

Observational future of cosmological scalar-tensor theoriesD. Alonso,¹ E. Bellini,^{1,2} P. G. Ferreira,¹ and M. Zumalacárregui^{3,4}¹*University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom*²*ICCUB, University of Barcelona, (IEEC-UB), Martí i Franquès, E08028 Barcelona, Spain*³*Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden*⁴*Berkeley Center for Cosmological Physics and University of California at Berkeley, California 94720, USA*

(Received 29 October 2016; published 3 March 2017)

The next generation of surveys will greatly improve our knowledge of cosmological gravity. In this paper we focus on how Stage IV photometric redshift surveys, including weak lensing and multiple tracers of the matter distribution and radio experiments combined with measurements of the cosmic microwave background will lead to precision constraints on deviations from general relativity. We use a broad subclass of Horndeski scalar-tensor theories to forecast the accuracy with which we will be able to determine these deviations and their degeneracies with other cosmological parameters. Our analysis includes relativistic effects, does not rely on the quasistatic evolution and makes conservative assumptions about the effect of screening on small scales. We define a figure of merit for cosmological tests of gravity and show how the combination of different types of surveys, probing different length scales and redshifts, can be used to pin down constraints on the gravitational physics to better than a few percent, roughly an order of magnitude better than present probes. Future cosmological experiments will be able to constrain the Brans-Dicke parameter at a level comparable to Solar System and astrophysical tests.

DOI: [10.1103/PhysRevD.95.063502](https://doi.org/10.1103/PhysRevD.95.063502)**I. INTRODUCTION**

The future of observational cosmology holds great promise. Not only will it be possible to refine our understanding of many of the ingredients of our current standard cosmological model (such as, for example, the geometry of space-time or the constituents of the mass density of Universe), but these observations will also allow us to explore truly foundational aspects in physics. In particular, it should be possible to test some of the most essential assumptions that go into constructing our working relativistic model of the Universe. One of the most fundamental assumptions is that gravity is perfectly described by Einstein's general theory of relativity.

General relativity (GR) has been exquisitely tested on noncosmological scales in the weak-field limit (see [1] for a recent review) and, more recently, has been shown to be broadly consistent with the first direct detection of gravitational waves from the merger of two black holes [2]. The common-sense approach has been to assume that we can safely extrapolate GR to cosmological scales and proceed to accurately calculate various properties of the Universe. This is, of course, an ambitious extrapolation, over 15 orders of magnitude in length scale and probing a completely new regime of gravitational potential and curvature [3]. Furthermore, there are a number of alternative theories that are

consistent with noncosmological tests but which can lead to a variety of cosmic histories [4]. Given all of this, it makes sense to construct cosmological tests of gravity which will be able to distinguish between GR and its alternatives.

The time is right to do so. A battery of large-scale surveys are planned to roll out over the next decade. More specifically we expect a new generation of photometric and spectroscopic galaxy redshift surveys, weak-lensing surveys, continuum surveys of radio galaxies, intensity mapping surveys of neutral hydrogen (HI), and a concerted campaign to map the temperature and polarization of the cosmic microwave background. These surveys will access a wide range of redshifts and scales with different sensitivities and will be affected by different systematic uncertainties. Cross-correlating these surveys will allow us to mitigate these effects and extract precise information on the morphology and evolution of large-scale structure. Indeed, as was shown in [5–8], combinations of various future data sets can be used to great effect to detect horizon-scale, general relativistic corrections to cosmological observables.

The next generation of surveys will deliver further tests of fundamental physics beyond the nature of gravity. For example, an important program in this direction is the determination of the mass scale of neutrinos, which is, arguably, easier to read off from their imprint on large-scale

structure and the cosmic microwave background (CMB) [9] than from laboratory experiments [10]. The minimal mass scale of neutrinos ($\sum m_{\nu,i} \gtrsim 0.06$ eV in the normal hierarchy) is set by oscillation experiments. Despite its smallness, it falls in the range of sensitivity for future experiments [11,12] and it is hence necessary to include it as a cosmological parameter. This is even more important in tests of gravity, whose signatures have been shown to be degenerate with neutrino masses in a range of models [13–15].

In this paper we look forward and forecast how this step changes in observational cosmology will affect our understanding of cosmological gravity. In broad terms, we will try to understand how our constraints on GR will improve over the next decade and how they will compare with constraints on other length scales. Ideally we would like to be completely agnostic about the theory of gravity. That is, we would like to make as few assumptions as possible about the space of theories within which we explore deviations. There has been great progress in characterizing gravity on linearly perturbed cosmological scales, from completely general approaches [16–24] to ones focused around scalar-tensor theories [25–29]. While we would like to be completely general, as a first step we will focus on a large, subclass of scalar-tensor theories described by the Horndeski action [30,31]. This will allow us to see how future constraints will improve our knowledge of what is currently a well-understood swathe of model space.

The Horndeski action encompasses all scalar-tensor theories (with one scalar field) that have second order equations of motion. It includes quintessence, K-essence, Jordan-Brans-Dicke [32] and its variants, and generalized covariant Galileons. Substantial effort has already gone into studying the cosmological dynamics of these theories [33] as well as attempts at constraining them [34–36] or forecasting future constraints from specific future surveys [37–39]. In this paper we will include key surveys that are planned to come online in the next decade or so. Considering this subclass of models will allow us to better understand the role that priors play in forecasting constraints and, in particular, how priors on the gravitational theories will interact with the priors on the usual cosmological parameters. It will also allow us to compare directly with constraints on smaller scales, i.e. with the fabled precision constraints from the Solar System, and binary systems [1].

This paper is structured as follows. In Sec. II we present Horndeski’s formulation of scalar-tensor theories and identify the class of theories we will be considering in our analysis. We take particular care in identifying three different levels of parametrizations which we will be forecasting for. In Sec. III we list the surveys we will be considering and the forecasting method we will be using; our focus is primarily on tomographic surveys

supplemented with intensity mapping experiments. We do briefly consider the added value of spectroscopic measurements of the baryon acoustic oscillation scale (BAO) and of the growth rate via redshift-space distortions. In Sec. IV we present our results, taking particular care to identify the optimal combination of surveys, identifying the degeneracies between parameters and contrasting the future with the current state of play. In Sec. V we discuss our findings.

II. HORNDESKI’S FORMULATION OF SCALAR-TENSOR THEORIES

A general scalar-tensor action, constructed from a metric, $g_{\mu\nu}$ and scalar field, ϕ was proposed in [30,31] and takes the following form:

$$S = \int d^4x \sqrt{-g} \left\{ \sum_{i=2}^5 \mathcal{L}_i[\phi, g_{\mu\nu}] + \mathcal{L}_M[g_{\mu\nu}, \varphi] \right\}, \quad (1)$$

where \mathcal{L}_M is the minimally coupled matter action (φ represents a general “matter” field) and the building blocks of the scalar field Lagrangian are

$$\begin{aligned} \mathcal{L}_2 &= K, \\ \mathcal{L}_3 &= -G_3 \square \phi, \\ \mathcal{L}_4 &= G_4 R + G_{4X} \{ (\square \phi)^2 - \nabla_\mu \nabla_\nu \phi \nabla^\mu \nabla^\nu \phi \}, \\ \mathcal{L}_5 &= G_5 G_{\mu\nu} \nabla^\mu \nabla^\nu \phi \\ &\quad - \frac{1}{6} G_{5X} \{ (\nabla \phi)^3 - 3 \nabla^\mu \nabla^\nu \phi \nabla_\mu \nabla_\nu \phi \square \phi \\ &\quad + 2 \nabla^\nu \nabla_\mu \phi \nabla_\nu^\nabla \phi \nabla^\mu \nabla_\alpha \phi \}. \end{aligned} \quad (2)$$

We have that K and G_A are functions of ϕ and $X \equiv -\nabla^\nu \phi \nabla_\nu \phi / 2$, and the subscripts X and ϕ denote derivatives. The four functions, K and G_A completely characterize this class of theories. It is possible to extend this action by including extra terms (to generate what has been dubbed the “beyond Horndeski” action [40–42]) “extended scalar-tensor theories” [43–45] or nonminimal coupling to matter [46].

From the point of view of cosmology, one can greatly reduce the number of degrees of freedom one needs to focus on to the homogeneous mode of the metric, $ds^2 = -dt^2 + a^2(t) d\vec{x}^2$, and scalar field, $\bar{\phi}(t)$, and their linear perturbations. In [47] it was shown that, on top of a convenient choice of the background expansion history (described by the effective equation of state, w , of the Horndeski field), one can completely characterize the evolution of linear perturbations in Horndeski cosmologies in terms of a set of free functions of time given by

$$\begin{aligned}
M_*^2 &\equiv 2(G_4 - 2XG_{4X} + XG_{5\phi} - \dot{\phi}HXG_{5X}), \\
HM_*^2\alpha_M &\equiv \frac{d}{dt}M_*^2, \\
H^2M_*^2\alpha_K &\equiv 2X(K_X + 2XK_{XX} - 2G_{3\phi} - 2XG_{3\phi X}) + 12\dot{\phi}XH(G_{3X} + XG_{3XX} - 3G_{4\phi X} - 2XG_{4\phi XX}) \\
&\quad + 12XH^2(G_{4X} + 8XG_{4XX} + 4X^2G_{4XXX}) - 12XH^2(G_{5X} + 5XG_{5\phi X} + 2X^2G_{5\phi XX}) \\
&\quad + 14\dot{\phi}H^3(3G_{5X} + 7XG_{5XX} + 2X^2G_{5XXX}), \\
HM_*^2\alpha_B &\equiv 2\dot{\phi}(XG_{3X} - G_{4\phi} - 2XG_{4\phi X}) + 8XH(G_{4X} + 2XG_{4XX} - G_{5\phi} - XG_{5\phi X}) + 2\dot{\phi}XH^2(3G_{5X} + 2XG_{5XX}), \\
M_*^2\alpha_T &\equiv 2X[2G_{4X} - 2G_{5\phi} - (\ddot{\phi} - \dot{\phi}H)G_{5X}].
\end{aligned} \tag{3}$$

Each of these functions can be tied to its own underlying physical aspect of scalar-tensor theories: M_*^2 and α_M are related to time variations in the background Newton's constant, α_K (dubbed “kineticity”) generalizes the canonical kinetic term of simple DE models while α_B (dubbed “braiding”) quantifies kinetic mixing between ϕ and the scalar perturbations of the metric. Finally, α_T is associated to modifications to the speed of propagation of tensor modes, but it is also responsible for anisotropic stress in the scalar sector, i.e. γ defined in Eq. (6).¹

Furthermore, one can infer from this parametrization, emergent scales which define how perturbations evolve. As originally discussed in [47] and further elaborated in [36], the “braiding scale,”

$$k_B = \frac{D}{\alpha_B^2} [(1 - \Omega_M)(1 + w) + 2(\alpha_M - \alpha_T)] + \frac{9}{2}\Omega_M, \tag{4}$$

(where the time-varying Planck mass has been absorbed into the definition of Ω_M and $D \equiv \alpha_K + \frac{3}{2}\alpha_B^2$) determines the scale dependence in the growth factor.

The set of α 's are a remarkably compact and efficient way of characterizing this large subclass of scalar-tensor theories. On the one hand, it greatly restricts the possible structure and evolution of non-GR cosmologies; this finite-dimensional subspace of theories can be easily parametrized by a handful of numbers characterizing their time evolution. One obvious, simplified, approach is to associate the onset of modifications to GR with the emergence of dark energy; indeed one of the motivations to consider such theories is a possible explanation for the late-time acceleration of the Universe. Thus a possible parametrization is

$$\alpha_X = f\left(\frac{\Omega_{\text{DE}}}{\Omega_{\text{DE}0}}\right), \tag{5}$$

where Ω_{DE} is the fractional energy density in dark energy (and the additional subscript “0” denotes its value today)

¹Both α_M and α_T affect the propagation of gravitational waves and can be constrained by the period decay in binary systems [48] and future detections from systems with electromagnetic or neutrino counterparts [49,50].

and $f(x)$ is a function such that $f \rightarrow 0$ as $x \rightarrow 0$. To describe gravity at low redshift we can Taylor expand this function and, for example, simply keep the leading order term, so that $f \simeq f'(0)x$. A recent analysis of this class of theories using this parametrization was undertaken in [36] where the uncertainty, from current data, on α_B and α_M was found to be $\mathcal{O}(1)$.

As stated, this is a simplified approach and does not necessarily capture the correct behavior of the α 's arising in Horndeski theory [51]. Simply put, if one were to start with the Horndeski action for the background evolution, choosing a physical range of initial conditions and free functions, one would then derive a set of theoretical priors for the time evolution of the α 's. One would then construct a parametrization of this prior space of α 's to be used in an accurate forecast analysis. This is analogous to what was shown for the case of the simple (w_0, w_a) parametrization of dark energy where a reasonably tight correlation between the two parameters exists already at the level of the priors [52]. Nevertheless, our simple parametrization does allow us to get a sense of how constraining future data sets will be in comparison with current data. In addition, this parametrization describes exactly the behavior on their attractor of simple shift-symmetric models as the imperfect fluid [53], and approximately more complicated models as the best fit of covariant Galileons without a cosmological constant [13,54].

Within this restricted class of models, we can also map directly onto a set of parameters which are of particular use in the quasistatic regime, when $c_s ak/H \gg 1$ (where k is the wave number of a given perturbation, a is the scale factor, H is the Hubble rate and c_s is the sound speed of the additional degree of freedom). The linear perturbation equations in the Newtonian gauge greatly simplify to

$$-k^2\Phi = 4\phi G\mu\bar{\rho}\delta, \quad \frac{\Phi}{\Psi} = \gamma, \tag{6}$$

where Φ and Ψ are the gravitational potentials, $\bar{\rho}$ is the background energy density of nonrelativistic matter with perturbations δ . It has been shown that the two quasistatic parameters—the modified Newton's constant, $G_{\text{eff}} = \mu G_0$

and the gravitational slip, γ —can be expressed as rational functions of k with time dependence uniquely determined by the α 's. While we will not actually forecast constraints on (μ, γ) in this paper, it is useful to assess what are current constraints. The most complete analysis of quasistatic parameters can be found in [35] where a set of somewhat paradoxical constraints were found. On the one hand, the current uncertainty in (μ, γ) is of $\mathcal{O}(1)$, although these parameters do suffer from a strong degeneracy; if we look at the direction orthogonal to that degeneracy, we can bring down the uncertainty to $\mathcal{O}(10^{-1})$. On the other hand, the use of these quasistatic parameters provided the first clear indication of a deviation from GR. While it is reasonably robust to different permutations of data sets, the authors of [35] have opted to claim that this is a signal of systematic effects in the data. Future data sets, such as the ones we consider should unambiguously resolve this issue.

So far we have explained how we will use the Horndeski parametrization to explore a broad subclass of theories. We can also do the converse and consider a very narrow subclass of well-known theories that are embedded in Horndeski. We will thus also focus on Jordan-Brans-Dicke (JBD) theories [32] which are described in terms of

$$S = \int d^4x \sqrt{-g} \left[\phi R - \frac{\omega_{\text{BD}}}{\phi} \nabla_\mu \phi \nabla^\mu \phi + V + \mathcal{L}_M[g_{\mu\nu}, \varphi] \right]. \quad (7)$$

When $V = 0$, these theories are completely described in terms of one parameter, ω_{BD} , and GR is recovered when $\omega_{\text{BD}} \rightarrow \infty$. This theory is the workhorse of modern tests of GR on a wide range of scales and allows us to compare the potential of future cosmological surveys with other astrophysical tests. This theory will trace out a particular slice in the space of α 's.

JBD theories are exquisitely constrained. The current tightest constraint on ω_{BD} is on astrophysical scales from the Shapiro time delay [55]: $\omega_{\text{BD}} > 4 \times 10^4$. The most up-to-date cosmological constraints using the Planck data are in [34] where it was found that $\omega_{\text{BD}} > 6 \times 10^2$. See also [56,57]). Naturally, the question arises whether it is possible to get cosmological constraints on ω_{BD} comparable to those found on astrophysical scales.

Many alternative theories of gravity are equipped with screening mechanisms, nonlinear effects by virtue of which the departures from GR fade in short scales or high-density environments [58]. This feature helps reconcile such theories with the stringent astrophysical and Solar System tests while leaving room for detectable signatures in the large-scale structure of the Universe. However, screening typically kicks in on small cosmological scales, suppressing the modified gravity effects relative to the linear prediction (see [59–61] and for an extensive cross comparison, see [62]). This feature of modified gravity has been studied in a model-by-model basis using nonlinear

techniques, but has been often ignored in forecasts for future experiments. We will use linear cosmological perturbation theory and hence the screening effects have to be included in a phenomenological fashion. We will model the small-scale recovery of GR through a scale dependence of the α s

$$\alpha_X(t) \rightarrow \alpha_X(t, k) = \alpha_X(t) S\left(\frac{k}{k_V}\right), \quad (8)$$

where $X = (M, B, K, T)$ and such that $S(x \ll 1) \approx 1$ and $S(x \gg 1) = 0$. We will model the screening using a Gaussian function $S = \exp(-\frac{1}{2}(k/k_V)^2)$ with a fiducial value $k_V = 0.1 h/\text{Mpc}$, in agreement with simulations of Vainshtein-screened models (cf. Fig. 4 of [60]). This is a minimal, one-parameter prescription. More sophisticated prescriptions can be added as needed based on model-specific studies. We note that this or other prescription needs to be introduced whenever screened scales are included in the analysis. If they can be properly modeled, nonlinear scales contribute greatly to constraint other cosmological parameters, but not accounting for screening can largely overestimate the surveys capacity to test gravity.

The situation with regards to baryonic physics is even more open: while there have been attempts at understanding the impact of, for example, active galactic nuclei feedback on specific models [63], a complete understanding is still lacking. In order to account for the effect of baryonic effects to some extent, we have used the model proposed by [64], determined by two quantities, M_c and η_b , parametrizing the mass dependence of the halo gas fraction and the ejected gas radius. For these we use the fiducial values $M_c = 1.2 \times 10^{14} M_\odot/h$ and $\eta_b = 0.5$. In Sec. IV C we will study the impact of the uncertainties on both screening mechanisms and baryonic physics by marginalizing over the values of these parameters as well as k_V .

III. STAGE IV COSMOLOGICAL SURVEYS

In the next decade, a number of astronomical facilities will cover a large portion of the sky visible from the Southern Hemisphere in multiple wavelengths and with almost perfect angular overlap. We will thus have at our disposal a large volume of the Universe where cosmological structures will have been covered by a large variety of observational probes. In this section we describe the forecasting method that we use to predict the combined constraining power of a representative subset of these experiments, the nature of the main cosmological probes pursued by them, and the main advantages of combining probes.

A. Forecasting formalism

Each probe a considered in this paper can be represented by a set of N_{maps}^a sky maps fully described by their spherical

harmonic coefficients $a_{\ell m}^{(a,i)}$ ($i \in [1, N_{\text{maps}}^a]$). These may correspond, for instance, to the perturbations in the number density of galaxies in a set of redshift bins or to the temperature and polarization perturbations of the CMB. Each sky map is labeled by a combined index $(a, i) \rightarrow A$, and to first approximation we will assume that the anisotropies $a_{\ell m}^A$ are Gaussianly distributed, such that the combined likelihood of the observed maps is given by

$$-2 \log p(\mathbf{a}_{\ell m}) = \sum_{\ell} \sum_{m=-\ell}^{\ell} \{ \mathbf{a}_{\ell m}^T \mathbf{C}_{\ell} \mathbf{a}_{\ell m} + \log[\det(2\pi\mathbf{C}_{\ell})] \}. \quad (9)$$

Here we have grouped all the sky maps into a vector $\mathbf{a}_{\ell m}$, and we have defined the power spectrum matrix \mathbf{C}_{ℓ} as the covariance of this vector:

$$\langle \mathbf{a}_{\ell m} \mathbf{a}_{\ell' m'}^{\dagger} \rangle = \delta_{\ell\ell'} \delta_{mm'} \mathbf{C}_{\ell}. \quad (10)$$

The information on cosmological parameters is encoded in the power spectrum. Expanding this likelihood around the maximum we can find the now usual expression for the Fisher matrix \mathbf{F} , describing the inverse of the covariance matrix of a set of parameters θ :

$$\mathbf{F}_{\mu\nu} = \sum_{\ell=2}^{\ell_{\text{max}}} f_{\text{sky}}(\ell + 1/2) \text{Tr}(\partial_{\mu} \mathbf{C}_{\ell} \mathbf{C}_{\ell}^{-1} \partial_{\nu} \mathbf{C}_{\ell} \mathbf{C}_{\ell}^{-1}), \quad (11)$$

where f_{sky} is the sky fraction covered by the considered probes, and ∂_{μ} implies differentiation with respect to the parameter θ_{μ} . The maximum multipole ℓ_{max} included in the constraints is a map-dependent choice, corresponding to the smallest scale for which sensible information can be extracted, and can be limited by noise and observational or theoretical systematics.

We will assume that the observed anisotropies contain both a cosmological signal and a noise component, $\mathbf{a}_{\ell m} = \mathbf{s}_{\ell m} + \mathbf{n}_{\ell m}$, and that both components are uncorrelated. Thus the power spectrum can also be decomposed into two components, $\mathbf{C}_{\ell} = \mathbf{C}_{\ell}^S + \mathbf{C}_{\ell}^N = \langle \mathbf{s}_{\ell m} \mathbf{s}_{\ell m}^{\dagger} \rangle + \langle \mathbf{n}_{\ell m} \mathbf{n}_{\ell m}^{\dagger} \rangle$. The form of the signal and noise components for each probe will be described below.

The partial derivatives in Eq. (11) were computed using finite central differences:

$$\partial_{\mu} f(\theta) = \frac{f(\theta_{\mu} + \delta\theta_{\mu}) - f(\theta_{\mu} - \delta\theta_{\mu})}{2\delta\theta_{\mu}}, \quad (12)$$

where the optimal value for the intervals $\delta\theta_{\mu}$ was found by iteratively halving the initial guesses until convergent results were found beyond the second significant digit. All power spectra were computed using a modified version

of `hi_class`² [65], a code based on the Cosmic Linear Anisotropy Solving System (CLASS) [66]. `hi_class` computes the evolution of linear cosmological observables without assuming the quasistatic approximation and includes relativistic corrections to galaxy clustering [67]. This ensures the validity of the computation on scales larger than the scalar field sound horizon [68] and all the way to ultralarge scales. In order to retain the correct large-scale shape, all power spectra were computed without adopting the Limber approximation. `hi_class` was run adding a constant value to the kineticity α_K , i.e. 10^{-4} , to protect the computation against numerical singularities in the evolution of the perturbations that happen when the kinetic term $D \rightarrow 0$. In addition, `hi_class` uses a set of precision parameters which is improved with respect to the ones in CLASS [66].

B. Cosmological surveys

1. Stage-4 CMB

The current CMB data sets, consisting of a combination of large-scale, full-sky experiments such as WMAP [69] and Planck [70], and high-resolution ground-based observatories (e.g. ACTPol [71], SPT-Pol [72]), are currently being upgraded through Stage 3 (S3) ground-based experiments (e.g. AdvACT [73] and SPT-3G [74]) with a larger numbers of detectors, multiple frequency channels, and covering larger sky fractions.

S3 experiments will eventually be superseded by a Stage 4 (S4) experiment, likely to be built by combining the observing power of different ground-based facilities, with similar potential for wide sky coverage and significantly lower noise levels. It is expected that S4 will cover $\sim 40\%$ of the sky, with a reduced noise level of around $\sigma_T = 1 \mu\text{K-arcmin}$ in temperature [75].

S4 will measure three main types of anisotropies: the primordial CMB perturbations in temperature and polarization ($a_{\ell m}^T$, $a_{\ell m}^E$ ³), as well as the reconstructed CMB lensing convergence $a_{\ell m}^K$. For the CMB anisotropies, the noise power spectrum is determined by two regimes. On small scales, the measurements are limited by the beam size of S4, which here we assume to be Gaussian with a width $\theta_{\text{FWHM}} = 3'$. In this regime the noise can be approximated as being white (before noise deconvolution [76]), and thus given by

$$N_{\ell}^{T,E} = \sigma_{T,E}^2 \exp[\ell(\ell + 1)\theta_{\text{FWHM}}^2/(8 \ln 2)], \quad (13)$$

with $\sigma_{T,E}^2$ in units of $\mu\text{K}^2 \text{ sr}$ (we assume $\sigma_E = \sigma_T \sqrt{2}$). On the other hand, ground-based facilities such as S4 are

²www.hiclass-code.net

³Although both the E - and B -modes contribute to the total polarized anisotropies, here we will set the primordial tensor perturbations (and hence the B -modes) to zero. The contribution of the lensing B -modes is therefore fully accounted for by considering the reconstructed CMB lensing convergence.

limited on large scales by a number of statistical and systematic uncertainties, such as the effect of atmospheric noise or ground pickup. For this reason we assume that S4 will not be able to probe the CMB anisotropies on scales $\ell < 30$, and in those multipoles we assume noise levels corresponding to Planck as given in [70]. Furthermore, we set the maximum multipoles for these probes at $\ell_{\max} = 3000$ in temperature and $\ell_{\max} = 5000$ in polarization.⁴

In order to compute the noise power spectrum for the lensing convergence, we assume a reconstruction process based on quadratic combinations of the lensed CMB maps [77], and estimate the reconstruction noise as detailed in [78]. In doing this we assume a minimum-variance noise achieved by combining the TT , TE , TB , EE and EB estimators, using only the multipole range $30 < \ell < 3000$.

It is worth noting that our formalism accounts for the nonzero correlation between overlapping CMB and large-scale structure measurements, both those caused by CMB lensing and by the late-time integrated Sachs-Wolfe effect.

2. The Large Synoptic Survey Telescope

Photometric redshift (photo- z) surveys have been proposed as a practical means to achieve simultaneously dense and deep galaxy catalogs. In this approach, the redshift of each galaxy is inferred from its flux in a small number of wide frequency bands, and although the corresponding redshift uncertainties prevent any efficient measurement of the radial clustering pattern, the large number density of tracers attainable by these surveys makes them ideal for weak-lensing studies as well as multitracer analyses.

Although large-scale photometric surveys, such as the Dark Energy Survey [79], are already underway, we will focus our discussion on the Large Synoptic Survey Telescope (LSST), a Stage-IV experiment aiming at covering $\sim 20,000$ deg² on the sky to with a magnitude limit $r \sim 27$. We will concentrate here on two of the main cosmological science cases covered by the LSST: galaxy clustering and cosmic shear.

Galaxy clustering.—The main cosmological observable in galaxy clustering studies is the shape of the power spectrum or two-point correlation function of the galaxy number density. In the case of LSST we will focus on a tomographic approach in which the galaxy sample is first separated into a number of bins of photo- z , and all auto- and cross-correlations between these bins are analyzed simultaneously.

Since the galaxy number density is known to be a biased tracer of the matter density field ($\delta_g \approx b\delta$), the uncertainties in the redshift- and possibly scale-dependent bias b prevent us from using the amplitude of perturbations in the galaxy overdensity to make an independent measurement of the

growth of structure. As a result, most of the information in the angular galaxy-galaxy power spectrum resides in robust features such as the angular BAO scale. Other “secondary” contributions to the power spectrum, such as that of redshift-space distortions can also be used to make a suboptimal measurement of the growth rate and, although subdominant, a number of relativistic contributions to the clustering pattern, such as that of lensing magnification [80–82] could potentially contain relevant information about deviations from GR. Further details can be found in [6]. These expressions are general for minimally coupled theories of gravity in the linear regime [54].

Under the assumption that galaxies are a Poisson sample of the underlying biased density field, we can model the noise power spectrum as a white component with an amplitude of $1/N_\Omega$, where N_Ω is the angular number density of sources in each redshift bin in units of sr⁻². As was done in [7], we separate the clustering sample into two subsamples of “red” (early-type, ellipticals, high-bias) and “blue” (late-type, disks, low-bias) galaxies. The specific models for the power spectrum, photo- z uncertainties, nuisance bias parameters and sample definitions are described in detail in [6].

Although a linear, scale-independent bias parameter is in most cases a good approximation to the relation between the galaxy and matter power spectra on large scales, on small nonlinear scales it becomes necessary to resort to more sophisticated models, which often prevents the use of the small-scale galaxy power spectrum for cosmology. In order to avoid these complications we define a minimum angular scale for each redshift bin down to which the corresponding map is used. At the mean redshift of each bin \bar{z} we start by defining a threshold comoving scale k_{\max} by requiring that the variance of the matter overdensity on larger scales be below a given threshold σ_{thr}^2 , i.e.,

$$\sigma_{\text{thr}}^2 = \frac{1}{2\pi^2} \int_0^{k_{\max}} dk k^2 P(k, \bar{z}). \quad (14)$$

We then translate this comoving scale into an angular multipole $\ell_{\max} = \chi(\bar{z})k_{\max}$. For our fiducial constraints we used a threshold variance of $\sigma_{\text{thr}} = 0.5$. This value corresponds to a cutoff scale $k_{\max} \sim 0.1h$ Mpc⁻¹ at $z = 0$, which is a conservative estimate of the scales up to which a reliable estimate of the covariance matrix of the matter power spectrum can be obtained using perturbation theory [83].

Cosmic shear.—The observed projected shapes of galaxies are distorted due to the weak gravitational lensing of the photons they emit caused by the intervening matter perturbations. Thus it is possible to probe those perturbations by analysing the correlated galaxy ellipticities. The constraining power of weak lensing can be summarized into the power spectrum of the traceless part of the cosmic

⁴The different small-scale cutoff in temperature and polarization is motivated by the contamination from extragalactic foregrounds on small scales in the temperature power spectrum.

shear tensor computed for galaxies lying in a set of photo- z bins (labeled here by an index i):

$$\begin{aligned}\gamma_i(\hat{\mathbf{n}}) &\equiv \int_0^{\chi_H} d\chi' W_\gamma^i(\chi) \delta\delta(\Phi + \Psi)(\chi, \hat{\mathbf{n}}), \\ W_\gamma^i(\chi) &\equiv \int_\chi^{\chi_H} d\chi' \frac{dp_i}{d\chi'} \left(\frac{\chi' - \chi}{\chi\chi'} \right),\end{aligned}\quad (15)$$

where $dp_i/d\chi$ is the selection function of the redshift bin, and δ is the spin-raising differential operator defined by their action on a spin- s function f_s :

$$\delta f_s(\hat{\mathbf{n}}) = -(\sin\theta)^s (\partial_\theta + i\partial_\phi / \sin\theta) (\sin\theta)^{-s} f_s(\hat{\mathbf{n}}). \quad (16)$$

We see that, at any particular redshift, weak-lensing traces the density perturbations integrated along the line of sight down to the observer weighed by the lensing kernel. Thus, good redshift precision is not necessary for this probe, since the large support of the lensing kernel erases all structure along the line of sight, making photometric redshift surveys ideal for this task.

Since the shear tensor is effectively estimated by averaging over the ellipticities of all galaxies laying in a given pixel, the noise in this estimate is directly proportional to the variance of the intrinsic galaxy ellipticities, and inversely proportional to the angular galaxy number density. We thus use a white noise model in which the noise power spectrum for the i th redshift bin is given by $N_\ell^i = \sigma_\gamma^2 / N_\Omega^i$, where σ_γ^2 is the per-component dispersion in the intrinsic galaxy ellipticities (including measurement noise), for which we use $\sigma_\gamma = 0.3$ [84].

Since weak lensing directly probes the matter perturbations, theoretical uncertainties on nonlinear scales are in principle far less cumbersome than in the case of galaxy clustering. However, even using numerical simulations and emulators [85], the uncertainties in the modeling of baryonic effects prevent an accurate description of the matter power spectrum on very small scales [86–89,90]. For these reasons, for weak lensing we use a small-scale cutoff $\ell_{\max} = 2000$, corresponding to $k_{\max} \sim 1$ at $z = 1$ (note that we will use a different prescription in Sec. IV C, with ℓ_{\max} defined as above for galaxy clustering).

Cosmic shear is not free from systematic uncertainties: shape-measurement uncertainties are known to affect the broadband shape of the lensing power spectrum, and multiplicative and additive bias parameters have been used to model the effects in the first cosmological lensing analyses [91–93]. Furthermore, intrinsic galaxy orientations are known to be correlated with each other, and there are hints that the local tidal field (i.e. the Hessian of the gravitational potential) is responsible for these intrinsic alignments, at least in the case of elliptical galaxies. This effect can also be modeled at the cost of introducing extra nuisance parameters. This work aims at forecasting the best achievable constraints on scalar-tensor theories, and

therefore we will ignore shape-measurement systematics, assuming that tight priors on the corresponding nuisance parameters can be found below the science requirements of LSST. We have also neglected the systematic uncertainties associated with the use of photometric redshifts, which would also affect galaxy clustering. We have, however, included the effect of intrinsic alignments following the so-called nonlinear alignment model [94], marginalizing over the alignment amplitude at $z \in \{0.5, 1.0, 1.5, 2.0\}$. We defer a more thorough study of the effect of systematic uncertainties on modified gravity constraints for future work.

3. The Square Kilometre Array

A wide-area radio astronomy facility such as the Square Kilometre Array (SKA) offers a large variety of unique as well as synergistic cosmological science cases, which have recently received a lot of attention in the literature [95–98]. Three types of cosmological surveys can be carried out in this range of frequencies:

- (i) In an *HI galaxy survey* [99], individual sources are spatially resolved, and their 21 cm, neutral hydrogen detected with significantly high signal-to-noise, thus producing a spectroscopic catalog of galaxies. The low intensity of this line makes it very challenging to reach high number densities and redshifts with this technique, and it is expected that only a hypothetical ‘‘Phase 2’’ of the SKA would be able to produce a HI catalog competitive with planned Stage-IV surveys such as Euclid [100].
- (ii) Dropping the requirement of resolving the HI line, a *continuum survey* [101] simply integrates the total radio flux of all sources in a wide frequency band. This allows the detection of much fainter sources, and a continuum survey with a flux limit of $5\mu\text{Jy}$ would be able to cover a wide range of redshifts ($z \lesssim 4$). The main drawback of this technique is the lack of redshift estimates for the detected sources, which in the best-case scenario makes them dependent on external data sets and optical follow-up, and in the worst renders them useless for cosmological studies. It is also worth noting that, given the good control over the point-spread function achievable with a radio interferometer, radio weak-lensing surveys from continuum catalogs could potentially be more robust against shape-measurement systematics than their optical counterparts [102].
- (iii) Conversely, the novel technique of *intensity mapping* [96] sacrifices angular resolution, avoiding the detection of individual sources by integrating the combined 21 cm emission in wider angular scales, thus producing three-dimensional maps of the distribution of neutral hydrogen in the Universe with good radial resolution. Intensity mapping is therefore complementary to the use of photometric redshift surveys in the coverage of the k_\perp - k_\parallel plane.

Given its complementarity with the science cases covered by LSST, we will focus here on intensity mapping only. The cosmological observable in this case is the HI antenna temperature measured in a set of frequency bins, related to the corresponding redshift as $\nu = \nu_{21 \text{ cm}}/(1+z)$. We will assume that the observations will be done using the SKA1-MID as described in [6,103] we will dub this data set SKA-IM. We consider the frequency band 350–1050 MHz, dividing it into 200 frequency channels, corresponding to a comoving width of $\sim 16 \text{ Mpc}/h$. The models used to describe the signal and noise power spectrum of this observable, including the models for the HI bias and background temperature, are presented in detail in [6]. We assume a 10,000-hour survey covering 40% of the sky carried out with a set of 200 15 m antennas in single-dish observation mode with a system temperature of 25 K. The noise power spectrum is modeled as white (before beam deconvolution):

$$N_\ell(\nu) = \sigma_N^2(\nu) \exp[\ell(\ell+1)\theta_{\text{FWHM}}^2(\nu)/(8 \ln 2)], \quad (17)$$

where the noise variance σ_N and beam width θ_{FWHM} are determined by the parameters listed above as described in [6]. The beam size of the SKA ($\theta_{\text{FWHM}} \sim 2^\circ$ at $z \sim 1$) is large enough that we do not need to impose a strict high- ℓ cut, since the measurements become noise dominated well before the scale of nonlinearities (in practice we impose a cut $\ell_{\text{max}} = 200$).

Since individual sources are not detected, the faint 21 cm emission needs to be separated from the much brighter (~ 5 orders of magnitude) diffuse galactic and extragalactic foregrounds. Although, given the smooth frequency dependence of the foregrounds, it should be possible to isolate the cosmological signal based on its different spectral properties (see [104,105]), foreground residuals will necessarily dominate the measurements on large radial scales, and it is expected that foreground contamination coupled with instrumental miscalibration will be the largest source of systematic uncertainties. Another cause of concern specific to single-dish observations is the effect of gain fluctuations in the time domain, which could be an important source of systematic uncertainties on large angular scales, although the effect may depend on the survey scanning strategy. As before, we will, for the most part, ignore these systematics in this work, in an attempt to present the best achievable constraints on scalar-tensor theories from future experiments.

4. Spectroscopic surveys

In our analysis we have not included constraints from wide spectroscopic such as DESI [106] or Euclid [100]. For this type of experiment, with good angular and radial resolution, the most efficient way to carry out Fisher forecasts is to use the Fourier coefficients of the galaxy overdensity $\delta_g(z, \mathbf{k})$ as an observable in a discrete set of

redshift bins within which evolution effects are effectively frozen. In this formalism it is however not straightforward to account for inter-bin correlations [107] and correlations with overlapping lensing and CMB experiments, a key aspect of our analysis. On the other hand, the option of modeling spectroscopic observations as a set of angular maps at different redshift becomes computationally intractable without losing radial information. Furthermore, given that the experiments listed above cover the main science cases that a spectroscopic survey would be able to approach (geometric radial and angular BAO measurements and redshift space distortions), we do not expect dramatic improvements in the final uncertainties due to the inclusion of spectroscopic data.

Nevertheless, and in order to estimate both the reach of future spectroscopic observations and the amount of information lost in our formalism, we have included constraints from an independent (i.e. uncorrelated) DESI-like spectroscopic survey using the expected uncertainties on the radial and angular BAO scales and the growth rate of structure estimated by [11]. Besides simplicity, lack of correlation is an optimistic assumption that lets us evaluate how much the constraints may improve by the addition of this type of survey. We do not include these additional constraints as part of our fiducial forecasts, but discuss their relevance in Sec. IV A.

IV. RESULTS

The class of theories we are considering can, as we have seen, be parametrized in terms of 5 free function of time, w , α_M , α_K , α_B and α_T (we fix the initial Planck mass, i.e. the integration constant needed to obtain M_*^2 from α_M , to 1). For the bulk of this analysis we assume that deviations from GR are intimately tied to the onset of accelerated expansion and hence, for now, we Taylor expand expression 5 and adopt a parametrization of the form:

$$\alpha_X = b_X + c_X \frac{\Omega_{\text{DE}}(z)}{\Omega_{\text{DE}}(z=0)}, \quad (18)$$

where Ω_{DE} is the fractional energy density in dark energy (or whatever is responsible for the onset of accelerated expansion) which itself depends on w . For our fiducial constraints we fix the early terms to zero ($b_X = 0$), and concentrate only on c_X . In a latter subsection we will consider b_X , as well as a different time dependence.

As a fiducial model we choose a point in the space of c_X that is close enough to Λ GR to be compatible within 1σ given our most optimistic constraints (this is done to avoid the numerical singularities at $c_X = 0$). The fiducial model we chose is $\{w = -1, c_K = 0.1, c_B = 0.05, c_M = -0.05, c_T = -0.05\}$. Beside these, we vary over the basic parameters of the flat Lambda cold dark model (Λ CDM) model: the dark matter and baryon densities ω_c and ω_b , the local expansion rate h , the amplitude and tilt of primordial scalar

TABLE I. 1σ constraints on the Horndeski parameters c_B and c_M , the dark-energy equation of state parameter w and the sum of neutrino masses $\sum m_\nu$ for different combinations of experiments. The last column shows the constraints assuming a modified gravity fiducial model given by the best fit in [36], in which case Λ CDM would be ruled out by more than 7σ from c_M alone. The corresponding values for the figure-of-merit defined in Sec. V are shown in the last column.

Case	$> \omega_{\text{BD}}$, 95% C.L.	$\sigma(c_B)$	$\sigma(c_M)$	$\sigma(c_T)$	$\sigma(c_K)$	$\sigma(w)$	$\sigma(\sum m_\nu)$ [meV]	FoM (c_B, c_M, c_T)
S4	2.9×10^3	0.796	0.746	1.26	4.9	0.112	71	1.3
LSST	1.2×10^4	0.193	0.089	0.205	8.8	0.016	45	61
SKA1-IM	9.5×10^3	13.3	6.0	8.6	106	0.018	74	1.0
S4 + LSST	1.3×10^4	0.169	0.072	0.179	3.5	0.011	22	88
S4 + SKA1-IM	1.0×10^4	0.305	0.238	0.786	3.5	0.0085	23	9.0
S4 + LSST + SKA1-IM	1.7×10^4	0.161	0.070	0.151	3.1	0.0069	15	121
S4 + LSST + SKA1-IM + Spec.	1.7×10^4	0.123	0.056	0.146	3.1	0.0061	13	143
Best fit of [36] as fiducial	not applicable	0.063	0.076	0.201	4.23	0.0059	13	...

fluctuations (A_s, n_s) and the optical depth to reionization τ . For these, we set their fiducial values to the best-fit cosmology of [9] (with $\tau = 0.06$ as per the latest measurement of [108]). Furthermore, we consider a single massive neutrino with a mass of 60 meV. When considering extended models with free early-time parameters, we set their fiducial values to $b_X = 0$.

A. Overall constraints

We begin by considering the combination of our three main data sets (S4, LSST and SKA1-IM) to identify the space of parameters on which the tightest constraints can be drawn. Table I summarizes the forecast constraints on the most relevant parameters, and Fig. 1 shows the covariance between them. The analysis of [36], gives us an idea of the overall structure of the constraints for fixed equation of state, $w = -1$, with c_M and c_B being more tightly constrained than c_K and c_T (note that in [36] the constraints on c_T are heavily affected by theoretical priors, such as stability conditions). Our results show that, while the future generation of surveys we consider will also be able to pin down c_T to a similar degree of precision, c_K will remain a highly uncertain parameter. Fortunately, c_K shows little or no degeneracy with any of the other Horndeski parameters, and therefore it can be marginalized over without degrading the constraints on $c_{B,M,T}$. This is explicitly shown in Table II, which summarizes the degeneracies on these parameters for fixed or marginalized c_K . We see, however, that the uncertainties on $c_{B,M,T}$ grow between 10% and 30% when considering an evolving dark energy component with $w \neq -1$.

We can explore the full set of relevant degeneracies in Fig. 1, which shows that c_B and c_M are tightly correlated along the direction $c_B \approx 2.5c_M$. The dark energy equation of state also shows significant degeneracy with all the Horndeski parameters, especially c_B and c_M , and our numerical results show also a non-negligible correlation with the fractional dark matter density ω_c . These

degeneracies are primarily driven by constraints on the growth of perturbations via the weak-lensing measurements in LSST (and less significantly through the redshift space distortion measurement from intensity mapping with SKA1). We can understand this from a brief analysis of the evolution equation for the growth rate, $f = d \ln \delta / d \ln a$:

$$\frac{df}{d \ln a} + qf + f^2 = \frac{3}{2} \Omega_M \frac{\mu}{\gamma}, \quad (19)$$

$$\text{where } q(a) = \frac{1}{2} [1 - 3w(a)(1 - \Omega_M(a))], \quad (20)$$

δ is the matter density contrast and Ω_M is the fractional matter density. Note that Ω_M and μ/γ are time dependent and, as μ/γ is a function of the α 's, we expect there to be a degenerate effect between these different parameters. The equation of state w will affect Ω_M as well as q leading to a further degeneracy. This also shows how crucial is to have precise distance measurements (specifically via the BAO) to be able to constrain this degeneracy.

While there is no significant degeneracy with the sum of neutrino masses [(through their effect on Ω_M , modifying Eq. (20)], it is worth noting that, as has been found in other works, we predict a $\sim 3\sigma$ measurement of the total neutrino mass (assuming the normal hierarchy lower bound) by combining S4 and low-redshift information from LSST [$\sigma(\sum m_\nu) = 22$ meV].⁵ The lack of degeneracy is in contrast with previous studies based on specific models such as $f(R)$ [14,15]. In this case the effect of neutrino masses (washing out perturbations on scales smaller than free-streaming scale) and modified gravity (enhanced growth due to the scalar force on scales shorter than the inverse mass of the field) can cancel partially. This cancellation is easier to achieve in restricted models [such

⁵This sensitivity could be increased to a $\sim 5\sigma$ detection if an external, cosmic-variance-limited measurement of τ could be carried out [$\sigma(\sum m_\nu) = 11$ meV] [12].

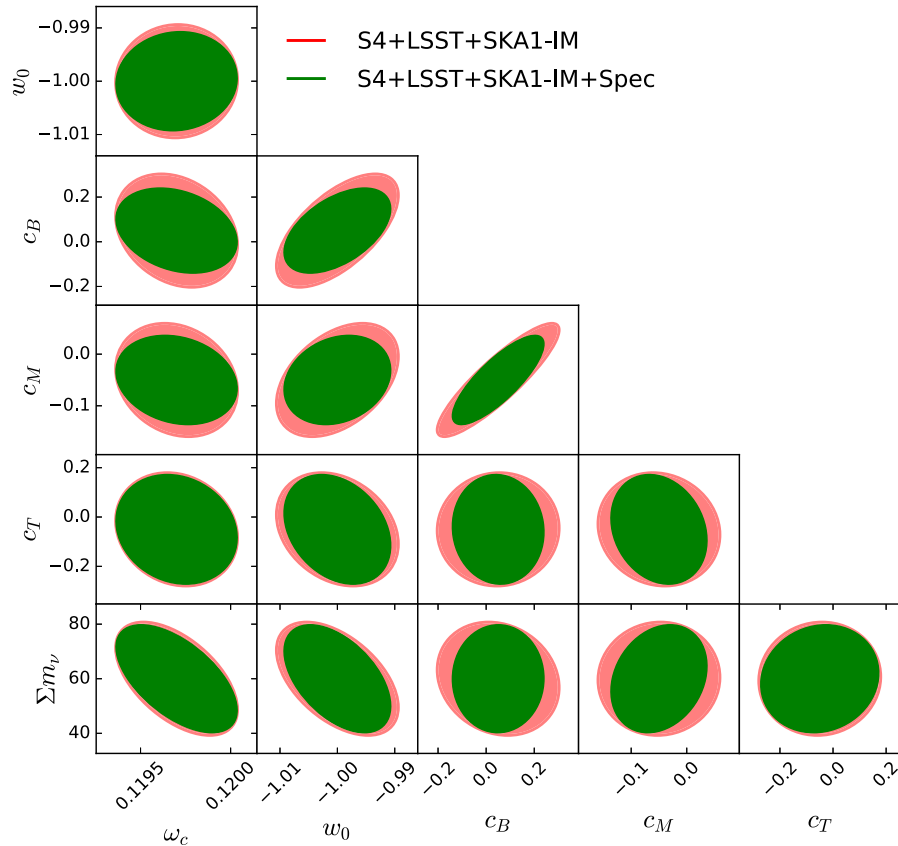


FIG. 1. Cosmological constraints for the combination of CMB S4 with LSST (galaxy clustering and shear) and an intensity mapping experiment carried out by Phase 1 of the SKA (red ellipses). The green ellipses show the additional constraining power achievable by combining these observations with measurements of the BAO scale and the growth rate of structure carried out by an independent DESI-like experiment.

as $f(R)$ due to the interdependences existing in the α functions (cf. [47]). In more general models such as the one considered here [Eq. (5)] most of the parameter space does not allow for this cancellation.

TABLE II. 1σ constraints for the Horndeski parameters for a combination of CMB-S4, LSST and SKA intensity mapping. Results are shown for c_B , c_M and c_T , with c_K and w fixed or marginalized over. Results for c_K are likewise shown for fixed and marginalized w .

Parameter	Fixed Parameters	68% Uncertainty
c_B	c_K, w	0.128
c_B	c_K	0.161
c_B	none	0.167
c_M	c_K, w	0.065
c_M	c_K	0.070
c_M	none	0.072
c_T	c_K, w	0.146
c_T	c_K	0.151
c_T	none	0.151
c_K	w	3.11
c_K	none	3.13

Figure 1 also shows, in green, the constraints achievable after adding independent BAO and growth rate information from a DESI-like spectroscopic survey. As discussed in Sec. III B 4, even though we observe a non-negligible improvement in the final uncertainties, the inclusion of spectroscopic information does not provide enough additional information to qualitatively change our results.

Finally, by comparing our forecast constraints with the analysis of [36] with current data, we find that it will be possible to constrain c_B and c_M a factor of ~ 5 better with next-generation surveys and, as reported above, c_T will be measured with similar accuracy. It is worth noting that the specific constraints on c_X depend mildly on the fiducial cosmology used in the forecast. The forecast uncertainties assuming the best-fit point of [36] are shown in the bottom row of Table I, and would correspond to a detection of deviations from Λ GR with a significance above 7σ .

B. Ultralarge scales and relativistic effects

It is also important to understand the dependence of the forecast constraints on the scales probed by each experiment. The interest is twofold: firstly, a number of relativistic effects are known to leave an imprint on

large-scale-structure observables on scales of the order of the horizon at the source redshift [7,54,109], and their purely relativistic nature suggests that they may be relevant in constraining deviations from GR. Secondly, a large fraction of the sources of systematic uncertainty that Stage-IV surveys will be sensitive to, such as the problem of CMB and radio foregrounds, or star contamination in galaxy surveys, affect the measurement of density fluctuations on large angular and radial scales. This analysis therefore helps us quantify the amount of information lost by the anticipated loss of sensitivity on large scales.

Figure 2 shows the 1σ uncertainties on the Horndeski parameters as a function of the minimum multipole (i.e. largest angular scale) included in the analysis. We see that, while most of the constraining power on c_K come from multipoles $\ell \lesssim 10$, it is the larger statistical power borne by small-scale perturbations (given the larger number of available small-scale modes) that drives the constraints on c_B , c_M and c_T . As shown in the Appendix, this is understandable, given that the leading order on subhorizon scales, i.e. $ak/H \gg 1$, of the quasistatic parameters does not depend on α_K .

It is possible to directly quantify the actual constraining power of the relativistic effects in large-scale structure [80–82]. In order to do so, we recomputed our forecasts after canceling the contribution from these effects to the total fluctuations on number counts; specifically we discard contributions from the integrated Sachs-Wolfe, the Doppler correction to the Kaiser effect, the gravitational time delay,

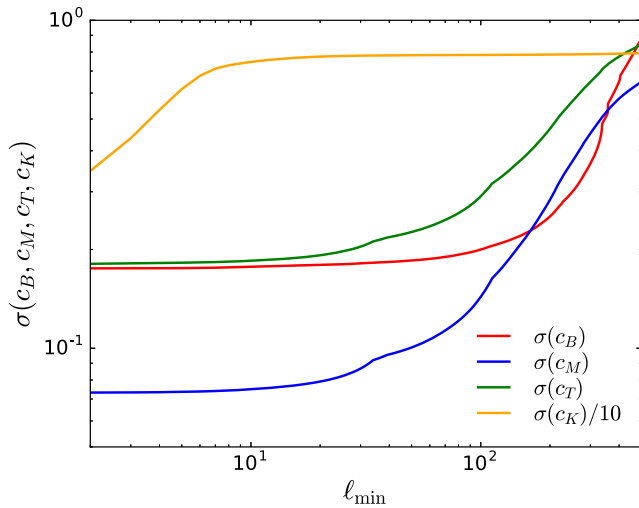


FIG. 2. Constraints on the Horndeski parameter as a function of the minimum multipole ℓ_{\min} included in the analysis. Except for c_K , most of the information on the Horndeski parameters is encoded in the smaller scales, due to their higher statistical weight. Furthermore, we do not find a significant improvement in the constraints by including relativistic effects in galaxy clustering, given their relatively small contribution to the galaxy power spectrum and their redundancy after including weak-lensing observations.

the effect of local gravitational potentials and the lensing magnification effect. Comparing the resulting uncertainties with our nominal results we only observe a negligible improvement in the final constraints of up to $\sim 1\%$, even in the case of c_K . Although this result may come as a surprise, given the direct relation of these effects with GR, it is actually understandable, given the small relative contribution of these effects to the overall fluctuations (e.g. see [109]), which is comparable to the effect of primordial non-Gaussianity on the large-scale clustering of a biased tracer with bias $b \sim 2$ and $f_{\text{NL}} \sim 1$.

Unlike other relativistic effects, the impact of lensing magnification correction on galaxy number counts persists on small scales. In Horndeski models it leads to distinct signatures on the cross-correlation between galaxies with large redshift separations [54]. Nonetheless, other probes of the gravitational potential such as shear measurements or CMB lensing probe the lensing potential with much higher signal-to-noise ratio. Hence, including or removing the lensing effect does not improve the sensitivity when clustering measurements are combined with lensing. For an LSST clustering-only forecast the constraints on the Horndeski parameters benefit from the inclusion of lensing magnification, with the 1-sigma bounds on c_B , c_M , c_T improving respectively by a factor 7, 2 and 3, but still considerably worse than the full LSST results including shear measurements.

C. Small scales

Smaller scales have smaller cosmic variance and hence much larger statistical weight than large scales. However it is on small scales that nonlinear gravity and baryonic effects will play a significant role. For example, at $z = 0$, nonlinear corrections to the power spectrum are of order 1% at $k \sim 0.01$ [110]. There is concerted effort under way to accurately model nonlinear corrections to sufficient precision so that one will be able to use small-scale modes in future analysis (attempts at doing this with current data can be found in [111] and a variety of approaches can be found in [112–114]). There have also been attempts at constructing phenomenological models which would include baryonic effects, including feedback [90,115].

While there is some hope that, within the context of standard models (with a cosmological constant and evolving under GR), it should be possible to harness some nonlinear modes, the case for models with modified gravity is less promising. A suite of N-body codes exists for particular subclasses of theories beyond general relativity, and there has been an attempt at cross-calibrating them to better than 1% on a wide range of scales [62]. Yet, the range of models which have been compared is still relatively restricted and is not sufficiently general that we have a firm understanding of the nonlinear regime.

A further complication of these models is the accurate modeling of nonlinear screening mechanisms. Lacking an

analytical treatment, we have opted modeled these effects as an scale dependent modulation of the “ α ” functions, as described in Eq. (8), to effectively suppress any deviation from GR in the nonlinear regime while retaining the information gain of nonlinear scales on the remaining cosmological parameters. Note that including nonlinear scales without accounting for screening could severely overestimate the constraining power of future surveys. Note that the inclusion of a finite cutoff k_V eliminates the effects of modified gravity on smaller scales, as expected due to screening mechanisms, but also introduces a feature in the matter power spectrum that can lead to a stronger measurement of the gravitational parameters. Although the signal-to-noise ratio is higher, this procedure actually slightly improves the constraints on the Horndeski parameters. The reason is that the screening cutoff effectively introduces a feature in the power spectrum that becomes sharper with stronger modifications of gravity. Similarly to the BAO feature, this feature is not degenerate with galaxy bias and its detection is more robust than the broadband amplitude. The fact that the constraints improve indicates that the existence of a screening feature effectively compensates the loss of signal-to-noise from weakening the modifications on smaller scales.

To explore the dependence of the constraints on the choice of smallest scale to include in the analysis, we have recomputed our forecasts for more optimistic and pessimistic values of σ_{thr} [see Eq. (14)]. Figure 3 shows the dependence of the final constraints on this choice. Even though the final uncertainties on e.g. c_B or c_T may decrease by a factor of ~ 2 with respect to the constraints found for our fiducial choice (displayed by a vertical dashed line), we find that significant information can still be gained

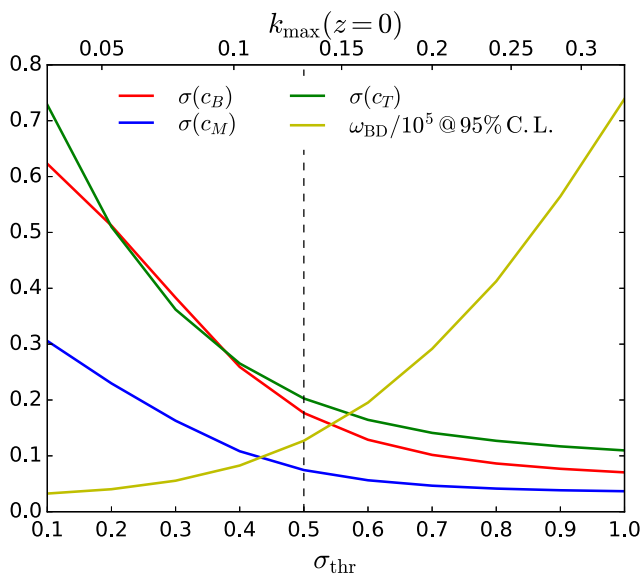


FIG. 3. Forecast constraints on c_M , c_B , c_T and ω_{BD} (the latter given as the upper bound at 95% C.L.) as a function of the maximum mode included, ℓ_{max} , set by maximum amplitude of clustering σ_{thr} as described in Sec. III.

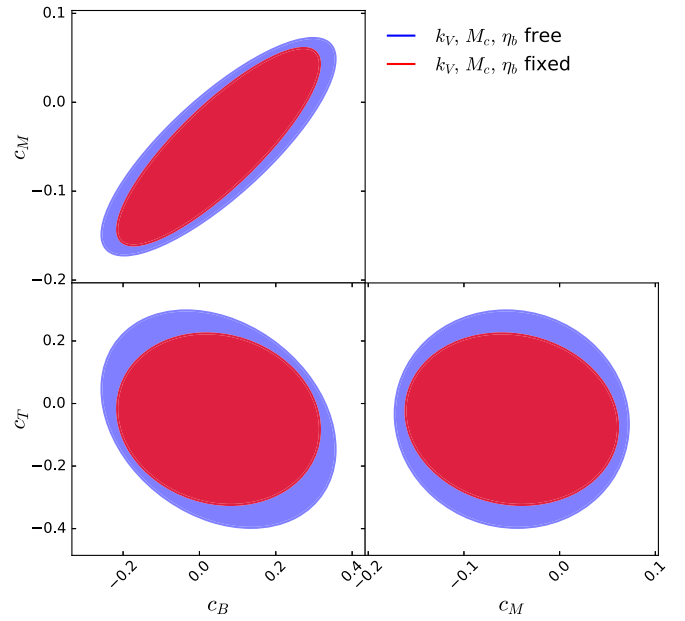


FIG. 4. Forecast constraints on c_M , c_B and c_T assuming perfect knowledge of baryonic physics and screening effects (red ellipses) and marginalizing over them (blue ellipses) as parametrized in Sec. II.

assuming a correct modeling of the matter power spectrum down to scales $k \sim 0.1 h \text{ Mpc}^{-1}$ at $z = 0$.

We have also quantified the impact of uncertainties on parametrization of the nonlinear screening and the effect of baryonic physics on the power spectrum by marginalizing over the values of M_c , η_b (see [64]) and k_V . Figure 4 shows the 1σ constraints on the Horndeski parameters assuming perfect knowledge of these parameters (red ellipses) and marginalizing over them (blue ellipses). The degradation in the final constraints is kept at a reasonable level due to the conservative cut on small scales used in this work.

D. The relative importance of the different surveys

One can broadly characterize a cosmological survey in terms of its ability to measure two main observables: the angular diameter distance relation and growth of structure. The former will be sensitive to the background expansion and therefore, the equation of state w , while the latter will depend on all parameters. It should also be possible to constrain the shape of the power spectrum as a function of redshift and thus pick out scale-dependent effects on the growth rate [e.g. the braiding scale in Eq. (4) and the k dependence in the quasistatic parameters, the Appendix]. But this will be, for now, a subdominant effect and, in the case of a galaxy redshift survey, is very sensitive to assumptions about bias.

At early times, a measurement of the primary anisotropies in the CMB will serve as an anchor for both distance measurements as well as for the growth rate: it fixes the angular diameter distance at $z \approx 1000$ as well as the overall

primordial amplitude of fluctuations. Per se it will not tightly constrain the gravitational parameters but it will play a crucial role in breaking degeneracies. In addition, S4 will supply us with a high-significance map of the projected matter fluctuations (with a radial kernel peaking around $z \sim 2$) via weak lensing. As such it will help to calibrate measurements of the growth rate at lower redshift as well as to pin down the neutrino mass.

Complementing early-time constraints from the CMB are late-time measurements of large-scale structure from galaxy clustering and weak lensing with LSST as well as intensity mapping with SKA. Specifically, an intensity mapping survey such as SKA1-IM will give us a biased measurement of the matter power spectrum as a function of redshift and therefore a distance measurement via the baryon acoustic oscillation features. To a lesser degree of importance, it will also give us a measurement of the growth rate via redshift-space distortion, but only if an independent measurement of the background HI temperature can be made. The LSST survey will have two complementary data sets. On the one hand it will supply us with a map of the galaxy distribution and therefore a measurement of the angular diameter distance as well as a low-significance measurement of redshift-space distortions over a range of redshifts. On the other hand it will supply us

with a tomographic set of weak-lensing maps which will give us an unbiased measurement of the growth of structure through the matter power spectrum.

In Fig. 5 we can see the important role that the complementarity between early-time and late-time constraints. It is already well established that the CMB and late time measurements (such as LSST) will combine to supply powerful constraints on the equation of state, the matter density and neutrino masses. The degeneracies broken by this combination thus improve the uncertainties on the gravitational parameters, as can be seen in the $\sim 30\%$ increase in the figure-of-merit shown in the last column of Table I between rows 2 and 4. There are two reasons for this. First of all, the CMB will anchor the distance measurement so that with LSST, it is possible to greatly reduce the uncertainty on w . This will feed into the degeneracy between c_M , c_B and w in the growth rate. Second, the CMB weak lensing will pin down the matter power spectrum, complementing the measurements of the growth rate via the LSST clustering survey and the matter power spectrum via the LSST weak-lensing survey.

As independent probes of geometry (through the BAO scale) and growth (through the power spectrum), it is also useful to look at the individual contributions of galaxy clustering and cosmic shear to the constraints. In Fig. 6 we

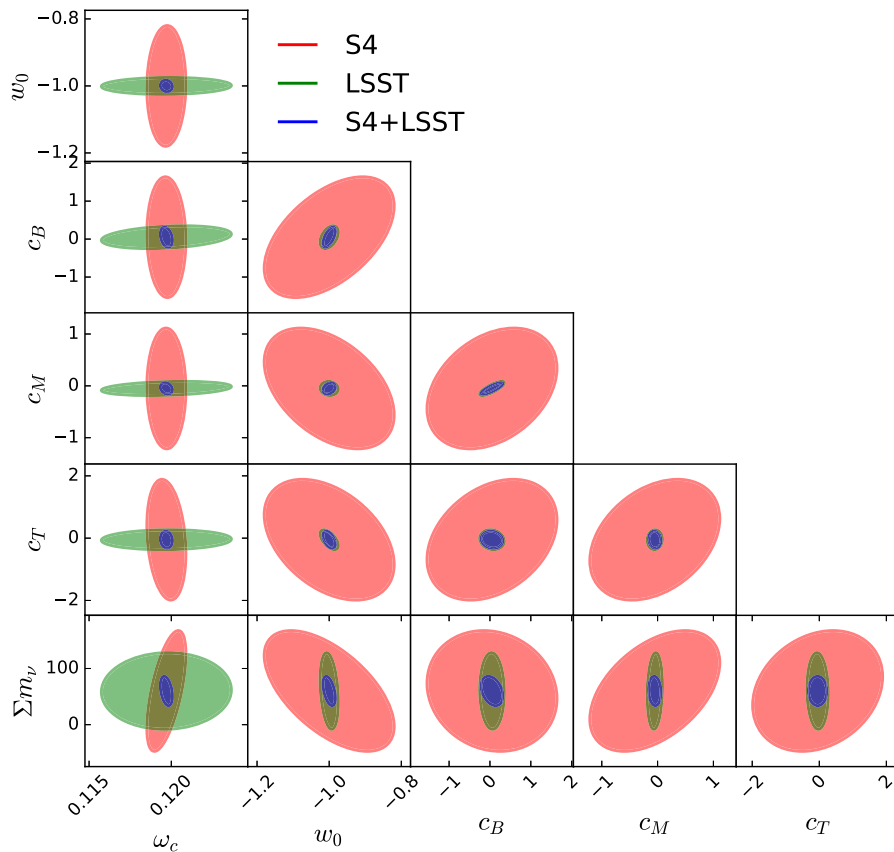


FIG. 5. 1σ cosmological constraints achievable by a Stage-4 CMB experiments (red), LSST (green) and the combination of the two (blue).

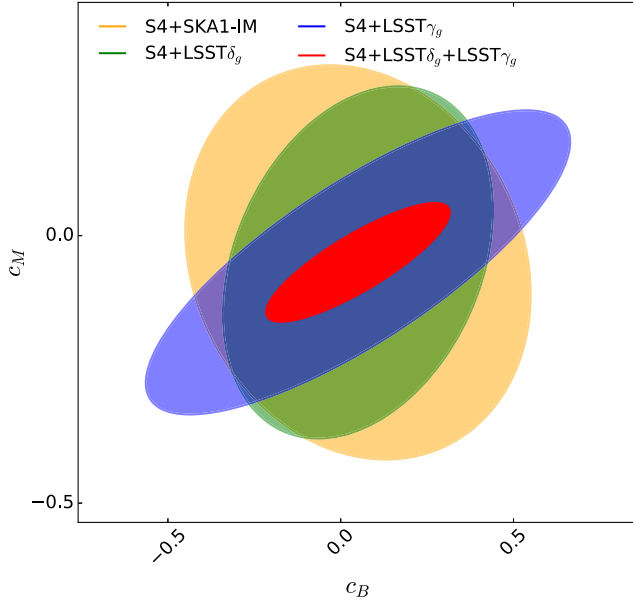


FIG. 6. Complementarity between clustering observables (SKA intensity mapping in orange and LSST clustering in green), providing measurements of expansion like the BAO scale, and weak lensing (LSST shear in blue), providing measurements of the growth of structure. The final constraints improve massively after combining both observables (red).

show how clustering and shear complement each other in the c_B - c_M plane. In particular, the ability of the shear catalogue to give us an unbiased measurement of the matter power spectrum as a function of time leads to an improved measurement of the standard cosmological parameters and the dark energy equation of state, and thus helps break some of the degeneracies that arise from galaxy clustering alone. The figure also shows how, unsurprisingly, the contributions the two clustering probes individually (intensity mapping and LSST clustering) are similar.

E. More complex time dependence

Throughout our analysis we have assumed a simplified fiducial model in which the time evolution of the α 's is tightly correlated with the emergence of an accelerated expansion phase. Given that our goal is to forecast how our constraints on the α 's will improve with Stage IV and given that current constraints are undertaken with such an assumption, we believe this is a sensible approach. The forecast constraints we obtain are meaningful and give us an idea of how much better constraints on Horndeski theory will be.

Nevertheless, we are well aware that, when focusing on specific models, our fiducial parametrization may not be ideal. It does correctly capture the time evolution for a large subclass of Horndeski models but it has been shown that for some subsets of model space, it is a poor approximation [51]. In fact, as we will see, in the case of the simplest nontrivial Horndeski model—Jordan-Brans-Dicke theory—the α 's are approximately constant and, thus, their evolution is

completely decoupled from the onset of accelerated expansion. Moreover, in models based on a covariant Lagrangian, such as Eq. (1), the background and perturbation evolution will depend on the same set of parameters, and their evolution will be related. In contrast, we are effectively decoupling the background [$w(z)$] from the growth [$\alpha_X(z)$], effectively granting the model more freedom to simultaneously fit different observables.

A favoured approach to achieve full generality is to construct an orthogonal basis which is completely agnostic about the time evolution of the α 's using, for example, principal component analysis or some form of eigenbasis [17,116,117]. This approach is systematic and, in some sense, complete, and will clearly be useful when actually analyzing the data. However it is very difficult to interpret forecasted errors on such parametrizations in a meaningful way and so, in what follows, we will choose an alternative parametrization which incorporates some features that escape our fiducial model. Doing so will allow us to assess how dependent our forecasts are on our assumptions.

A first, straightforward modification is to incorporate a constant term that allows us to forecast constraints from epochs prior to the accelerated expansion era. This allows us to generalize the idea of early dark energy [118] and to place bounds on the properties of gravity at early times. Models that attempt to unify cosmic acceleration in inflation and the late universe will be severely constrained by the data if they present with residual departures from GR between those stages. Thus, we start by reintroducing the constant term in our parametrization, given by b_X in Eq. (18). This term modifies the dynamics throughout cosmic history (see [119,120] for the effect of modifications only on the matter era and [121] for CMB effects in a reduced parameterization). For small early-time departures from GR the super-horizon adiabatic perturbations evolve as in the standard case, allowing us to set the standard initial conditions [122]. We shall see, when assessing constraints on Jordan-Brans-Dicke theories, that indeed these early-time, nontrivial values of α are present.

Additionally, we have also considered an alternative parametrization where the transition between the early-time and late-time terms is characterized by a redshift z_H and a transition width Δz_H :

$$\alpha_X = b_X + (c_X - b_X) \frac{1 - \tanh[(z - z_H)/(2\Delta z_H)]}{1 + \tanh[z_H/(2\Delta z_H)]}. \quad (21)$$

By marginalizing over z_H and Δz_H , this model thus allows us to assess the impact of uncertainties on the precise time evolution of the Horndeski parameters on the final constraints on their early-time and late-time values. Here we will use fiducial values for these parameters $z_H = 0.5$, $\Delta z_H = 0.5$. The corresponding time evolution of the α 's roughly mimics that of $\propto \Omega_{DE}$, although the effects of the

TABLE III. 1σ constraints on the Horndeski parameters c_X (late-time) and b_X (early-time). Columns 2 and 3 give the errors for the parametrization proportional to the normalized dark energy density, Eq. (18). Columns 4 and 5 are for the time parametrization given by Eq. (21), described in Sec. IV E. In the latter case, forecasts include marginalization over the time-dependence parameters z_H and Δz_H [see Eq. (21)].

Term	$\propto \Omega_{\text{DE}}$		$\propto \tanh$	
	σ_b (early)	σ_c (late)	σ_b (early)	σ_c (late)
α_K	0.29	3.5	0.15	1.6
α_B	1.3×10^{-3}	0.17	2.3×10^{-3}	0.21
α_M	3.6×10^{-5}	0.072	3.6×10^{-5}	0.013
α_T	0.12	0.18	2.7×10^{-2}	0.070

late-time modifications are more prominent for a slightly larger period of time.

Columns 2 and 4 in Table III show the forecast 1σ uncertainties from the combination of S4 and LSST for the early-time parameters for both parametrizations, while the late-time constraints are shown in columns 3 and 5. We find that early-time modification are much more tightly constrained than late-time modifications. This is to be expected, as early modifications will have an impact on CMB observables and induce modification to the growth rate which will then persist over a much longer period of time. Furthermore, the effects of late and early times are almost entirely uncorrelated. This can be seen if we compare the fourth row of Table I with the fourth column of Table III, where the uncertainty on c_X account for marginalization over b_X .

It is also worth noting that, while the achievable uncertainties on both the early and late-time parameters are similar in both models, the constraints improve for the new parametrization (see below for the particular case of c_B). The reason for this is twofold: first, the effect of the late-time terms is present for a larger period of time in this model. Second, the time derivative of the α 's may be more pronounced in this model (depending on the choice of Δz_H), and since this time derivative enters in the evolution equations the constraints are tighter.

Finally, Fig. 7 shows the effects of marginalizing over the time-evolution parameters z_H and Δz_H (without any prior) on the final constraints on the late-time parameters. While the effect on c_M and c_T is small, the uncertainty on c_B increases by a factor ~ 3 . This is likely because the perturbation equations depend on the time derivative of α_B (as well as α_K) and hence changing the width of the transition is more degenerate with c_B than α_M , α_T , whose time derivative does not enter the perturbation equations (cf. Appendix A2 in Ref. [65]). This shows that, when constraining general classes of models, it will be important to account for possible uncertainties in the time evolution of the modifications, although tight constraints can still be achieved after this marginalization.

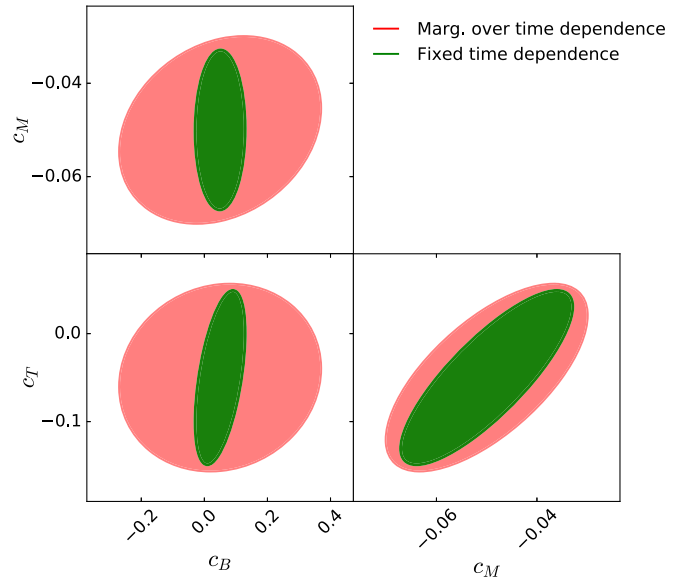


FIG. 7. 68% constraints for the Horndeski parameters for a fixed time dependence (green) and marginalized over the time-dependence parameters z_H and Δz_H [see Eq. (21)].

F. Constraints on the Jordan-Brans-Dicke theory

We now focus on a very specific class of models: Jordan-Brans-Dicke theory [32]. One can think of this as a very restricted set of priors on the α 's in the Horndeski theory. In fact we can express the α 's in terms of the background scalar field evolution, $\phi(t)$ and the Brans-Dicke parameter, ω_{BD} as

$$\begin{aligned}
 \alpha_M &= \frac{d \ln \phi}{d \ln a}, \\
 \alpha_B &= -\alpha_M, \\
 \alpha_K &= \omega_{\text{BD}} \alpha_M^2, \\
 \alpha_T &= 0.
 \end{aligned} \tag{22}$$

In principle it should be possible to further restrict the dependence of the α 's on ϕ by using the fact that $\phi(t)$, in certain situations, will lock onto a tracking solution. In an Einstein-de Sitter universe we have $\phi = \phi_0 a^{1/(\omega_{\text{BD}}+1)}$ [123] where $\phi_0 = (2\omega_{\text{BD}} + 4)/(2\omega_{\text{BD}} + 3)$ and therefore $\alpha_M = 1/(\omega_{\text{BD}} + 1)$. In other words, the Jordan-Brans-Dicke theory corresponds to a point in the space defined by the α 's.

In the presence of dark energy, characterized by an equation of state w , there will be very slight deviations from the attractor due to late-time acceleration of the scalar field [124], although it may still be possible to construct an accurate, perturbative solution to the time evolution of ϕ that only depends on ω_{BD} and w . If that is the case, then the α will be uniquely determined by ω_{BD} and w . Note that, for simplicity, we will set $w = -1$ in this section.

We carry out a Fisher matrix analysis using $1/\omega_{\text{BD}}$ as a parameter which we vary around the fiducial point

$\omega_{\text{BD}} = 10^5$ (we verified that the forecasts do not depend on this choice). Assuming that the uncertainty on $1/\omega_{\text{BD}}$ is Gaussianly distributed, it is straightforward to translate the 1σ uncertainty on $1/\omega_{\text{BD}}$ into a 95% C.L. lower bound for ω_{BD} . The second column in Table I shows the forecast values of this lower bound for a variety of combinations of data sets, and Figure 8 displays the forecast 1D distribution of $1/\omega_{\text{BD}}$ for a similar range of observations. For a start we can see that the combined, most optimistic constraint is remarkable: we will be able to place a lower bound, $\omega_{\text{BD}} > 1.7 \times 10^4$, which is comparable to current millisecond pulsar constraints and just under current constraints from the Solar System.

The fact that different data sets probe very different regimes plays a crucial role in our findings. Unlike in the case of our simplified parametrization of the α 's (in terms of c_M , etc), in this case, for a given value of ω_{BD} , the α 's are nonzero throughout cosmic history. We see then that combining the CMB S4 data, which primarily constrains the α 's at $z \sim 1000$ with late-time constraints from LSST galaxy and weak-lensing surveys will improve the constraint on ω_{BD} by a factor of 4 or more. It is also interesting to note that, independently, both SKA-IM and LSST give us a bound on ω_{BD} of order 10^4 . Given that each of these data sets will be affected by their own set of systematics, this will allow us a stringent cross-check on the bounds.

Finally, it is important to check how our ability to model nonlinear scales will affect our constraint. In Fig. 3 we show the projected lower bound on ω_{BD} as function of the maximum wave number (or density threshold). Reducing what corresponds to k_{max} at $z = 0$ from 0.1 Mpc^{-1} (our fiducial value) to 0.05 Mpc^{-1} degrades the constrain on ω_{BD} by more than a factor of 3. If cosmological constraints are to be competitive with other, more local, constraints, it will be important to model nonlinearities accurately for Jordan-Brans-Dicke theory; given its dependence on one (constant) parameter this should be a much simpler problem than in the

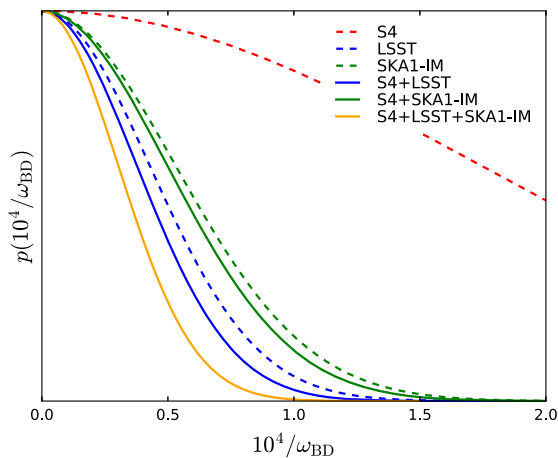


FIG. 8. Marginalized distribution for the inverse of the Brans-Dicke parameter ω_{BD} for different combinations of experiments. The constraints are mostly driven by late-time probes.

general Horndeski scenario. Moreover, this theory does not have a screening mechanism and hence reliably modeling nonlinear scales significantly improves the constraints, as can be seen in Fig. 3.

V. DISCUSSION

In this paper we have looked forward at what might be achievable with future surveys, focusing on constraints on scalar-tensor theories described by a broad subclass of Horndeski theories as well as a very specific model: Jordan-Brans-Dicke theory. Our results have been enlightening and we can summarize them as follows:

- (i) Constraints on the Horndeski parameters will greatly improve. But, given the weak effect α_K has on the data, we only expect precise constraints on α_M , α_B and α_T .
- (ii) We find that, in the best-case scenario, through a combination of Stage-IV CMB, photometric and spectroscopic surveys, constraints on α_M and α_B will improve (relative to current constraints) by a factor of 5 to $\sigma(c_M) \approx 0.06$ and $\sigma(c_B) \approx 0.12$. Furthermore, α_T will join the ranks of the well-constrained parameters, with $\sigma(c_T) \approx 0.15$.
- (iii) Relativistic effects and, more generally, ultralarge-scale modes, will have negligible statistical weight on the constraints on α 's compared to smaller-scale fluctuations.
- (iv) There are correlations between the gravitational parameters, $\alpha_{B,M,T}$ and the more conventional cosmological parameters, w and ω_c , which can be understood in terms of their joint effects on the growth of structure.
- (v) There is no apparent degeneracy between modified gravity and the sum of neutrino masses.
- (vi) The combined Stage-IV surveys will lead to a lower bound on the Brans-Dicke parameter $\omega_{\text{BD}} > 1.7 \times 10^4$, which is comparable with current millisecond pulsar constraints and of the same order of current Solar System constraints.
- (vii) It will be important to find an accurate method for modeling nonlinearities, although restricting the analysis to purely linear modes still leads to moderately tight constraints. Focusing on the Jordan-Brans-Dicke theory, we find that the degradation in the lower bound of ω_{BD} can be substantial, even if we discard a moderate number of nonlinear modes.
- (viii) The model for the time evolution of the α 's can have a substantial impact in the final constraints. While the fiducial model we use here captures the correct behavior at low redshift for a large subclass of the Horndeski theories, the accuracy of the forecasts and future constraints would be greatly improved with a parametrization which reflects the correct, theoretically sound, underlying behavior of the parameters.

There are a few comments we can make, building on our conclusions. For a start, the fact that ultralarge scales play a

negligible role in constraining the gravitational parameters lends credence to the approach of [120,125]. There, the authors suggest that, by looking at the values of the quasistatic parameters, it is possible to efficiently rule out large swathes of the Horndeski family. In a sense, by establishing a clear classification of the effects of Horndeski theories on the effective Newton's constant and the gravitational slip, it is possible to make definitive quantitative statements about the values and ranges of the α 's which are observationally viable.

It is interesting to compare different combination of surveys and their ability to constrain the gravitational parameters. We can define a figure of merit (FoM) from the Fisher matrix⁶

$$\text{FoM}_S = [\det(F^{-1})_S]^{-1/\dim(S)}, \quad (23)$$

where S indicates the subspace of parameters that enter the figure of merit (all others marginalized), which we have chosen to be the set of currently constrainable Horndeski parameters $\{c_B, c_M, c_T\}$ [36]. We do not include the expansion rate of the Universe (captured by the commonly used $w - w_a$ FoM) to focus instead on the impact of gravity on LSS. The values for this figure of merit forecast for the different experiments discussed in this work are shown in the last column of Table I. It clearly displays the benefit of combining different probes to test gravity: the figure of merit for S4 + LSST + SKA1-IM increases by 2 orders of magnitude relative to single-experiment S4 and SKA1-IM, and roughly doubles with respect to LSST-only. Further addition of a DESI-like spectroscopic survey increases this FoM by about 15%. These values show a very significant improvement with respect to current constraints. By restricting ourselves to the $c_B - c_M$ plane, constrained by [36], we observe an increase of a factor ~ 40 in the figure of merit between current constraints and those achievable by the combination of our three main experiments.

We have restricted ourselves to minimally coupled, strictly Horndeski terms. It is straightforward to enlarge our parameter space, as has been done in [38] and [128]. In fact, it is now possible to go beyond scalar-tensor theories and consider vector-tensor (such Einstein-Aether or Maxwell-Proca) or tensor-tensor (such as bigravity models) self-consistently on linear scales [129,130]. Furthermore, we have not included spectroscopic redshift surveys in a completely consistent way. As explained above, a completely consistent yet computational tractable way of including future redshift surveys is still lacking although, with inexorable improvement of computational speed and memory, may be achievable within the next few years. However, the result that ultralarge scales do not play a

significant role in the constraints means that it should be possible to construct an approximate, but far more efficient, forecasting apparatus using three-dimensional power spectra (as opposed to angular cross-power spectra).

Our results are also incomplete with regards to the treatment of systematic uncertainties. Even though, as stated in the text, the availability of cross-correlations between different surveys and probes ensures a robust self-calibration of sources of systematic uncertainty such as the effect of intrinsic alignments, a full treatment of shape-measurement systematics in weak lensing, or a complete characterization of photo- z uncertainties would allow us to place forecasts on scalar-tensor theories at the same level as the work currently done for standard cosmological scenarios. We leave this for future work, but expect these effects to somewhat weaken our forecasts.

On the other hand, even though we have considered most of the relevant overlapping Stage-IV surveys, thus exploiting the complementary coverage of the same patch of the sky by multiple tracers of the matter density field, there are other cosmological probes that might be used to further constrain gravity. Notable examples are peculiar velocities (nearby, via Tully-Fisher measurements or distant, via the kinetic-Sunyaev-Zeldovich effect), cluster number counts (using optical, x-ray or Sunyaev-Zeldovich measurements) and distance measurements from supernovae and strong lenses. All of these will add further statistical weight to our forecast constraints and could reduce the uncertainties on the gravitational parameters.

What is clear is that the future of gravitational physics in the context of observational cosmology is very promising. Over the next decade, constraints on cosmological scales will match in precision those on astrophysical scales. This will place GR as one of the most thoroughly tested theories in the physical canon.

ACKNOWLEDGMENTS

We acknowledge discussions with T. Baker, E. Calabrese, M. Lagos, N. Lima, L. Lombriser, S. Nesseris, J. Noller, and A. Silvestri. We are grateful to A. Avilez and C. Skordis for help in testing our implementation of Brans-Dicke models. D. A. is supported by the European Research Council (ERC), Science Technology Facilities Council (STFC), and the Beecroft Trust. P. G. F. acknowledges support from STFC, the Beecroft Trust, the Higgs Centre at the University of Edinburgh, and ERC. E. B. is supported by the ERC and the Beecroft Trust.

APPENDIX: THE QUASISTATIC LIMIT

Consider the two quasistatic parameters

$$\begin{cases} \gamma = \gamma_0(t) + \gamma_1(t)\left(\frac{H}{k}\right)^2 \\ G_{\text{eff}} = G_0(t) + G_1(t)\left(\frac{H}{k}\right)^2. \end{cases}$$

⁶Our definition ensures that the FoM scales as a surface, rather than the multi-dimensional volume of the parameter space (cf. [126,127]). This facilitates comparison of FoMs involving different number of parameters.

We can relate them to the Horndeski parameters as follows. At leading order we have

$$\gamma_0 \equiv \frac{\beta_3}{\beta_1 + \beta_2}, \quad (\text{A1})$$

$$G_0 \equiv \frac{2(\beta_1 + \beta_2)}{M_*^2[2\beta_1 + (2 - \alpha_B)\beta_2]}. \quad (\text{A2})$$

At second order in (\mathcal{H}/k) we have

$$\frac{\gamma_{\text{N1}}}{\alpha_M - \alpha_T} \equiv 3\alpha_B\beta_2 \left[1 + \frac{H''}{aHH'} \right] \frac{H'}{aH^2} \quad (\text{A3})$$

$$- 3\beta_1 \left[\beta_3 - 2(\alpha_M - \alpha_T) \frac{H'}{aH^2} \right] \quad (\text{A4})$$

$$- \alpha_K \frac{\beta_3}{aH} \left[\frac{H'}{H} + \frac{\alpha'_K}{\alpha_K} \right] \quad (\text{A5})$$

$$- (3\alpha_B + \alpha_K)\beta_3 \left[3 + \alpha_M + \frac{H'}{aH^2} \right], \quad (\text{A6})$$

$$\gamma_{\text{D1}} \equiv (\beta_1 + \beta_2)^2,$$

$$G_{\text{N1}} \equiv \frac{2\alpha_B\gamma_{\text{N1}}}{\alpha_M - \alpha_T} - 2(6 - 6\alpha_B - \alpha_K)(\beta_1 + \beta_2)\beta_3 \quad (\text{A7})$$

$$+ 6(\beta_1 + \beta_2)\beta_2 \left[3 \frac{(\rho_m + p_m)}{M_*^2 H^2} + (2 - \alpha_B) \frac{H'}{aH^2} \right], \quad (\text{A8})$$

$$G_{\text{D1}} \equiv M_*^2[2\beta_1 + (2 - \alpha_B)\beta_2]^2, \quad (\text{A9})$$

where $G_1 \equiv G_{\text{N1}}/G_{\text{D1}}$, $\gamma_1 \equiv \gamma_{\text{N1}}/\gamma_{\text{D1}}$ and

$$\beta_1 \equiv -\frac{3(\rho_m + p_m)}{H^2 M_*^2} + \frac{\alpha'_B H - (2 - \alpha_B)H'}{aH^2}, \quad (\text{A10})$$

$$\beta_2 \equiv \alpha_B(1 + \alpha_T) + 2(\alpha_M - \alpha_T), \quad (\text{A11})$$

$$\beta_3 \equiv (1 + \alpha_T)\beta_1 + (1 + \alpha_M)\beta_2. \quad (\text{A12})$$

Here ' represents the derivative with respect to conformal time and $\mathcal{H} = aH$.

-
- [1] E. Berti *et al.*, *Classical Quantum Gravity* **32**, 243001 (2015).
- [2] B. P. Abbott *et al.* (Virgo Collaboration and LIGO Scientific Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [3] T. Baker, D. Psaltis, and C. Skordis, *Astrophys. J.* **802**, 63 (2015).
- [4] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, *Phys. Rep.* **513**, 1 (2012).
- [5] J. Yoo, N. Hamaus, U. Seljak, and M. Zaldarriaga, *Phys. Rev. D* **86**, 063514 (2012).
- [6] D. Alonso, P. Bull, P. G. Ferreira, R. Maartens, and M. Santos, *Astrophys. J.* **814**, 145 (2015).
- [7] D. Alonso and P. G. Ferreira, *Phys. Rev. D* **92**, 063525 (2015).
- [8] J. Fonseca, S. Camera, M. Santos, and R. Maartens, *Astrophys. J.* **812**, L22 (2015).
- [9] P. Collaboration *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A13 (2016).
- [10] K. A. Olive *et al.* (Particle Data Group), *Chin. Phys. C* **38**, 090001 (2014).
- [11] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, and A. Slosar, *J. Cosmol. Astropart. Phys.* **05** (2014) 023.
- [12] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, *Phys. Rev. D* **92**, 123535 (2015).
- [13] A. Barreira, B. Li, C. Baugh, and S. Pascoli, *J. Cosmol. Astropart. Phys.* **08** (2014) 059.
- [14] M. Baldi, F. Villaescusa-Navarro, M. Viel, E. Puchwein, V. Springel, and L. Moscardini, *Mon. Not. R. Astron. Soc.* **440**, 75 (2014).
- [15] M. Baldi and F. Villaescusa-Navarro, *arXiv:1608.08057*.
- [16] M. A. Amin, R. V. Wagoner, and R. D. Blandford, *Mon. Not. R. Astron. Soc.* **390**, 131 (2008).
- [17] L. Pogosian, A. Silvestri, K. Koyama, and G.-B. Zhao, *Phys. Rev. D* **81**, 104023 (2010).
- [18] R. Bean and M. Tangmatitham, *Phys. Rev. D* **81**, 083534 (2010).
- [19] J. N. Dossett, M. Ishak, and J. Moldenhauer, *Phys. Rev. D* **84**, 123001 (2011).
- [20] M. Kunz and D. Sapone, *Phys. Rev. Lett.* **98**, 121301 (2007).
- [21] T. Baker, P. G. Ferreira, C. Skordis, and J. Zuntz, *Phys. Rev. D* **84**, 124018 (2011).
- [22] T. Baker, P. G. Ferreira, and C. Skordis, *Phys. Rev. D* **87**, 024015 (2013).
- [23] R. A. Battye and J. A. Pearson, *J. Cosmol. Astropart. Phys.* **07** (2012) 019.
- [24] R. A. Battye and J. A. Pearson, *J. Cosmol. Astropart. Phys.* **03** (2014) 051.
- [25] P. Creminelli, G. D'Amico, J. Norena, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **02** (2009) 018.

- [26] G. Gubitosi, F. Piazza, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **02** (2013) 032.
- [27] J. K. Bloomfield, E. E. Flanagan, M. Park, and S. Watson, *J. Cosmol. Astropart. Phys.* **08** (2013) 010.
- [28] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **08** (2013) 025.
- [29] J. Gleyzes, D. Langlois, and F. Vernizzi, *Int. J. Mod. Phys. D* **23**, 1443010 (2014).
- [30] G. W. Horndeski, *Int. J. Theor. Phys.* **10**, 363 (1974).
- [31] C. Deffayet, G. Esposito-Farese, and A. Vikman, *Phys. Rev. D* **79**, 084003 (2009).
- [32] C. Brans and R. H. Dicke, *Phys. Rev.* **124**, 925 (1961).
- [33] A. De Felice, T. Kobayashi, and S. Tsujikawa, *Phys. Lett. B* **706**, 123 (2011).
- [34] A. Avilez and C. Skordis, *Phys. Rev. Lett.* **113**, 011101 (2014).
- [35] P. Collaboration *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A14 (2016).
- [36] E. Bellini, A. J. Cuesta, R. Jimenez, and L. Verde, *J. Cosmol. Astropart. Phys.* **02** (2016) 053; **06** (2016) E01.
- [37] F. Piazza, H. Steigerwald, and C. Marinoni, *J. Cosmol. Astropart. Phys.* **05** (2014) 043.
- [38] J. Gleyzes, D. Langlois, M. Mancarella, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **02** (2016) 056.
- [39] J. S. Y. Leung and Z. Huang, *Int. J. Mod. Phys. D* **0**, 1750070 (2017).
- [40] M. Zumalacarregui and J. Garcia-Bellido, *Phys. Rev. D* **89**, 064046 (2014).
- [41] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Phys. Rev. Lett.* **114**, 211101 (2015).
- [42] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **02** (2015) 018.
- [43] D. Langlois and K. Noui, *J. Cosmol. Astropart. Phys.* **02** (2016) 034.
- [44] M. Crisostomi, K. Koyama, and G. Tasinato, *J. Cosmol. Astropart. Phys.* **04** (2016) 044.
- [45] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, *J. High Energy Phys.* **12** (2016) 100.
- [46] J. Gleyzes, D. Langlois, M. Mancarella, and F. Vernizzi, *J. Cosmol. Astropart. Phys.* **08** (2015) 054.
- [47] E. Bellini and I. Sawicki, *J. Cosmol. Astropart. Phys.* **07** (2014) 050.
- [48] J. Beltran Jimenez, F. Piazza, and H. Velten, *Phys. Rev. Lett.* **116**, 061101 (2016).
- [49] L. Lombriser and A. Taylor, *J. Cosmol. Astropart. Phys.* **03** (2016) 031.
- [50] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler, and M. Zumalacarregui, [arXiv:1608.01982](https://arxiv.org/abs/1608.01982).
- [51] E. V. Linder, G. Sengor, and S. Watson, *J. Cosmol. Astropart. Phys.* **05** (2016) 053.
- [52] D. J. E. Marsh, P. Bull, P. G. Ferreira, and A. Pontzen, *Phys. Rev. D* **90**, 105023 (2014).
- [53] O. Pujolas, I. Sawicki, and A. Vikman, *J. High Energy Phys.* **11** (2011) 156.
- [54] J. Renk, M. Zumalacarregui, and F. Montanari, *J. Cosmol. Astropart. Phys.* **07** (2016) 040.
- [55] B. Bertotti, L. Iess, and P. Tortora, *Nature (London)* **425**, 374 (2003).
- [56] C. Umilt, M. Ballardini, F. Finelli, and D. Paoletti, *J. Cosmol. Astropart. Phys.* **08** (2015) 017.
- [57] M. Ballardini, F. Finelli, C. Umilt, and D. Paoletti, *J. Cosmol. Astropart. Phys.* **05** (2016) 067.
- [58] A. Joyce, B. Jain, J. Khoury, and M. Trodden, *Phys. Rep.* **568**, 1 (2015).
- [59] B. Li, A. Barreira, C. M. Baugh, W. A. Hellwing, K. Koyama, S. Pascoli, and G.-B. Zhao, *J. Cosmol. Astropart. Phys.* **11** (2013) 012.
- [60] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, and S. Pascoli, *J. Cosmol. Astropart. Phys.* **10** (2013) 027.
- [61] B. Li, G.-B. Zhao, and K. Koyama, *J. Cosmol. Astropart. Phys.* **05** (2013) 023.
- [62] H. A. Winther *et al.*, *Mon. Not. R. Astron. Soc.* **454**, 4208 (2015).
- [63] E. Puchwein, M. Baldi, and V. Springel, *Mon. Not. R. Astron. Soc.* **436**, 348 (2013).
- [64] A. Schneider and R. Teyssier, *J. Cosmol. Astropart. Phys.* **12** (2015) 049.
- [65] M. Zumalacarregui, E. Bellini, I. Sawicki, and J. Lesgourgues, [arXiv:1605.06102](https://arxiv.org/abs/1605.06102).
- [66] D. Blas, J. Lesgourgues, and T. Tram, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [67] E. Di Dio, F. Montanari, J. Lesgourgues, and R. Durrer, *J. Cosmol. Astropart. Phys.* **11** (2013) 044.
- [68] I. Sawicki and E. Bellini, *Phys. Rev. D* **92**, 084061 (2015).
- [69] C. L. Bennett *et al.*, *Astrophys. J. Suppl. Ser.* **208**, 20 (2013).
- [70] Planck Collaboration *et al.*, *Astron. Astrophys.* **571**, A16 (2014).
- [71] S. Naess *et al.*, *J. Cosmol. Astropart. Phys.* **10** (2014) 007.
- [72] R. Keisler *et al.*, *Astrophys. J.* **807**, 151 (2015).
- [73] S. W. Henderson *et al.*, *J. Low Temp. Phys.* **184**, 772 (2016).
- [74] B. A. Benson *et al.*, in Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, *Proceedings of the SPIE*, Vol. 9153 (SPIE-International Society for Optical Engineering, Bellingham, WA, 2014), p. 91531P.
- [75] K. N. Abazajian *et al.*, [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [76] T. Louis *et al.*, [arXiv:1610.02360](https://arxiv.org/abs/1610.02360).
- [77] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006).
- [78] W. Hu and T. Okamoto, *Astrophys. J.* **574**, 566 (2002).
- [79] The Dark Energy Survey Collaboration, [arXiv:astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346).
- [80] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, *Phys. Rev. D* **80**, 083514 (2009).
- [81] A. Challinor and A. Lewis, *Phys. Rev. D* **84**, 043516 (2011).
- [82] C. Bonvin and R. Durrer, *Phys. Rev. D* **84**, 063505 (2011).
- [83] I. Mohammed, U. Seljak, and Z. Vlah, *Mon. Not. R. Astron. Soc.* **466**, 780 (2017).
- [84] LSST *et al.* (Science Collaboration), [arXiv:0912.0201](https://arxiv.org/abs/0912.0201).
- [85] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, *Astrophys. J.* **715**, 104 (2010).
- [86] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov, *Astrophys. J.* **672**, 19 (2008).
- [87] M. P. van Daalen, J. Schaye, C. M. Booth, and C. Dalla Vecchia, *Mon. Not. R. Astron. Soc.* **415**, 3649 (2011).

- [88] M. P. van Daalen, J. Schaye, I. G. McCarthy, C. M. Booth, and C. Dalla Vecchia, *Mon. Not. R. Astron. Soc.* **440**, 2997 (2014).
- [89] W. A. Hellwing, M. Schaller, C. S. Frenk, T. Theuns, J. Schaye, R. G. Bower, and R. A. Crain, *Mon. Not. R. Astron. Soc.* **461**, L11 (2016).
- [90] J. Harnois-Déraps, L. van Waerbeke, M. Viola, and C. Heymans, *Mon. Not. R. Astron. Soc.* **450**, 1212 (2015).
- [91] C. Heymans *et al.*, *Mon. Not. R. Astron. Soc.* **432**, 2433 (2013).
- [92] The Dark Energy Survey Collaboration *et al.*, *Phys. Rev. D* **94**, 022001 (2016).
- [93] H. Hildebrandt *et al.*, [arXiv:1606.05338](https://arxiv.org/abs/1606.05338).
- [94] C. M. Hirata and U. Seljak, *Phys. Rev. D* **70**, 063526 (2004).
- [95] L. D. Ferramacho, M. G. Santos, M. J. Jarvis, and S. Camera, *Mon. Not. R. Astron. Soc.* **442**, 2511 (2014).
- [96] P. Bull, P. G. Ferreira, P. Patel, and M. G. Santos, *Astrophys. J.* **803**, 21 (2015).
- [97] D. Bacon *et al.*, *Proc. Sci.*, AASKA14 (2015) 145 [[arXiv:1501.03977](https://arxiv.org/abs/1501.03977)].
- [98] T. Kitching *et al.*, *Proc. Sci.*, AASKA14 (2015) 146 [[arXiv:1501.03978](https://arxiv.org/abs/1501.03978)].
- [99] S. Yahya, P. Bull, M. G. Santos, M. Silva, R. Maartens, P. Okouma, and B. Bassett, *Mon. Not. R. Astron. Soc.* **450**, 2251 (2015).
- [100] R. Laureijs *et al.* (EUCLID Collaboration), [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [101] M. Jarvis *et al.*, *Proc. Sci.*, AASKA14 (2015) 18 [[arXiv:1501.03825](https://arxiv.org/abs/1501.03825)].
- [102] I. Harrison, S. Camera, J. Zuntz, and L. Brown, *Mon. Not. R. Astron. Soc.* **463**, 3674 (2016).
- [103] M. Santos *et al.*, *Proc. Sci.*, AASKA14 (2015) 19 [[arXiv:1501.03989](https://arxiv.org/abs/1501.03989)].
- [104] L. Wolz, F. B. Abdalla, C. Blake, J. R. Shaw, E. Chapman, and S. Rawlings, *Mon. Not. R. Astron. Soc.* **441**, 3271 (2014).
- [105] D. Alonso, P. Bull, P. G. Ferreira, and M. G. Santos, *Mon. Not. R. Astron. Soc.* **447**, 400 (2015).
- [106] D. Schlegel *et al.*, [arXiv:1106.1706](https://arxiv.org/abs/1106.1706).
- [107] A. Bailoni, A. Spurio Mancini, and L. Amendola, [arXiv:1608.00458](https://arxiv.org/abs/1608.00458).
- [108] Planck Collaboration *et al.*, *Astron. Astrophys.* **596**, A107 (2016).
- [109] T. Baker and P. Bull, *Astrophys. J.* **811**, 116 (2015).
- [110] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, [arXiv:1602.00674](https://arxiv.org/abs/1602.00674).
- [111] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, and M. White, *Mon. Not. R. Astron. Soc.* **444**, 476 (2014).
- [112] M. Crocce and R. Scoccimarro, *Phys. Rev. D* **73**, 063519 (2006).
- [113] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, *J. High Energy Phys.* **09** (2012) 082.
- [114] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, *J. Cosmol. Astropart. Phys.* **07** (2016) 052.
- [115] A. Mead, J. Peacock, C. Heymans, S. Joudaki, and A. Heavens, *Mon. Not. R. Astron. Soc.* **454**, 1958 (2015).
- [116] A. Hojjati, G.-B. Zhao, L. Pogosian, A. Silvestri, R. Crittenden, and K. Koyama, *Phys. Rev. D* **85**, 043508 (2012).
- [117] K. N. Abazajian *et al.*, [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [118] V. Pettorino, L. Amendola, and C. Wetterich, *Phys. Rev. D* **87**, 083009 (2013).
- [119] N. A. Lima, V. Smer-Barreto, and L. Lombriser, *Phys. Rev. D* **94**, 083507 (2016).
- [120] L. Perenon, C. Marinoni, and F. Piazza, *J. Cosmol. Astropart. Phys.* **01** (2017) 035.
- [121] P. Brax, C. van de Bruck, S. Clesse, A.-C. Davis, and G. Sculthorpe, *Phys. Rev. D* **89**, 123507 (2014).
- [122] L. Amendola, E. Bellini, J. Lesgourgues, V. Pettorino, I. Sawicki, and M. Zumalacarregui (to be published).
- [123] H. Nariai, *Prog. Theor. Phys.* **42**, 742 (1969).
- [124] N. A. Lima and P. G. Ferreira, *J. Cosmol. Astropart. Phys.* **01** (2016) 010.
- [125] L. Pogosian and A. Silvestri, *Phys. Rev. D* **94**, 104014 (2016).
- [126] S. Nesseris and J. Garcia-Bellido, *J. Cosmol. Astropart. Phys.* **08** (2013) 036.
- [127] A. Alarcon, M. Eriksen, and E. Gaztañaga, [arXiv:1609.08510](https://arxiv.org/abs/1609.08510).
- [128] G. D'Amico, Z. Huang, M. Mancarella, and F. Vernizzi, [arXiv:1609.01272](https://arxiv.org/abs/1609.01272).
- [129] M. Lagos, T. Baker, P. G. Ferreira, and J. Noller, *J. Cosmol. Astropart. Phys.* **08** (2016) 007.
- [130] M. Lagos and P. G. Ferreira, *J. Cosmol. Astropart. Phys.* **01** (2017) 047.