PAPER

Misinterpreting modified gravity as dark energy: a quantitative study

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Misinterpreting modified gravity as dark energy: a quantitative study

Yuewei Wen,[a] **Eva Nesbit,**[b,c] **Dragan Huterer**[a,d,e] **and Scott Watson**[c]

[a]Department of Physics, University of Michigan,
450 Church St., Ann Arbor, MI 48109-1040 U.S.A.
[b]Department of Physics, Kenyon College,
Gambier, Ohio 43022 U.S.A.
[c]Department of Physics, Syracuse University,
Syracuse, NY 13244 U.S.A.
[d]Leinweber Center for Theoretical Physics, University of Michigan,
450 Church St., Ann Arbor, MI 48109-1040 U.S.A.
[e]Max-Planck-Institut für Astrophysik,
Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
E-mail: ywwen@umich.edu, ehnesbit@syr.edu, huterer@umich.edu, gswatson@syr.edu

Received November 9, 2021
Accepted February 11, 2022
Published March 18, 2022

Abstract. Standard cosmological data analyses typically constrain simple phenomenological dark-energy parameters, for example the present-day value of the equation of state parameter, w_0, and its variation with scale factor, w_a. However, results from such an analysis cannot easily indicate the presence of modified gravity. Even if general relativity does not hold, experimental data could still be fit sufficiently well by a phenomenological w_0w_aCDM, unmodified-gravity model. Hence, it would be very useful to know if there are generic signatures of modified gravity in standard analyses. Here we present, for the first time to our knowledge, a quantitative mapping showing how modified gravity models look when (mis)interpreted within the standard unmodified-gravity analysis. Scanning through a broad space of modified-gravity (Horndeski) models, and assuming a near-future survey consisting of CMB, BAO, and SNIa observations, we report values of the best-fit set of cosmological parameters including (w_0, w_a) that would be inferred if modified gravity were at work. We find that modified gravity models that can masquerade as standard gravity lead to very specific biases in standard-parameter spaces. We also comment on implications for measurements of the amplitude of mass fluctuations described by the parameter S_8.

Keywords: dark energy theory, modified gravity, Machine learning

ArXiv ePrint: 2111.02866

 https://doi.org/10.1088/1475-7516/2022/03/042

JCAP03(2022)042

Contents

1 Introduction

Overwhelming observational evidence for the current acceleration of the universe presents one of the most outstanding theoretical challenges in all of cosmology and physics [1, 2]. The physical mechanism for the apparent acceleration remains fundamentally mysterious. It could be given by the presence of the cosmological-constant term in Einstein's equations, but the tiny size of the constant presents an apparently insurmountable challenge [3, 4]. A number of dark energy models beyond the cosmological constant have been proposed as well [5]. Similarly, the accelerated expansion could be that gravity is modified on large scales [6–8], but thus far there is no direct evidence for such a modification.

The difficulty with studying modified-gravity models with data is that the space of possibilities is enormous. There are many completely distinct classes of models to modify gravity and, in each, a large number of possible parameterizations. Constraining any *one* of those modified-gravity model parameterizations with large-scale structure also presents a challenge, for the following reasons: i) modified-gravity-model predictions for nonlinear clustering are, with a few exceptions, not available at all; and ii) the linear-theory predictions generally need to be validated by (modified-gravity) N-body simulations, as e.g. galaxy bias in these models may differ from that in standard gravity (for example [9, 10]). Tests of modified gravity with the cosmic microwave background (CMB) are a little easier as one only needs linear-theory predictions and there is no galaxy bias, but the large scale of possible modified-gravity theories still presents a major obstacle.

As a consequence of these challenges, the majority of confrontations of theory with data has not encompassed models of modified gravity. Instead, most analyses consider simple phenomenological descriptions of the dark-energy sector, such as the model with a cosmological constant (ΛCDM), and that with constant dark-energy equation of state parameter w (wCDM) [11]. Also popular is the time-varying parameterization of the dark-energy equation of state [12] that allows for the dynamics, $w(a) = w_0 + w_a(1 - a)$, where a is the

JCAP03(2022)042

scale factor and w_0 and w_a parameters to be constrained by the data. Modified gravity has typically been constrained only for very specific models (e.g. Σ, μ parameterizations of the gravitational potentials, [13–15]). There have been attempts to constrain individual modified-gravity models [16–39] or even reconstruct the temporal behavior of certain models [40, 41], but canvassing the space of modified-gravity theories is challenging because that space is extremely large and difficult to constrain with currently available cosmological surveys.

In this paper we aim to answer a fundamental question:

What happens when the data is analyzed assuming smooth dark energy and the universe is dominated by modified gravity?

Such a scenario will clearly lead to an overall biased estimate of the inference of the cosmological model; see for example figure 1 in ref. [42]. Yet it would be very useful to know if modified-gravity theories lead to *generic* shifts in the cosmological parameters relative to their true values. For example, it could be that a departure of the equation of state w relative to its ΛCDM value of -1 indicates modified gravity. Or, that the currently observed Hubble tension — the discrepancy between measurements of H_0 from the distance ladder and the CMB — is a signature of modified gravity (something that a number of papers in the literature have explored, e.g. [43, 44]). It would be extremely useful to have knowledge of whether there are any *generic* parameter shifts that modified gravity typically induces if analyzed assuming the standard unmodified model.

To address the highlighted question above, we opt for a forward-modeling approach. We wish to generate a large number of modified-gravity models, coming perhaps from different *classes* of such models, and compute the cosmologically observable quantities. We then analyze those observables using some assumed future data, consisting of the cosmic microwave background, baryon acoustic oscillations, and type Ia supernova (these data are further discussed in section 3). Crucially, when analyzing these data we assume unmodified-gravity, i.e. the ΛCDM or the w_0w_aCDM model. We can thus assess the bias in all cosmological parameters, relative to their true values, due to the fact that data were analyzed using a wrong model. We then iterate the procedure many times. This informs us about what range of values for the standard (unmodified-gravity) cosmological parameters are inferred when the universe is subject to modified gravity.

One important decision in this procedure is to choose a general framework of modified-gravity theories from which to sample individual models. Here we opt to utilize a familiar approach from particle physics (and, as of recently, cosmology) — the Effective Field Theory (EFT). Here our approach is to utilize the EFT of Dark Energy (EFTDE) [45–48], where (universality) classes of models are established through a grouping of terms in the fundamental Lagrangian. This has the advantage that instead of considering one particular model at a time, one can consider an entire class of models with similar properties. One example of such a universality class in the EFTDE are the Horndeski models of modified gravity. In fact, here we will focus our investigation on the Horndeski sub-class of EFTDE models as described in section 2.1 below.

Our procedure in this paper also includes a solution to a pesky technical problem: how to fit the eight-dimensional w_0w_aCDM models to each of the thousands of EFTDE models. This is computationally expensive because traditional Boltzmann-Einstein equation solvers used for this purpose such as `CAMB` are slow for what we are trying to do here. We thus employ and adapt an existing emulator package to speed up this fitting process. This development enables us to obtain our numerical results with relatively modest computer resources. Most

JCAP03(2022)042

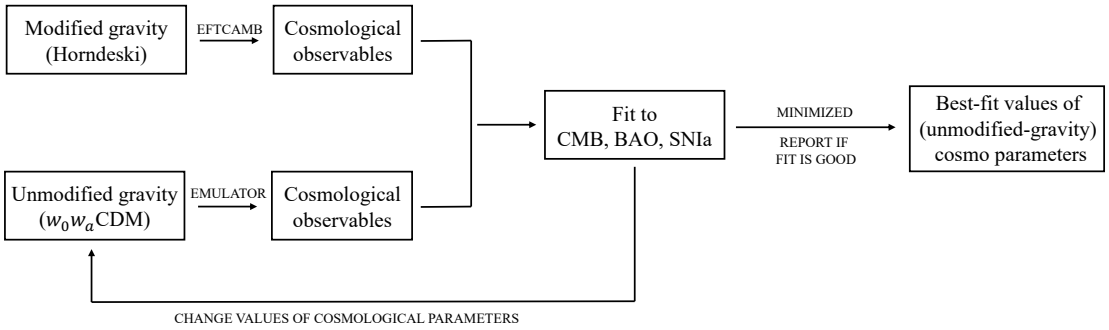

Figure 1. A schematic describing our pipeline to interpret and fit a modified gravity data vector with an (unmodified-gravity) dark energy model. We show the complete procedure for a single Horndeski data vector, corresponding to one point in our final best-fit parameter values in the plots that follow. We repeat this procedure for thousands of Horndeski models.

readily available cosmological emulators for the CMB power spectrum (such as [49] and [50]) function for a fixed set of parameters — usually the standard six cosmological parameters — while our methodology of setting up the emulator allows a much greater freedom in including parameters.

The paper is organized as follows. Section 2 is divided into two parts and gives an overview of our overall methodology. The first half explains how we select a subset of Horndeski gravity models and compute cosmological observable quantities from them. The second half goes over methods (including a brief introduction on the emulation technique) used to reinterpret the data vectors generated by Horndeski models by fitting them with an unmodified-gravity w_0waCDM model. Section 3 introduces the cosmological probes and assumed future experiment data used in the fitting process. Section 4 discusses and summarizes the results. We conclude in section 5.

2 Methodology overview

As discussed in section 1, we generate the data vector assuming a modified-gravity model, but analyze it assuming unmodified gravity in the w_0w_aCDM model. Specifically, for each Horndeski data vector, we generate a CMB angular power spectrum predicted by this theory through a package `EFTCAMB`,[1] and also generate predictions for BAO and SNIa. Then, we fit to this synthetic data with w_0w_aCDM cosmological models. We record the best-fit parameters of such w_0w_aCDM model, and move on to the next iteration, selecting a new EFTDE model. Figure 1 shows our approach schematically.

We now describe the key pieces of our approach: the modified-gravity theory to generate fake data, and the unmodified-gravity theory to analyze it with. For both modified and unmodified-gravity aspects of our analysis, we also discuss the numerical tools that enable the feasibility of our analysis.

2.1 Generating data: modified gravity

Inspired by the EFT formalism for Inflation by Cheung et al. [51], the EFTDE provides a universal description for all viable dark energy and modified gravity models [45–47]. Working

[1] https://eftcamb.github.io.

JCAP03(2022)042

in unitary gauge, the EFTDE action takes the form [47],

$$S = \int d^4x\sqrt{-g}\Big[\frac{1}{2}m_0^2\Omega(t)R - \Lambda(t) - c(t)g^{00} + \frac{M_2^4(t)}{2}(\delta g^{00})^2 \\ -\frac{\bar{M}_1^3(t)}{2}\delta K\delta g^{00} - \frac{\bar{M}_2^2(t)}{2}\delta K^2 - \frac{\bar{M}_3^2(t)}{2}\delta K_\nu{}^\mu\delta K_\mu{}^\nu \\ +\frac{\hat{M}^2(t)}{2}\delta R^{(3)}\delta g^{00} + m_2(t)\partial_i g^{00}\partial^i g^{00} + \mathcal{L}_m\Big], \tag{2.1}$$

where $\delta g^{00} = g^{00} + 1$ is the perturbation to the time component of the metric, $R^{(3)}$ is the perturbation to the spatial component, and $\delta K_{\mu\nu}$ is the perturbation of the extrinsic curvature. The background evolution depends on three functions, $c(t)$, $\Lambda(t)$, and $\Omega(t)$. Two of the three can be constrained using the Einstein equations and are equivalent to the energy density and pressure. The third function, $\Omega(t)$, parameterizes the effect of modified gravity [47]. In what follows we will take $\Omega = 1$, thus explicitly fixing the background to ΛCDM.[2] The rest of the EFT functions describe perturbations about this background and correspond to observables that we are interested in when comparing to observations. For a summary of all models included in this very general formalism, refer to table 1 in [52]. Again, we note that the EFTDE includes such well-known simpler models as DGP and $f(R)$ (see [46, 47] for a discussion).

Here we specialize in a very broad subset of models captured by the EFTDE approach — Horndeski models (for a general review of this class of models see [53] and references therein). These models have been of particular interest because even if one does not take the EFTDE approach they have stable, second order equations of motion, leading to a well defined Cauchy problem and viable models of modified gravity. However, within the EFTDE approach, this is guaranteed from the outset. This universality class of models is obtained when the following relations are imposed on EFTDE functions

$$2\hat{M}^2 = \bar{M}_2^2 = -\bar{M}_3^2; \qquad m_2 = 0. \tag{2.2}$$

We will be interested in the linear-theory predictions of Horndeski models as given by the EFTCAMB code [54]. There, the EFTDE is described in terms of dimensionless parameters γ_i defined as

$$\gamma_1 = \frac{M_2^4}{m_0^2H_0^2}, \quad \gamma_2 = \frac{\bar{M}_1^3}{m_0^2H_0}, \quad \gamma_3 = \frac{\bar{M}_2^2}{m_0^2}, \\ \gamma_4 = \frac{\bar{M}_3^2}{m_0^2}, \quad \gamma_5 = \frac{\hat{M}^2}{m_0^2}, \quad \gamma_6 = \frac{m_2^2}{m_0^2}. \tag{2.3}$$

In terms of these new variables, the Horndeski models are obtained from the full EFTDE with these conditions

$$2\gamma_5 = \gamma_3 = -\gamma_4; \;\; \gamma_6 = 0. \tag{2.4}$$

Our approach is therefore to canvass through the possible Horndeski models by varying $\gamma_i(t)$ for $i = 1, 2, 3$.

There is an important caveat to our assumptions about the Horndeski parameter space. It has been argued that there exists a strong additional constraint on the parameter γ_3, based on the comparison of the speed of light and gravitational-wave speed of propagation from

[2] m_0 is the mass scale of the theory and is equivalent to $m_{\rm pl}$ when $\Omega(t) = 1$.

JCAP03(2022)042

the event GW170817 discovered by LIGO (see e.g. [55]). Because γ_3 is related to the speed of the gravitational wave c_T (see e.g. [52] and references therein), such a constraint would impose a strong prior that γ_3 is very close to zero. However, there are various theoretically motivated possible exceptions to this constraint [56–58]. With that in mind, and to make our analysis broadly applicable and not tied to specific theoretical models, we opt to keep γ_3 as a free parameter without any gravitational-wave-inspired prior. [To reinsert this prior, one could simply inspect and analyze our results evaluated for the small range of γ_3 around zero, although of course such an analysis will necessarily have a lower statistics than one where the γ_3 prior has been assumed from the beginning.]

In our approach, we require Horndeski models to successfully reproduce an approximate ΛCDM background and then focus on the connection between the perturbations and observations. That is, we take the equation of state to be near that of a pure cosmological constant (always with $w > -1$), which in terms of the EFTDE parameters corresponds to a nearly vanishing value of the parameter c, Λ nearly constant and Ω close to unity in the EFTDE. This is a subset of Horndeski models, but corresponds to those consistent with a viable alternative to ΛCDM as required by data. Our approach is similar to that of the EFT of inflation where one assumes an inflationary background and then focuses on the perturbations (observables) [51].

With the background constrained to a ΛCDM universe, we now consider allowed variations in the perturbations of our Horndeski models. Recall that there are three free time-dependent EFTDE functions in Horndeski gravity, $\gamma_i(t)$ for $i = 1, 2, 3$. The first task is to parametrize the time-dependence of these functions, which we take as

$$\gamma_i(a) = \gamma_{i,0} a, \tag{2.5}$$

reproducing the CMB power spectra that are closest to current observations.

Next, we determine the range of the coefficients $\gamma_{i,0}$. In section 2.2, we describe how we set a 5σ requirement for each unmodified-gravity model as to be a good fit for the Horndeski model. By phenomenologically studying sample fits to various Horndeski models, we determine that the Horndeski parameter space restricted to the range

$$\gamma_{1,0} \leq 1; \quad \gamma_{2,0} \leq 0.1; \quad \gamma_{3,0} \leq 0.06, \tag{2.6}$$

encompasses models that are sufficiently in correspondence to unmodified-gravity models, using criteria that we now describe. In figure 2, we display how different values of the Horndeski parameter $\gamma_{3,0}$, which is of particular interest to our analysis in the later sections, impact the peak height of CMB temperature power spectrum.

2.2 Analyzing data: unmodified gravity

Our main goal is to fit simulated modified-gravity data using standard dark energy (unmodified-gravity) models. To be as general as possible, we fit $w_0 w_a$CDM cosmological models to the data, with parameters

$$\{p_i\} = \left\{\omega_b, \omega_c, H_0, \ln(10^{10} A_s), n_s, \tau_{\rm reio}, w_0, w_a\right\}, \tag{2.7}$$

where $\omega_b \equiv \Omega_b h^2$ is the physical baryon density, $\omega_c \equiv \Omega_c h^2$ is the physical cold dark matter density, H_0 is the Hubble constant, A_s is the amplitude of the primordial power spectrum at pivot wave number $k_{\rm piv} = 0.05\,{\rm Mpc}^{-1}$, n_s is the scalar spectral index, $\tau_{\rm reio}$ is the optical

JCAP03(2022)042

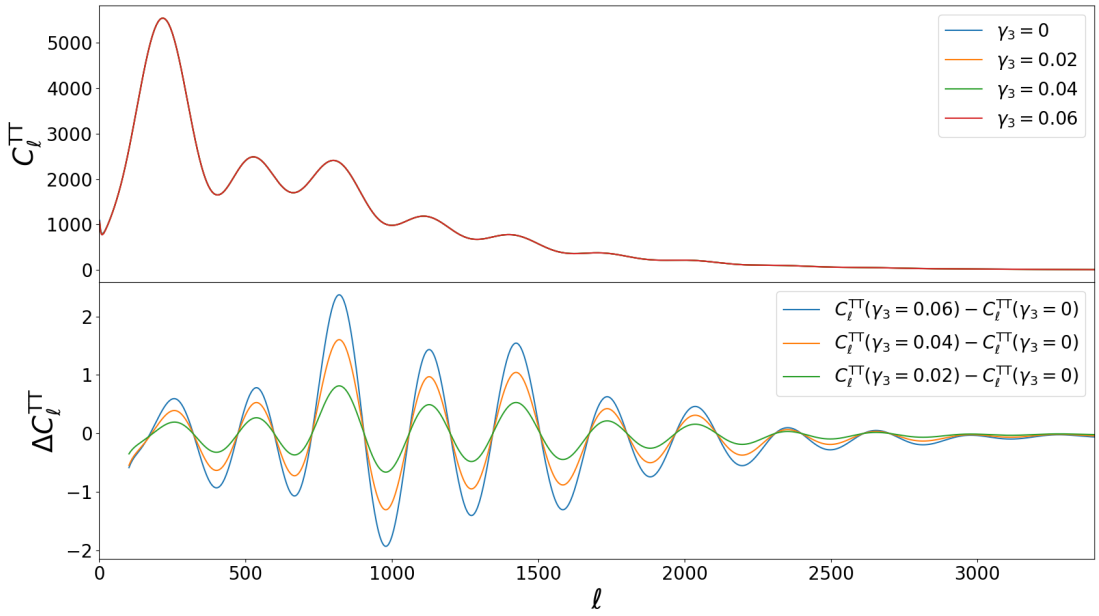

Figure 2. CMB temperature power spectrum generated from EFTCAMB with varied values of the Horndeski parameter $\gamma_{3,0}$ in eq. (2.6) while fixing $\gamma_{1,0} = \gamma_{2,0} = 0$. Increasing the value of $\gamma_{3,0}$ makes the peaks higher and troughs lower.

depth to reionization, and (w_0, w_a) are the parameters describing the dark energy equation of state.

For each Horndeski data vector generated using `EFTCAMB` with assumptions as described in section 2.1, we need to find the best-fit $w_0 w_a$CDM model. We thus need to be able to produce the supernova and BAO observables (distances and the Hubble parameters) and the CMB angular power spectrum in $w_0 w_a$CDM models many times for a single Horndeski model. Calculating distances is straightforward, while the CMB temperature and polarization angular power spectra are typically obtained using the standard Boltzmann-Einstein solver `CAMB`. Here we employ an emulator due to computational cost reasons explained above.

Given a single Horndeski data vector and predictions from unmodified-gravity models, we minimize the total χ^2, defined as a sum of chi-squareds for each cosmological probe in section 3 and thus find the best-fitting parameters. To carry out chi-squared minimization in our eight-dimensional parameter space given in eq. (2.7), we adopt `iminuit`.[3] This optimizer allows us to restart the minimization process from the ending point of the last minimization, re-doing the minimization five times for each EFTDE model to improve the result. The allowed ranges for each parameter to explore is set to be 5% smaller than the parameter range specified in table 1.

As alluded to in section 2.1, we wish to only use reasonably good fits to our Horndeski data vectors, as an analysis resulting in a bad fit to the data would simply not be allowed to proceed in a realistic situation. To that end, we only accept best-fit $w_0 w_a$CDM models that have a minimized χ^2 within 5σ of the expectation for a chi-square distribution of N_{dof} degrees of freedom. Our simulated cosmological data, described below in section 3, have

[3] https://iminuit.readthedocs.io/en/stable.

JCAP03(2022)042

Parameter	Fiducial value	Parameter range
$\Omega_b h^2$	0.02222	(0.02147, 0.02297)
$\Omega_c h^2$	0.1197	(0.1137, 0.1257)
A_s	2.196×10^{-9}	$(1.132 \times 10^{-9}, 2.703 \times 10^{-9})$
H_0	67.5	(64.8, 70.2)
n_s	0.9655	(0.9445, 0.9865)
τ	0.06	(0.0235, 0.0965)
w_0	-1	$(-1.5, -0.5)$
w_a	0	$(-0.5, 0.5)$

Table 1. Fiducial values of cosmological parameters and their ranges used in training the emulator.

$N_{\rm dof} = 7492$.[4] Recall that our simulated Horndeski data vectors are noiseless, so that a perfect fit would have $\chi^2 = 0$. With this information, the "5-σ" limit to a cosmological fit corresponds to chi-square limit of

$$\chi^2 < 650 \quad \text{(acceptable fit)}. \tag{2.8}$$

If the best fit to a given Horndeski model is worse than this, we judge that such a model would not be interpreted as a viable cosmological model. We also exclude results for models where one or more parameters reach the upper or lower bounds of their respective parameter range given in table 1 as it indicates that this model cannot be fitted by a $w_0 w_a$CDM model within the range of current measurements well; this affects about 21 percent of Horndeski models that we considered.

In our model-fitting procedure, the main challenge is the significant computational cost. Consider that `CAMB`[5] takes about 1.5 second to produce a $w_0 w_a$CDM CMB angular power spectrum. For a single Horndeski model, the minimizer requires of order 1,000 $w_0 w_a$CDM model evaluations, and our overall goal is to produce results for 10,000 or more Horndeski models. To addrress this challenge we constructed an emulator to generate model predictions for $w_0 w_a$CDM cosmologies. An emulator is essentially an interpolator. Given a set of grid points in an N-dimensional parameter space and corresponding outcomes evaluated at these points, the emulator interpolates to produce an expected outcome on arbitrary points off the grid (but still within its boundaries). In our case, the grid is the eight-dimensional parameter space listed in eq. (2.7). Since the spectrum is obtained through interpolation, and not from solving the Boltzmann-Einstein equation, this method generates spectra much faster. The emulator we developed builds on the `EGG`[6] package.

- *Parameter ranges:* the prior range for each of the first six parameters in eq. (2.7) is set to $\pm 5\sigma$ around their fiducial values, where σ is the 68% marginalized error on each corresponding parameter from the Planck 2018 analysis using the `Plik` likelihood [59]. For the two dark energy parameters w_0 and w_a, we adopt ranges $-1.5 \leq w_0 \leq -0.5$ and $-0.5 \leq w_a \leq 0.5$. A summary of all parameter ranges are in table 1.

[4] We used 3×2500 multipoles from temperature and polarization spectra respectively as our data, and it was constrained by 8 parameters as listed in table 1.

[5] https://camb.info.

[6] https://github.com/lanl/EGG.

JCAP03(2022)042

- *Parameter grid values:* a uniform grid is not ideal as, for a reasonable number of values in each parameter, it leads to a large number of grid points and slow emulator training. Therefore, we employ the Latin Hypercube sampling (l.h.s.) which is known to be very efficient for emulators [60]. The points in l.h.s. are stratified along the direction of each axis in a multi-dimensional space. This design is mathematically equivalent to forming a $n \times m$ matrix such that every column of this matrix is a unique permutation of $\{1, \ldots, n\}$. There are a number of strategies to design an LHS,[7] and the one we use is provided by a python package `pyDOE`.[8] This package allows us to specify the number of parameters and the number of grid points with much greater flexibility.

- *Training:* to "train" an emulator is to assign the corresponding outcomes to the grid points. Here, we use `CAMB` to calculate the CMB temperature and polarization angular power spectra (TT, EE, and TE) and assign them to the corresponding grid points. During training, the emulator uses a Markov chain Monte Carlo type process to find and optimize an interpolative function that describes the nonlinear relationship between the grid points and their corresponding CMB power spectra.

- *Testing emulator's performance:* the performance of an interpolation under a given LHS setup can be determined quantitatively by comparing the interpolated power spectrum at an arbitrary point in parameter space with the one generated directly by `CAMB`. Adopting a test similar to the one used in [49], we randomly selected 100 points from the allowed parameter space in table 1 and calculated the fractional difference between the angular power spectrum interpolated by the emulator and the power spectrum generated by `CAMB`. For the temperature power spectrum, the emulator's fractional errors within the first and third quartile are 0.3% for multipoles $\ell > 8$. For the polarization power spectra EE and TE, the fractional errors are 0.5% for $\ell > 25$ and 3.5% for $\ell > 55$ respectively.

  The performance of the interpolation is mostly determined by the number of grid points in the LHS design and the number of MCMC iterations when training the emulator. A larger number of grid points and a higher number of steps in the MCMC-type process during training would both improve the performance of interpolation, but at the cost of a slower evaluation per model. In this work, we use 570 grid points and 1000 iterations. With the current setup, each interpolation takes about 0.3 seconds to finish, which is five times faster than using `CAMB`.

  Under this setup, we further validate the emulator's accuracy by testing its ability to recover the input values used to generate the data. Further details are explained in appendix A.

[7] We did not opt for the commonly used orthogonal-array Latin hypercube (OALH) design. This is because using OALH, one relies on the existing library of orthogonal arrays, and the latter does not offer much flexibility to change the number of parameters and the number of samples (i.e. grid points). Specifically, there exist only a few available orthogonal arrays for an eight-dimensional parameter space, and the allowed sample numbers for these arrays are too low for our purposes. The strategy we adopt, as discussed in the text, is not as optimal as the OALH design in its coverage of the parameter space, but its performance can be easily improved through increasing the number of grid points.

[8] `Designs of Experiments for Python`, https://pythonhosted.org/pyDOE/randomized.html#latin-hypercube.

JCAP03(2022)042

Probes	Experiment	Measurements	Details	Data error
CMB	Stage-4	angular power spectrum C_ℓ	from $l = 2$ to $l = 2500$ in temperature (TT) and polarization (EE and TE) spectra	eq. (3.4)
SNIa	WFIRST	apparent magnitude $m(z)$	16 effective supernovae in redshift bins of size 0.1 from $z = 0.1$ to $z = 1.6$	eq. (3.9)
	Pan-STARRS1	apparent magnitude $m(z)$	870 supernovae from $z = 0.00508$ to $z = 1.06039$	ref. [61]
BAO	DESI	angular diameter distance $D_A(z)$ Hubble parameter $H(z)$	13 redshift bins of size 0.1 from $z = 0.65$ to $z = 1.85$	ref. [62]

Table 2. A summary table of the probes and data sets used to determine the best-fit parameters for a certain EFTDE model.

JCAP03(2022)042

3 Simulated data

In this section, we will discuss the probes and experiment specifics we used to determine the best-fit values of dark energy parameters w_0 and w_a.

We use cosmic microwave background, baryon acoustic oscillations (BAO), and type Ia supernovae (SN Ia) as our data. In this first paper on the topic, we opt not to use weak gravitational lensing or galaxy clustering. As mentioned in the introduction, this is due to the significant additional complexity in modeling clustering, which for starters one typically needs to restrict to linear scales only in modified-gravity models as obtaining reliable nonlinear predictions is very challenging. It is our goal to set up a robust proof-of-principle analysis pipeline with the CMB, BAO and SN Ia alone. In a future publication, we will add the galaxy clustering and weak lensing (and, ideally, the full "3x2" pipeline that also includes galaxy-galaxy lensing).

A summary of the probes used can be seen in table 2. We now describe them in more detail.

3.1 CMB

We assume a CMB survey modeled on expectations from CMB-S4 [63]. The survey covering 40% of the sky, with other specifications given below. We utilize scales out to maximum multipole $\ell_{\rm max} = 2500$, consistent with the cutoff in Planck 2018 results [59]. Assuming a Gaussian likelihood $\mathcal{L}$, the chi squared, $\chi^2 \equiv -2 \ln \mathcal{L}$, is given by

$$\chi^2_{\rm CMB} = \sum_{\ell=2}^{\ell=2500} \left(\mathbf{C}_\ell^{\rm data} - \mathbf{C}_\ell^{\rm th}\right)^T \, {\rm Cov}_\ell^{-1} \left(\mathbf{C}_\ell^{\rm data} - \mathbf{C}_\ell^{\rm th}\right), \tag{3.1}$$

where $\mathbf{C}_\ell^{\rm th}$ is the data-vector corresponding to theory ($w_0 w_a$CDM) prediction, and $\mathbf{C}_\ell^{\rm data}$ are the data which, recall, are produced assuming the EFT model. Both the theory and the data $\mathbf{C}_\ell$ are composed of parts corresponding to temperature-temperature (TT), temperature-polarization (TE), and polarization-polarization (EE) correlations:

$$\mathbf{C}_\ell \equiv \begin{pmatrix} C_\ell^{TT} \\ C_\ell^{EE} \\ C_\ell^{TE} \end{pmatrix}. \tag{3.2}$$

The overall covariance matrix Cov_ℓ is diagonal between the different multipoles. At each multipole, the covariance for the data vector $\mathbf{C}_\ell^{\rm data}$ is given by (e.g. [63])

$$\mathrm{Cov}_\ell = \frac{2}{(2\ell+1)f_{\rm sky}} \times \begin{pmatrix} (\tilde{C}_\ell^{TT})^2 & (\tilde{C}_\ell^{TE})^2 & \tilde{C}_\ell^{TT}\tilde{C}_\ell^{TE} \\ (\tilde{C}_\ell^{TE})^2 & (\tilde{C}_\ell^{EE})^2 & \tilde{C}_\ell^{EE}\tilde{C}_\ell^{TE} \\ \tilde{C}_\ell^{TT}\tilde{C}_\ell^{TE} & \tilde{C}_\ell^{EE}\tilde{C}_\ell^{TE} & \frac{1}{2}[(\tilde{C}_\ell^{TE})^2 + \tilde{C}_\ell^{TT}\tilde{C}_\ell^{EE}] \end{pmatrix}. \tag{3.3}$$

The elements of this covariance matrix are explicitly

$$\begin{aligned} \tilde{C}_\ell^{TT} &= C_\ell^{TT} + N_\ell^{TT} \\ \tilde{C}_\ell^{EE} &= C_\ell^{EE} + N_\ell^{EE} \\ \tilde{C}_\ell^{TE} &= C_\ell^{TE}, \end{aligned} \tag{3.4}$$

and the noise terms are

$$\begin{aligned} N_\ell^{TT} &= \Delta_T^2 \exp\left[\frac{\ell(\ell+1)\theta^2_{\rm FWHM}}{8\ln 2}\right] \\ N_\ell^{EE} &= 2 \times N_\ell^{TT}, \end{aligned} \tag{3.5}$$

where $\Delta_T = 1\,\mu K$, $\theta_{\rm FWHM} = 8.7 \times 10^{-4}$ radians, and assume $f_{\rm sky} = 0.4$, using the specifics of the Stage-4 experiment [63].

We generate the data vector $\mathbf{C}_\ell^{\rm data}$ (for each ℓ) using `EFTCAMB`, for a given cosmological model as discussed in section 2.1. This is an important step, as CMB is the only part of our simulated data that is directly affected by modified gravity.

We generate *noiseless* data vectors — that is, the final $\mathbf{C}_\ell^{\rm data}$ used in the likelihood are precisely centered on theory, with no stochastic noise. This assumption is justified because we are not interested in statistical errors on the infered parameters, but rather only at the best-fit parameters (for a given simulated Horndeski model). Had we included stochastic noise, we could have still obtained the results that we are after, but it would have required running a number of statistical realizations of data vectors for a given Horndeski model in order to account for stochasticity in the data.

3.2 SNIa

Type Ia supernovae (SNIa) are sensitive to distances alone. Because in our generated data we fix the background cosmology to ΛCDM and only vary the perturbations according to modified gravity, SNIa data vector is not directly sensitive to modified gravity. Nevertheless, SNIa are very useful in pinning down the cosmological parameters and breaking degeneracies between them, and thus helping isolate the effects of modified gravity on data analyzed assuming w_0w_aCDM.

Assuming again a gaussian likelihood, the chi squared for SNIa measurements is determined by

$$\chi^2_{\rm SN}(\{p_i\}, \mathcal{M}) = (\mathbf{m}^{\rm data} - \mathbf{m}^{\rm th})^T \mathrm{Cov}^{-1} (\mathbf{m}^{\rm data} - \mathbf{m}^{\rm th}),$$

where $\mathbf{m}^{\rm data}$ is the apparent magnitude of simulated data which is calculated based on the cosmology in each fit to the Horndeski model. The theoretical magnitude $\mathbf{m}^{\rm th}$ is, conversely, calculated based on the fiducial w_0w_aCDM cosmological model:

$$m^{\rm th}(z) = 5\log_{10}[H_0 d_L(z, \{p_i\})] + \mathcal{M} \tag{3.6}$$

JCAP03(2022)042

where d_L is the luminosity distance, and $\mathcal{M} = M - 5\log_{10}(H_0 \times 1\text{Mpc}) + 25$ is a nuisance parameter that always needs to be marginalized over in a SNIa analysis. We can analytically marginalize over $\mathcal{M}$ and obtain a marginalized effective χ^2

$$\chi^2_{\text{SN, marg}} = a - \frac{b^2}{c}, \tag{3.7}$$

where

$$\begin{aligned} a &= (\mathbf{m} - \mathbf{m}^{th})^T \text{Cov}^{-1} (\mathbf{m} - \mathbf{m}^{th}) \\ b &= \mathbf{1}^T \text{Cov}^{-1} (\mathbf{m} - \mathbf{m}^{th}) \\ c &= \mathbf{1}^T \text{Cov}^{-1} \mathbf{1}, \end{aligned} \tag{3.8}$$

where $\mathbf{1}$ is a unit vector.

We employed the SNIa redshift bins and the covariance matrix as forecasted for the WFIRST satellite [64]. The covariance matrix is diagonal between different bins, and is calculated as a combination of systematic and statistical errors. In a given redshift bin,

$$\sigma_{\text{tot}} = (\sigma^2_{\text{sys}} + \sigma^2_{\text{stat}})^{1/2}, \tag{3.9}$$

where

$$\begin{aligned} \sigma_{\text{sys}} &= \frac{0.01(1+z)}{1.8} \\ \sigma_{\text{stat}} &= \frac{(\sigma^2_{\text{meas}} + \sigma^2_{\text{int}} + \sigma^2_{\text{lens}})^{1/2}}{N^{1/2}_{\text{SN}}}. \end{aligned} \tag{3.10}$$

Here, $\sigma_{\text{meas}} = 0.08$, $\sigma_{\text{int}} = 0.09$, $\sigma_{\text{lens}} = 0.07z$, and N_{SN} is the number of supernovae in that redshift bin.

We have also incorporated redshift bins and the corresponding covariance matrix from measurements at low redshift by Pantheon dataset [61], which includes 870 supernovae. The covariance matrix for this data set is diagonal, and the error at each redshift is given by Pantheon as well.

3.3 BAO

Baryon acoustic oscillations (BAO) — wiggles in the matter power spectrum due to photon-baryon oscillations prior to recombination — are a powerful cosmological probe. Much like SNIa, they probe geometry, and are sensitive to the angular-diameter distance $D(z)$ and Hubble parameter $H(z)$ evaluated at the redshift of tracer galaxies in question. Often, the general analysis of the BAO provides precisely these "compressed quantities" for one or more effective redshifts, which in turn can be used to constrain a cosmological model.

Here we assume the $D(z)$ and $H(z)$ measurements that are forecasted to be measured DESI experiment [62]. The measurements of both the distances and the Hubble parameters are each reported separately in 13 redshift bins; we thus organize these measurements in data vectors $\mathbf{D}$ and $\mathbf{H}$ that each have 13 elements. As before, we generate synthetic noiseless data ($\mathbf{D}^{\text{data}}$ and $\mathbf{H}^{\text{data}}$) assuming Horndeski models, and analyze it using theoretically computed quantities ($\mathbf{D}^{\text{th}}$ and $\mathbf{H}^{\text{th}}$) that assume the $w_0 w_a$CDM model.

The goodness-of-fit for BAO is written down in a similar way as for the CMB and SNIa

$$\begin{aligned} \chi^2_{\text{BAO}}(\{p_i\}) = &(\mathbf{D}^{\text{data}} - \mathbf{D}^{\text{th}})^T \text{Cov}_D^{-1} (\mathbf{D}^{\text{data}} - \mathbf{D}^{\text{th}}) \\ &+ (\mathbf{H}^{\text{data}} - \mathbf{H}^{\text{th}})^T \text{Cov}_H^{-1} (\mathbf{H}^{\text{data}} - \mathbf{H}^{\text{th}}), \end{aligned} \tag{3.11}$$

JCAP03(2022)042

where Cov_D and Cov_D are respectively the 13×13 covariance matrices for the distance and Hubble parameter measurements, which are diagonal. We adopt these matrices also from DESI forecasts [62].

4 Results and discussions

Our results are summarized in figure 3. Here we show the eight-dimensional space of w_0w_aCDM models that were fit to Horndeski data vectors. Each point corresponds to values of the best-fit w_0w_aCDM model for a given Horndeski model. We show results for a total of 15186 Horndeski data vectors which passed our criteria laid out in section 2.2. We show all possible 2D planes of cosmological parameters, as well as histograms of the distributions in each parameter on the diagonal. The axis limits are chosen so that they indicate the range within which each parameter is allowed to vary during the minimization. The grey crosshair in each panel indicates our fiducial cosmology (see table 1), which corresponds to the background cosmology we set in all our Horndeski models.

Note specifically that figure 3 does *not* show any kind of parameter constraint — that is, no "error bars" are represented here. Rather, in each parameter panel of the figure, the distribution of points relative to the crosshair demonstrates how values of the respective parameters shift relative to their true values when modified gravity (Horndeski) theories are incorrectly interpreted as dark energy (w_0w_aCDM). Recall also that these fits are only performed for w_0w_aCDM models that are decent fits to Horndeski data vectors, judged by the criterion in eq. (2.8), mimicking the decision point that would be applied in an analysis of real data. Finally, the density of points in figure 3 is not particularly important, as it merely reflects the metric on our prior in the space of models (e.g. the fact that we used a flat prior in the parameters γ_i rather than, say, a log prior). What we are interested instead is the overall extents and shapes of the clouds of points.

The most apparent observation from figure 3 is that the biases in w_0w_aCDM parameters, relative to their true values, carve out very specific directions in the parameter space. Table 3 summarizes the directions in which the parameters are shifted. The specific shifts are generally unsurprising, as we would guess that there exist specific degeneracies between Horndeski models and w_0w_aCDM parameters where the former can be interpreted as the latter. Nevertheless, the precision to which the w_0w_aCDM biases are carved out in their respective parameter spaces is remarkable.

The next most noticeable feature of our results are the branchings in the w_0w_aCDM parameter biases. In other words, biases in the parameters trace out multiple (two or three) directions in several 2D parameter planes. This indicates multiple degeneracy directions between shifts in the w_0w_aCDM space and Horndeski models. A very general quantitative expectation for this multi-modality is difficult to establish, but we have nevertheless explored this in some detail. We found that the value of the parameter w_0 — dark energy equation of state value today — is a good predictor for the branchings. Specifically, we found that modified-gravity models that are best fit with, respectively, $w_0 < -1.05$ and $w_0 > 0.97$, lead to two prominent branches that are evident in a number of 2D planes, and that are labeled with green and red points respectively in figure 3. Conversely, models fit with $-1.05 < w_0 < -0.975$, labeled with black points, form the "core" of the distribution, at the nexus of the two branches.

Closing the analysis of figure 3, note that the overall biases in the standard-model parameters are, very roughly, comparable to the current statistical uncertainties in these

JCAP03(2022)042

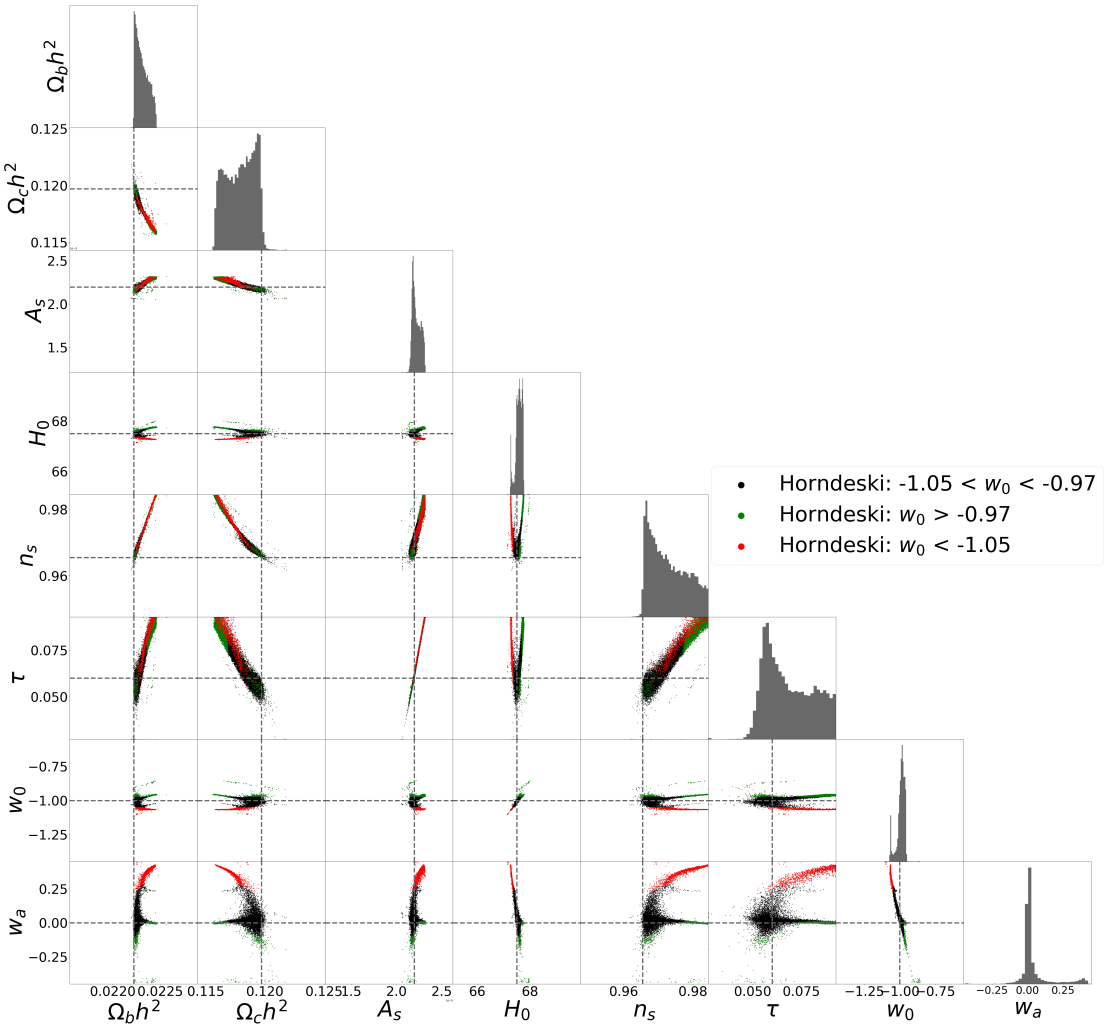

Figure 3. Best-fit values and histograms of cosmological parameters and dark energy sector parameters obtained from fitting to 15186 Horndeski models with a $w_0 w_a$CDM cosmology. Branches shown in the panels along the rows of H_0, w_0 and w_a can be separated by values of w_0, as indicated by red and green points.

parameters. For example, the range of the scalar spectral index, roughly $[0.96, 0.98]$, is somewhat larger than its present statistical uncertainty, while that in the Hubble constant, $[66.86, 68.43]$, is also somewhat larger than the constraints from Planck 2018 analysis [59]. This is not particularly surprising as we have only shown models whose fit to Horndeski data vectors is "good" as quantified in terms of near-future experimental errors. Nevertheless, this tells us that future constraints on these parameters will likely favor a subset of models shown in figure 3. Future data may thus indicate whether a specific sub-class of modified-gravity models lurks in the data.

Of particular interest to cosmologists is the measured value of the equation-of-state parameters (w_0, w_a). Can these measured values indicate the presence of modified gravity? To help answer this question, we enlarge and display figure 3's w_0-w_a plane in the top panels of figure 4. First, note that the w_0 and w_a values of best-fit unmodified-gravity

JCAP03(2022)042

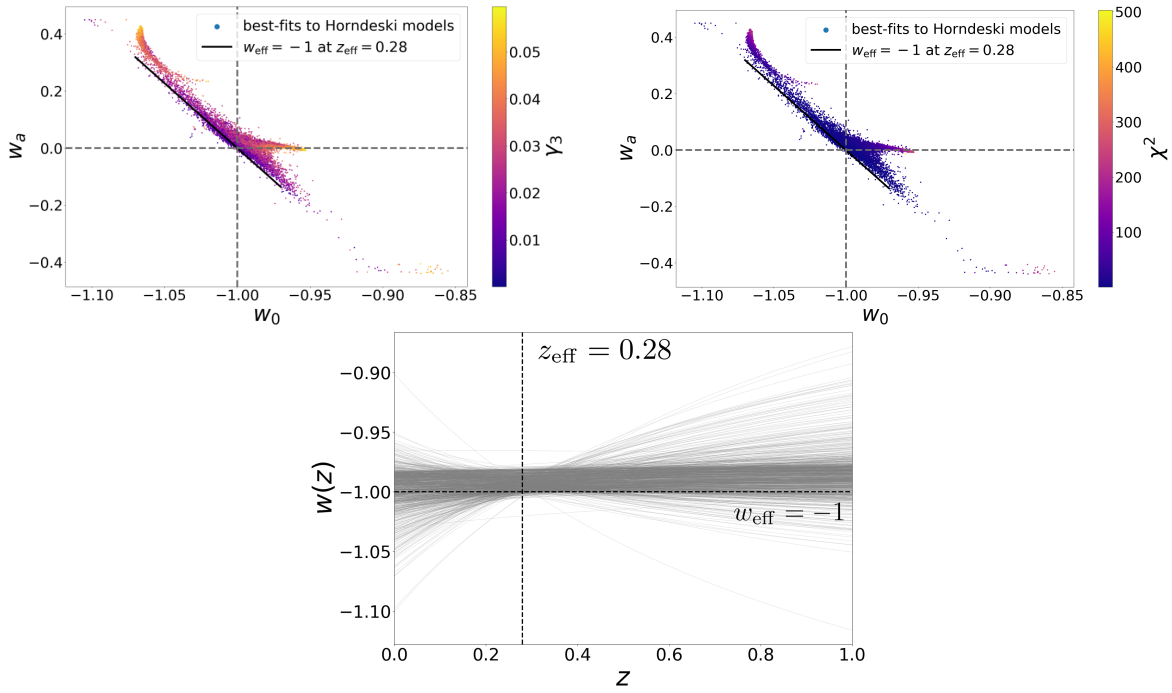

Figure 4. *Top left panel:* the w_0-w_a plane from figure 3, where each point is colored by the γ_3 function of the corresponding Horndeski data vector that was fitted with a $w_0 w_a$CDM cosmology. *Top right panel:* the same as the left panel but each point is colored by the χ^2 value that quantifies the difference between the best-fit power spectrum and the Horndeski data vector. *Bottom panel:* equation of state $w(z)$ for 1000 randomly selected models (corresponding to a subset of points in the purple-pink region in the left panel). Notice that the equation-of-state curves intersect around an effective redshift $z_{\rm eff} = 0.28$, at the value of the effective equation of state typically slightly larger than $w_{\rm eff} = -1$.

models are mutually highly correlated. This is entirely expected, as the physically relevant quantity is $w(z)$ at the redshift where best constrained by the data — the effective, or "pivot" redshift [65, 66]. In fact, it turns out that our range of Horndeski models given by eq. (2.6), the largely one-dimensional direction of best-fit models in w_0-w_a plane is

$$w_{\rm eff} = w_0 + w_a(1 - a_{\rm eff}) \simeq -1 \tag{4.1}$$

with the effective scale factor $a_{\rm eff} = 0.78$ or redshift $z_{\rm eff} = 0.28$.

Therefore, the best-fit models do allow variation in w_0 and w_a, but constrained so that the two parameters combine to produce a constant $w(z)$ at some effective redshift. To illustrate this, the black line in both of the top panels of figure 4 follows combinations of w_0 and w_a that give $w_{\rm eff} \equiv w(z_{\rm eff}) = -1$ at effective redshift $z_{\rm eff} = 0.28$ based on eq. (4.1). [Note that most best-fit models are actually slightly above the black dashed line, indicating that $w_{\rm eff}$ is slightly larger than -1.] The linear relation in eq. (4.1) is not unexpected, as it is really the "physical" value of the equation of state $w_{\rm eff}$ at some redshift $z_{\rm eff}$ to which the theory is most sensitive. In our scenario, dark-energy parameters preferentially follow the relation in eq. (4.1) so as to fit our SNIa and BAO data which are generated using $w = -1$, even while individually departing from the cosmological-constant values of $(w_0, w_a) = (-1, 0)$ in order to fit the CMB data which are generated using Horndeski. Specifically, w_0 and w_a

JCAP03(2022)042

obey a definite one-parameter family of curves for a fixed value of the distance to the last-scattering surface, which the CMB data constrain particularly well [67]. The value of the pivot value that the analysis reports to us, $z_{\rm eff} = 0.28$, merely reflects the typical redshift to which cosmological SNIa, BAO and CMB data are most sensitive [65].

We shed more light on what best-fit (w_0, w_a) values are favored as fits to Horndeski models in the bottom panel of figure 4. Here, each curve represents the function $w(z)$ (in the w_0, w_a model) for each corresponding (purple or pink-colored) point in figure 4. Notably, most best-fit $w(z)$ curves intersect around the effective redshift $z_{\rm eff} = 0.28$, the value that is indicated with a vertical black dashed line.

It is instructive to look at the overall extent of the distribution of models in the top panels of figure 4. The coverage of the w_0-w_a "island" is highly non-uniform, with more models with a positive w_a than negative. In the top left panel, we obtain additional information by plotting the γ_3 parameter from eq. (2.3) for each model, which dominates how far that Horndeski data vector's departure from our background ΛCDM cosmology is. As expected, lower values of γ_3 (i.e. models that resemble the ΛCDM background most) forms the core of the distribution, while models with higher values of γ_3 have larger deviations in (w_0, w_a) and tend to either aggregate in the branch favoring a higher value of w_0 and w_a around zero, or at the upper left tip which favors the lowest values of w_0 but the highest ones of w_a.

In the top right panel, we also color the points in the w_0-w_a plane with their associated values of χ^2 that quantify the difference between the input Horndeski data vector and the data vector corresponding to the best-fit w_0w_aCDM model. The core of the distribution in w_0-w_a is made up of models with a low value of χ^2; these are the models that can be fit well with a w_0w_aCDM cosmology. As in the left panel, models aggregating in the branch on the right, or at the tip on the upper left, are fit less well with a w_0w_aCDM cosmology.

The top panels of figure 4 also show a branching in the distribution of models in the w_0-w_a plane, though weaker than the more prominent ones in the full 8D parameter space seen in figure 3. We did not pursue understanding this feature, given hat it is not extended, and probably encodes subtle correlations between dark energy parameters (w_0, w_a) and Horndeski model parameters when the former are enforced to fit the latter.

Finally, we ask what implications are on two of the most readily measured parameters by lensing surveys — Ω_M and $S_8 \equiv \sigma_8(\Omega_M/0.3)^{0.5}$. Note that the values of these two parameters measured in lensing surveys and the CMB are typically interpreted within the context of the flat ΛCDM cosmological model. Therefore, to infer σ_8 from our set of simulated Horndeski data vectors, we now enforce a fit of modified gravity with a ΛCDM cosmology rather than w_0w_aCDM. We thus fix $w_0 = -1$ and $w_a = 0$, and vary the six other parameters listed in eq. (2.7) to find the best-fit ΛCDM model. Then, we use `CAMB` to calculate the value of σ_8 and the corresponding S_8 for each best-fit ΛCDM model.

We plot ΛCDM's best-fit (Ω_M, S_8) pair for each Horndeski model in figure 5. Each point is colored by the γ_3 parameter (as defined in eq. (2.3)) for each Horndeski model we fitted to. As before, the cross-hairs denote the fiducial, input values of these parameters. In this case, we do not observe a particularly narrow region, or multiple branches, in the best-fit $\Omega_M - S_8$ plane. Rather, we see a near-universal shift to lower values of the best-fit Ω_M, and also a preferential shift toward lower S_8. As the Horndeski model deviates more from general relativity when γ_3 is larger, we observe a shift in Ω_m towards lower values. It is known that Horndeski models can generally accommodate *both* a larger and a smaller amplitude of structure formation relative to the standard model with the same background parameters. However, we need to remember that the CMB measurements at large scales,

JCAP03(2022)042

	w_0w_aCDM		ΛCDM	
Compared to fiducial value	% Larger	% Smaller	% Larger	% Smaller
$\Omega_b h^2$	99.7	0.3	99.9	0.1
$\Omega_c h^2$	2.8	97.2	1.1	98.9
A_s	62.3	37.7	35.2	64.8
H_0	78.6	21.4	99.2	0.8
n_s	99.2	0.8	99.97	0.03
τ	67.5	32.5	41.2	58.8
w_0	73.0	27.0	N/A	N/A
w_a	78.7	21.3		
Ω_m	N/A	N/A	0.9	99.1
S_8			0.7	99.3
$A_s e^{-2\tau}$	6.1	93.9	12.2	87.8

Table 3. Summary of the trends in the inferred cosmological parameters when modified-gravity (Horndeski) models are interpreted within the context of unmodified gravity — either in w_0w_aCDM or ΛCDM cosmology. For each parameter, we show the percentage of best-fit values larger/smaller than the true (input) value. Parameters whose best-fit values are overwhelmingly shifted in the same direction are highlighted in red.

which fit the Integrated Sachs-Wolfe plateau, lie below the ΛCDM prediction. With the newfound parametric freedom in Horndeski models, it appears that the spectral index n_s increases to lower the large-scale power, and in turn lowers Ω_m and σ_8 (with which n_s is negatively correlated) to preserve the good fit at intermediate and smaller scales. This explains why we find the preferentially low Ω_m and σ_8 values in Horndeski models.

We also investigated the biases that one would observe on all six base cosmological parameters when interpreting modified gravity with a ΛCDM cosmology. The results are displayed in figure 6, which contains all possible 2D planes and histograms of cosmological parameters. The grey crosshair again indicates the unbiased, fiducial value of a parameter. In every panel, each point represents a parameter's relative shift or bias resulting from misinterpreting one of the 16769 modified gravity models with dark energy. Here, we observe a shift towards a uniform direction among four of the six parameters, $\Omega_b h^2$, $\Omega_c h^2$, H_0 and n_s, which are listed in table 3. The degenerate combination of $A_s e^{-2\tau}$ also mostly shifts towards a value smaller than the fiducial one.

5 Conclusion

In this work we address the question of how analyses that fit standard cosmological models (say ΛCDM or w_0w_aCDM) to data may show hints of modified gravity. Assume for the moment that modified gravity is at work. In a realistic situation, it is entirely plausible that a standard, unmodified-gravity model is a good fit to the data, so that we cannot immediately rule it out and claim evidence for modified gravity. This scenario, however, will generally

JCAP03(2022)042

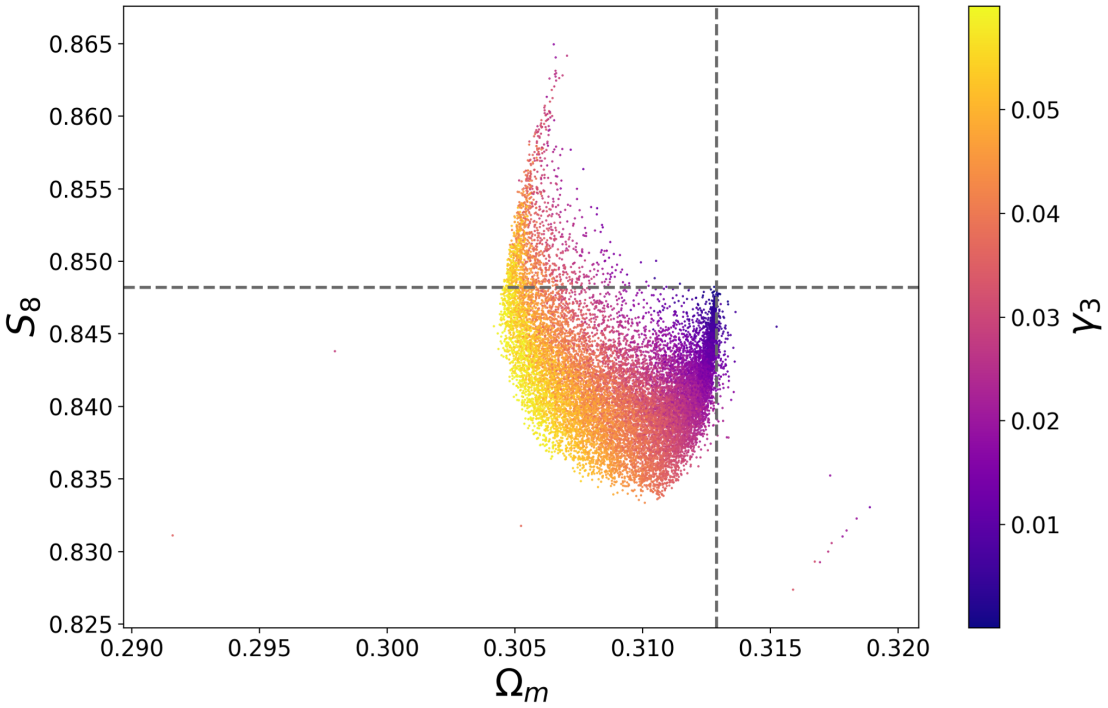

Figure 5. Best-fit values of Ω_m and S_8 derived from fitting 16769 Horndeski data vectors with a ΛCDM cosmological model. Each point is colored by the γ_3 parameter for each Horndeski model as defined in eq. (2.3).

lead to shifts in the (standard-model, unmodified-gravity) parameter values relative to their true values. And such shifts, interpreted together and in relation to other measurements in cosmology that depend on different kinds of data, may reveal the presence of modified gravity.

In this paper, we quantitatively investigate these parameter biases in scenarios when modified gravity is misinterpreted as a standard model. Specifically, we establish the link between modified-gravity models and shifts in the standard cosmological parameters. To scan through a broad range of modified-gravity model, we focus on the Horndeski universality class of models, whose phenomenological predictions (on linear scales) are produced by the code `EFTCAMB` [54]. Horndeski models allow a separate specification of the cosmological theory background and perturbations. For simplicity, we assume a cosmological-constant background for the Horndeski models (in agreement with the most recent cosmological data to date), and vary the perturbations, allowing the full freedom of Horndeski models. We fit these models with simulated future data consisting of CMB temperature and the polarization power spectra, BAO data, and type Ia supernova data. We restrict the analysis to only those Horndeski models whose simulated data vectors are well fit by the w_0w_aCDM model. In doing this we mimic a realistic situation where one would only proceed with the interpretation of model fits in scenarios where the goodness of fit passes some threshold.

We report the best-fit values of the standard cosmological parameters for each Horndeski model that passes the aforementioned cuts. We find that the distribution of the best-fit values cover remarkably tight regions in the standard eight-dimensional parameter space (figure 3). These regions are largely linear, though on occasion carve out multi-pronged directions in

JCAP03(2022)042

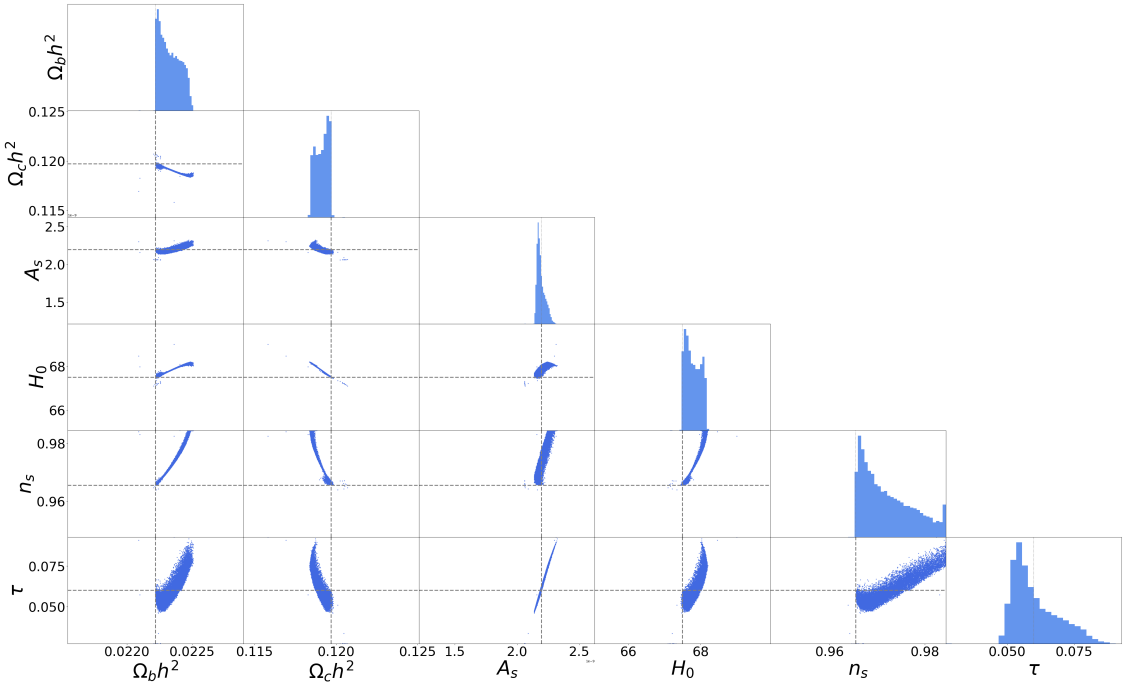

Figure 6. Best-fit values and histograms of cosmological parameters obtained from fitting to 16769 Horndeski models with a ΛCDM cosmology.

the 2D parameter spaces. These tight correlations in standard parameter best-fits imply that even general classes of modified-gravity models register as specific deviations (from true values) in the unmodified-gravity parameters. This is good news; for example, a deviation in standard parameters that does *not* lie in one of these directions would indicate that systematic errors, rather than modified gravity, may be the cause of such unexpected shifts. Hence it should be possible to spot such signatures of systematic errors in future data.

Focusing now on the equation-of-state parameter values that are best fits to Horndeski models, we find that, even though significant deviations in both w_0 and w_a are allowed, they obey a tight mutual relation (figure 4). Specifically, most Horndeski models are fit with an effective equation of state of $w(z_{\rm eff}) \simeq -1$, evaluated at the effective redshift of $z_{\rm eff} = 0.28$. This can be taken as a very generic prediction of the perturbations provided by the large class of modified-gravity models that we study, given a ΛCDM background as stipulated above. This prediction, along with those on all other parameters specified in figure 3, will be sharply tested using upcoming cosmological data.

We finally study the implications of our result to the currently much debated tension between constraints on the S_8 parameter obtained from lensing probes and CMB measurements. Assuming now the ΛCDM model (in which the S_8 tension is usually framed), we find that Horndeski models typically predict a lower S_8, and near-universally a lower Ω_M, than the truth when the latter two are inferred assuming the ΛCDM model. Because the only direct probe of S_8 that we assumed was the CMB, this implies that CMB's S_8 value is preferentially low when Horndeski data are analyzed assuming the ΛCDM model. This should be compared to the prediction from applying the same pipeline to lensing data, something we plan to do in a future work.

JCAP03(2022)042

Acknowledgments

We would like to thank Marco Raveri and Alessandra Silvestri for useful conversations, Anqi Chen for her initial collaboration, and Earl Lawrence for assistance on the usage of the emulator. We also thank Eric Linder for useful comments on an earlier draft of this paper. The work of EN and SW was supported in part by NSF grant 1813834 and NSF grant PHY-1748958. They would also like to thank the Kavli Institute for Theoretical Physics and the Aspen Center for Physics for hospitality. YW and DH have been supported by NSF under contract AST-1812961. DH has additionally been supported by DOE under Contract No. DE-FG02-95ER40899, and he thanks the Humboldt Foundation for support via the Friedrich Wilhelm Bessel award. We acknowledge the Syracuse University HTC Campus Grid and NSF award ACI-1341006 for the use of computing resources.

A Fitting error

Here, we illustrate the extent of uncertainty in our process of finding best-fits. In each panel of figure 7, there are 93 blue points, each generated from fitting the 8 standard cosmological parameters to the fiducial cosmology listed in table 1. The dim light grey, green and red points in the background are the same as the corresponding points in figure 3, and in both figures they denote the best-fit parameter values to Horndeski data vectors. For a perfect fitting process, the blue points should all coincide with the grey cross-hair, which indicates the fiducial values of each parameter. Our fitting error, as indicated by level of scatter among the blue points and the historgrams, is small compared to both the best-fits to Horndeski data vectors and the parameters' allowed ranges of variation.

JCAP03(2022)042

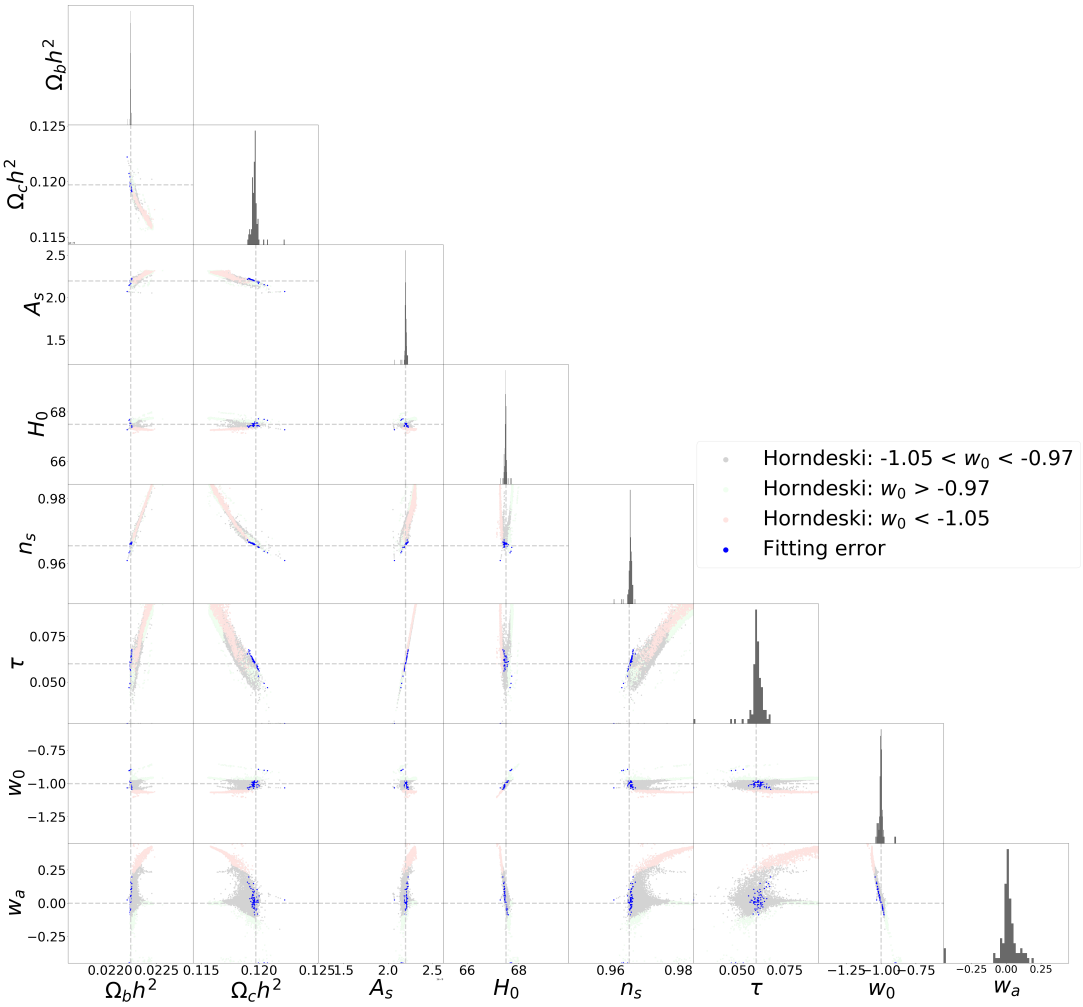

Figure 7. A test of the performance of the emulator and our minimization tool, `iminuit`. The test consists of 93 separate fits of cosmological parameters to the same CMB power spectrum generated by `CAMB` (with input parameter values as in table 1); each fit starts from a different, randomly chosen, starting point in parameter space. The best-fits parameter values are plotted as the blue points in each panel (superimposed to results from figure 3). The histograms on the diagonal show the distribution of the recovered values for the corresponding parameter on the vertical axis. These results demonstrate that the emulator and `iminuit` successfully and accurately recover the input cosmological parameter values, which are shown by the cross-hair in each panel.

References

[1] J. Frieman, M. Turner and D. Huterer, *Dark energy and the accelerating universe*, *Ann. Rev. Astron. Astrophys.* **46** (2008) 385 [arXiv:0803.0982] [INSPIRE].

[2] D. Huterer and D.L. Shafer, *Dark energy two decades after: observables, probes, consistency tests*, *Rept. Prog. Phys.* **81** (2018) 016901 [arXiv:1709.01091] [INSPIRE].

[3] S. Weinberg, *The cosmological constant problem*, *Rev. Mod. Phys.* **61** (1989) 1 [INSPIRE].

[4] S.M. Carroll, *The cosmological constant*, *Living Rev. Rel.* **4** (2001) 1 [astro-ph/0004075] [INSPIRE].

JCAP03(2022)042

[5] E.J. Copeland, M. Sami and S. Tsujikawa, *Dynamics of dark energy*, *Int. J. Mod. Phys. D* **15** (2006) 1753 [hep-th/0603057].

[6] T. Clifton, P.G. Ferreira, A. Padilla and C. Skordis, *Modified gravity and cosmology*, *Phys. Rept.* **513** (2012) 1 [arXiv:1106.2476] [INSPIRE].

[7] A. Joyce, B. Jain, J. Khoury and M. Trodden, *Beyond the cosmological standard model*, *Phys. Rept.* **568** (2015) 1 [arXiv:1407.0059] [INSPIRE].

[8] A. Silvestri and M. Trodden, *Approaches to understanding cosmic acceleration*, *Rept. Prog. Phys.* **72** (2009) 096901 [arXiv:0904.0024] [INSPIRE].

[9] C. Arnold, M. Leo and B. Li, *Realistic simulations of galaxy formation in $f(R)$ modified gravity*, *Nature Astron.* **3** (2019) 945 [arXiv:1907.02977] [INSPIRE].

[10] M.A. Mitchell, C. Arnold and B. Li, *A general framework to test gravity using galaxy clusters – V. A self-consistent pipeline for unbiased constraints of f(R) gravity*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4157 [arXiv:2107.14224] [INSPIRE].

[11] M.S. Turner and M.J. White, *CDM models with a smooth component*, *Phys. Rev. D* **56** (1997) R4439 [astro-ph/9701138] [INSPIRE].

[12] E.V. Linder, *Exploring the expansion history of the universe*, *Phys. Rev. lett.* **90** (2003) 4 [astro-ph/0208512].

[13] P. Zhang, M. Liguori, R. Bean and S. Dodelson, *Probing gravity at cosmological scales by measurements which test the relationship between gravitational lensing and matter overdensity*, *Phys. Rev. Lett.* **99** (2007) 141302 [arXiv:0704.1932] [INSPIRE].

[14] S.F. Daniel and E.V. Linder, *Constraining cosmic expansion and gravity with galaxy redshift surveys*, *JCAP* **02** (2013) 007 [arXiv:1212.0009] [INSPIRE].

[15] L. Pogosian and A. Silvestri, *What can cosmology tell us about gravity? Constraining Horndeski gravity with Σ and μ*, *Phys. Rev. D* **94** (2016) 104014 [arXiv:1606.05339] [INSPIRE].

[16] P. Zhang, *Testing $f(R)$ gravity against the large scale structure of the universe*, *Phys. Rev. D* **73** (2006) 123504 [astro-ph/0511218] [INSPIRE].

[17] R. Caldwell, A. Cooray and A. Melchiorri, *Constraints on a new post-general relativity cosmological parameter*, *Phys. Rev. D* **76** (2007) 023507 [astro-ph/0703375] [INSPIRE].

[18] J. Guzik, B. Jain and M. Takada, *Tests of gravity from imaging and spectroscopic surveys*, *Phys. Rev. D* **81** (2010) 023503 [arXiv:0906.2221] [INSPIRE].

[19] R. Bean and M. Tangmatitham, *Current constraints on the cosmic growth history*, *Phys. Rev. D* **81** (2010) 083534 [arXiv:1002.4197] [INSPIRE].

[20] G.-B. Zhao et al., *Probing modifications of General Relativity using current cosmological observations*, *Phys. Rev. D* **81** (2010) 103510 [arXiv:1003.0001] [INSPIRE].

[21] R. Reyes et al., *Confirmation of general relativity on large scales from weak lensing and galaxy velocities*, *Nature* **464** (2010) 256 [arXiv:1003.2185] [INSPIRE].

[22] S.F. Daniel et al., *Testing general relativity with current cosmological data*, *Phys. Rev. D* **81** (2010) 123508 [arXiv:1002.1962] [INSPIRE].

[23] S.F. Daniel and E.V. Linder, *Confronting general relativity with further cosmological data*, *Phys. Rev. D* **82** (2010) 103523 [arXiv:1008.0397] [INSPIRE].

[24] G.-B. Zhao, R.G. Crittenden, L. Pogosian and X. Zhang, *Examining the evidence for dynamical dark energy*, *Phys. Rev. Lett.* **109** (2012) 171301 [arXiv:1207.3804] [INSPIRE].

[25] M. Raveri, B. Hu, N. Frusciante and A. Silvestri, *Effective field theory of cosmic acceleration: constraining dark energy with CMB data*, *Phys. Rev. D* **90** (2014) 043513 [arXiv:1405.1022] [INSPIRE].

[26] E. Bellini, A.J. Cuesta, R. Jimenez and L. Verde, *Constraints on deviations from ΛCDM within Horndeski gravity*, *JCAP* **02** (2016) 053 [*Erratum ibid.* **06** (2016) E01] [arXiv:1509.07816] [INSPIRE].

JCAP03(2022)042

[27] A. Hojjati et al., *Searching for scalar gravitational interactions in current and future cosmological data*, *Phys. Rev. D* **93** (2016) 043531 [arXiv:1511.05962] [INSPIRE].

[28] V. Salvatelli, F. Piazza and C. Marinoni, *Constraints on modified gravity from Planck 2015: when the health of your theory makes the difference*, *JCAP* **09** (2016) 027 [arXiv:1602.08283] [INSPIRE].

[29] S. Joudaki et al., *KiDS-450: testing extensions to the standard cosmological model*, *Mon. Not. Roy. Astron. Soc.* **471** (2017) 1259 [arXiv:1610.04606] [INSPIRE].

[30] E.-M. Mueller et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: constraining modified gravity*, *Mon. Not. Roy. Astron. Soc.* **475** (2018) 2122 [arXiv:1612.00812] [INSPIRE].

[31] G.-B. Zhao et al., *Dynamical dark energy in light of the latest observations*, *Nature Astron.* **1** (2017) 627 [arXiv:1701.08165] [INSPIRE].

[32] A. Amon et al., *KiDS+2dFLenS+GAMA: testing the cosmological model with the E_G statistic*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 3422 [arXiv:1711.10999] [INSPIRE].

[33] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [arXiv:1807.06209] [INSPIRE].

[34] DES collaboration, *Dark energy survey year 1 results: constraints on extended cosmological models from galaxy clustering and weak lensing*, *Phys. Rev. D* **99** (2019) 123505 [arXiv:1810.02499] [INSPIRE].

[35] J. Noller and A. Nicola, *Cosmological parameter constraints for Horndeski scalar-tensor gravity*, *Phys. Rev. D* **99** (2019) 103502 [arXiv:1811.12928] [INSPIRE].

[36] A. Spurio Mancini et al., *KiDS + GAMA: constraints on horndeski gravity from combined large-scale structure probes*, *Mon. Not. Roy. Astron. Soc.* **490** (2019) 2155 [arXiv:1901.03686] [INSPIRE].

[37] EBOSS collaboration, *Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory*, *Phys. Rev. D* **103** (2021) 083533 [arXiv:2007.08991] [INSPIRE].

[38] KIDS collaboration, *KiDS-1000 Cosmology: constraints beyond flat ΛCDM*, *Astron. Astrophys.* **649** (2021) A88 [arXiv:2010.16416] [INSPIRE].

[39] DES collaboration, *Probing gravity with the DES-CMASS sample and BOSS spectroscopy*, *Mon. Not. Roy. Astron. Soc.* **509** (2021) 4982 [arXiv:2104.14515] [INSPIRE].

[40] M. Raveri, *Reconstructing gravity on cosmological scales*, *Phys. Rev. D* **101** (2020) 083524 [arXiv:1902.01366] [INSPIRE].

[41] L. Pogosian, M. Raveri, K. Koyama, M. Martinelli, A. Silvestri and G.-B. Zhao, *Imprints of cosmological tensions in reconstructed gravity*, arXiv:2107.12992 [INSPIRE].

[42] D. Huterer and E.V. Linder, *Separating dark physics from physical darkness: minimalist modified gravity vs. dark energy*, *Phys. Rev. D* **75** (2007) 023519 [astro-ph/0608681] [INSPIRE].

[43] M.-X. Lin, M. Raveri and W. Hu, *Phenomenology of modified gravity at recombination*, *Phys. Rev. D* **99** (2019) 043514 [arXiv:1810.02333] [INSPIRE].

[44] M. Braglia, M. Ballardini, F. Finelli and K. Koyama, *Early modified gravity in light of the H_0 tension and LSS data*, *Phys. Rev. D* **103** (2021) 043528 [arXiv:2011.12934] [INSPIRE].

[45] M. Park, K.M. Zurek and S. Watson, *A unified approach to cosmic acceleration*, *Phys. Rev. D* **81** (2010) 124008 [arXiv:1003.1722] [INSPIRE].

[46] G. Gubitosi, F. Piazza and F. Vernizzi, *The effective field theory of dark energy*, *JCAP* **02** (2013) 032 [arXiv:1210.0201] [INSPIRE].

[47] J.K. Bloomfield, E.E. Flanagan, M. Park and S. Watson, *Dark energy or modified gravity? An effective field theory approach*, *JCAP* **08** (2013) 010 [arXiv:1211.7054] [INSPIRE].

[48] N. Frusciante and L. Perenon, *Effective field theory of dark energy: a review*, *Phys. Rept.* **857** (2020) 1 [arXiv:1907.03150] [INSPIRE].

JCAP03(2022)042

[49] M.D. Schneider, O. Holm and L. Knox, *Intelligent design: on the emulation of cosmological simulations*, *Astrophys. J.* **728** (2011) 137 [arXiv:1002.1752] [INSPIRE].

[50] A. Spurio Mancini, D. Piras, J. Alsing, B. Joachimi and M.P. Hobson, *CosmoPower: emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys*, *Mon. Not. Roy. Astron. Soc.* **511** (2022) 1771 [arXiv:2106.03846] [INSPIRE].

[51] C. Cheung, P. Creminelli, A.L. Fitzpatrick, J. Kaplan and L. Senatore, *The effective field theory of inflation*, *JHEP* **03** (2008) 014 [arXiv:0709.0293] [INSPIRE].

[52] E.V. Linder, G. Sengör and S. Watson, *Is the effective field theory of dark energy effective?*, *JCAP* **05** (2016) 053 [arXiv:1512.06180] [INSPIRE].

[53] T. Kobayashi, *Horndeski theory and beyond: a review*, *Rept. Prog. Phys.* **82** (2019) 086901 [arXiv:1901.07183] [INSPIRE].

[54] B. Hu, M. Raveri, N. Frusciante and A. Silvestri, *EFTCAMB/EFTCosmoMC: numerical notes v3.0*, arXiv:1405.3590 [INSPIRE].

[55] C.D. Kreisch and E. Komatsu, *Cosmological constraints on Horndeski gravity in light of GW170817*, *JCAP* **12** (2018) 030 [arXiv:1712.02710] [INSPIRE].

[56] C. de Rham and S. Melville, *Gravitational rainbows: LIGO and dark energy at its cutoff*, *Phys. Rev. Lett.* **121** (2018) 221101 [arXiv:1806.09417] [INSPIRE].

[57] L. Amendola, D. Bettoni, G. Domènech and A.R. Gomes, *Doppelgänger dark energy: modified gravity with non-universal couplings after GW170817*, *JCAP* **06** (2018) 029 [arXiv:1803.06368] [INSPIRE].

[58] R.A. Battye, F. Pace and D. Trinh, *Gravitational wave constraints on dark sector models*, *Phys. Rev. D* **98** (2018) 023504 [arXiv:1802.09447] [INSPIRE].

[59] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [arXiv:1807.06209] [INSPIRE].

[60] K. Heitmann, D. Higdon, M. White, S. Habib, B.J. Williams and C. Wagner, *The Coyote Universe II: cosmological models and precision emulation of the nonlinear matter power spectrum*, *Astrophys. J.* **705** (2009) 156 [arXiv:0902.0429] [INSPIRE].

[61] PAN-STARRS1 collaboration, *The complete light-curve sample of spectroscopically confirmed SNe Ia from Pan-STARRS1 and cosmological constraints from the combined pantheon sample*, *Astrophys. J.* **859** (2018) 101 [arXiv:1710.00845] [INSPIRE].

[62] DESI collaboration, *The DESI Experiment Part I: science, targeting, and survey design*, arXiv:1611.00036 [INSPIRE].

[63] X. Li, N. Weaverdyck, S. Adhikari, D. Huterer, J. Muir and H.-Y. Wu, *The quest for the inflationary spectral runnings in the presence of systematic errors*, *Astrophys. J.* **862** (2018) 137 [arXiv:1806.02515] [INSPIRE].

[64] R. Hounsell et al., *Simulations of the WFIRST supernova survey and forecasts of cosmological constraints*, *Astrophys. J.* **867** (2018) 23 [arXiv:1702.01747] [INSPIRE].

[65] D. Huterer and G. Starkman, *Parameterization of dark-energy properties: a principal-component approach*, *Phys. Rev. Lett.* **90** (2003) 031301 [astro-ph/0207517] [INSPIRE].

[66] E.V. Linder, *Biased cosmology: pivots, parameters, and figures of merit*, *Astropart. Phys.* **26** (2006) 102 [astro-ph/0604280] [INSPIRE].

[67] E.V. Linder, *The Mirage of $w = -1$*, arXiv:0708.0024 [INSPIRE].

JCAP03(2022)042