PHYSICAL REVIEW D **95,** 023518 (2017)

Challenges in connecting modified gravity theory and observations

Eric V. Linder

Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, California 94720, USA

(Received 19 July 2016; published 27 January 2017)

Cosmic acceleration may be due to modifications of cosmic gravity and to test this we need robust connections between theory and observations. However, in a model independent approach like effective field theory or a broad class like Horndeski gravity, several free functions of time enter the theory. We show that simple parametrizations of these functions are unlikely to be successful; in particular the approximation $\alpha_i(t) \propto \Omega_{\rm de}(t)$ drastically misestimates the observables. This holds even in simple modified gravity theories like $f(R)$. Indeed, oversimplified approximations to the property functions $\alpha_i(t)$ can even miss the signature of modified gravity. We also consider the question of consistency relations and the role of tensor (gravitational wave) perturbations.

DOI: 10.1103/PhysRevD.95.023518

I. INTRODUCTION

The origin of cosmic acceleration is an extraordinary mystery in modern physics. The observation of cosmic acceleration [1–4] must be connected to some fundamental theory beyond the current standard model of particle physics, but we do not know whether its origin lies in the structure of the quantum vacuum or an extension to Einstein's theory of gravitation. Considerable progress has occurred in the last decade in exploring aspects of modified gravity [5–7] but the ability to connect theory and observations in a manner not highly dependent on a specific model is lacking in essential aspects.

Here we examine the challenges for such a connection, and caution against oversimplification. Modified gravity is a much more complex arena than scalar field dark energy, with its one free function of time [e.g. the equation of state $w(a)$]. In large part this is because of the role played by perturbations and the tensor sector.

To begin, consider the case for cosmic acceleration not arising from modified gravity, e.g. quintessence dark energy. Here we also have challenges in connecting essential theory to observations, with perhaps the most information arising from the thawing vs freezing classification of scalar fields [8]. This at least describes the steepness of the potential relative to the Hubble friction, and has distinct implications for whether the theory is approaching or departing from a cosmological constantlike (or sometimes de Sitter) state. Beyond that, the expansion history from dark energy, whether from quintessence or modified gravity, is extremely accurately characterized phenomenologically by two numbers [9], w_0 and w_a, measures of the present and time variation of the dark energy equation of state. Indeed, this characterization has been shown valid to the 0.1% level in the observables of distances and Hubble parameters for a wide range of quintessence, k-essence, modified gravity, etc. models.

On the cosmic structure, i.e. perturbative, side of observations, dark energy not arising from modified gravity (or nonstandard couplings) has little to add: quintessence perturbations are small inside the Hubble scale and k-essence (noncanonical kinetic energy model) perturbations have little observational effect since they are suppressed by equations of state near $w = -1$, as observations indicate. For modified gravity effects, a successful, if limited, phenomenological parametrization is the gravitational growth index γ [10], again accurately describing observables at the subpercent level for a variety of modified gravity models [11]. However, this has very restricted interpretable connection to fundamental theory. Better (pseudo)observables include effects not only on growth of structure, but on the deflection of light. These come from the nonrelativistic and relativistic modified Poisson equations [12], and can be written as effective gravitational coupling strengths $G_{\rm matter}(k,a)$ and $G_{\rm light}(k,a)$, where we explicitly show their scale dependence (e.g. on the Fourier wave number k) and time dependence (e.g. on the cosmic scale factor a). The gravitational growth index γ is directly related to $G_{\rm matter}$ in the scale independent limit.

To connect with theory, however, we need to relate the scale and time dependences of these "observables" (we henceforth refer to these quantities $G_{\rm matter}$, $G_{\rm light}$, and their ratio, related to the gravitational slip function, as observables because, while not directly observable, they are so closely connected to observations, i.e. structure growth and gravitational lensing) to the theory—or at least to phenomenological property functions α_i [13].

The functional form of the scale dependence is a ratio of $k^0 + k^2$ polynomials in many cases (see [14–16] and the especially clear [17], but see [18,19] for exceptions), and one simplification is that on scales below the Hubble scale (or more generally the sound horizon or braiding scale [13]) the scale dependence (k^2 terms) is subdominant and one essentially has purely functions of time.

 © 2017 American Physical Society

While this seems to be considerable progress, the problem is that in order to know whether $G_{\rm matter}$ and $G_{\rm light}$ have any simple parametrization of their time dependence one has to evaluate them from the underlying theory, ideally in as model independent a fashion as possible. An excellent framework for this is the effective field theory (EFT) of dark energy [20–23]. Within EFT at quadratic (lowest) order there are seven free functions of time, and within the Horndeski class of gravity there are four free functions. The challenge of connecting such theory functions to realistically parametrized observables was highlighted in [23], which examined various limits. Here we go into greater depth and quantify the problems with oversimplification of the parametrization.

In Sec. II we start with the workhorse of modified gravity, $f(R)$ theory. This corresponds to only one independent free function of time in the EFT formalism and so is a basic place to start in assessing parametrizations. We expand in Sec. III to the four functions of Horndeski gravity and examine the motivation for a potential simplification in the early time (matter dominated) limit, deriving the asymptotic behavior of the property functions and observables. In Sec. IV we identify how extending these limiting behaviors to the epoch of cosmological structure observations raises foundational issues, and we quantify the dramatic deviations that actually arise in generic circumstances. Section V discusses the reasons why simplified parametrizations appear generally unviable, and solving a problem like modified gravity is so difficult. We conclude in Sec. VI with some thoughts on further progress, while the Appendix explores the possibility of proving broad consistency relations that an entire class of modified gravity theories must obey.

II. A ONE FUNCTION CASE: $f(R)$ GRAVITY

There are four free functions of time within the Horndeski class of gravity theories [apart from the Hubble expansion itself, $H(a)$ or $a(t)$, which can also be phrased in terms of an effective dark energy equation of state $w(a)$]. It is convenient to take these functions to be treated in terms of property functions [13]. The property functions describe the structure of the scalar kinetic sector of the theory via the kineticity α_K, the tensor sector via speed of tensor perturbation propagation $\alpha_T = c_T^2 - 1$, the mixing of the scalar and tensor sectors via the braiding α_B, and the running of the Planck mass α_M. Translations between these and the EFT functions and the observable functions are given in, e.g., [23]. Explicit expressions for α_i in terms of Horndeski functions are given in [13], and in terms of covariant Galileon κ's in, e.g., [24].

In specific theories some of these functions can be 0 and some can be redundant. In general relativity all are 0. We can start by considering the simplest nontrivial situation where the theory has one independent free function of time; $f(R)$ is one such theory, with the only nonzero property function being $\alpha_M = -\alpha_B$ [13]. Note that α_M is closely related to the $f(R)$ function $B(a)$, the square of the effective Compton wavelength of the scalaron in units of the Hubble scale.

The property function $\alpha_M = d\ln M_\star^2/d\ln a$ describes the running of the Planck mass $M_\star$. Note that the strength of gravity is proportional to $M_\star^{-2}$ just as Newton's constant $G_N = M_{\rm Pl}^{-2}$. We can write

$$M_\star^2(a) = M_{\rm Pl}^2 e^{\int_0^a d\ln a' \alpha_M(a')} \tag{1}$$

$$= M_{\rm Pl}^2 e^{\int_0^a d\ln a' \Omega_{\rm de}(a')[\alpha_M(a')/\Omega_{\rm de}(a')]}, \tag{2}$$

where the second line is in a form suggestive of an approximation where $\alpha_M(a) \propto \Omega_{\rm de}(a)$.

We can immediately see the unfortunate consequences of such a proportionality approximation if it holds into the late Universe. As dark energy dominates, $\Omega_{\rm de} \to 1$, the quantity in brackets remains constant, but a is unbounded. Thus the running Planck mass either goes to infinity or to 0, depending on the sign of the proportionality constant. Indeed we see in the next section that if we match the early time behavior to evaluate the constant, then it is negative, forcing $M_\star^2 \to 0$ at late times. Thus the strength of gravity blows up to infinity in this approximation.

More quantitatively, if we take a ΛCDM background expansion history, or a constant w dark energy equation of state, we can do the integral analytically to find

$$M_\star^2(a) = m_p^2\left(1 + \frac{\Omega_{{\rm de},0}}{\Omega_{m,0}} a^{-3w}\right)^{\bar\alpha_M/(-3w)}, \tag{3}$$

where $\bar\alpha_M$ denotes the proportionality constant and subscripts 0 indicate the present densities of dark energy and matter (and $w = -1$ for the cosmological constant case). As a gets large, $M_\star^2$ is driven to 0 or infinity, depending on the sign of $\bar\alpha_M$.

The way out of this unphysical catastrophe is to break the proportionality

$$\alpha_i^{\rm prop} = \bar\alpha_i \Omega_{\rm de}(a) \tag{4}$$

at some epoch. Indeed, physically we know this must happen: as the Universe approaches a de Sitter state the running of the Planck mass must freeze, i.e. $\alpha_M \to 0$.

Let us explore through the exact numerical solution of the $f(R)$ gravity model when the approximation that the property function (deviation from general relativity) is proportional to the effective dark energy, which we abbreviate as prop, breaks down. Figure 1 shows the numerical solutions for $\alpha_M(a)$ and $\alpha_M(a)/\Omega_{\rm de}(a)$, for the exponential $f(R)$ gravity model with $c = 3$, compatible with current observations, given in [25].

We see that even for the simple $f(R)$ model that $\alpha_M \propto \Omega_{\rm de}(a)$ is a poor approximation. Indeed, $\alpha_M \approx 10^{-10}$ at

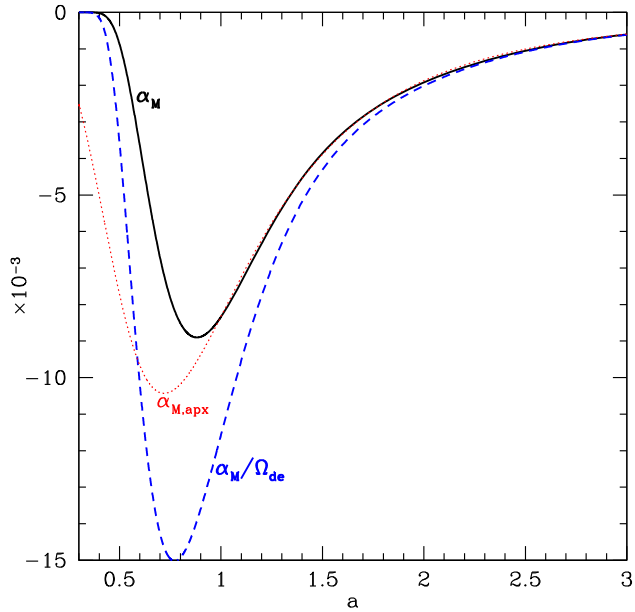

FIG. 1. The property function α_M (solid black) and its ratio relative to the effective dark energy density, $\alpha_M/\Omega_{\rm de}$ (dashed blue), is plotted vs scale factor a. It is well behaved and physical, with the expected early and late time limits. Note that the quantity $\alpha_M/\Omega_{\rm de}$ is definitely not well approximated by a constant in the recent Universe (or the late time Universe). We also show the result (dotted red) when calculating α_M using Eq. (5). It gives an excellent approximation at late times but not at early times; the text explains why any function of $\Omega_{\rm de}$ is likely to fail during the observational epoch.

$a = 0.3$ while $\alpha_M \approx 10^{-2}$ at $a = 1$, while $\Omega_{\rm de}$ only changes by 1 order of magnitude over this range. This should be no surprise: $f(R)$ gravity involves a function of the Ricci scalar R, which has a steep time dependence. Indeed, we want $f(R)$ to restore to general relativity rapidly in the high curvature regime. Due to this very steep dependence, it is difficult to see that any reasonable, model independent function of $\Omega_{\rm de}(a)$ will approximate α_M during the observable epoch $z \approx 0$–3 ($a \approx 0.25$–1).

Recall that prop, i.e. that $\alpha_i(a)/\Omega_{\rm de}(a) =$ constant, has the problematic feature that it forces $M_\star^2 \to 0$ at late times, with the consequence that the strength of gravity, $G_{\rm eff}$, blows up to infinity. Let us attempt to heal this pathology at least.

We take as our ansatz instead

$$M_\star^2 = m_p^2[1 + \mu\Omega_{\rm de}(a)], \tag{5}$$

where μ is a constant. That is, instead of α_i deviating from general relativity at early times proportional to the effective dark energy density, instead it is the running Planck mass that has such a linear deviation. So at early times the Planck mass restores to general relativity, and at late times it freezes to a constant. The latter is what we expect physically in the de Sitter phase. Moreover, now $M_\star^2(a)$ is a function of the background expansion only at that scale factor, rather than an integral over all past history as in the prop case.

From this we find that

$$\alpha_M(a) \equiv \frac{d\ln M_\star^2}{d\ln a} = \frac{-3\mu w(a)\Omega_{\rm de}(a)[1-\Omega_{\rm de}(a)]}{1+\mu\Omega_{\rm de}(a)}, \tag{6}$$

where $w(a)$ is the effective dark energy equation of state function. At early times, for many modified gravity models $w(a)$ is constant so to first order in $\Omega_{\rm de}(a)$ we do have $\alpha_{M,\rm early} \propto \Omega_{\rm de}(a)$. At late times, in the de Sitter phase $\Omega_{\rm de} \to 1$ and hence $\alpha_M \to 0$, exactly as physically expected. Equation (6) for α_M, coming from Eq. (5), is plotted as the dotted red curve in Fig. 1.

While Eq. (5) leads to a remarkably good approximation to α_M for times after the present, it too fails at early times. Indeed this rapid evolution for the gravitational coupling strength was discussed in terms of the "paths of gravity"—the phase space diagram of the gravitational strength—in [26].

Thus, while we managed to remove a pathology and found a parametrization suitable for the latter half of evolution, we still do not see the way to a reliable parametrization for $\alpha_i(a)$ for observational data, even in this simplest case of a single free function of time in the modified gravity theory.

III. EARLY TIME LIMIT

Let us back up and understand why the simplified parametrization $\alpha_i(a) \propto \Omega_{\rm de}(a)$ seemed to be an attractive first attempt. We focus on what physics can lead to such a relation, and what physics breaks it.

In the early time limit we expect general relativity to be an excellent description of gravity, as seen from observational constraints from primordial nucleosynthesis and the cosmic microwave background recombination epoch. Not only should deviations from general relativity be small, but also any contributions of the effective dark energy density—i.e. observations indicate that the Universe was matter dominated (including radiation dominated). The impact of this on the behavior of modified gravity functions was discussed qualitatively by [23] in terms of all the EFT functions being of the same order, as well as showing how this arises within the specific case of covariant Galileons. Some quantitative behaviors for the time dependence of the observables and the property functions $\alpha_i(t)$ were also derived in [24,27].

Within the framework of EFT, or the property functions (we now consider all four within the full Horndeski class as independent functions of time), each function is made up of an array of terms from the theory Lagrangian (see, e.g., [28]). That is, each contains terms depending on different numbers of times the field enters and different numbers of derivatives. Since each term has different dependences on

the Hubble parameter H/H_0, which is large at early times, generically one term dominates at early times. However, as just stated, this term generally contributes to all the property functions, the observables, and the effective dark energy density. Since these are thus proportional to each other and the dark energy density, one has

$$\alpha_{i,\mathrm{early}} \propto \Omega_{\mathrm{de}}(a). \tag{7}$$

We make this relation explicit in the following, and derive the constants of proportionality for various Horndeski cases. However, we emphasize strongly that Eq. (7) is only the early time limit—some specific conditions for when this proportionality breaks down are given in [23] and we elaborate on them here (in particular see Sec. IV), as well as show when this whole ansatz is invalid even at early times.

To calculate the early time relations, recall that for Horndeski gravity the Lagrangian consists of a sum of terms with the scalar field ϕ (and its derivatives) entering two through five times. The prefactors of these operators are functions G_i, with $i = 2, \ldots 5$, and their derivatives, and these Horndeski functions depend on ϕ and its kinetic energy $X = \dot{\phi}^2/2$, i.e. $G_i(\phi, X)$. The early time behavior of G_i is determined by the leading order "pole" behavior, e.g. the lowest power of X (or ϕ) that enters. Thus we treat the early time limit in terms of $G_i \propto X^n$ or ϕ^m. (In the uncoupled covariant Galileon case of Horndeski gravity, G_2, $G_3 \propto X$ and G_4, $G_5 \propto X^2$, with G_4 also having a constant part.)

We can use the generalized Klein-Gordon equation for the scalar field evolution to define

$$\beta \equiv \frac{2\ddot{\phi}}{H\dot{\phi}} = \frac{\dot{X}}{HX}. \tag{8}$$

In the early time limit, β goes to a constant. We can evaluate this constant using that in general the G_5 term dominates at early times due to the number of products of the (large) Hubble parameter from its associated operators. Thus initially we consider $G_5 \propto X^n$ (we consider powers of ϕ later). In this case

$$\beta = \frac{-3(1 + \dot{H}/H^2)(2n+1)}{2n + 1 + (n-1)[5 + 2(n-1)(n-2)]}. \tag{9}$$

Note that since terms like $n - 2$ come from G_{5XX}, i.e. two derivatives with respect to X, if $n = 0$ these terms will not actually exist. For a background equation of state w_b, then $\dot{H}/H^2 = -(3/2)(1 + w_b)$, i.e. $-3/2$ for nonrelativistic matter domination.

Using the known expressions for the property functions α_j in terms of G_i, and for Ω_{de} in terms of G_i, we can solve for the early time limits of the property functions,

$$\frac{\alpha_B}{\Omega_{\mathrm{de}}} \to \frac{3(2n+1)}{2n+3} \longrightarrow \frac{15}{7}, \tag{10}$$

$$\frac{\alpha_K}{\Omega_{\mathrm{de}}} \to \frac{6[7n - 4 + 2(n-1)(n-2)]}{2n+3} \longrightarrow \frac{60}{7}, \tag{11}$$

$$\frac{\alpha_M}{\Omega_{\mathrm{de}}} \to \frac{-3}{2n+3}\left[\frac{\dot{H}}{H^2} + \beta\left(n + \frac{1}{2}\right)\right] \longrightarrow \frac{-9}{56}, \tag{12}$$

$$\frac{\alpha_T}{\Omega_{\mathrm{de}}} \to \frac{-3}{2}\frac{\beta - 2}{2n+3} \longrightarrow \frac{15}{56}. \tag{13}$$

Here the short arrow denotes the early time limit, and the long arrow denotes the further specialization to the covariant Galileon case (where $\beta = 3/4$ in nonrelativistic matter domination). These constants agree with our numerical computation of the full evolution. Note that the first two lines do not depend on β while the last two lines do; one can use Eq. (9) to write those expressions wholly in terms of n.

For the metric

$$ds^2 = -(1 + 2\Psi)dt^2 + a^2(t)(1 - 2\Phi)d\vec{x}^2, \tag{14}$$

we can consider the effective gravitational coupling strengths appearing in the modified Poisson equations for nonrelativistic and relativistic particles, e.g. galaxies and light,

$$\nabla^2\Psi = 4\pi a^2 G_{\mathrm{eff}}^{\Psi}\rho_m\delta_m, \tag{15}$$

$$\nabla^2(\Psi + \Phi) = 8\pi a^2 G_{\mathrm{eff}}^{\Psi+\Phi}\rho_m\delta_m. \tag{16}$$

The quantity G_{eff}^{Ψ} is also called G_{matter} and the quantity $G_{\mathrm{eff}}^{\Psi+\Phi}$ is also called G_{light}.

The property functions can then be propagated to these and other "observables" such as the gravitational slip η and tensor wave speed c_T [24]; using the above early time limits, and specializing to the covariant Galileon limit for simplicity,

$$G_{\mathrm{eff,early}}^{\Psi} = 1 + \frac{759}{224}\Omega_{\mathrm{de}}, \tag{17}$$

$$\eta_{\mathrm{early}} = 1 + \frac{111}{32}\Omega_{\mathrm{de}}, \tag{18}$$

$$c_{T,\mathrm{early}}^2 = 1 + \frac{15}{56}\Omega_{\mathrm{de}}. \tag{19}$$

Now let us consider the case where the ϕ dependence of G_5 is the dominant contribution. In this case one can readily find that

$$\frac{\alpha_B}{\Omega_{\mathrm{de}}} \to \frac{4}{3}, \tag{20}$$

$$\frac{\alpha_K}{\Omega_{\rm de}} \to 2, \tag{21}$$

$$\frac{\alpha_M}{\Omega_{\rm de}} \to -\frac{1}{3}\left[\beta + (m-1)\frac{\dot\phi}{H\phi}\right] \longrightarrow -\frac{\beta}{3}, \tag{22}$$

$$\frac{\alpha_T}{\Omega_{\rm de}} \to \frac{2}{3}. \tag{23}$$

Here the long arrow denotes specialization to the derivatively coupled covariant Galileon, where $G_5 \sim c_G\phi$. However in this case β is no longer given by Eq. (9). Instead, $\beta = 6w_b$. Indeed, from Eq. (50) of [27] we find that in the nonrelativistic matter early time limit with the c_G term dominating, $X = \text{constant}$ and hence $\beta = 0$, so $\alpha_M = 0$.

One can carry out the same analysis if another term than the expected G_5 dominates the Horndeski Lagrangian at early times.

The basic rule is that as long as the same term dominates for both α_i and $\Omega_{\rm de}$, one obtains their proportionality in the early time limit. The next section goes beyond the early time limit, but first we should look for any exceptions to the early time proportionality arising from a mismatch between the terms entering $\Omega_{\rm de}$ and each α_i. We find that indeed α_B, α_K, and α_T all lack certain terms that $\Omega_{\rm de}$ has, while α_M has a term that $\Omega_{\rm de}$ lacks.

For example, α_B is lacking the term $G_{3\phi}$ that $\Omega_{\rm de}$ has, so if the theory is arranged (possibly fine-tuned) to make this dominant at early times, then $\alpha_B/\Omega_{\rm de} \to 0$. A similar situation occurs for α_K and α_T when $G_{4\phi}$ is dominant (and for α_T when any G_3 term dominates). These results have important implications in the next section.

For α_M, there is an extra term involving $\dot\phi^3 G_{5\phi\phi}$. For a leading order behavior of $G_5 \sim \phi^m$, this term involves $m(m-1)(\dot\phi^3/\phi^2)G_5$ and so one can have

$$\frac{\alpha_M}{\Omega_{\rm de}} \sim \frac{\dot\phi}{H\phi} \quad \text{or} \quad \left(\frac{\dot\phi}{H\phi}\right)^2, \tag{24}$$

whichever is dominant, unless $m = 0, 1$. The ratio therefore in general either goes to 0, if $\dot\phi/(H\phi)$ is small, giving a similar problem as with the other α_i, or diverges, if $\dot\phi/(H\phi)$ is large, giving a new problem.

Thus, prop proportionality is not even guaranteed at early times, while we found in Sec. II that it fails during the observational epoch in the recent Universe for the well-known, simple case of $f(R)$ gravity. We investigate further in the next section.

IV. PARAMETRIZING PROPERTY FUNCTIONS

A. Limits to linear proportionality

The early time limit is the only case where the behavior of the property functions, and observables, can be calculated analytically. As seen in the previous section, this showed that at early times α_i was almost always proportional to $\Omega_{\rm de}$. However, there are three important caveats:

(i) We saw in Sec. III that for some theories the constant of proportionality was either 0 or infinity. These give behavior at later times that is clearly invalid (or trivial) within the proportionality approximation.

(ii) In the de Sitter limit one must have $\alpha_M = 0$ so proportionality must break down for this function.

(iii) The behavior actually deviates from the early time asymptote at quite early times.

The discussion in [23] makes clear that linear proportionality breaks down not when $\Omega_{\rm de}$ becomes appreciable compared to unity (i.e. at redshift $z \lesssim 1$) but when H/H_0 is no longer much greater than 1, i.e. at redshift $z \approx 10$ (for a ΛCDM background, $H/H_0 = 20$ at $z = 10$). Remember, the physics comes from the interrelation of the multiple terms in the Lagrangian with different powers of H. Furthermore, we see that even at $z = 10$, the observable functions $G_{\rm matter}$ and $G_{\rm light}$ calculated from the combinations of the property functions are poorly approximated by using a linear $\alpha_i \propto \Omega_{\rm de}$ relation. Note though that the very useful modified gravity Boltzmann code `hi_class` [29] provides Eq. (4) as a default parametrization.

B. Tracker trajectories

Can we force the linear proportionality to hold longer, despite the strong physical basis for the breakdown? If we can freeze the relation between the Lagrangian terms, preventing their natural evolution relative to each other, then this may be possible (although such a construction certainly breaks model independence and raises the specter of fine-tuning). While the terms differ in powers of H, they also differ in powers of other dynamical variables, and so one could impose conditions on their combination to force the terms in lockstep.

This is what "tracker" models do: they fix $H^2X = \text{constant}$ for all times. When this holds, then all the Horndeski G_i are the same order, despite their differing powers of H/H_0. This certainly gives up model independence by narrowing to a specific subclass.[1] Moreover this then implies that $H^2\rho_{\rm de} = \text{constant}$ [31]. That is a dramatic imposition.

Recall that

$$\Omega_{\rm de}(a) \propto \frac{\rho_{\rm de}}{H^2}, \tag{25}$$

so the tracker condition forces

[1]Note that [30] claims that all other regions of covariant Galileons are observationally unviable but their analysis only concerns cubic Galileons, which have only a single term other than the kinetic one and so lack the freedom of the full covariant Galileon, let alone the Horndeski class.

$$\Omega_{\rm de}(a) \propto \left(\frac{H(a)}{H_0}\right)^{-4}. \tag{26}$$

By contrast,

$$\Omega_\Lambda(a) \propto \frac{\rho_\Lambda}{H^2} \propto \left(\frac{H(a)}{H_0}\right)^{-2}. \tag{27}$$

Figure 2 illustrates the implications of this. The tracker model is more fine-tuned than even a cosmological constant, by $(H_0/H)^2$. We see that if one extended this behavior back to the Planck scale, then the tracker model has a fine-tuning of 10^{-240} in contrast to the cosmological constant's 10^{-120}. Some articles in the literature, e.g. [30,32], use this condition back to $z = 10^{14}$, where Fig. 2 shows the fine-tuning is at the level of 10^{-104}, or 10^{52} times more severe than the cosmological constant at that redshift. These results agree completely with Fig. 11 of [30], which only plots the density back to $a = 10^{-3}$.

The physics behind the approach to the $H^2X =$ constant behavior is the same as that causing cosmic acceleration. One can immediately recognize that when this is written in the form $H^2\rho_{\rm de} =$ constant this is merely the de Sitter attractor, when H and $\rho_{\rm de}$ become constant. The natural epoch for the approach to the tracker behavior of $H^2X =$ constant is simply $z \approx 0$; the general physics says that, barring fine-tuning, one would expect the behavior not to hold before the present. This is borne out by numerical solution of the evolution equations.

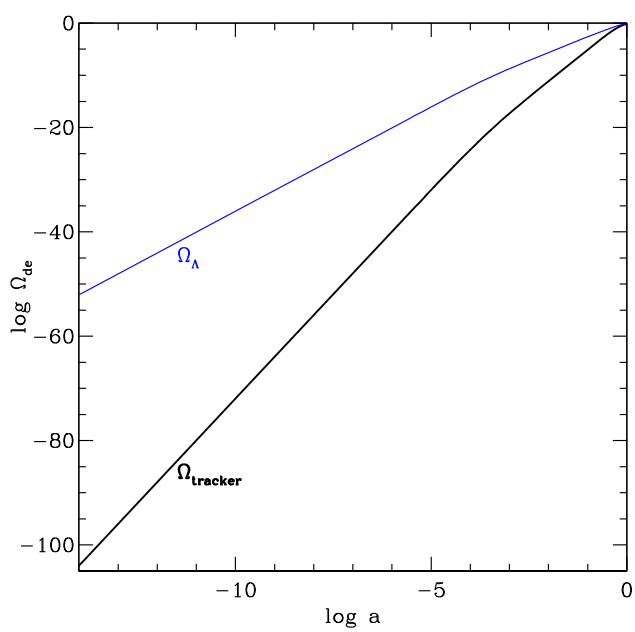

FIG. 2. The fractional energy density in dark energy is plotted vs the log of the scale factor from $a = 10^{14}$ to the present, for the tracker assumption (thick, black curve) and the cosmological constant (thin, blue curve). Assuming that the behavior follows the tracker at early times (rather than only at late times as in the generic physics scenario) imposes severe fine-tuning, well beyond the fine-tuning of the cosmological constant problem.

Suppose one did allow the severe fine-tuning required by Fig. 2. This not only gives up model independence by narrowing to a highly specific subregion of theory space, but also has theoretical issues: the condition $H^2X =$ constant [which motivates Eq. (4)] forces the tensor perturbation propagation speed (for the uncoupled or derivatively coupled Galileon) to be less than the speed of light, $c_T^2 < 1$, leading to a gravi-Cherenkov catastrophe [33–36].

Thus the means of preventing the early breakdown of linear proportionality by forcing $H^2X =$ constant does not appear to be a generally viable solution.

C. Numerical solutions of evolution

Let us examine the exact numerical solutions of the property functions and observable functions to investigate the question of reasonable parametrizations. The relations between the property and observable functions are [13]

$$\frac{G_{\rm eff}^\Phi}{G_N} = \frac{2m_p^2}{M_\star^2}\frac{[\alpha_B(1+\alpha_T)+2(\alpha_M-\alpha_T)]+\alpha_B'}{(2-\alpha_B)[\alpha_B(1+\alpha_T)+2(\alpha_M-\alpha_T)]+2\alpha_B'}, \tag{28}$$

where prime denotes $d/d\ln a$. The gravitational slip $\bar\eta = G_{\rm matter}/G_{\rm light} = G_{\rm eff}^\Psi/G_{\rm eff}^{\Psi+\Phi}$ [note that $\eta = G_{\rm eff}^\Psi/G_{\rm eff}^\Phi = \bar\eta/(2-\bar\eta)$] is given by

$$\bar\eta = \frac{(2+2\alpha_M)[\alpha_B(1+\alpha_T)+2(\alpha_M-\alpha_T)]+(2+2\alpha_T)\alpha_B'}{(2+\alpha_M)[\alpha_B(1+\alpha_T)+2(\alpha_M-\alpha_T)]+(2+\alpha_T)\alpha_B'}, \tag{29}$$

and the tensor wave speed is

$$c_T^2 = 1+\alpha_T. \tag{30}$$

For definiteness in the numerical exploration, we here work with uncoupled covariant Galileon gravity; recall that we showed the results for $f(R)$ gravity in Sec. II. We calculate for the models of Figs. 6 and 4 of [27] (slightly adjusting c_2 to obtain $\Omega_{\rm de,0} = 0.713$), which exhibit very different de Sitter limits for the observable functions. We call these case 1 and case 2. Figure 3 plots the property functions $\alpha_i(a)$, divided by $\Omega_{\rm de}(a)$ to examine whether such a ratio is really constant for all times. The vertical shaded region highlights the region $z = 0$–3 where cosmic structure data exist, and we are particularly interested in an accurate parametrization (taking the constant of proportionality to be a fit parameter, rather than fixed to its analytic, early time value). The quantities $\alpha_i/\Omega_{\rm de}$ are certainly seen to be not well approximated as constant (regardless of the value of the constant), especially over this range.

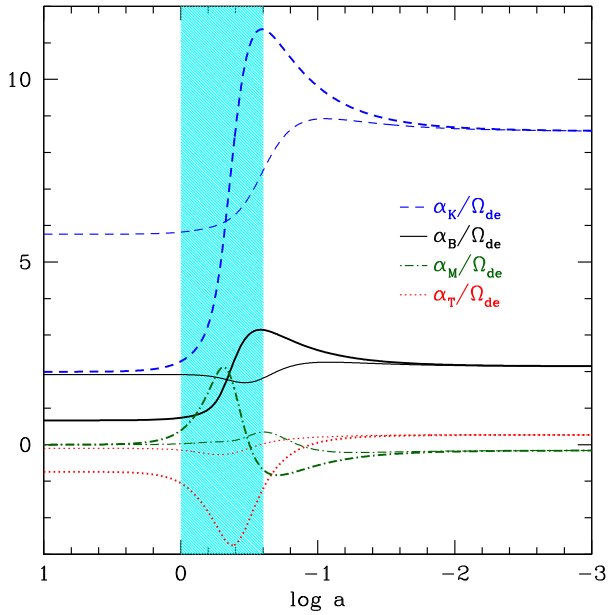

FIG. 3. The time dependence of the property functions divided by the effective dark energy, $\alpha_i(a)/\Omega_{de}(a)$, is exhibited for exact solutions corresponding to case 1 (the thicker curves, with larger variation) and case 2 (the thinner curves, with smaller variation). They are not constant, as the prop approximation of Eq. (4) assumes. Such an approximation is particularly inaccurate during the key observational epoch $z = 0$–3 ($\log a \approx -0.6$–0, shaded).

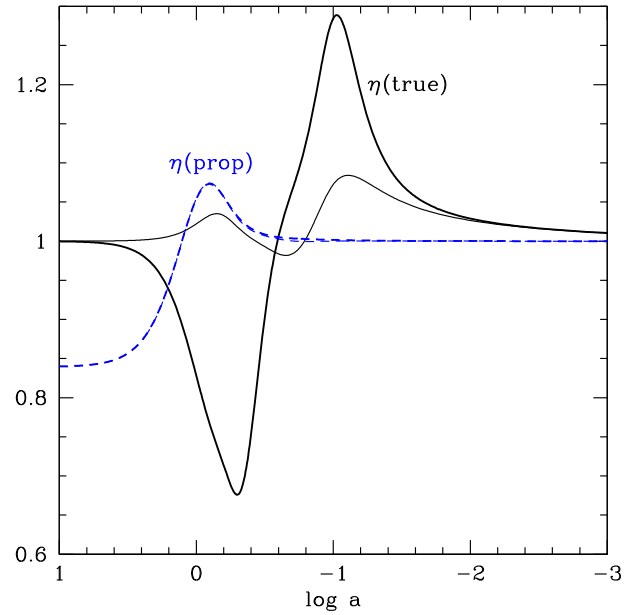

FIG. 4. The time dependence of the gravitational slip η is compared for the exact solution (solid curves) and the $\alpha_i \propto \Omega_{de}(t)$ approximation (dashed curves). The prop approximation gives completely different and inaccurate results for the physics. While the true behavior depends on the differing parameters of the theory (shown for the same two cases as Fig. 3), the prop approximation has nearly identical behavior since the models have nearly the same expansion history. The prop approximation also underestimates the deviation from general relativity, possibly missing detection of modified gravity, and has an incorrect de Sitter asymptote.

Trying to define a constant of proportionality by taking the early time (high redshift) value—where proportionality does hold—can even give the wrong sign during the observational epoch: see the α_M and α_T curves. Note that the physics of Horndeski gravity requires $\alpha_M = 0$ in the de Sitter limit, while the linear proportionality approximation of Eq. (4) violates this for any nonzero constant. Finally, defining the constant of proportionality by an average over cosmic history (hence not obeying either the early time or late time limits) could greatly reduce the sensitivity to observing deviations from general relativity, as we see next.

In Fig. 4 we calculate the gravitational slip observable function η. In addition to the numerical solution we show the predictions for the same models using the linear proportionality approximation. Note that because the background expansion histories for cases 1 and 2 are very similar, the prop approximation, which is a function only of the background, delivers nearly the same observable function for each. However the true solutions show highly differing behaviors for the two cases. Moreover, at $z \gtrsim 1$, prop shows almost no deviation from general relativity (below 1% for $z > 2$ and below 3% for $z = 1$–2), while the true solutions have considerably larger deviations. Thus, using the linear proportionality assumption can miss even quite dramatic signatures of modified gravity. Finally, as we saw for α_M, the prop approximation does not give the physically required de Sitter property that $\eta = 1$ for Horndeski gravity.

Next we consider the gravitational coupling $G_{\rm eff}$. The left panel of Fig. 5 shows the true, numerical solutions and prop predictions for the two cases. The right panel zooms in on the detail within the observational epoch. Again we see that prop almost entirely misses the modified gravity signal, cannot distinguish between the two different cases, and has a pathological late time limit where $G_{\rm eff} \to -\infty$.

V. FITTING MODIFIED GRAVITY

If linear proportionality as a method for parametrizing the time dependence of the property functions is not generically valid, is there some other low dimensional (few parameter) approximation?

Figures 3–5 exhibit the challenge of parametrizing modified gravity with a simple time dependence for either the property functions or observational functions. Even at early times when α_i/Ω_{de} does not appear to be far from constant, the small deviations have a large impact on the observables. For example, at $a = 0.1$ ($z = 9$) the property functions α_i/Ω_{de} deviate from prop by 5%, 4%, 9%, and 24% for subscripts B, K, M, and T. Deviations from general relativity in the observable functions η and $G_{\rm eff}^{\Phi}$ of 7% and 32% respectively—just for case 2, the smaller variation

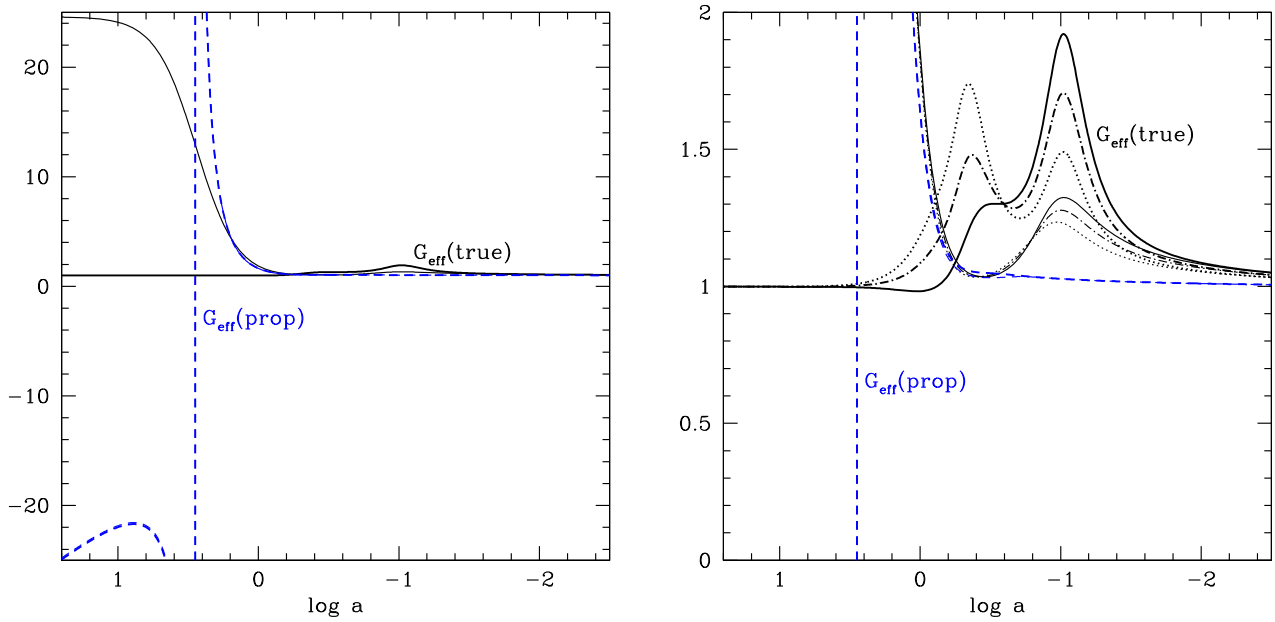

FIG. 5. The time dependence of the gravitational coupling $G_{\rm eff}$ is compared for the exact solution (solid curves) and the $\alpha_i \propto \Omega_{\rm de}(t)$ approximation (dashed curves). The left panel shows the global behavior of $G_{\rm eff}^{\Phi}$ while the right panel zooms in on the detail around the observational epoch and also shows $G_{\rm matter}$ (dotted curves) and $G_{\rm light}$ (dot-dashed curves). The linear proportionality approximation gives completely different and inaccurate results for the physics. While the true behavior depends on the parameters of the theory (shown for the same two cases as Fig. 3), the prop approximation has nearly identical behavior since the models have nearly the same expansion history. The prop approximation also underestimates the deviation from general relativity during the observable epoch, possibly missing detection of modified gravity, and has an incorrect, and divergent, de Sitter asymptote.

case—are missed by the constant proportionality approximation. Recall that $\eta = 1$ to within 1% for $z > 2$ according to the prop approximation.

Note that any attempt to make the property functions $\alpha_i(a)$ follow the effective dark energy density—whether through linear proportionality or a more complicated function—has a disadvantage from a physics perspective. One hallmark of modified gravity is that growth does not follow expansion, so attempting to make the growth purely a function of expansion does not seem to follow this. The four (or more) free functions of time within EFT are in addition to $H(a)$, or $\Omega_{\rm de}(a)$, and it restricts highly the physics if they are all forced to be strictly dependent on it.

Within the observational epoch the behavior tends to be quite complicated. Also, note that in general we need to know values of the property functions or observable functions at all times before the present. The quantity $M_\star^2$ that enters the gravitational strength $G_{\rm eff}$, and hence the growth, requires an integral over all past history (hence the deviations discussed above at $z \geq 2$ are important for lower redshift observations as well). Thus even a three parameter parametrization for each α_i such as using binned values for $z \in [0, 1]$ and $[1, 3]$ and constant proportionality at earlier times does not work well. Approximate functional forms, bins, or principal components all fall short because of the complexity of the relation between the theory and observables; as stated in [23], these relations are at best the ratios of sums of products of ratios of sums of functions.

The failure of parametrizations should not be a huge surprise. The degrees of freedom in a general model are too manifold. For example, although all Horndeski models with a de Sitter late time behavior have the same background expansion and $\eta = 1$ there, the values of $G_{\rm eff}$ can widely vary, as seen in Fig. 5. Similarly, while any Horndeski models with the same dominant function, e.g. $G_5(\phi, X)$, and functional form at early times have the same values of $\alpha_i/\Omega_{\rm de}$ then, at observable times the interplay between all the terms is important and cannot be made model independent. Modified gravity cannot be forced into a few simple numbers without restricting to a specific model or perhaps the benefit of some new theoretical insight.

Even for the simplest case of one function of time, as seen in Fig. 1 for $f(R)$ gravity, the form of the numerical solutions gives no expectation that a simple low order polynomial can capture the richness of the theory, let alone be model independent. We emphasize that this case was wholly observationally viable, so the complicated time dependence is not a matter of a bizarre area of model space, but rather is generic.

VI. CONCLUSIONS

Modified gravity leading to cosmic acceleration is a much richer field than envisioned even a few years ago. The early models like Dvali-Gabadadze-Porrati (DGP) gravity

with a single number (the crossover scale) or $f(R)$ gravity with a time dependent scalaron mass as described by a single power law index of scale factor have much less freedom compared to even the quite restricted covariant Galileon theories with constant coefficients, let alone the Horndeski class or EFT with their several free functions of time.

This complexity, in both the theory and its connections to observables, means that accurate approximations to the observables—being ratios of sums of products of ratios of sums of functions—are rare. We derive analytic limits in the early time, matter dominated regime for general classes of Horndeski gravity, and show under what conditions they appear.

These early time approximations, however, break down dramatically even at redshifts $z \approx 10$, let alone in the heart of the observable epoch. Even percent level deviations in the property functions $\alpha_i(a)$ can lead to large misestimations in observable properties. In particular, we demonstrate that taking them proportional to the effective dark energy density, $\alpha_i(a)/\Omega_{\rm de}(a) \propto$ constant can lead to unphysical behavior and fine-tuning and can miss significant signatures of departure from general relativity. This last property is perhaps most damaging: misestimation could just give a false alert, but lack of an alert misses essential physics [37].

To meet the challenge of connecting theory and observations, we need some parametrization that can prove itself accurate on at least broad swathes of theories in the literature. The numerical solutions we have shown for $f(R)$ and covariant Galileon gravity, demonstrating the complexity of the evolution, indicate that this may be a difficult task. In a real sense this is no surprise: the hallmark of modified gravity is that the physics of growth does not simply follow the expansion history, e.g. $\Omega_{\rm de}(a)$.

If a nearer term goal is merely an alert that general relativity may not be matching observations, then bins in scale and time of $G_{\rm matter}$ and $G_{\rm light}$, proposed in [38,39], work well. Moreover, they would give some indication of how the breakdown occurs, i.e. the trend in space and time variation. While the lack of an elegant parametrization such as exists for the background expansion (e.g. dark energy equation of state) or even simple linear growth (e.g. the gravitational growth index) is disappointing, it also points out the richness of the problem of modified gravity. In Appendix we comment on a conjecture for a general consistency relation between observables that could apply to wide classes of modified gravity theories.

At the same time, we should seek new gravitation theories that are neither overly simplified and so lacking model independence nor complicated but observationally unviable.

ACKNOWLEDGMENTS

I thank Tessa Baker, Alexandre Barreira, Jerome Gleyzes, and Miguel Zumalacárregui for helpful discussions. Some of this work was performed at the Aspen Center for Physics, which is supported by NSF Grant No. PHY-1066293. This work is supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award No. DE-SC-0007867 and Contract No. DE-AC02-05CH11231.

APPENDIX: CONSISTENCY RELATIONS

As an alternative method to that taken in the main text, a broader though less detailed approach is to consider the observables without any parametrization. Are there any properties of them, or relations between them, for which an observational constraint could rule out an entire class of theories? One interesting conjecture, recently put forth by [40], was that the deviation from general relativity of one of the observables (i.e. $G_{\rm matter}$ or $G_{\rm light}$, which they call μ and Σ), either positive or negative at some instant of time, could not have a deviation of opposite sign in the other quantity. The intriguing concept is that an observational violation of such a consistency relation would effectively rule out all Horndeski models. Unfortunately no proof is given but rather an assertion of likeliness. Let us briefly examine the expressions for the deviations and see if such unlikelihood is obvious.

To make the expressions as simple as possible, consider a subclass of Horndeski theory called covariant Galileons. If the consistency relation is not obvious for the simple case, then any obviousness for the general Horndeski class should be more difficult to see. The expressions for $G_{\rm matter}$ and $G_{\rm light}$ are given in [27] (there called $G_{\rm eff}^{\Psi}$ and $G_{\rm eff}^{\Psi+\Phi}$). If the deviations from general relativity $G_{\rm matter} - 1$ and $G_{\rm light} - 1$ have the same sign (note that all gravitational couplings are here normalized by the strength in general relativity, i.e. Newton's constant), then their ratio must be positive,

$$R \equiv \frac{G_{\rm light} - 1}{G_{\rm matter} - 1} > 0. \tag{A1}$$

Writing this out for covariant Galileons,

$$R = \frac{\kappa_6(2\kappa_3 + \kappa_4) - \kappa_1(2\kappa_1 + \kappa_5) - \kappa_5(\kappa_4\kappa_1 - \kappa_5\kappa_3) + \kappa_4(\kappa_4\kappa_6 - \kappa_5\kappa_1)}{4(\kappa_3\kappa_6 - \kappa_1^2) - \kappa_5(\kappa_4\kappa_1 - \kappa_5\kappa_3) + \kappa_4(\kappa_4\kappa_6 - \kappa_5\kappa_1)}, \tag{A2}$$

where κ_i in turn are given by sums of many terms (see [27]). It does not appear obvious that $R > 0$ is required, or even highly likely (and of course the relation in the full Horndeski class is even more complicated).

While a specific model with $R < 0$ is not easily identified, consider instead

$$R_\Phi \equiv \frac{G_{\rm light} - 1}{G_{\rm eff}^\Phi - 1} > 0, \tag{A3}$$

i.e. where $G_{\rm eff}^\Phi$ rather than $G_{\rm eff}^\Psi$ is used in the denominator. A violation of this relation is shown in Sec. IV, and the expression for R_Φ is no more complex or substantially different from that for R.

The consistency relation in terms of $G_{\rm matter}$–$G_{\rm light}$ may indeed hold, but nothing in the equations obviously seems to require this. A firm proof of a consistency condition such as conjectured in [40] would be highly interesting.

[1] S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999).
[2] A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998).
[3] M. J. Mortonson, D. H. Weinberg, and M. White, arXiv: 1401.0046.
[4] J. Frieman, M. Turner, and D. Huterer, Annu. Rev. Astron. Astrophys. **46**, 385 (2008).
[5] A. Joyce, B. Jain, J. Khoury, and M. Trodden, Phys. Rep. **568**, 1 (2015).
[6] K. Koyama, Rep. Prog. Phys. **79**, 046902 (2016).
[7] A. Joyce, L. Lombriser, and F. Schmidt, Annu. Rev. Nucl. Part. Sci. **66**, 95 (2016).
[8] R. R. Caldwell and E. V. Linder, Phys. Rev. Lett. **95**, 141301 (2005).
[9] R. de Putter and E. V. Linder, J. Cosmol. Astropart. Phys. 10 (2008) 042.
[10] E. V. Linder, Phys. Rev. D **72**, 043529 (2005).
[11] E. V. Linder and R. N. Cahn, Astropart. Phys. **28**, 481 (2007).
[12] E. Bertschinger, Phil. Trans. R. Soc. A **369**, 4947 (2011).
[13] E. Bellini and I. Sawicki, J. Cosmol. Astropart. Phys. 07 (2014) 050.
[14] E. Bertschinger and P. Zukin, Phys. Rev. D **78**, 024015 (2008).
[15] A. De Felice, T. Kobayashi, and S. Tsujikawa, Phys. Lett. B **706**, 123 (2011).
[16] G.-B. Zhao, H. Li, E. V. Linder, K. Koyama, D. J. Bacon, and X. Zhang, Phys. Rev. D **85**, 123546 (2012).
[17] A. Silvestri, L. Pogosian, and R. V. Buniy, Phys. Rev. D **87**, 104015 (2013).
[18] S. A. Appleby, A. De Felice, and E. V. Linder, J. Cosmol. Astropart. Phys. 10 (2012) 060.
[19] T. Baker, P. G. Ferreira, C. D. Leonard, and M. Motta, Phys. Rev. D **90**, 124030 (2014).
[20] G. Gubitosi, F. Piazza, and F. Vernizzi, J. Cosmol. Astropart. Phys. 02 (2013) 032.
[21] J. K. Bloomfield, E. E. Flanagan, M. Park, and S. Watson, J. Cosmol. Astropart. Phys. 08 (2013) 010.
[22] J. Gleyzes, D. Langlois, and F. Vernizzi, Int. J. Mod. Phys. D **23**, 1443010 (2014).
[23] E. V. Linder, G. Sengör, and S. Watson, J. Cosmol. Astropart. Phys. 05 (2016) 053.
[24] E. V. Linder, Phys. Rev. D **90**, 083536 (2014).
[25] E. V. Linder, Phys. Rev. D **80**, 123528 (2009).
[26] E. V. Linder, Phil. Trans. R. Soc. A **369**, 4985 (2011).
[27] S. Appleby and E. V. Linder, J. Cosmol. Astropart. Phys. 03 (2012) 043.
[28] J. Bloomfield, J. Cosmol. Astropart. Phys. 12 (2013) 044.
[29] M. Zumalacárregui, E. Bellini, I. Sawicki, and J. Lesgourgues, arXiv:1605.06102.
[30] A. Barreira, B. Li, C. M. Baugh, and S. Pascoli, J. Cosmol. Astropart. Phys. 08 (2014) 059.
[31] A. De Felice and S. Tsujikawa, Phys. Rev. Lett. **105**, 111301 (2010).
[32] J. Renk, M. Zumalacárregui, and F. Montanari, J. Cosmol. Astropart. Phys. 07 (2016) 040.
[33] C. M. Caves, Ann. Phys. (N.Y.) **125**, 35 (1980).
[34] G. D. Moore and A. E. Nelson, J. High Energy Phys. 09 (2001) 023.
[35] J. W. Elliott, G. D. Moore, and H. Stoica, J. High Energy Phys. 08 (2005) 066.
[36] R. Kimura and K. Yamamoto, J. Cosmol. Astropart. Phys. 07 (2012) 050.
[37] R. Munroe, https://xkcd.com/937.
[38] S. F. Daniel, E. V. Linder, T. L. Smith, R. R. Caldwell, A. Cooray, A. Leauthaud, and L. Lombriser, Phys. Rev. D **81**, 123508 (2010).
[39] S. F. Daniel and E. V. Linder, Phys. Rev. D **82**, 103523 (2010).
[40] L. Pogosian and A. Silvestri, Phys. Rev. D **94**, 104014 (2016).