic and potential energy densities. For our illustration, we take $\Lambda(a) = \Lambda_0(1 + \lambda(a))$, with $\lambda(a) = -0.1a$, and solve the background equations. Panel (a) of Fig. 1 shows the DE EoS $w_{\mathrm{DE}}(z)$, while panel (b) shows the matter, radiation, and DE energy fractions as a function of redshift. For this model, we see that $w_{\mathrm{DE}} \approx 0$ during the matter and radiation eras, before evolving towards -1 at low redshift. This behavior is quite generic and can be understood as follows. The effective DE EoS is given by

$$w_{\mathrm{DE}}(a) = \frac{\Lambda}{2c - \Lambda} = \frac{K - P}{\rho_{\mathrm{DE}}(a)}, \qquad (4)$$

where K and P are the quintessence kinetic and potential energy densities, and $\rho_{\mathrm{DE}} \equiv 3m_0^2 H^2 - \rho_M - \rho_r$ is the total energy density excluding matter and radiation. In typical quintessence models, the field is slowly rolling at late times (low redshifts), with $P > K$, but retains a small nonzero kinetic term; hence, $w_{\mathrm{DE}} \gtrsim -1$. The energy density of the scalar field with $w_{\mathrm{DE}} \geq -1$ increases with redshift, unless the kinetic energy of the field is tuned to be exactly zero (which would be the case of the cosmological constant). This drives $w_{\mathrm{DE}} \to 0$ at high redshift, as seen in Fig. 1. Note that this transition to the dustlike EoS happens when the DE component is subdominant and in such a way that it cannot become dominant again in the past. Thus, the impact of DE on the expansion rate and the clustering effects would be limited at early times.

Next, let us consider a model in which both Ω and Λ can vary in time. For illustration, we take $\Omega(a) = 0.2(a - 1)/[1 + 0.7(a - 1)]$ and $\lambda(a) = 0.3 + 1.0(a - 1)$.

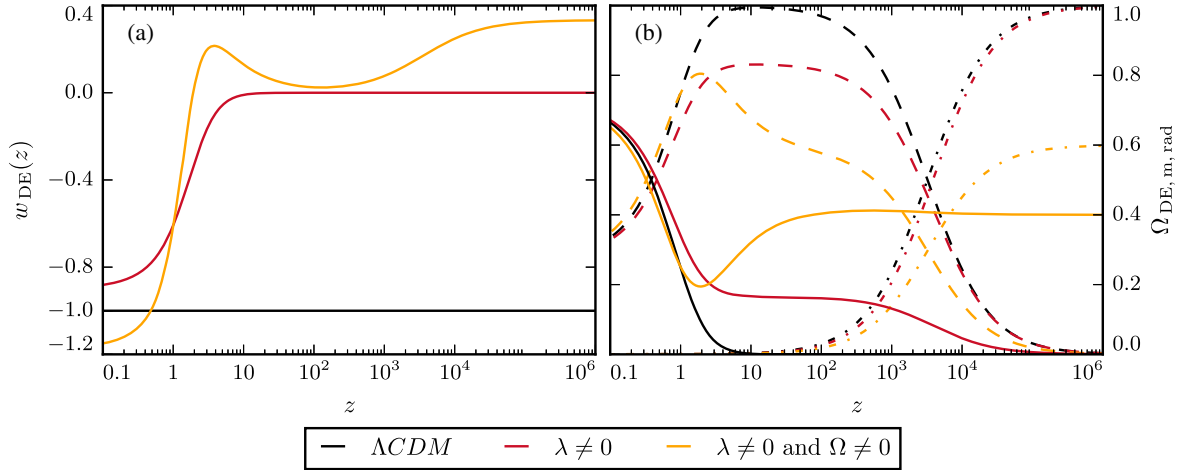
FIG. 1. The evolution of the effective DE EoS, $w_{\mathrm{DE}}(z)$ (*panel a*) and the matter, radiation, and effective DE densities (*panel b*) in the ΛCDM model, in a model with a nonconstant Λ, and a model with a nonconstant Λ and $\Omega \neq 0$. In *panel b*, the solid lines show the effective DE density, dashed lines the dark matter energy density, dot-dashed lines the radiation density.

As can be seen from Fig. 1, there are two key differences from the Λ-only model. Firstly, w_{DE} has a tracking behavior at early times, with $w_{\mathrm{DE}} \sim 0$ during the matter era and $w_{\mathrm{DE}} \sim 1/3$ in the radiation era. Secondly, w_{DE} crosses -1 at late times, when dark energy starts becoming important. Both features are quite representative of nonminimally coupled models. To understand the tracking behavior, observe that changing the value of Ω changes the gravitational coupling, while keeping the densities of the matter and the radiation species the same. In this case, the effective DE fluid compensates for the difference between the expected and the actual gravitational contributions of matter and radiation. The fluid must therefore take on an EoS that tracks the EoS of the dominant species at any given time.

The fact that the effective w_{DE} can cross -1 in nonminimally coupled models is well known [42]. To see this, consider the form of the effective w_{DE} in the simple case of constant Λ and Ω,

$$w_{\mathrm{DE}}(a) = \frac{\Lambda - \Omega P_m}{2c - \Lambda - \Omega \rho_m}, \qquad (5)$$

where P_m and ρ_m are the matter pressure and energy densities. It is clear from the new term in the denominator that the coupling to matter can act to decrease the energy density of the effective dark energy fluid, pushing w_{DE} below -1. While the expression for w_{DE} in the case of varying Λ and Ω is more complicated, the physics is essentially the same.

III. METHODS

This section explains how we build the ensemble of viable Horndeski models, and the way in which we store and present the priors on $w_{\mathrm{DE}}(z)$.

A. Sampling the ensemble of viable scalar field dark energy models

A particular model realization is defined by specifying the EFT functions that appear in action (1). Ideally, one would adopt the most general functional forms possible for these functions, to avoid unduly focusing on only a few—potentially unrepresentative—corners of the space of possible models. In practice, this means choosing a parametrization that is broad enough to represent many different functions of time, while having sufficiently few parameters that it can be efficiently sampled using a Monte Carlo algorithm. To strike a balance between generality and tractability, we will use a number of different series expansions in the scale factor, a.

The first such expansion is a truncated Taylor polynomial, defined by

$$f(a) = \sum_{n=0}^{N} \frac{\alpha_n}{n!}(a - a_0)^n, \qquad (6)$$

where N is the order at which we choose to truncate the expansion, a_0 is the point around which we expand, and $\{\alpha_n\}$ is a set of coefficients to be drawn from a random distribution. Given α_n coefficients with identical prior distributions, such a parametrization would favor the lower order terms.

The second expansion that we consider is a polynomial expansion,

$$f(a) = \sum_{n=0}^{N} \alpha_n(a - a_0)^n. \qquad (7)$$

This differs from the Taylor expansion in the absence of the $n!$ term that suppresses higher order terms. Depending on how the α_n coefficients are drawn, this will allow rapid variations in $f(a)$ to arise more easily.

Finally, we will also use a Padé expansion,

$$f(a) = \frac{\sum_{n=0}^{N} \alpha_n (a - a_0)^n}{1 + \sum_{m=1}^{M} \beta_m (a - a_0)^m}, \qquad (8)$$

where the truncation order is now set by both N and M. This expansion is well suited for describing models in which the functions transition from one value at small a to another at large a.

We consider these three expansions around $a_0 = 0$ and $a_0 = 1$ to represent models that exhibit "thawing" and "freezing" behaviors respectively. We also explore several truncation orders, ranging from zero-order expansions, where all the three expansions reduce to constants, to ninth-order expansions that should allow very general model behaviors to be captured. We progressively raise the truncation order, obtaining results on the DE EoS prior for each order in turn. We find that the results stabilize beyond the fifth order, changing very little as the expansion order is further increased.

The prior distributions for all the coefficients $\{\alpha_n\}$ and $\{\beta_m\}$, that define particular realizations of the EFT functions, are chosen to be uniform in the range $[-1, 1]$. We checked that extending the ranges of the flat priors does not affect the results.

In an effort to make our prior on the DE equation of state as conservative as possible, we adopted weak priors on cosmological parameters even though some of them are actually very well constrained by cosmological experiments. The relative density of radiation is fixed by the CMB temperature, while the matter density today (the sum of the baryon and the CDM contributions) was drawn from the range $\Omega_m \in [0, 1]$. We assume the Universe to be flat and the sum of neutrino masses to be $\sum_\nu m_\nu = 0.06$ eV. The present-day DE density (a free parameter that is not fixed by flatness in the models we considered) was allowed to span $\Omega_{DE} \in [0, 1]$. The Hubble constant was allowed to vary in the range $H_0 \in [20, 100]$ km/s/Mpc. In addition, we impose a weak prior on the *present* value of the gravitational coupling, allowing variations of no more than 10%, and the *present* value of the speed of gravitational waves, allowing variations of no more than 20%. The latter two conditions help to exclude models that are in obvious contradiction with laboratory experiments and observations of the nearby universe [43–46].

When deriving the probability distributions of w_{DE} values at different z, we also take into account mild background data constraints to remove histories that are obviously ruled out. Namely, we use information from baryon acoustic oscillation (BAO) measurements [47–50], estimates of the Hubble constant [51], and supernovae distance measures [52], but with a significantly enlarged covariance (by a factor of 4) to avoid the biasing of our results by tensions between these data sets. When using the SN measurements, we fix the stretching parameter and the color-at-maximum-brightness parameter to their best fit ΛCDM values. The enlarged data covariance mitigates possible biases induced by this choice.

Let us stress that we do not use this information when deriving the covariance of the w_{DE} bins, so that it can be used in reconstructions of $w_{DE}(z)$ as a purely theoretical prior.

We explore the space of cosmological parameters and $\{\alpha_n\}$ coefficients using a Monte Carlo sampling procedure. For each sample, after solving the background equations (with initial conditions defined at the present day; see the Appendix), we check the stability of the corresponding model and, if accepted, compute $w_{DE}(z)$. To ensure good coverage, we enforce a minimum number of 10^5 accepted samples. Depending on the acceptance rate, this results in $\sim 10^6$–10^8 total samples.

While we have chosen several different ways to parametrize the behavior of the EFT functions, we can marginalize over these choices if desired. Given a set of models $\mathcal{M}$ corresponding to the different available options, we can compute

$$P(\vec{w}) = \sum_{\mathcal{M}} P(\vec{w}|\mathcal{M})P(\mathcal{M}), \qquad (9)$$

where $P(\mathcal{M})$ is the prior probability of each of the choices. Note that we are not interested in the normalization of $P(\vec{w})$, which can be recovered *a posteriori*, so there is no need to impose the constraint $\sum_{\mathcal{M}} P(\mathcal{M}) = 1$. We will assume uninformative priors on the choice of parametrization, so that $P(\mathcal{M})$ is equal for all models. Since we have generated Monte Carlo samples of $P(\vec{w}|\mathcal{M})$ with the same number of accepted points, we can then obtain $P(\vec{w})$ simply by merging the Monte Carlo samples from all parametrizations and reweighting them with their respective acceptance rates.

B. The priors on $w_{DE}(z)$

Given an expansion history from a particular model realization, we store the corresponding $w_{DE}(z)$ in 100 linearly spaced bins in scale factor that cover the redshift range $z \in [0, 6]$. From the Monte Carlo samples of the binned EoS, $w_{DE}(z_i) = w_i$, we compute the mean and the covariance,

$$\bar{w}_i = \frac{1}{N_{\text{samp}}} \sum_{\text{samples}} w_i,$$

$$C_{ij} = \frac{1}{N_{\text{samp}} - 1} \sum_{\text{samples}} (w_i - \bar{w}_i)(w_j - \bar{w}_j). \qquad (10)$$

One can also define the normalized correlation matrix as

$$\mathcal{C}_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}. \qquad (11)$$

Under the assumption of Gaussianity, the prior covariance matrix can be turned into a prior probability distribution and used in reconstructions of $w_{\rm DE}(z)$ from data, following the procedure suggested in [15].

While having a numerically obtained discrete prior covariance matrix C_{ij} may be sufficient for many practical applications, it is useful to also have an analytical expression characterizing the correlation of $w_{\rm DE}(z)$ between an arbitrary pair of redshifts. To this end, following [13,15], we can introduce a two-point correlation function $C(a, a')$ defined as

$$C(a, a') \equiv \langle [w_{\rm DE}(a) - \bar{w}_{\rm DE}(a)][w_{\rm DE}(a) - \bar{w}_{\rm DE}(a')] \rangle. \quad (12)$$

As in the discrete case, Eq. (11), this prior covariance can be expressed in terms of the normalized correlation, $\mathcal{C}(a, a')$, equal to unity for $a = a'$, and the autocorrelation (or variance), $C(a) \equiv C(a, a)$, i.e.,

$$C(a, a') = \sqrt{C(a)C(a')}\mathcal{C}(a, a'). \quad (13)$$

In Sec. IV, we will obtain analytical forms of $C(a, a')$ for different classes of models by separately fitting $\mathcal{C}(a_i, a_j)$ and $C(a_i)$ to the numerically obtained C_{ij} and C_{ii}.

As derived quantities, in Sec. IV, we also compute projections of the binned $w_{\rm DE}(z)$ model onto parameters of the CPL parametrization,

$$w_{\rm DE}(a) \approx w_0 + w_a(1 - a), \quad (14)$$

where w_a is obtained from

$$w_a = -\frac{dw_{\rm DE}}{da}\bigg|_{a=1}. \quad (15)$$

This, amongst other things, allows us to compare our results to those in [22], where priors on (w_0, w_a) were derived for quintessence by sampling a broad range of scalar field potentials.

IV. RESULTS

In this section, we present the results of our Monte Carlo exploration of allowed DE EoS histories in scalar-tensor theories. We first discuss the probability density functions (PDFs) of $w_{\rm DE}$ at different z, and then the covariance of $w_{\rm DE}(z)$ and the corresponding analytical fitting formulas.

A. The ensemble of $w_{\rm DE}(z)$ functions

By following the procedure for sampling the ensemble of scalar field theories and calculating the corresponding expansion histories described earlier, we obtain an ensemble of functions $w_{\rm DE}(z)$. Figure 2 (left column) shows the probability density of $w_{\rm DE}$ as a function of z, derived from the Monte Carlo sample of two classes of models: the

minimally coupled quintessence field, and the full Horndeski theory. These distributions have been marginalized over cosmological parameters, boundary conditions, EFT function parametrizations, orders of the Taylor expansion, and freezing/thawing behaviors (values of a_0). In addition, the mild observational constraints discussed in the previous section have been applied. The figures are shaded according to probability density, and show the mean $w_{\rm DE}(z)$, as well as the 68%, 95%, and 99% C.L. contours.

Despite the large amount of freedom in the specification of the models, there is a remarkably well-defined structure to the $w_{\rm DE}(z)$ prior probability density. The overall behavior is similar between the two model classes, despite the significant increase in model freedom in going from the quintessence (Λ-only) class to the most general Horndeski class. In both cases, the EoS transitions from $w_{\rm DE} \approx 0$ deep in the matter-dominated regime to ~ -1 at low redshift. The high-redshift behavior is quite generic, for the reasons given in Sec. II A: only quintessence models with kinetic energy fine-tuned to be close to zero can remain potential dominated (and thus have $w \ll 0$) at early times, while the tracking behavior is common in nonminimally coupled models, giving $w_{\rm DE} \approx w_m = 0$ in the matter-dominated regime. The low-redshift behavior is typical of freezing models, where the kinetic energy of the scalar field always remains subdominant and is significantly driven by the mild data constraints. These downweight all of the $w_{\rm DE}(z)$ functions that result in nonaccelerating cosmologies with $w_{\rm DE} > -1/3$ at $z = 0$.

There are, however, some significant qualitative differences between the two model classes. As shown in panel 1(a) of Fig. 2, all models in the quintessence class respect $w_{\rm DE} \geq -1$ at all times, as expected. This bound is not respected in nonminimally coupled models, which have a significant probability density of models with $w_{\rm DE} < -1$ at low redshift. These "phantom" models are relatively common, with 40% of viable Horndeski models having $w_{\rm DE} < -1$ at $z = 0$.

Another feature worthy of note is the lack of models with $w_{\rm DE} \simeq -1 = $ const., i.e., a cosmological constantlike evolution at all redshifts. While values of $w_{\rm DE} \approx -1$ are very common at low redshift, they are rather unlikely in both model classes at $z \simeq 2$. This can be explained dynamically, by the tracking and kinetic energy tuning arguments above, but has interesting observational implications. While the fractional DE density is low at $z \gtrsim 2$, making measurements of the EoS more difficult, our results suggest that EoS constraints in this regime may actually be the most powerful for distinguishing a cosmological constant from dynamical DE.

The right-hand column of Fig. 2 shows the projection of the $w_{\rm DE}(z)$ curves onto the w_0 and w_a parameters. The results for the quintessence class are shown in panel 1(b). This should be compared with Fig. 1 of [22], which also shows the (w_0, w_a) distribution for a
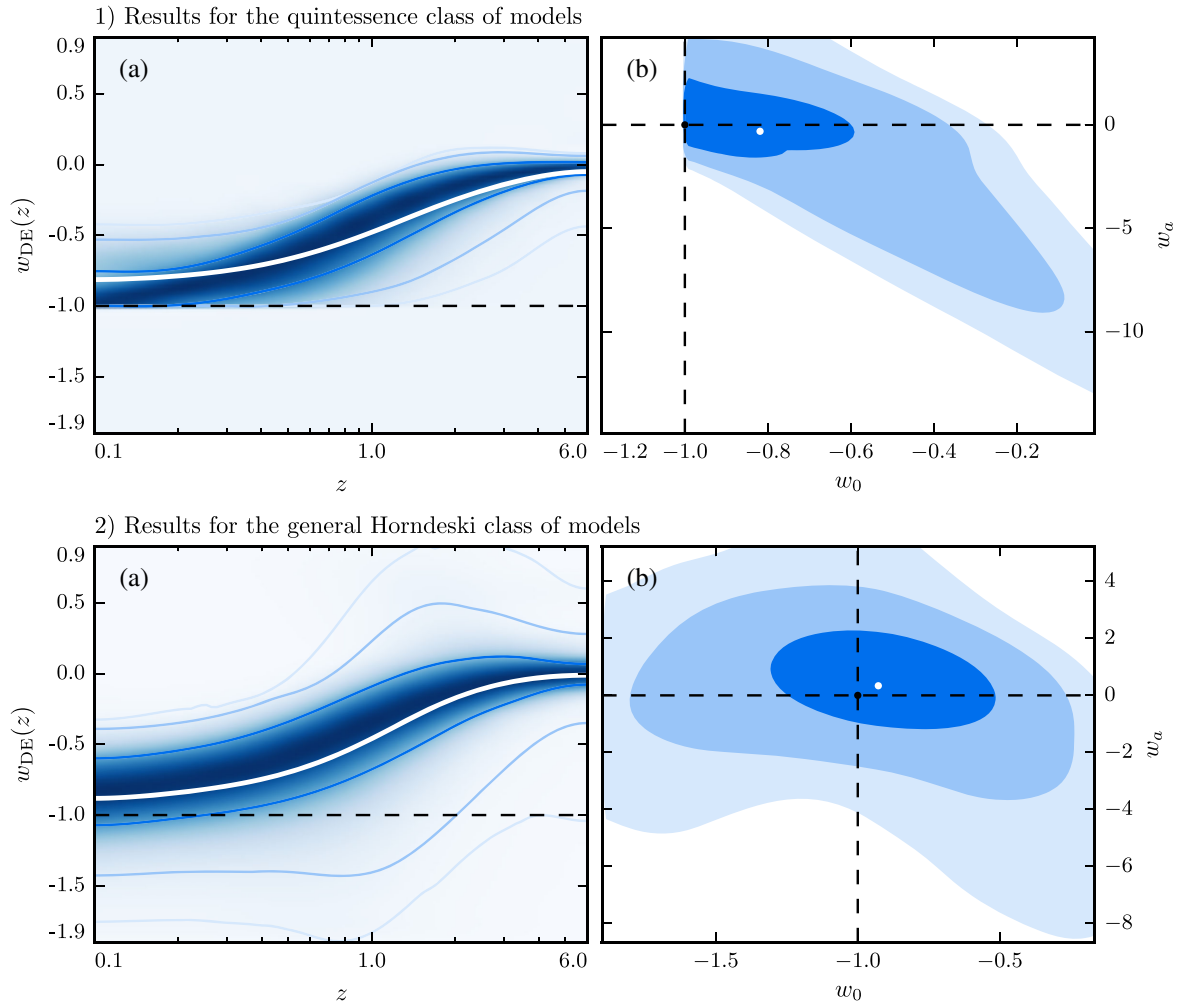
1) Results for the quintessence class of models

2) Results for the general Horndeski class of models

FIG. 2. The ensemble of effective DE equations of state, $w_{\mathrm{DE}}(z)$, for the two model classes, with mild data constraints applied. *Panel a, all figures*: probability density of the effective DE EoS as a function of redshift. *Panel b, all figures*: probability density for the projection onto the $w_0 - w_a$ parameters. In all panels, the white lines/points show the mean, and contours show the 68%, 95%, and 95% C.L. intervals, the blue shading in the left-hand column shows the probability density, while the black point and dashed line represent the ΛCDM model.

Monte Carlo-generated sample of quintessence models. The orientation of the prior is roughly the same, with most models following a track that goes from $(w_0, w_a) = (-1, 0)$ towards more positive w_0 and more negative w_a. This is in part due to the $w_{\mathrm{DE}} \geq -1$ restriction that applies to quintessence models, which excludes a substantial region of the (w_0, w_a) plane.

The distribution shown in the panel 1(b) of Fig. 2 is quite broad, with a significant fraction of models having $w_a > 0$. In contrast, the models of [22] occupied a much narrower track that is almost entirely at $w_a < 0$. Part of the reason for this is that the Monte Carlo method in [22] used randomly generated scalar field *potentials* to define the quintessence models, while we are working directly in terms of randomly generated EFT functions. This effectively leads to a different weighting of the models; simple random distributions over the parametrized potentials may translate to nontrivial distributions over the EFT functions. In this case, the

weighting obtained by parametrizing the theories at the EFT function level seems to make it easier to obtain thawing models (c.f. [53], who also find relatively more models with $w_a > 0$). Also, the models in [22] were selected to be "physically motivated," in the sense that the forms of their potentials were inspired by high-energy theory, like moduli fields and axions. These models form a more restricted class than the general, phenomenological EFT approach considered here. In particular, treating $\Lambda(a)$, which represents a particular evolution history for the quintessence Lagrangian, as being *a priori* arbitrary, amounts to a significantly more agnostic approach, leading to a broader allowed range of (w_0, w_a) values.

The (w_0, w_a) distribution for the Horndeski class is broader and much less constrained, primarily as a result of the broadening of the $w_{\mathrm{DE}}(z)$ distributions and the ability of the models to extend to $w_{\mathrm{DE}} < -1$. The bulk of the models remain within the vicinity of the $(-1, 0)$ point in both

classes, although there is a very slight shift towards less negative w_0 and positive w_a for the full Horndeski class. This is despite the fact that very few of the models have equations of state that resemble a cosmological constant out to high redshift. In fact, the mean functional form in the left-hand panels approximately satisfies $w|_{z=0} \to -1$ and $w|_{z\to\infty} \to 0$ in both model classes, which would imply $w_0 = -1$ and $w_a = +1$. These values are within the 68% region for all three classes, but shifted slightly from the mean (shown by the white points in Fig. 2). This is, at least in part, because both w_0 and w_a have been evaluated directly at $z = 0$. It is clear from Fig. 2 however that the derivative of $w_{\text{DE}}(z)$ is not constant with redshift; *i.e.*, the (w_0, w_a) parametrization cannot fully reproduce its true redshift evolution when defined in this way. This means that we can trust the projection onto the CPL parametrization, as an approximation to our results, only at late times.

In addition to considering the quintessence subclass of Horndeski theories, we have separately studied the case of the GBD models, *i.e.*, models with a canonical scalar field kinetic term. Within the EFT framework, they are described by $\Lambda(a) \neq$ const. and $\Omega(a) \neq 0$, with $c(a)$ derived from $\{\Lambda, \Omega\}$, and the other functions set to zero. We find the

results for the PDF of $w_{\text{DE}}(z)$ in this case are practically the same as in the full Horndeski case. On one hand, this is expected, since only $\Lambda(a)$ and $\Omega(a)$ affect the background evolution. On other hand, the higher order functions figure in the stability conditions, which do affect the acceptance rate in the Monte Carlo sampling. We find therefore that the stabilization effect of higher order functions has a limited impact on the shape of the prior.

In the analysis presented so far, we have applied mild observational constraints to ensure that the models are broadly consistent with the properties of the real Universe (see Sec. III). The results are qualitatively similar even if these constraints are relaxed, with the only significant difference being the presence of trajectories that remain above $w_{\text{DE}} = -1/3$ at low redshift.

B. Theoretical priors on correlations of $w_{\text{DE}}(z)$

In this section, we present the covariance between values of w_{DE} at different redshifts, obtained from averaging over the ensemble of $w_{\text{DE}}(z)$ functions from each of the model classes. In doing so, we do not employ any observational constraints, as even relatively mild constraints result in correlations induced by the data overwhelming those
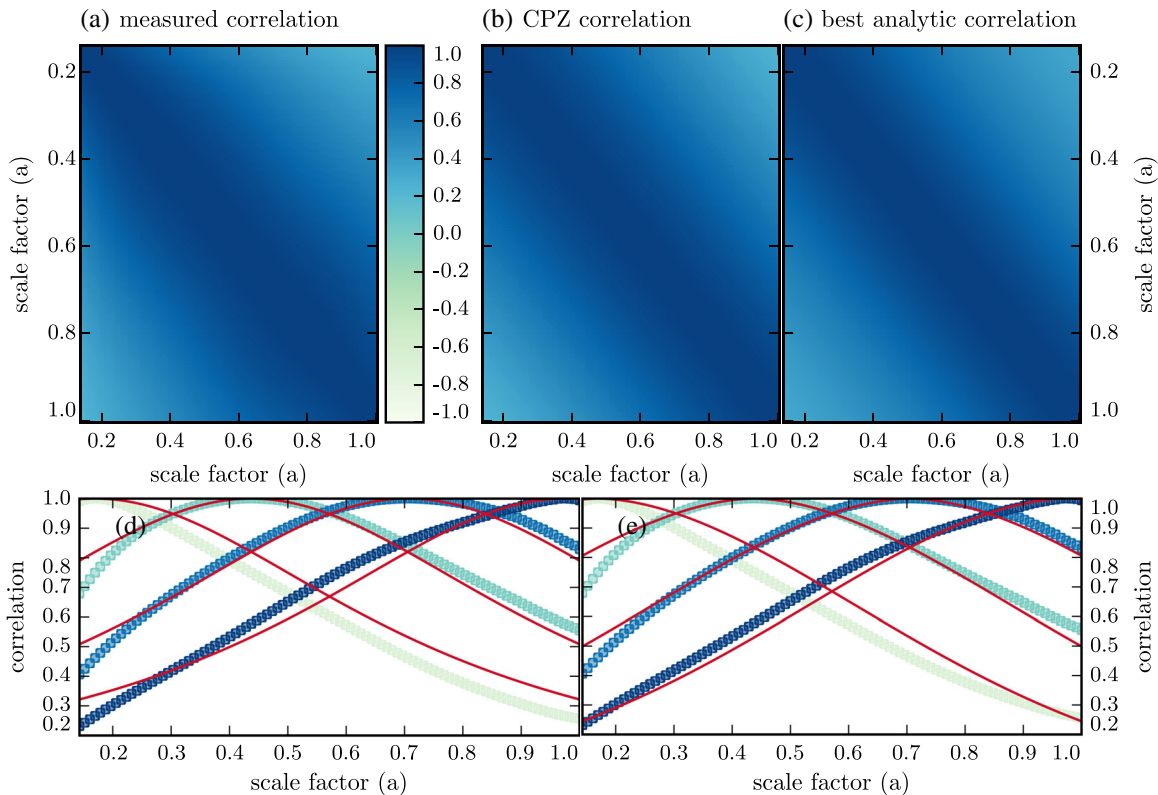

FIG. 3. Correlation results for the average over all quintessence models, without the mild data constraints. *Panel a*: measured EoS correlation as a function of the scale factor. *Panel b*: best fit CPZ form [13] (fixed power law) analytic prediction of the EoS correlation. See details in the text. *Panel c*: best overall analytic prediction of the EoS correlation. See details in the text. *Panel d, e*: correlation of the EoS at a single scale factor against all other scale factors, showing the measured correlation (dots) and the predicted correlation (continuous line) for the CPZ form and the best analytic model, respectively. Different colors correspond to different scale factors.
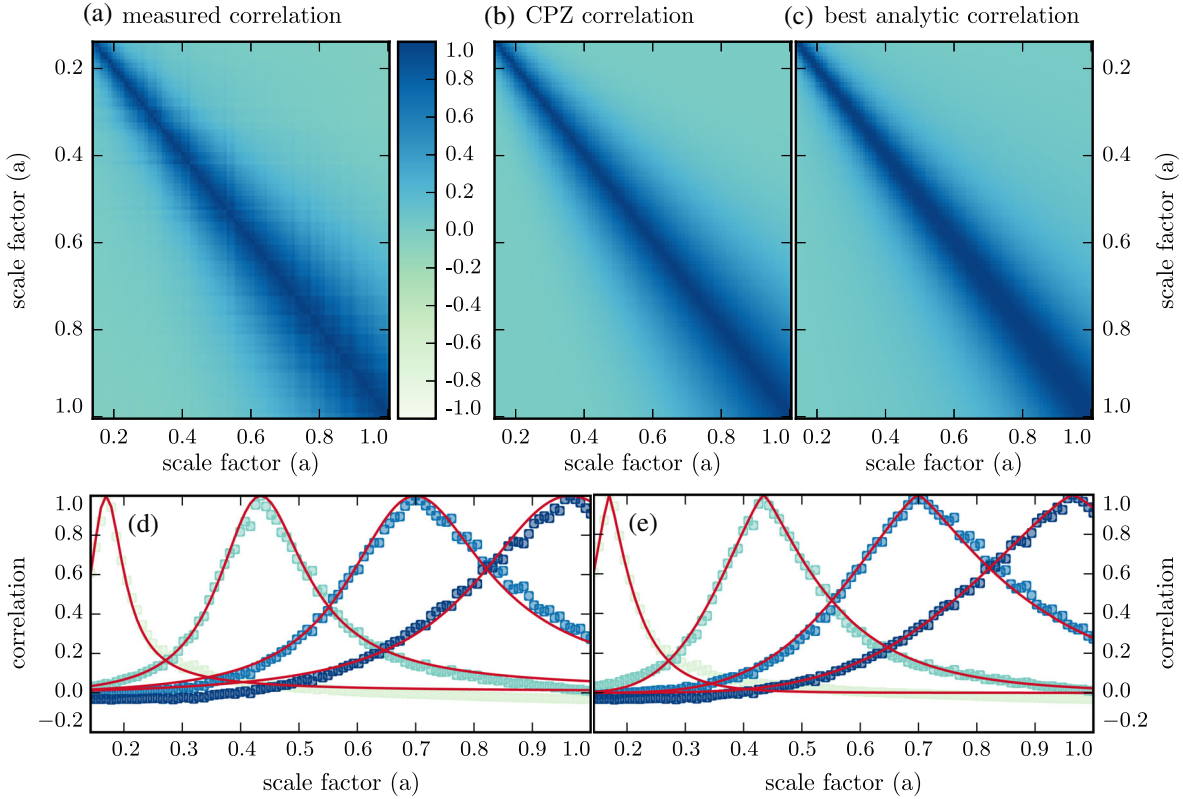
FIG. 4. Correlation results for the average over all Horndeski models, without the mild data constraints. *Panel a*: measured EoS correlation as a function of the scale factor. *Panel b*: best fit CPZ form [13] (fixed power law) analytic prediction of the EoS correlation. See details in the text. *Panel c*: best overall analytic prediction of the EoS correlation. See details in the text. *Panel d, e*: correlation of the EoS at a single scale factor against all other scale factors, showing the measured correlation (dots) and the predicted correlation (continuous line) for the CPZ form and the best analytic model, respectively. Different colors correspond to different scale factors.

coming from theory, effectively hiding the information that we want to extract.

We separate the covariance into the normalized correlation matrix, $\mathcal{C}(a_i, a_j)$, and the autocorrelation (or variance), $C(a_i)$, as in Eq. (12). The top left panels in Figs. 3 and 4 show the numerically obtained $\mathcal{C}(a_i, a_j)$ at binned values of a for the quintessence and the general Horndeski classes, respectively. The points in Fig. 5 show the variance $C(a_i)$ for the two models.

As mentioned in Sec. III B, having analytical expression for the correlation prior can simplify practical applications of our priors to reconstructing $w_{DE}(z)$ from data using methods similar to those in [16,54]. We can obtain them by fitting simple functional forms to the numerically obtained discrete covariance matrices.

First, we perform a least-squares fit to the autocorrelation $C(a_i)$ using the following functional forms:
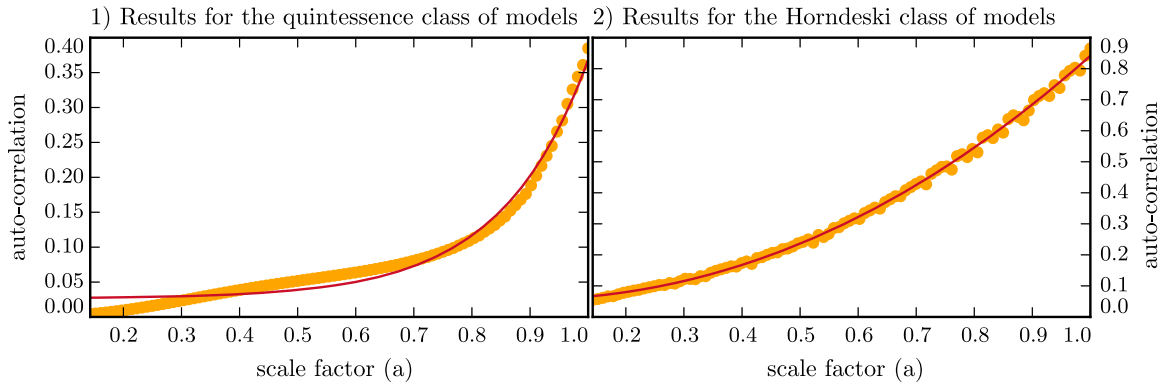(i) a Taylor expansion, $C(x) = \alpha + \beta(x - x_0)$;

FIG. 5. Autocorrelation results for the quintessence (left panel) and Horndeski (right panel) model classes, without the mild data constraints. The dots represent the measured EoS autocorrelation while the continuous line is the best-fit parametrization.

(ii) an exponential, $C(x) = \alpha + \beta \exp[\gamma(x - x_0)]$;
(iii) a power law, $C(x) = \alpha + \beta(x - x_0)^\gamma$;

where x is the evolution variable for which we try three different choices: the scale factor, $x = a$; the redshift, $x = z$; and the number of e-folds, $x = \ln(a)$. We also try two choices of the reference point, x_0, in each case: for $x = a$ and $x = \ln(a)$, we either fix $a_0 = 1$ or allow a_0 to be free, and for $x = z$, we either fix $z_0 = 0$ or allow z_0 to be free. After fitting all of these different parametrizations to the measured autocorrelation function for each model class, we select the ones that result in the smallest residuals in each case. The choices that give the best fits are summarized in Table I and plotted in Fig. 5 as a function of the scale factor. One can see that the exponential fitting formula with $x = \ln a$ and $a_0 = 1$ works very well in the Horndeski case, while the exponential form with $x = a$ and $a_0 = 1$ works best in the quintessence case. The analysis for the GBD class of models gives results that are almost indistinguishable from the Horndeski case.

As expected, the autocorrelation (variance) is generally bigger in the Horndeski case, as these models permit more freedom than quintessence. In both classes, the autocorrelation is small at high redshifts, where the field is expected to have little or no dynamics, and so tends to result in similar values of w_{DE} in all models. The spread is much more significant at low redshifts however, where a greater range of dynamical behaviors is realized. The autocorrelation of the quintessence models is clipped by the hard bound of $w_{DE} \geq -1$, reducing the allowed variation from model to model.

Next, we use a similar procedure to find the analytical form of $C(a, a')$. We consider three fitting forms. One is a generalized version of the CPZ parametrization [13],

$$C(x, y) = \frac{1}{1 + (|x - y|/\xi)^n}, \qquad (16)$$

where x and y are either the scale factor, the redshift or $\ln a$, ξ is a parameter defining the correlation time scale, and n is

a free parameter. We also separately fit to (16) with $n = 2$, which is the CPZ form. In addition, we consider an exponentially decaying correlation,

$$C(x, y) = \exp[-(|x - y|/\xi)^n], \qquad (17)$$

where the physical meaning of the parameters is the same as in the previous parametrization (16). We then evaluate the least-squares distance between $C(a_i, a_j)$ and the numerically found C_{ij}, and find the parameters that minimize it, as well as the fitting form that results in the smallest residuals for each class of models.

We present the best fit analytical forms of $C(a, a')$ and the corresponding residuals in Table I, along with best fit parameters of the CPZ parametrization. The two top right panels of Figs. 3 and 4 show the best fit CPZ and overall best fit $C(a, a')$ for the quintessence and Horndeski models, respectively. To have a more detailed visual check of the goodness of fit, in the bottom panel of Figs. 3 and 4, we also plot $C(a, a')$ as a function of a' for several fixed values of a. One can see that the exponentially decaying parametrization works marginally better than the CPZ form, as it has more freedom built in. Overall, the analytical expressions work well in reproducing the correlation in different models. We also note that the best fit correlation length, ξ, is roughly the same in both parametrizations for each model.

One can clearly see that the correlation $C(a, a')$ is long ranged for quintessence at all a and is slowly decaying with $|a - a'|$. In Horndeski models, on the other hand, the correlation is generally shorter ranged and decays faster at early times compared to late times. This is expected since Horndeski models have more freedom in the choice of expansion histories compared to quintessence; hence, values of w_{DE} at different times are less correlated.

It is also interesting to note that the autocorrelation, $C(a)$, depends linearly on a for quintessence, while it scales as a^2 in the Horndeski case. Also, the correlation $C(a, a')$ scales as a power of $|a - a'|$ for quintessence, while in the case of Horndeski, it scales as a power of $|\ln a - \ln a'|$. This is due to the fact that, in quintessence, w_{DE} is solely a property of DE, which is only minimally coupled to other fluids. In nonminimally coupled models, on the other hand, w_{DE} is an effective quantity, determined largely by how the other fluids scale with redshift. This is true especially at higher redshifts, where the variance and the correlation of $w_{DE}(z)$ in Horndeski models is set by the dynamics of the matter fluid, which has a fixed time-dependence and uniformly distributed random amplitude set by Ω_M.

TABLE I. Summary of the autocorrelation and correlation function fits.

	Best-fit autocorrelation	Residuals
Quintessence	$0.03 + 0.3 \exp[6.5(a - 1)]$	0.01
GBD	$0.05 + 0.8 \exp[1.8 \ln a]$	0.007
Horndeski	$0.05 + 0.8 \exp[2 \ln a]$	0.007
	Best-fit correlation	Residuals
Quintessence	$\exp[-(|\delta a|/0.7)^{1.8}]$	9
GBD	$\exp[-(|\delta \ln a|/0.3)^{1.3}]$	6
Horndeski	$\exp[-(|\delta \ln a|/0.3)^{1.2}]$	6
	Best-fit correlation (fixed CPZ)	Residuals
Quintessence	$(1 + (|\delta a|/0.6)^2)^{-1}$	11
GBD	$(1 + (|\delta \ln a|/0.2)^2)^{-1}$	12
Horndeski	$(1 + (|\delta \ln a|/0.2)^2)^{-1}$	13

C. A note on practical applications of the prior

As mentioned earlier, one can use the prior covariance obtained above to aid reconstruction of $w_{DE}(z)$ from data using the method developed in [15] and applied in [16,54].

There, a piecewise (binned) $w_{DE}(z_i) = w_i$ is fit to data along with other model parameters, $\vec{p}$, using the usual Markov Chain Monte Carlo method for sampling the posterior PDF, $\mathcal{P}(\{w_i\}, \vec{p})$. According to Bayes' theorem, the posterior PDF is the product of the likelihood and the prior PDFs,

$$\mathcal{P}(\{w_i\}, \vec{p}) \propto \mathcal{L}(\{w_i\}, \vec{p}) \times \mathcal{P}_{\text{prior}}(\vec{p}) \times \mathcal{P}_{\text{prior}}(\{w_i\}), \tag{18}$$

where the prior on $\{w_i\}$ is obtained, under the assumption of Gaussianity, from the covariance matrix C_{ij} discussed in the previous section,

$$\mathcal{P}_{\text{prior}}(\{w_i\}) \propto e^{-\sum_{ij}(w_i - \bar{w}_i)[C^{-1}]_{ij}(w_j - \bar{w}_j)/2}. \tag{19}$$

This prior probability explicitly depends on the average EoS, $\{\bar{w}_i\}$. We note that this is *not* the mean $w_{DE}(z)$ shown as white lines in panels 1(a) and 2(a) of Fig. 2. The PDFs obtained in that figure used mild data constraints that pushed the mean closer to the observationally favored $w_{DE} \sim -1$ region.

In practice, *any* reasonable choice for the fiducial $\bar{w}_i$, *e.g.*, any choice within the 1σ band of the EoS shown in Fig. 2, should be more or less equally acceptable. The primary purpose of the prior is to add curvature to the PDF along otherwise flat directions in the parameter space, thus eliminating the degeneracies between w_i and helping the Markov Chain Monte Carlo to converge. A reconstruction that is highly sensitive to the choice of $\bar{w}_i$ should not be trusted.

Other options, discussed in [15], include using the so-called "running average", or marginalizing over $\bar{w}_i$ altogether. We refer the reader to [15,16,54] for further details on the method and its application.

V. SUMMARY

We have derived theoretical priors on the effective DE EoS within the Horndeski class of scalar-tensor theories, which includes all models with a single scalar field that have second order equations of motion. We separately considered the widely studied subclasses of the minimally coupled scalar field, or quintessence [8,23], and models with the canonical form of the scalar field kinetic energy, *i.e.*, the generalized Brans-Dicke (GBD) models [33,34]. Overall, we find that the covariance of $w_{DE}(z)$ in GBD is indistinguishable from that of the general Horndeski case.

Our priors on $w_{DE}(z)$ are stored in the form of a covariance matrix for binned w_{DE}, which can be projected onto priors on the parameters of any specific parametrization, such as CPL.

We found that there are notable differences between the case of the minimally coupled quintessence and the non-minimally coupled models, both in the mean values and the covariance of binned w_{DE}. We found simple analytical

forms for the correlation function, describing the correlation of w_{DE} at different redshifts, that fit our numerical results well. These should simplify the practical application of the priors to reconstructions of $w_{DE}(z)$ from data. Such a reconstruction will be the subject of future work.

APPENDIX: THE BACKGROUND SOLUTION IN EFT

Here, we discuss in detail the procedure for solving for the cosmological background evolution given a choice of EFT functions. In what follows, the overdot represents a derivative with respect to the conformal time, the accent mark represents a derivative with respect to the scale factor, and the subscript m indicates the sum over all particle species: CDM, baryons, photons, and massless and massive neutrinos. The Friedmann equations for EFT are given by

$$\mathcal{H}^2 = \frac{a^2}{3m_0^2(1+\Omega)}(\rho_m + 2c - \Lambda) - \mathcal{H}\frac{\dot{\Omega}}{1+\Omega}, \tag{A1}$$

$$\dot{\mathcal{H}} = -\frac{a^2}{6m_0^2(1+\Omega)}(\rho_m + 3P_m)$$
$$- \frac{a^2(c+\Lambda)}{3m_0^2(1+\Omega)} - \frac{\ddot{\Omega}}{2(1+\Omega)}. \tag{A2}$$

Changing the time coordinate from conformal time to the scale factor, Eq. (A1) can be recast as

$$\frac{ca^2}{m_0^2} = \frac{3}{2}(1 + \Omega + a\Omega')\mathcal{H}^2 - \frac{1}{2}\frac{a^2\rho_m}{m_0^2} + \frac{1}{2}\frac{\Lambda a^2}{m_0^2}. \tag{A3}$$

Combining this with Eq. (A2) and introducing $y \equiv \mathcal{H}^2$, we obtain

$$\left(1 + \Omega + \frac{1}{2}a\Omega'\right)\frac{dy}{d\ln a} + (1 + \Omega + 2a\Omega' + a^2\Omega'')y$$
$$+ \left(\frac{P_m a^2}{m_0^2} + \frac{\Lambda a^2}{m_0^2}\right) = 0, \tag{A4}$$

which is the differential equation we solve to find $\mathcal{H}(a)$. The initial conditions for this equation, along with the current values of the model parameters, can be derived from the Friedmann constraint (A1) computed today,

$$\frac{c_0}{m_0^2} = \frac{3}{2}\mathcal{H}_0^2(1 + \Omega_0 + \Omega_0' - \Omega_m^0 - \Omega_\Lambda^0), \qquad \text{(A5)}$$

where we have defined

$$\Omega_m^0 \equiv \frac{1}{3}\frac{\rho_m^0}{\mathcal{H}_0^2 m_0^2} \quad ; \quad \Omega_\Lambda^0 \equiv -\frac{1}{3}\frac{\Lambda_0}{\mathcal{H}_0^2 m_0^2}. \qquad \text{(A6)}$$

A few comments regarding this equation are in order. The definition of matter densities follows from what we can directly measure, at least in principle. The quantities ρ_m^0 are in fact the densities that can be measured with a non-gravitational experiment by an observer that is assuming that all these species are moving on geodesics of the metric. These quantities are combined in the Eq. (A5) that is usually used as the flatness constraint. Normally, when assuming flatness, we use this to compute the value of Ω_Λ^0 given the value of Ω_m. Here, the situation is slightly different, as both Ω_Λ^0 and Ω_m^0 can be chosen arbitrarily and constitute two of the theory parameters that we are going to sample. On the other hand, the flatness constraint is satisfied, once the present day value of the gravitational constant is fixed, by a suitable choice of the present day value of c.

Once we have a solution of Eq. (2), we can deduce the effective DE pressure and density from the standard form of the Friedmann equations,

$$\mathcal{H}^2 = \frac{a^2}{3m_0^2}(\rho_m + \rho_\nu + \rho_{\text{DE}}),$$

$$\dot{\mathcal{H}} = -\frac{a^2}{6m_0^2}(\rho_m + \rho_\nu + \rho_{\text{DE}} + 3P_m + 3P_\nu + 3P_{\text{DE}}),$$

$$\text{(A7)}$$

and compute the effective DE EoS as

$$w \equiv \frac{P_{\text{DE}}}{\rho_{\text{DE}}} = \frac{-2\dot{\mathcal{H}} - \mathcal{H}^2 - P_m a^2/m_0^2}{3\mathcal{H}^2 - \rho_m a^2/m_0^2}. \qquad \text{(A8)}$$

[1] A. G. Riess *et al.* (Supernova Search Team), Astron. J. **116**, 1009 (1998).
[2] S. Perlmutter *et al.* (Supernova Cosmology Project), Astrophys. J. **517**, 565 (1999).
[3] S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989).
[4] C. P. Burgess, in *Proceedings, 100th Les Houches Summer School: Post-Planck Cosmology* (Oxford University Press, Oxford, 2015), p. 149.
[5] G. Dvali, S. Hofmann, and J. Khoury, Phys. Rev. D **76**, 084006 (2007).
[6] C. Charmousis, E. J. Copeland, A. Padilla, and P. M. Saffin, Phys. Rev. Lett. **108**, 051101 (2012).
[7] N. Kaloper and A. Padilla, Phys. Rev. Lett. **112**, 091304 (2014).
[8] R. R. Caldwell, R. Dave, and P. J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).
[9] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).
[10] M. Chevallier and D. Polarski, Int. J. Mod. Phys. D **10**, 213 (2001).
[11] E. V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).
[12] V. Sahni and A. Starobinsky, Int. J. Mod. Phys. D **15**, 2105 (2006).
[13] R. G. Crittenden, L. Pogosian, and G.-B. Zhao, J. Cosmol. Astropart. Phys. 12 (2009) 025.
[14] A. Albrecht *et al.*, arXiv:0901.0721.
[15] R. G. Crittenden, G.-B. Zhao, L. Pogosian, L. Samushia, and X. Zhang, J. Cosmol. Astropart. Phys. 02 (2012) 048.
[16] G.-B. Zhao, R. G. Crittenden, L. Pogosian, and X. Zhang, Phys. Rev. Lett. **109**, 171301 (2012).
[17] A. Shafieloo, A. G. Kim, and E. V. Linder, Phys. Rev. D **85**, 123530 (2012).
[18] M. Seikel, C. Clarkson, and M. Smith, J. Cosmol. Astropart. Phys. 06 (2012) 036.
[19] R. A. Daly and S. G. Djorgovski, Astrophys. J. **597**, 9 (2003).
[20] A. Shafieloo, U. Alam, V. Sahni, and A. A. Starobinsky, Mon. Not. R. Astron. Soc. **366**, 1081 (2006).
[21] R. Crittenden, E. Majerotto, and F. Piazza, Phys. Rev. Lett. **98**, 251301 (2007).
[22] D. J. E. Marsh, P. Bull, P. G. Ferreira, and A. Pontzen, Phys. Rev. D **90**, 105023 (2014).
[23] B. Ratra and P. J. E. Peebles, Phys. Rev. D **37**, 3406 (1988).
[24] G. Gubitosi, F. Piazza, and F. Vernizzi, J. Cosmol. Astropart. Phys. 02 (2013) 032.
[25] J. K. Bloomfield, E. E. Flanagan, M. Park, and S. Watson, J. Cosmol. Astropart. Phys. 08 (2013) 010.
[26] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, J. Cosmol. Astropart. Phys. 08 (2013) 025.
[27] J. Bloomfield, J. Cosmol. Astropart. Phys. 12 (2013) 044.
[28] F. Piazza and F. Vernizzi, Classical Quantum Gravity **30**, 214007 (2013).
[29] J. Gleyzes, D. Langlois, and F. Vernizzi, Int. J. Mod. Phys. D **23**, 1443010 (2014).
[30] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, Phys. Rev. Lett. **114**, 211101 (2015).

[31] G. W. Horndeski, Int. J. Theor. Phys. **10**, 363 (1974).
[32] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, Phys. Rev. D **84**, 064039 (2011).
[33] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[34] S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity* (Addison Wesley, Reading, MA, 2004).
[35] R. Kase and S. Tsujikawa, Int. J. Mod. Phys. D **23**, 1443008 (2014).
[36] N. Frusciante, M. Raveri, D. Vernieri, B. Hu, and A. Silvestri, Phys. Dark Universe **13**, 7 (2016).
[37] N. Frusciante, G. Papadomanolakis, and A. Silvestri, J. Cosmol. Astropart. Phys. 07 (2016) 018.
[38] B. Hu, M. Raveri, A. Silvestri, and N. Frusciante, Phys. Rev. D **91**, 063524 (2015).
[39] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, Phys. Rev. D **89**, 103530 (2014).
[40] M. Raveri, B. Hu, N. Frusciante, and A. Silvestri, Phys. Rev. D **90**, 043513 (2014).
[41] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, arXiv:1405.3590.
[42] S. Das, P. S. Corasaniti, and J. Khoury, Phys. Rev. D **73**, 083509 (2006).
[43] L. Perenon, F. Piazza, C. Marinoni, and L. Hui, J. Cosmol. Astropart. Phys. 11 (2015) 029.
[44] C. M. Caves, Ann. Phys. (N.Y.) **125**, 35 (1980).
[45] G. D. Moore and A. E. Nelson, J. High Energy Phys. 09 (2001) 023.
[46] J. Beltran Jimenez, F. Piazza, and H. Velten, Phys. Rev. Lett. **116**, 061101 (2016).
[47] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, Mon. Not. R. Astron. Soc. **416**, 3017 (2011).
[48] W. J. Percival *et al.* (SDSS Collaboration), Mon. Not. R. Astron. Soc. **401**, 2148 (2010).
[49] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, A. J. Cuesta, K. T. Mehta, and E. Kazin, Mon. Not. R. Astron. Soc. **427**, 2132 (2012).
[50] L. Anderson *et al.*, Mon. Not. R. Astron. Soc. **427**, 3435 (2012).
[51] A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, and R. Chornock, Astrophys. J. **730**, 119 (2011); Astrophys. J. **732**, 129(E) (2011).
[52] N. Suzuki *et al.*, Astrophys. J. **746**, 85 (2012).
[53] D. Huterer and H. V. Peiris, Phys. Rev. D **75**, 083503 (2007).
[54] G.-B. Zhao *et al.*, Nat. Astron. **1**, 627 (2017).