Will Multiple Probes of Dark Energy find Modified Gravity?

Charles Shapiro[1], Scott Dodelson[2,3,4], Ben Hoyle[5], Lado Samushia[1,6], Brenna Flaugher[2]

[1] *Institute of Cosmology and Gravitation, Portsmouth, PO1 3FX, United Kingdom*

[2] *Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510*

[3] *Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637*

[4] *Kavli Institute for Cosmological Physics, Chicago, IL 60637*

[5] *Institut de Ciencies del Cosmos, Barcelona, Spain and*

[6] *National Abastumani Astrophysical Observatory,*

Ilia State University, 2A Kazbegi Ave, GE-0160 Tbilisi, Georgia

(Dated: July 23, 2010)

One of the most pressing issues in cosmology is whether general relativity (GR) plus a dark sector is the underlying physical theory or whether a modified gravity model is needed. Upcoming dark energy experiments designed to probe dark energy with multiple methods can address this question by comparing the results of the different methods in constraining dark energy parameters. Disagreement would signal the breakdown of the assumed model (GR plus dark energy). We study the power of this consistency test by projecting constraints in the $w_0 - w_a$ plane from the four different techniques of the Dark Energy Survey in the event that the underlying true model is modified gravity. We find that the standard technique of looking for overlap has some shortcomings, and we propose an alternative, more powerful *Multi-dimensional Consistency Test*. We introduce the methodology for projecting whether a given experiment will be able to use this test to distinguish a modified gravity model from GR.

PACS numbers: 95.35.+d; 95.85.Pw

I. INTRODUCTION

General relativity (GR) is currently a bad fit to cosmological data unless a new substance, so-called *dark energy*, is invoked. If GR really is an incomplete or incorrect theory and we are tasked with identifying the correct model, a major hurdle will be determining how to confront upcoming data sets in the absence of a well-understood model. What new parameters should be introduced and fit for when, e.g., data on weak gravitational lensing or galaxy clusters are analyzed? Several authors have addressed this question [1–5], and it has recently become possible to test GR using survey data [6–9].

Here we address a slightly less ambitious question: using multiple cosmological probes, how can we determine whether cosmic acceleration is driven by dark energy or modified gravity (MG)? One approach is to analyze the data assuming that GR is correct and see whether the constraints on dark energy parameters from different probes overlap [10, 11]. Non-overlapping constraints would be a strong signal that the underlying parameterization is wrong; i.e, that GR+dark energy cannot account for the data and that a modified theory of gravity is called for. A similar approach is to look at parameter constraints coming from separate dynamical effects such as the cosmic expansion or perturbation growth [12]. Here we explore the former method in depth in the context of a concrete example.

Ishak et al. showed that, in principle, non-overlapping dark energy parameter constraints obtained from multiple experiments is a signature of MG [11]. In particular, they found that dark energy parameters obtained from a space-based supernova survey and a space-based weak lensing survey will not agree if the Universe is in fact described by the Dvali-Gabadadze-Porrati (DGP) braneworld model [13]. We reexamine this general method with our own example, assuming that the universe is governed by a toy MG model and considering projections from the upcoming Dark Energy Survey (DES). We present the projected constraints from all four DES probes in the plane of dark energy parameters w_0 and w_a, where the dark energy equation of state is assumed to be $w = w_0 + w_a(1 - a)$ and a is the scale factor of the universe. This straightforward plot is not the most powerful way to combine probes, so we introduce a more quantitative formalism that should be useful for future attempts in this direction. The formalism assigns a χ^2 for the combined probes which can be interpreted in the usual fashion so that a "bad" χ^2 corresponds to disagreement among the probes, and therefore a quantitative assessment of how well the model of GR+dark energy works.

Section II discusses modified gravity models in general and details the modified gravity model we adopt as our working example. Section III then presents the DES projections in the (w_0, w_a) plane along with a description of the shortcomings of this approach. In Section IV, we present a more quantitative approach (see also [14]), which we call the *Multi-dimensional Consistency Test* (MCT), illustrate how to obtain MCT projections, deal with the issue of degenerate directions, and finally conclude by applying this formalism to DES for the model under study.

II. PERTURBATIONS IN MODIFIED GRAVITY

The metric in the class of modified gravity models we consider retains its standard GR form

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

where a is the cosmic scale factor and Φ and Ψ are the scalar gravitational potentials. Hu & Sawicki [4] proposed introducing two functions which parameterize deviations from GR:

$$
\begin{aligned}
g &\equiv \frac{\Phi + \Psi}{\Phi - \Psi} \\
f &\equiv \frac{8\pi G \rho_m a^2 \delta}{k^2 (\Phi - \Psi)} - 1
\end{aligned}
\tag{2}
$$

where ρ_m is the background matter density. In GR, the two scalar potentials are equal and opposite and Poisson's equations holds, so both f and g vanish. In modified gravity models, though, both f and g can be non-vanishing and are functions of both wavenumber k and scale factor a.

Other parameterizations can be rewritten in terms of f and g. For example, instead of g, $\eta \equiv -\Phi/\Psi$ is often used, so that $g = (\eta - 1)/(\eta + 1)$. Zhao et al. [5] introduce a function which governs the peculiar velocity potential

$$\mu \equiv -\frac{k^2 \Psi}{4\pi G \rho_m a^2 \delta}. \tag{3}$$

In terms of f and g,

$$\mu = \frac{1 - g}{1 + f}, \tag{4}$$

so $\mu = 1$ in standard GR. Another common way of parameterizing the effects of modified gravity was introduced by Linder [1, 2] and treats the growth factor of matter perturbations

$$\frac{d\ln\delta}{d\ln a} = \Omega_m(a)^\gamma \tag{5}$$

where $\Omega_m(a) \equiv \Omega_{m,0}/[H(a)/H_0]^2$. For a wide variety of models, taking γ to be a constant works well and is an appealing way to confront data. In the context of the Hu & Sawicki formalism, γ is *not* another new parameter, but rather is governed by the growth equation which becomes

$$\Omega_m(a)^{\gamma-1}\left[(1 - 2\gamma)\frac{d\ln H}{d\ln a} - 3\gamma + 2\right] + \Omega_m(a)^{2\gamma-1} = \frac{3}{2}\mu = \frac{3(1 - g)}{2(1 + f)} \tag{6}$$

where $H(a)$ is the expansion rate. In standard GR, one finds that $\gamma \simeq 0.55$. A choice of γ corresponds to a choice of f and g. Therefore, one cannot choose all three independently.

In this work, we choose as the underlying "true" model a toy model in which $\gamma = 0.68$ and $f = 0$. This is approximately the case in the Dvali-Gabadadze-Porrati (DGP) braneworld model [1, 13], though we will not be using that model specifically and only mention it here as a motivation. The background expansion in DGP is different from that of ΛCDM, but we have chosen our model to have $w = -1$. Thus, the only observable differences between our toy model and ΛCDM will enter via the growth function, which determines the normalization of the linear matter power spectrum as a function of redshift. This is a mild modification to GR since scale-dependent growth modifications are indeed possible, and it has been shown that scale-dependent MG will be easier to detect [5]. In order to maintain consistency, our choices of γ and f determine g via (6). Fig. 1 shows the resulting $g(z)$; note that constant γ is not consistent with a constant g model. It is worth noting here that the MG model will produce *more* structure at early times than ΛCDM for fixed fluctuation amplitude today σ_8. This is because structure grows more slowly in the MG model, therefore the Universe was clumpier (closer to its present state) in the past.

III. DES PROJECTIONS FOR THE DARK ENERGY EQUATION OF STATE

A. Parameter Misestimation

For concreteness, we focus on the upcoming Dark Energy Survey, which will probe dark energy using thousands of Type Ia supernovae (SN), many high redshift clusters (CL), the Baryon Acoustic Oscillation (BAO) scale, and
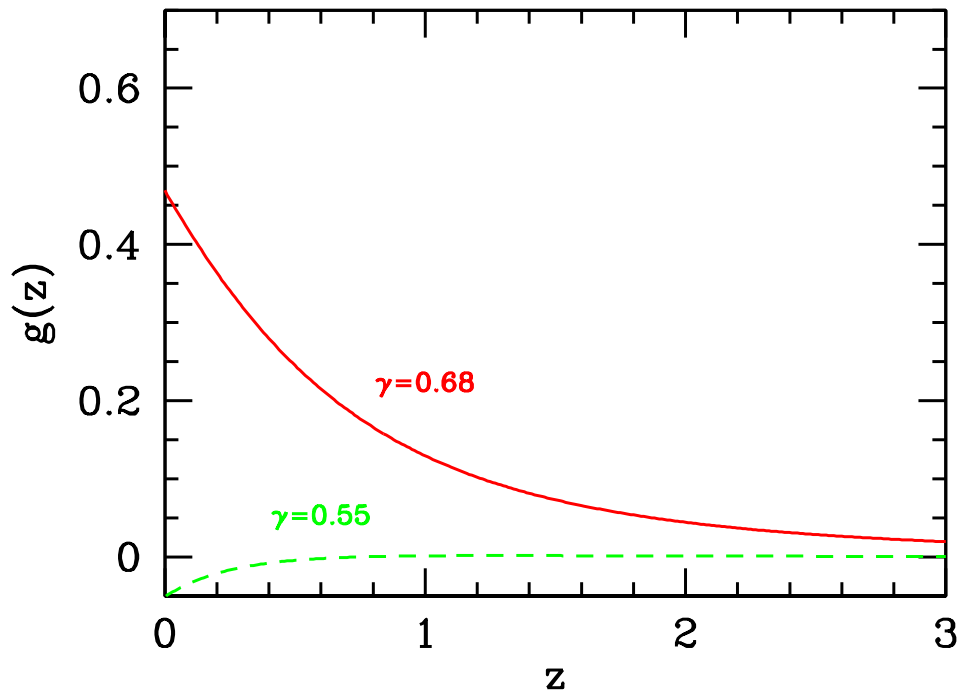
FIG. 1: The post-Friedmann parameter g as a function of redshift. Standard general relativity corresponds to $g = 0$ or $\gamma \simeq 0.55$ (dashed curve). The fiducial value used in this paper is shown by the solid curve corresponding to a model in which $\gamma = 0.68$.

the cosmic shear signal extracted from weakly lensed shapes of millions of background galaxies (WL). There is a well-established formalism, the Fisher matrix approach, for projecting constraints from experiments such as DES. Ordinarily, one forms the Fisher matrix from which contours representing the 68% confidence region in the (w_0, w_a) plane, say, can be drawn. These contours are drawn centered on the assumed underlying model. The Fisher matrix formalism is valid when the joint likelihood function of the cosmological parameters is a gaussian.

Here we are after something a little different. We want to determine not only how large the error contours will be, but also where they would be centered in the event that an incorrect model is used to analyze the data. Our assumed true model is a toy modified gravity model with $\gamma = 0.68$ and the same cosmic expansion as ΛCDM. We want to know what answer an analyst would get if s/he fit for dark energy parameters while assuming that GR was correct ($\gamma = 0.55$). What values of w_0 and w_a would be obtained? There is a simple extension [15] of the Fisher formalism which provides an answer to this question. It involves three steps:

1. Calculate the difference in the quantity to be measured (e.g., a power spectrum) in the true model and in the fitted model. Call this ΔP_i where subscript i labels the bins in which it is measured

2. Calculate the Fisher matrix for the parameters λ_α to be fit to the data. For a single experiment,

$$F_{\alpha\beta} = \sum_{ij}(\text{Cov}^{-1})_{ij}\frac{\partial P_i}{\partial \lambda_\alpha}\frac{\partial P_j}{\partial \lambda_\beta} \tag{7}$$

where P_i is the observed quantity in bin i, and $(\text{Cov})_{ij}$ is the covariance matrix for bins i and j. The covariance should be calculated using the cosmological model assumed to be true, while the derivatives should be calculated using the model we will fit. Priors may be added to the Fisher matrix if they come from probes which are robust in the case of MG (e.g. a prior on the Hubble constant, h).

3. To first order in ΔP_j, the parameter λ_α will be mis-estimated, or biased, by an amount

$$\Delta\lambda_\alpha = \sum_{\beta}(F^{-1})_{\alpha\beta}\sum_{ij}(\text{Cov}^{-1})_{ij}\frac{\partial P_i}{\partial \lambda_\beta}\Delta P_j \tag{8}$$
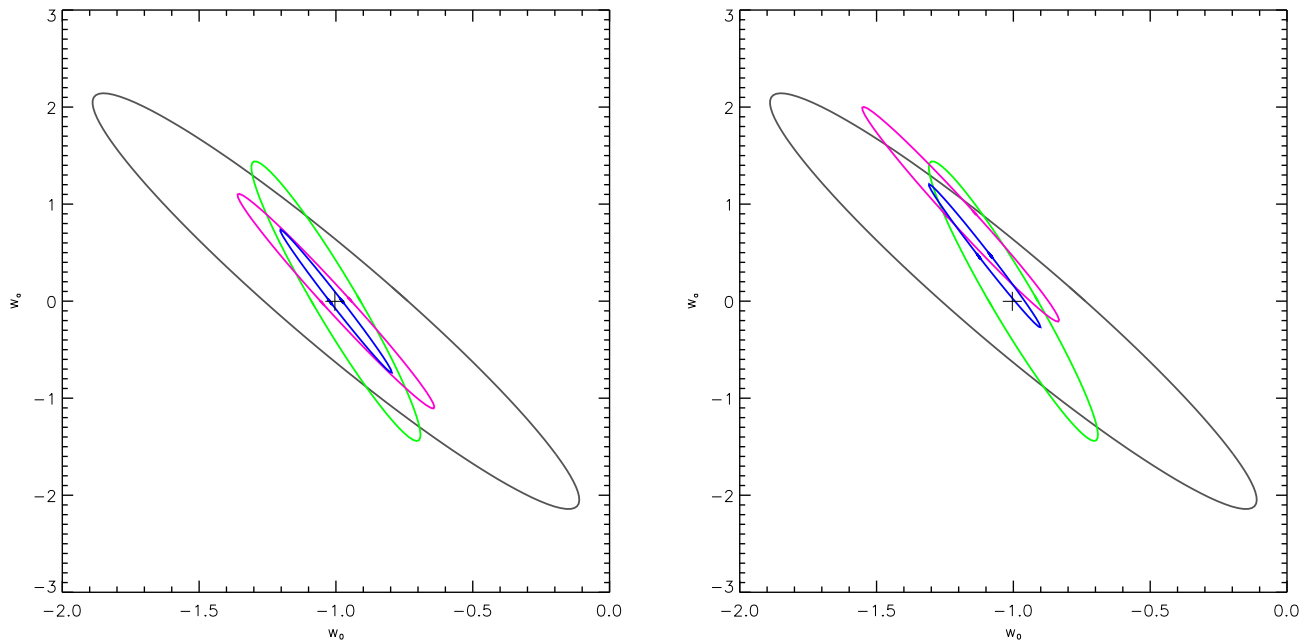
where $F_{\alpha\beta}$ includes any priors.

FIG. 2: LEFT: Forecasted 1σ constraints on dark energy parameters from the DES probes, including only statistical errors and assuming ΛCDM as the true model. From the largest to the smallest ellipse, the probes considered are baryon acoustic oscillations (black), supernovae (green), cluster counts (magenta), and weak lensing (blue). Each constraint is combined with a prior expected from Planck CMB measurements; additionally, the supernovae constraint includes an 8% prior on H_0.

RIGHT: Same as LEFT but now the true model is assumed to be our toy modified gravity model with $\gamma = 0.68$. Shown are the forecasted constraints when we incorrectly attempt to fit a GR+dark energy model to the data. The center of the weak lensing ellipse has moved to (w_0, w_a)=(-1.1, 0.47) while the cluster counts ellipse has moved to (-1.19, 0.90). The probes are *seemingly* consistent, but we discuss the problems with this interpretation in Section III B.

Our first task then is to determine the expected values of the measurements for the four probes in the assumed modified gravity model and compare those to the predictions in standard GR+Dark energy. We consider a set of 8 standard cosmological parameters with fiducial values $\{w_0, w_a, \Omega_{DE}, \Omega_k, h, \Omega_b, n_s, \sigma_8\} = \{-1, 0, 0.73, 0, 0.72, 0.046, 1, 0.8\}$ where Ω_k is the curvature density, h is the Hubble constant in units of 100 km/s/Mpc, Ω_b is the baryon density, n_s is the slope of the primordial spectrum, and σ_8 normalizes the matter power spectrum at $z = 0$. For each probe, we then compute the constraints including projected priors from the Planck satellite [see e.g. 16]. We include only statistical errors in the projections for each experiment, therefore our parameter constraints will be optimistic but sufficient for our goal, which is to compare methods of testing GR.

For two probes, supernovae and BAO, the answer is simple: these probes are sensitive only to background geometry which is assumed identical in our MG and GR models, so the predictions for the distance moduli (from supernovae) and correlation function peak (due to BAO) are identical to standard GR and $\Delta P = 0$. The projected contours therefore are centered on the point in parameter space corresponding to the fiducial values. The only work that needs to be done is to determine the Fisher matrix which delineates the allowed region. This has been done before; here we simply reproduce these results, shown projected onto the (w_0, w_a) plane in Figure 2. The CMB is mostly insensitive to our choice of MG since γ only determines structure growth in the late Universe. The CMB power spectrum is in fact affected by gravity modifications via the late Integrated Sachs-Wolfe effect [17, 18] and gravitational lensing, but we ignore these effects, which should only reduce our sensitivity to MG. Our Planck prior is therefore unchanged between the GR and MG cases. Only the weak lensing and cluster predictions are significantly changed when comparing GR to our toy MG model. Details on these calculations and Fisher matrix calculations for all probes are provided in the appendix.

B. Initial Results and Shortcomings

The right side of Figure 2 suggests that if the underlying model were modified gravity with $\gamma = 0.68$, then the allowed "1σ" regions from the 4 probes of DES are expected to overlap in the (w_0, w_a) plane. They would show some disagreement but certainly not enough to reject the notion that dark energy and GR are the correct description of nature. This is a somewhat disappointing conclusion as the two models make very different predictions, so it behooves us to re-examine the proposed methodology. Perhaps the models can be distinguished with a more powerful statistical probe.

The ambiguous conclusion stems from the approach of plotting the constraints from the four probes in the (w_0, w_a) plane and seeing if they overlap. There are three shortcomings of this approach. First, the approach is not quantitative; it is visually appealing (do the contours overlap?) but does not yield a statistical conclusion. Second, each of the 4 constraints are obtained with the Planck prior added in. This prior therefore is used multiple times. To the extent that one can obtain quantitative conclusions from observing the (non-)overlap of the contours, those conclusions will be incorrect due to the redundant information. Finally, the allowed regions are actually 8 dimensional. It is very possible that the allowed regions do **not** overlap in this 8D space but do in the projections onto the 2D subspace. Ishak et al. [11] did find significant disagreement in the 2D subspace; however, compared to this work, their test-case considered more futuristic experiments (space-based supernovae and weak lensing) and a MG model more dissimilar to GR (DGP versus our simple tweak of the growth index). They also exploited the ISW effect on CMB anisotropy, which we have ignored. Furthermore, we have included a recent prescription by Hu & Sawicki [4], which forces the matter power spectrum in MG to agree with GR on non-linear scales (see (B13) and preceding text); failure to account for this limiting behavior causes an over-estimation of the difference between the weak lensing observables in GR and MG. Each of these differences should lead to a less optimistic result in our case.

In the next section, we introduce a new approach for comparing the 4 DES probes and show that this approach is much more successful at distinguishing our toy MG model from the canonical GR model.

IV. MULTI-DIMENSIONAL CONSISTENCY TEST

A. Searching for Parameter Tension

Consider a simple example with one free parameter λ and two probes. To determine if the probes are consistent we form

$$\chi^2(\lambda) = \sum_{i=1}^{2} \left(\lambda - \lambda^{(i)}\right) \frac{1}{[\sigma^{(i)}]^2} \left(\lambda - \lambda^{(i)}\right) \tag{9}$$

where $\lambda^{(i)}$ is the best fit value of the parameter from an analysis of probe i, and $\sigma^{(i)}$ is the error on probe i. A simple way to see if these two probes are consistent is to minimize the χ^2 with respect to λ. The value of λ at this minimum is then the best fit value, and the value of χ^2 there quantifies the goodness of fit in the standard way. There are two probes and one parameter corresponding to one degree of freedom, $\nu = 1$. So a $\chi^2_{\min} = 10$, say, means that the fit is quite bad because the expectation value is $\langle \chi^2_{\min} \rangle = \nu$ for ν degrees of freedom. The difference, $\Delta\chi^2 = 9$, tells us that the probes are inconsistent with 99.7% confidence.

How can we forecast the expected tension between the 2 probes if the assumed model is incorrect? In that case, it will not necessarily be true that $\langle \lambda^{(1)} \rangle = \langle \lambda^{(2)} \rangle$. Then the usual result, $\langle \chi^2_{\min} \rangle = \nu$, will not hold: the expected value of $\chi^2_{\min}$ will be larger, corresponding to a worse fit: $\langle \chi^2_{\min} \rangle = \nu + B$ with $B > 0$. A signature then of an incorrect assumed model is a value of B large relative to ν. I.e., the large $\chi^2_{\min}$ means that the probes are inconsistent in their measurements of λ. For example, if $B = 9$ and $\nu = 1$, we would conclude that the 2 probes are inconsistent with 99.7% confidence. Thus, if we wish to forecast the tension among several probes, the tension parameter B is an appropriate quantity to calculate.

Let us now generalize this 1-parameter/2-probe example to a situation where there are M parameters and N probes. Specifically, we are considering 8 cosmological parameters and 5 probes – the four of DES plus one from Planck. Suppose that probe i returns a best fit set of parameters $\lambda_\alpha{}^{(i)}$ with a covariance matrix $C^{(i)}_{\alpha\beta}$, and for now we assume that each of these is invertible so none of the probes is plagued with any degeneracy in parameter space (we will relax this assumption later). Let λ_α be a random point in cosmological parameter space. Then, a simple statistic [14] which speaks to the agreement of the probes is

$$\chi^2(\lambda_\alpha) = \sum_i \sum_{\alpha\beta} (\lambda_\alpha - \lambda_\alpha{}^{(i)}) \left[C^{(i)}\right]^{-1}_{\alpha\beta} (\lambda_\beta - \lambda_\beta{}^{(i)}). \tag{10}$$

If the likelihood from each individual probe is Gaussian in parameter space and if the assumed model is correct, then the measured value of this χ^2 will be drawn from a distribution with mean and variance equal to $(N-1)M$. Excessively large values of χ^2 would falsify the underlying model.

B. Predicting Parameter Tension

We are now in a position to project how powerful a given set of probes will be when it comes to falsifying a model. Namely, we can compute the expectation value of χ^2 as defined in Eq. (10) if the true model were MG and determine by how much it exceeds $(N-1)M$. First, we minimize χ^2 and compute the value of λ at its minimum:

$$\lambda_\alpha^{\min} = \sum_{\beta,\gamma} \left[\sum_j (C^{(j)})^{-1} \right]_{\alpha\beta}^{-1} \sum_i (C^{(i)})^{-1}{}_{\beta\gamma} \lambda_\gamma{}^{(i)}. \tag{11}$$

We want to insert this into Eq. (10) and take the expectation value. To do this, we will set $\langle (C^{(i)})^{-1} \rangle = F^{(i)}$, the Fisher matrix for probe i, and

$$\left\langle \lambda_\alpha{}^{(i)} \lambda_\beta{}^{(j)} \right\rangle = \overline{\lambda}_\alpha{}^{(i)} \overline{\lambda}_\beta{}^{(j)} + \delta_{ij} (F^{(i)})_{\alpha\beta}^{-1} \tag{12}$$

where $\overline{\lambda}_\alpha{}^{(i)} \equiv \left\langle \lambda_\alpha{}^{(i)} \right\rangle$ is the expected outcome of the ith experiment. This assumes that the errors on the various probes are uncorrelated with each other, which is not strictly true since, for example, we expect a stronger weak lensing signal when we see more clusters, but it is a reasonable approximation [19, 20]. Using these leads to

$$\left\langle \chi_{\min}^2 \right\rangle = (N-1)M + \sum_i \sum_{\alpha\beta} y_\alpha{}^{(i)} y_\beta{}^{(i)} (F^{(i)})_{\alpha\beta}^{-1} - \sum_{\alpha\beta} Y_\alpha Y_\beta (G^{-1})_{\alpha\beta} . \tag{13}$$

where

$$y_\alpha{}^{(i)} \equiv \sum_\beta F_{\alpha\beta}{}^{(i)} \overline{\lambda}_\beta{}^{(i)} \tag{14}$$

$$Y_\alpha \equiv \sum_i y_\alpha{}^{(i)} \tag{15}$$

$$G_{\alpha\beta} \equiv \sum_i F_{\alpha\beta}{}^{(i)}. \tag{16}$$

If all the probes are expected to return the same parameter (the assumed model is correct), then the two last terms on the right in Eq. (13) cancel, and $\left\langle \chi_{\min}^2 \right\rangle = (N-1)M$. It makes sense therefore to subtract off a fiducial parameter set:

$$\Delta\lambda_\alpha{}^{(i)} \equiv \lambda_\alpha{}^{(i)} - \lambda_\alpha^{\text{fid}} \tag{17}$$

so that

$$\Delta y_\alpha{}^{(i)} \equiv \sum_\beta F_{\alpha\beta}{}^{(i)} \Delta\overline{\lambda}_\beta{}^{(i)} \tag{18}$$

$$\Delta Y_\alpha \equiv \sum_i \Delta y_\alpha{}^{(i)}. \tag{19}$$

Then, the expected value of χ^2 at the minimum depends only on these differences. In particular,

$$\left\langle \chi_{\min}^2 \right\rangle = (N-1)M + B \tag{20}$$

with

$$B \equiv \sum_i \sum_{\alpha\beta} \Delta y_\alpha{}^{(i)} \Delta y_\beta{}^{(i)} (F^{(i)})_{\alpha\beta}^{-1} - \sum_{\alpha\beta} \Delta Y_\alpha \Delta Y_\beta (G^{-1})_{\alpha\beta}. \tag{21}$$

Our desired result – the projection for the excess χ^2 due to inconsistency in the probes – has been reduced to a calculation of B. The ingredients of this calculation, $\Delta y^{(i)}$, are directly related to the errors in our theoretical predictions. By (8) and (18), we have

$$\Delta y_\alpha{}^{(i)} = \sum_{\beta\gamma} (\text{Cov}^{-1})^{(i)}_{\beta\gamma} \frac{\partial P_\gamma{}^{(i)}}{\partial \lambda_\alpha} \Delta P_\beta{}^{(i)} \tag{22}$$

so that $\Delta \overline{y}^{(i)} = \Delta \overline{Y} = 0$ when we fit the correct model (or even an incorrect model as long as it does not produce tension in the probes) to the data.

Using the χ^2 probability distribution function for ν degrees of freedom, we can use B to quantify the implied goodness-of-fit: we compute the probability of finding a worse $\chi^2_{\min}$ than its expected value $P(\chi^2_{\min} > \langle \chi^2_{\min} \rangle ; \nu)$ with $\nu = MN - M$ and $\langle \chi^2_{\min} \rangle = \nu + B$. As an example, with $N = 5$ and $M = 8$, we have $\nu = 32$. If we have chosen an incorrect model so that $B = 14.2$, then $P(\chi^2_{\min} > \langle \chi^2_{\min} \rangle ; \nu) = 0.046$, meaning we predict that the constraints from the 5 probes will be inconsistent at the 95.4% or "2σ" level.

C. Parameter Degeneracies

In practice, Fisher matrices can be singular due to an experiment's insensitivity to a parameter or combination of parameters (parameter degeneracies). For example, because SN and BAO cannot measure non-geometric parameters $(\Omega_b, n_s, \sigma_8)$, the corresponding rows/columns of the SN and BAO Fisher matrices are exactly zero (and the covariance matrix contains some corresponding infinite values). Also, the CMB cannot jointly constrain $(w_0, w_a, \Omega_{de}, \Omega_k)$ since these parameters essentially affect only one observable – the position of the 1st CMB peak – although there is some extra sensitivity to these parameters via the late ISW effect. Such insensitivity yields a singular and hence non-invertible Fisher matrix, or if the matrix is effectively singular (ill-conditioned), then its inverse formally exists but will be numerically unstable. These degeneracies represent more than a computational issue: they correspond to error ellipsoids which are infinite in some directions in parameter space, and hence parameter constraints cannot be inconsistent in those directions. E.g. we clearly do not expect the SN constraint to be inconsistent with any of the other probes in the σ_8 direction. Subsequently, it is not sensible to count σ_8 as a degree of freedom of the SN error ellipsoid.

To handle degenerate Fisher matrices, we "clean" them using singular value decomposition. For each Fisher matrix, we find a unitary matrix U such that

$$F = U^T \Lambda U \tag{23}$$

where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of F. We replace the smallest elements of Λ with zeros (some are already zero) and recompute F using the above equation; this procedure negligibly changes the elements of F provided we only remove eigenvalues which are much smaller than the largest eigenvalue. We compute the inverse of F via

$$F^{-1} \equiv U^T \Lambda^{-1} U \tag{24}$$

where Λ^{-1} is diagonal and we have set

$$(\Lambda^{-1})_{\alpha\alpha} = \begin{cases} 1/\Lambda_{\alpha\alpha} & \text{for } \Lambda_{\alpha\alpha} \neq 0 \\ 0 & \text{for } \Lambda_{\alpha\alpha} = 0 \end{cases} . \tag{25}$$

Let $S^{(i)}$ be the number of eigenvalues of $F^{(i)}$ which are zero (the nullity of the matrix). Starting again from (13), one now finds that

$$\langle \chi^2_{\min} \rangle = (N-1)M - \sum_i S^{(i)} + B \tag{26}$$

with B unchanged from (21). Effectively, the number of degrees of freedom ν has been reduced by the total number of parameters that each probe cannot constrain. When very small eigenvalues of each $F^{(i)}$ are set to zero, we find that B changes negligibly. This means that no significant tension is expected among the probes in these highly degenerate directions in parameter space, so we are indeed wrong to count them among the degrees of freedom. By setting the eigenvalues to zero, we force ourselves to evaluate tension only among the degrees of freedom where we expect tension. We excise the 3 most serious degeneracies from both the CMB and WL Fisher matrices, and we excise the 4 most serious degeneracies from the SN, BAO and CL Fisher matrices. Thus, when all of these probes are combined, the total number of degrees of freedom is $\nu = 5 \times 8 - 8 - 18 = 14$.

D. Results

We suppose that we will mistakenly fit an 8-parameter ΛCDM model to data in a Universe described by our toy modified gravity model. The MG model has a scale-independent linear growth history given by $\gamma = 0.68$, and all ΛCDM parameters of the model are otherwise identical. In Table I, we show our estimates of the subsequent tension among the DES probes and Planck, assuming only statistical errors in the DES probes. As described in §IV C, when a probe's individual Fisher matrix contains a significant parameter degeneracy, we do not count the corresponding direction in parameter space as a source of tension or a degree of freedom.

WL	CL	SN	BAO	CMB	ν	B	$P(\chi^2_{\min} > \nu + B; \nu)$
	$\checkmark$		$\checkmark$	$\checkmark$	5	2.06	0.2164
	$\checkmark$	$\checkmark$		$\checkmark$	5	1.67	0.2466
	$\checkmark$	$\checkmark$	$\checkmark$		4	0.02	0.4030
	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	9	3.02	0.2121
$\checkmark$			$\checkmark$	$\checkmark$	6	2.08	0.2326
$\checkmark$		$\checkmark$		$\checkmark$	6	1.96	0.2414
$\checkmark$		$\checkmark$	$\checkmark$		5	0.75	0.3313
$\checkmark$		$\checkmark$	$\checkmark$	$\checkmark$	10	2.21	0.2715
$\checkmark$	$\checkmark$			$\checkmark$	6	6.71	0.0478
$\checkmark$	$\checkmark$		$\checkmark$		5	0.61	0.3462
$\checkmark$	$\checkmark$		$\checkmark$	$\checkmark$	10	8.23	0.0512
$\checkmark$	$\checkmark$	$\checkmark$			5	2.29	0.2003
$\checkmark$	$\checkmark$	$\checkmark$		$\checkmark$	10	9.22	0.0376
$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$		9	2.76	0.2271
$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	14	15.58	0.0087

TABLE I: Expected tension between dark energy probes, assuming that our toy MG model is correct but we fit our GR+dark energy model to the data. We try different probe combinations – check-marks denote included probes. The total degrees of freedom, ν, counts the number of non-degenerate parameter combinations measured by each probe, minus the one 8-parameter set used to minimize χ^2. The expected parameter tension, B, is calculated in (21) and vanishes when all probes are expected to have the same best-fit parameter set. We expect that $\left\langle \chi^2_{\min} \right\rangle = \nu + B$, and this too-high value of $\chi^2_{\min}$ will cause us to interpret the constraints to be inconsistent (non-overlapping) when B is large. $P(\chi^2_{\min} > \nu + B; \nu)$ is the expected goodness-of-fit of the probes, i.e. the probability that our probes, in a truly GR universe, would yield parameter constraints with more tension than the tension we predict due to fitting an incorrect model.

The last row of Table I shows that when all DES probes are combined with Planck, we expect their parameter constraints to have a goodness-of-fit of 0.0087. That is, the overlap of the error contours in the 8D parameter space will be so poor, we would interpret them to be inconsistent at about the 99.1% level. Using only WL, CL and CMB, we still expect these probes to disagree at about the 95% level. We predict no significant tension among the probes when either WL or CL is excluded. Our result is an improvement over the method of looking for overlap in the $w_0 - w_a$ plane (see Figure 2) since it leverages tension among all parameters, and does not use the Planck prior multiple times. We remind the reader that our forecasts assume only statistical errors in the DES probes. Accounting for systematic errors would degrade parameter constraints, leading to more overlap, i.e. less tension. At the same time, we have chosen a modest toy MG model which differs from ΛCDM only in the linear growth of perturbations. Other MG models could easily produce more tension.

Ours is an encouraging result in the search for modified gravity: it means that an inconsistency among the DES probes is, in principle, a useful diagnostic for identifying a wrong cosmological model. Of course, if such an inconsistency were found, each probe's working group would revisit its pipeline, looking for systematic errors. If tension persisted, then the evidence for inconsistency would be strengthened.

V. CONCLUSIONS

The consistency of different dark energy probes promises to be a powerful tool in the quest to distinguish dark energy from modified gravity. Here we have illustrated that, using the Multi-dimensional Consistency Test (MCT), future probes from the Dark Energy Survey will be able to rule out standard (GR+dark energy) if the true gravity

model is only a modest modification of GR. Carrying out the MCT once the data are in reduces to computing the χ^2 of Eq. (10) while properly accounting for degeneracies as described in §IV C. Although we have not explored this in detail here, projections of the MCT might make a useful metric for future surveys when trying to understand their constraining power towards modified gravity models, complementary to the figures of merit for dark energy [21].

[1] E. V. Linder, Phys. Rev. **D72**, 043529 (2005), [astro-ph/0507263].
[2] E. V. Linder and R. N. Cahn, Astropart. Phys. **28**, 481 (2007), [astro-ph/0701317].
[3] P. Zhang, M. Liguori, R. Bean and S. Dodelson, Phys. Rev. Lett. **99**, 141302 (2007), [0704.1932].
[4] W. Hu and I. Sawicki, Phys. Rev. **D76**, 104043 (2007), [0708.1190].
[5] G.-B. Zhao, L. Pogosian, A. Silvestri and J. Zylberberg, Phys. Rev. Lett. **103**, 241301 (2009), [0905.1326].
[6] D. Rapetti, S. W. Allen, A. Mantz and H. Ebeling, ArXiv e-prints (2009), [0911.1787].
[7] S. F. Daniel *et al.*, ArXiv e-prints (2010), [1002.1962].
[8] R. Reyes *et al.*, Nature**464**, 256 (2010), [1003.2185].
[9] L. Lombriser, A. Slosar, U. Seljak and W. Hu, 1003.3009.
[10] J. Annis *et al.*, astro-ph/0510195.
[11] M. Ishak, A. Upadhye and D. N. Spergel, Phys. Rev. D**74**, 043513 (2006), [arXiv:astro-ph/0507184].
[12] J. Zhang, L. Hui and A. Stebbins, ApJ**635**, 806 (2005), [arXiv:astro-ph/0312348].
[13] G. Dvali, G. Gabadadze and M. Porrati, Physics Letters B **485**, 208 (2000), [arXiv:hep-th/0005016].
[14] G. Bernstein and D. Huterer, MNRAS**401**, 1399 (2010), [0902.2782].
[15] L. Knox, R. Scoccimarro and S. Dodelson, Phys. Rev. Lett. **81**, 2004 (1998), [astro-ph/9805012].
[16] J. Dick, DETFast (2006), [http://www.physics.ucdavis.edu/DETFast/].
[17] A. Lue, R. Scoccimarro and G. Starkman, Phys. Rev. D**69**, 044005 (2004), [arXiv:astro-ph/0307034].
[18] F. Schmidt, M. Liguori and S. Dodelson, Phys. Rev. D**76**, 083518 (2007), [0706.1775].
[19] M. Takada and S. Bridle, New Journal of Physics **9**, 446 (2007), [0705.0163].
[20] C. Shapiro and S. Dodelson, Phys. Rev. D**76**, 083515 (2007), [0706.2395].
[21] A. J. Albrecht *et al.*, astro-ph/0609591.
[22] A. Refregier, A. Amara, T. Kitching and A. Rassat, 0810.1285.
[23] H. A. Feldman, N. Kaiser and J. A. Peacock, ApJ**426**, 23 (1994), [arXiv:astro-ph/9304022].
[24] M. Tegmark, Physical Review Letters **79**, 3806 (1997), [arXiv:astro-ph/9706198].
[25] H. Seo and D. J. Eisenstein, ApJ**598**, 720 (2003), [arXiv:astro-ph/0307460].
[26] H. Seo and D. J. Eisenstein, ApJ**665**, 14 (2007), [arXiv:astro-ph/0701079].
[27] B. A. Bassett, Y. Fantaye, R. Hlozek and J. Kotze, ArXiv e-prints (2009), [0906.0974].
[28] A. Jenkins *et al.*, Mon. Not. Roy. Astron. Soc. **321**, 372 (2001), [astro-ph/0005260].
[29] D. J. Eisenstein and W. Hu, ApJ**511**, 5 (1999), [arXiv:astro-ph/9710252].
[30] R. A. Battye and J. Weller, Phys. Rev. **D68**, 083506 (2003), [astro-ph/0305568].
[31] R. K. Sheth and G. Tormen, Mon. Not. Roy. Astron. Soc. **329**, 61 (2002), [astro-ph/0105113].
[32] B. M. Schaefer and K. Koyama, Mon. Not. Roy. Astron. Soc. **385**, 411 (2008), [0711.3129].
[33] T. Kobayashi and H. Tashiro, MNRAS**398**, 477 (2009), [0903.3738].
[34] F. Schmidt, A. Vikhlinin and W. Hu, Phys. Rev. **D80**, 083505 (2009), [0908.2457].
[35] A. Lue, R. Scoccimarro and G. D. Starkman, Phys. Rev. **D69**, 124015 (2004), [astro-ph/0401515].
[36] K. C. Chan and R. Scoccimarro, Phys. Rev. **D80**, 104005 (2009), [0906.4548].
[37] F. Schmidt, Phys. Rev. **D80**, 043001 (2009), [0905.0858].
[38] R. D. Blandford, A. B. Saust, T. G. Brainerd and J. V. Villumsen, MNRAS**251**, 600 (1991).
[39] N. Kaiser, ApJ**388**, 272 (1992).
[40] S. Dodelson, *Modern cosmology* (Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2, 2003, XIII + 440 p., 2003).
[41] M. Bartelmann and P. Schneider, Phys. Rep.**340**, 291 (2001), [arXiv:astro-ph/9912508].
[42] Z. Ma, W. Hu and D. Huterer, ApJ**636**, 21 (2006), [arXiv:astro-ph/0506614].
[43] R. E. Smith *et al.*, MNRAS**341**, 1311 (2003), [arXiv:astro-ph/0207664].

[44] C. M. Will, Living Reviews in Relativity **9**, 3 (2006), [arXiv:gr-qc/0510072].
[45] K. Koyama, A. Taruya and T. Hiramatsu, Phys. Rev. D**79**, 123512 (2009), [0902.0618].
[46] H. Oyaizu, M. Lima and W. Hu, Phys. Rev. D**78**, 123524 (2008), [0807.2462].
[47] E. Beynon, D. J. Bacon and K. Koyama, MNRAS**403**, 353 (2010), [0910.1480].
[48] D. H. Rudd, A. R. Zentner and A. V. Kravtsov, ApJ**672**, 19 (2008), [arxiv:astro-ph/0703741].
[49] A. R. Zentner, D. H. Rudd and W. Hu, Phys. Rev. D**77**, 043507 (2008), [0709.4029].
[50] A. P. Hearin and A. R. Zentner, Journal of Cosmology and Astro-Particle Physics **4**, 32 (2009), [0904.3334].
[51] Y. P. Jing, P. Zhang, W. P. Lin, L. Gao and V. Springel, ApJ**640**, L119 (2006), [arXiv:astro-ph/0512426].
[52] M. Takada and B. Jain, MNRAS**348**, 897 (2004), [arXiv:astro-ph/0310125].

Appendix A: Fisher Matrices for Unmodified Probes

1. Supernova

Given the number of supernovae N and a projected uncertainty on the magnitude of each supernova, we can form the Fisher matrix,

$$F_{\alpha\beta} = \sum_{i=1}^{N} \frac{\partial \mu_i}{\partial \lambda_\alpha} \frac{\partial \mu_i}{\partial \lambda_\beta} \frac{1}{\sigma_i^2} \tag{A1}$$

where μ_i is the distance modulus for a supernova at redshift z_i and p_α are the parameters of interest. If we bin the supernovae in redshift bins of width $\Delta z = 0.1$, each with an expected N_a such that

$$\sum_a N_a = N, \tag{A2}$$

then the Fisher matrix becomes:

$$F_{\alpha\beta} = \sum_{a=1}^{13} N_a \frac{\partial \mu_a}{\partial \lambda_\alpha} \frac{\partial \mu_a}{\partial \lambda_\beta} \frac{1}{\sigma_a^2} \tag{A3}$$

where the first bin is at $z = 0.05$ and the last at $z = 1.25$. We set $\sigma = 0.12$ in low redshift ($z < 0.8$) bins and $\sigma = 0.24$ in the higher bins. Following the DES team, we also introduce a systematic floor, so that the error in any bin is $\sigma_{\rm bin}^2 = \sigma_a^2/N_a + \sigma_{\rm floor}^2$ with $\sigma_{\rm floor} = 0.02$.

The derivative of the distance modulus to bin a is

$$\frac{\partial \mu_a}{\partial \lambda_\alpha} = \frac{5}{\ln(10)} \frac{\partial \ln d_L(z_a)}{\partial \lambda_\alpha}. \tag{A4}$$

For small values of the curvature, the luminosity distance can be approximated as

$$d_L(z) \simeq (1+z) \left[\chi(z) + \frac{\Omega_k H_0^2 \chi^3(z)}{6} \right] \tag{A5}$$

where

$$\chi(z) = 3000\,{\rm Mpc} \int_0^{z_a} \frac{dz}{E(z)} \tag{A6}$$

with

$$
\begin{aligned}
E(z) &\equiv h \left[\Omega_m (1+z)^3 + \Omega_X \exp \left\{ 3 \int_0^z \frac{dz'}{1+z'} (1 + w_0 + w_a \frac{z'}{1+z'}) \right\} + \Omega_k (1+z)^2 \right]^{1/2} \\
&= h \left[\Omega_m (1+z)^3 + \Omega_X \exp \left\{ 3(1 + w_0 + w_a) \ln(1+z) - 3 w_a z/(1+z) \right\} + \Omega_k (1+z)^2 \right]^{1/2}.
\end{aligned}
\tag{A7}
$$

We do not observe distance moduli directly but rather the peak apparent magnitude of the supernovae in a particular band, $m(z) = M_0 + \mu(z)$, where M_0 is a supernova's absolute magnitude. To account for uncertainty in the absolute magnitudes of type 1A supernovae, we include M_0 as a nuisance parameter in the SN Fisher matrix; M_0 is completely degenerate with the Hubble constant h. The SN matrix can be combined with other matrices by padding the others with a row and column of zeros, adding all matrices together, and then marginalizing over M_0.

2. Baryon Acoustic Oscillations

Galaxy surveys can measure a power spectrum of the galaxy distribution $P(k)$ in a thin redshift bin around $k = k_\mathrm{n}$ with an expected covariance matrix

$$C_\mathrm{mn} = 2 \frac{P(k_\mathrm{m})P(k_\mathrm{n})}{V_\mathrm{n} V_\mathrm{eff}} \delta_\mathrm{m,n}, \tag{A8}$$

where

$$V_\mathrm{eff} = \int \left[\frac{\bar{n}(\mathbf{r})P(k)}{1 + \bar{n}(\mathbf{r})P(k)} \right]^2 d^3\mathbf{r} = \left[\frac{nP(k)R(\mu)}{nP(k)R(\mu) + 1} \right] V_0 \tag{A9}$$

is an effective volume utilized for the measurement, V_0 is the total volume of the survey, $R = (1 + \beta\mu^2)^2$ describes the effect of linear redshift space distortions, and $V_\mathrm{n} = \frac{d^3\mathbf{k}_\mathrm{n}}{(2\pi)^3}$ is the volume of a shell in the Fourier space [for details see, 23]. Equation (A9) holds if the number density of galaxies is constant within the volume.

If we have a model of how the power spectrum depends on the cosmological parameters of interest λ_i and if we assume that the errors on the power spectrum measurements are uncorrelated random Gaussian variables, then by propagating correlated errors given by the covariance matrix in Eq. (A8) we can get a Fisher matrix of cosmological parameters

$$F_\mathrm{ij} = \frac{1}{2} \int_{k_\mathrm{min}}^{k_\mathrm{max}} \frac{d^3\mathbf{k}}{(2\pi)^2} \frac{\partial P(k)}{\partial \lambda_i} \frac{\partial P(k)}{\partial \lambda_i} V_\mathrm{eff}, \tag{A10}$$

[for details see 24]. We want to forecast cosmological constraints from a measured position of the baryon acoustic peak in the three dimensional power spectrum. We will follow the formalism of Seo and Eisenstein [25, 26].

The baryonic part of the power spectrum can be modeled as

$$P_\mathrm{b}(k) \sim \frac{\sin ks_0}{ks_0} \exp\left[-\left(\frac{k}{k_\mathrm{silk}} \right)^{1.4} \right], \tag{A11}$$

where s_0 is the sound horizon at recombination and k_silk is the Silk damping scale. The Silk damping scale can be accurately fit by

$$k_\mathrm{silk} = 1.6(\Omega_\mathrm{b}h^2)^{0.52}(\Omega_\mathrm{m})^{0.73}[1 + (10.4\Omega_\mathrm{m}h^2)^{-0.095}]h^{-1}\,(h\,\mathrm{Mpc}^{-1})\,. \tag{A12}$$

As in [26] we will multiply Eq. (A11) by additional Gaussian functions to account for the erasure of information due to nonlinear evolution and photometric redshifts. The final P_b is given by

$$\begin{aligned} P_\mathrm{b}(k) &= \sqrt{8\pi^2} A_0 P_{0.2} \frac{\sin\left[\left(k_{||}^2 s_{||}^2 + k_\perp^2 s_\perp^2 \right)^{1/2} \right]}{\left(k_{||}^2 s_{||}^2 + k_\perp^2 s_\perp^2 \right)^{1/2}} \exp\left[-\left(\frac{k}{k_\mathrm{silk}} \right)^{1.4} \right] \\ &\times \exp\left[-k^2(1 - \mu^2)\Sigma_\perp^2 - k^2\mu^2\Sigma_{||}^2 \right] \exp\left[-k^2\mu^2\Sigma_\mathrm{z}^2 \right], \end{aligned} \tag{A13}$$

where $s_{||}$ and $s_\perp$ are sound horizon scales measured along and across the line of sight, μ is cosine of the along the line of sight, $P_{0.2}$ is galaxy power spectrum at $k = 0.2h\,\mathrm{Mpc}^{-1}$, A_0 is a normalization factor, $\Sigma_{||}$ and $\Sigma_\perp$ model the loss of information due to nonlinear growth and Σ_z models the loss of information in radial direction due to photometric redshifts. As in [26] we will use numerical values $\Sigma_{||} = \Sigma_0 G(1 + f)$ and $\Sigma_\perp = \Sigma_0 G$, where G is the growth function, $f = d\ln G/d\ln a$ and $\Sigma_0 = 11.0\,h^{-1}\,\mathrm{Mpc}$ for the cosmology with $\sigma_8 = 0.8$. When the physical value of the sound horizon is known to high precision from CMB measurements the errors on $s_{||}$ and $s_\perp$ are equivalent to the errors on angular and radial distances D_A and H.

Derivatives of Eq. (A13) with respect to $s_{||}$ and $s_\perp$ are

$$\frac{\partial P_{\rm b}(x)}{\partial \ln s_{||}} = \frac{\partial P_{\rm b}(x)}{\partial \ln x} f_{||}(\mu), \tag{A14}$$

$$\frac{\partial P_{\rm b}(x)}{\partial \ln s_{\perp}} = \frac{\partial P_{\rm b}(x)}{\partial \ln x} f_{\perp}(\mu), \tag{A15}$$

where $x = \left(k_{||}^2 s_{||}^2 + k_{\perp}^2 s_{\perp}^2 \right)^{1/2}$, $f_{||} = \mu^2$ and $f_{\perp}(\mu) = 1 - \mu^2$. For most wavenumbers of interest ks_0 is large and the sinusoidal terms oscillate rapidly. We will use this fact to replace $\cos^2(ks_0)$ term by its rms value of $1/2$ and to drop terms proportional to $\sin^2(ks_0)$ when computing derivatives in Eqs. (A14)–(A15). Using Eq. (A10) with derivatives in Eqs. (A14)–(A15) results in a Fisher matrix on radial and angular distances $H(z)$ and $D_A(z)$

$$
\begin{aligned}
F_{\rm ij} &= V_0 A_0^2 \int_0^1 d\mu f_{\rm i}(\mu) f_{\rm j}(\mu) \\
&\times \int_0^{\infty} k^2 dk \frac{\exp\left[-2(k/k_{\rm silk})^{1.4}\right]}{[P(k)/P_{0.2} + (nP_{0.2}R(\mu))^{-1}]^2} \\
&\times \exp\left[-k^2(1-\mu^2)\Sigma_{\perp}^2 - k^2\mu^2\Sigma_{||}^2\right] \exp\left[-k^2\mu^2\Sigma_{\rm z}^2\right]
\end{aligned} \tag{A16}
$$

The Fisher matrix for $H(z)$ and $D_A(z)$ computed from Eq. (A16) can be easily transformed into a Fisher matrix on cosmological parameters w_0, $w_{\rm a}$, $\Omega_{\rm m}$, Ω_{Λ} and h [see, e.g., 27].

3. CMB

Throughout, we compute Fisher matrices for the parameter set $\{w_0, w_a, \Omega_{\rm DE}, \Omega_k, h, \Omega_b, n_s, \sigma_8\}$. However, our prior from Planck CMB measurements comes in the form $\{w_0, w_a, \Omega_{\rm DE}, \Omega_k, \Omega_m h^2, \Omega_b h^2, n_s, \ln P\}$ where $\ln P$ is the log of the amplitude of primordial fluctuations [16]. We transform the prior matrix to our preferred format using a simple chain rule. Starting from (7),

$$F_{\alpha\beta} = \sum_{ij}(C^{-1})_{ij}\frac{\partial P_i}{\partial \lambda_\alpha}\frac{\partial P_j}{\partial \lambda_\beta} \tag{A17}$$

$$= \sum_{\mu\nu}\sum_{ij}(C^{-1})_{ij}\frac{\partial P_i}{\partial \lambda_\mu'}\frac{\partial P_j}{\partial \lambda_\nu'}\frac{\partial \lambda_\mu'}{\partial \lambda_\alpha}\frac{\partial \lambda_\nu'}{\partial \lambda_\beta} \tag{A18}$$

$$= \sum_{\mu\nu}F_{\mu\nu}'\frac{\partial \lambda_\mu'}{\partial \lambda_\alpha}\frac{\partial \lambda_\nu'}{\partial \lambda_\beta} \tag{A19}$$

$$\tag{A20}$$

Thus we express one set of parameters, λ_α', in terms of the others, λ_β, and compute the necessary matrix of partial derivatives.

Appendix B: Fisher Matrices for Modified Probes

1. Clusters

DES will optically detect thousands of galaxy clusters out to a redshift of 1.6. The clusters will also be detected in microwaves by the South Pole Telescope (SPT) via the Sunyaev-Zeldovich (SZ) effect, allowing tighter constraints on the cluster masses. Binning the detected clusters by redshift, the observables for our cluster experiment will be the total number of clusters in each bin, above a given mass threshold which allows detection by SPT.

To conform with DES predictions, we assume the cluster mass function of Jenkins et al. [28] with the parameters in their equation (B4). Let $n(M, z)$ be the comoving number density of clusters with mass M at redshift z. We take

$$n(M, z) = -\frac{\rho_{c0}}{M}\frac{d \ln \sigma_M}{d \ln M}f(M, z) \tag{B1}$$

$$f(M, z) = 0.316\exp\left(-1|\log\left(\sigma_M(z)^{-1}\right) + 0.67|^{3.82}\right), \tag{B2}$$
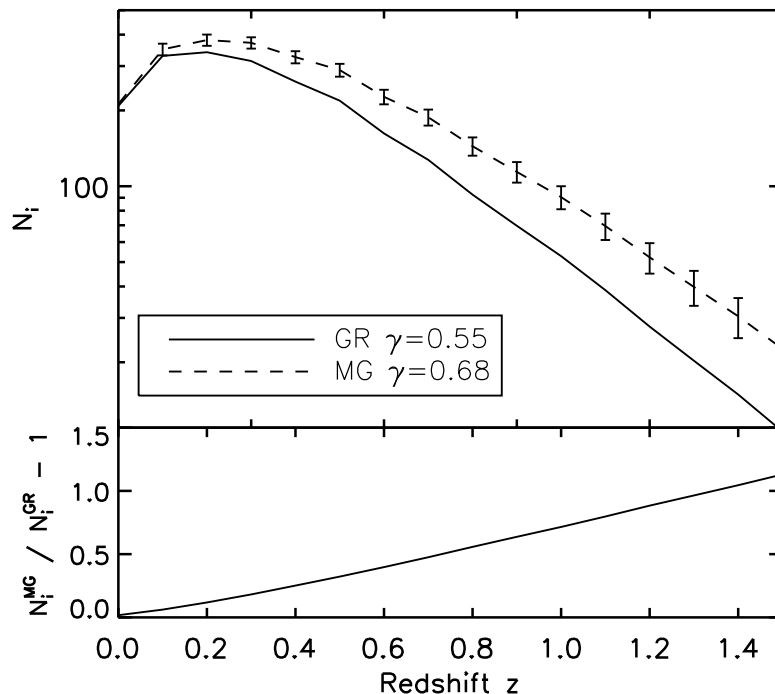
FIG. 3: The SPT+DES number of detectable clusters N_i in redshift bins of width $\Delta z = 0.1$ along with projected error bars in each bin. Solid curve assumes GR $\gamma = 0.55$, and the dashed curve assumes modified gravity with $\gamma = 0.68$. We plot the ratio of the two curves in the lower panel.

Here, ρ_{c0} is the critical density today and σ_M is the RMS of the matter density field, smoothed by a top-hat filter of radius R where $R^3 \equiv 3M/4\pi\rho_{c0}$. The RMS is calculated using linear theory:

$$\sigma_R(z)^2 = \int d\ln k \; \frac{k^3 P_{\mathrm{lin}}(k;z)}{2\pi^2} J_1^2(kR) \; , \tag{B3}$$

where J_1 is the Bessel function of the first kind and $P_{\mathrm{lin}}(k;z)$ is the linear matter power spectrum computed from the fitting formula of Eisenstein and Hu [29]. The minimum mass limit of clusters detectable by SPT, $M_{\mathrm{lim}}(z)$, was calculated in [30]. Let N_i denote the total number of clusters above this mass limit in the ith redshift bin. It is given by

$$N_i = 4\pi f^{\mathrm{sky}} \int_{z_i}^{z_{i+1}} dz \, \frac{\chi(z)^2}{H(z)} \int_{M_{\mathrm{lim}}(z)}^{\infty} dM \, n(M,z) \tag{B4}$$

where χ is comoving distance, z_i denotes the lower edge of the ith bin, and $f^{\mathrm{sky}} = 0.125$ is the sky coverage of the overlapping DES+SPT survey. We show N_i as a function of redshift for $\gamma = 0.55$ and $\gamma = 0.68$ in Figure 3. The clusters are divided into 16 redshift bins of width $\Delta z = 0.1$, assuming perfect measurements of their redshift and mass.

Our toy MG model only differs from ΛCDM via the linear density growth, parametrized by γ in (5). In this case, computing the effect of MG on clusters is easily implemented by computing the linear growth function as a function of z for $\gamma = 0.68$ and then using it to normalize $P_{\mathrm{lin}}(k;z)$ in (B3). A more general MG model could change the dynamics of collapsing halos, such as the halo formation time or the critical overdensity for halo collapse [31]. Studies of DGP [32, 33] and $f(R)$-gravity [34] have shown that these models alter the critical spherical overdensity, δ_c, by only 1-2% relative to ΛCDM. Furthermore, changes in halo formation times are already incorporated into the GR spherical collapse mass function of Sheth and Tormen [31], which has been shown to fit simulations well [e.g. 35–37]. Therefore our assumption, that cluster numbers depend primarily on the linear growth function, is realistic.

Assuming that the error on the number of clusters in the ith redshift bin is dominated by counting error, the covariance between bins is

$$\mathrm{Cov}[N_i, N_j] = \delta_{ij} N_i \; . \tag{B5}$$

The resulting projections in the dark energy parameter plane are shown in Figs. 2 when the underlying model is ΛCDM. Fig. 5 shows the projected constraints if the underlying model is MG with $\gamma = 0.68$. Fig. 5 shows that to get the extra clusters that this MG model would produce, the fitter favors a larger w, thereby introducing more dark energy at early times. This tends to leave σ_8 closer to its present value, so mimics the increase in cluster abundance at high z.

2. Lensing

We begin with a review of relevant weak lensing equations [for a full introduction see e.g. 38–41]. The lensing convergence at a particular sky position, $\kappa(\boldsymbol{\theta})$, can be expressed as the matter density contrast, $\delta(\boldsymbol{x})$, projected over comoving distance, χ, along the line of sight:

$$\kappa_i(\boldsymbol{\theta}) = \int_0^\infty d\chi\, \delta(\boldsymbol{\theta}\chi, \chi)\, W_i(\chi). \tag{B6}$$

The subscript i denotes the redshift bin from which source galaxies have been selected, and $W_i(\chi)$ is the lensing kernel for that bin, defined below. The cosmic convergence power spectra and cross spectra, $C_{l;ij}$, are defined as

$$\langle \tilde{\kappa}_i(\boldsymbol{l})\tilde{\kappa}_j(\boldsymbol{l'}) \rangle \equiv (2\pi)^2 \delta^2(\boldsymbol{l} + \boldsymbol{l'})\, C_{l;ij} \tag{B7}$$

where $\boldsymbol{l}$ is the Fourier conjugate to $\boldsymbol{\theta}$, and we are working in the small angle limit, rather than decomposing the fields into spherical harmonics. Note that δ^2 is a 2-dimensional Dirac delta function and the angle brackets denote an ensemble average.

Measurements of galaxy ellipticities allow us to estimate the cosmic shear power and cross spectra, not the convergence, but these are equal at first order in the gravitational potentials. For N redshift bins, the weak lensing observables are the $N(N+1)/2$ total spectra for a given l. The leading order calculation of $C_{l;ij}$ is

$$C_{l;ij} = \int_0^\infty \frac{d\chi}{d_A(\chi)^2} W_i(\chi) W_j(\chi) P_\delta(k; \chi) \tag{B8}$$

where the function d_A modifies distances in a curved Universe. For $K \equiv (H_0^2 \Omega_k)^{-1}$,

$$d_A(\chi) \equiv \begin{cases} \chi & \text{for } K = 0 \quad \text{(flat)} \\ |K|^{-1/2} \sin(|K|^{1/2}\chi) & \text{for } K > 0 \quad \text{(closed)} \\ |K|^{-1/2} \sinh(|K|^{1/2}\chi) & \text{for } K < 0 \quad \text{(open)} \end{cases} . \tag{B9}$$

$P_\delta(k; \chi)$ is the 3D matter power spectrum for $k = \ell/d_A(\chi)$ at a distance χ, accounting for the growth of structure. Equation (B8) uses the Limber approximation, which assumes that the only matter density modes $\tilde{\delta}(\boldsymbol{k})$ contributing to the lensing signal are those modes with $\boldsymbol{k}$ transverse to the line of sight.

The lensing kernel is given by

$$W_i(\chi) = \frac{W_0}{n^{\text{gal}}_i} \frac{d_A(\chi)}{a(\chi)} \int_\chi^\infty d\chi_s\, p_i(z) \frac{dz}{d\chi_s} \frac{d_A(\chi_s - \chi)}{d_A(\chi_s)} \tag{B10}$$

with $W_0 = \frac{3}{2}\Omega_m H_0^2$. Here, $p_i(z)$ is the true (spectroscopic) distribution of galaxies in the ith redshift bin, and n^{gal}_i is the total projected number density of galaxies in that bin. Binning is done according to the galaxies' photometric redshifts (or "photo-z"s), which we take to be unbiased estimators of the true redshifts with gaussian scatter σ_z. To avoid degrading parameter constraints, it is actually more important to have a small uncertainty in the bias and scatter of the photo-zs rather than a small scatter [42].

The redshift distribution of source galaxies in the survey is taken to be

$$\frac{dn^{\text{gal}}}{dz} \propto z^2 \exp[-(z/z_0)^{1.5}]. \tag{B11}$$

The median redshift of this distribution is $z_{\text{med}} \approx z_0\sqrt{2}$, and the distribution is normalized so that the total projected number density of galaxies is

$$\int_0^\infty dz\, \frac{dn^{\text{gal}}}{dz} = \sum_i n^{\text{gal}}_i \equiv n^{\text{gal}}. \tag{B12}$$

For DES, we take $z_{\mathrm{med}} = 0.68$, $n^{\mathrm{gal}} = 12$, and $\sigma_z = 0.08$. The galaxies are divided into 7 photo-z bins with bin boundaries defined by z=(0, 0.37, 0.51, 0.63, 0.75, 0.89, 1.10, 2.0), which give them approximately equal projected number densities.

Equation (2) describes the relation between the lensing potential $\Phi - \Psi$ and the matter density δ in a modified gravity scenario. Since our modified gravity model has $f = 0$, this relation is unchanged from GR, and the weak lensing power spectrum can be computed from the matter power spectrum via (B8). We need to compute the non-linear matter power spectrum in both the GR and MG cases. To compute the shape of the linear power spectrum in GR, we use the fitting formula of Eisenstein and Hu [29]. The redshift dependence of the linear power is given by the growth function, which can be easily adapted to incorporate the growth index, γ, in the MG case. We compute the nonlinear matter power spectrum in GR via the `Halofit` fitting formula of Smith et al. [43] [1], which takes the linear power spectrum as its input. Solar system observations indicate that that GR is valid on small non-linear scales, therefore a viable MG theory must agree with GR in that limit [44, and references therein]. `Halofit` does not impose such behavior on the non-linear power spectrum, therefore it must be adapted in the MG case. One way to accomplish this is to compute the linear power spectra in the GR and MG cases and then apply `Halofit` to both; a realistic non-linear spectrum for MG can then be constructed by interpolating between the two results so that we enforce the GR prediction for density modes that have gone very non-linear. Such an interpolation was proposed by [4]:

$$P_{\mathrm{MG}}(k,z) = \frac{P_{\mathrm{MGX}}(k,z) + c_{\mathrm{nl}}(z)\Sigma^2(k,z)P_{\mathrm{GR}}(k,z)}{1 + c_{\mathrm{nl}}(z)\Sigma^2(k,z)}, \tag{B13}$$

where P_{GR} is the non-linear power in GR and P_{MGX} comes from the naive application of a non-linear fitting formula to the linear MG power spectrum. The interpolation is moderated by two functions obtained from perturbation theory [45]:

$$\Sigma^2(k,z) = \left(\frac{k^3}{2\pi^2}P_{\mathrm{lin}}(k,z)\right)^{\alpha_1}, \quad c_{\mathrm{nl}}(z) = A(1+z)^{\alpha_2} \tag{B14}$$

Here, $P_{\mathrm{lin}}(k,z)$ is the linear power spectrum in MG and the parameters α_1, α_2, and A are calibrated by N-body simulations [37, 46]. For our working model, we take $A = 0.3$, $\alpha_1 = 1$ and $\alpha_2 = 0.16$, which are their values in DGP, ignoring their weak dependence on Ω_M and σ_8. We remind the reader that we are not using DGP specifically, particularly since we are keeping the background expansion of ΛCDM. The ability of future weak lensing surveys to measure these nonlinear parameters was investigated by [47].

The differences between the GR and MG predictions for $C_{l;ii}$ are plotted in Figure 4 for a few redshift bins. These differences are the ΔP_i defined above, although for clarity, here we have divided them by the GR prediction. As expected, there is more power in the MG gravity model at early times. We impose a conservative cutoff of $l \leq 1000$ since smaller scales contain non-linear and baryonic effects which are not completely understood even in GR, and we don't want incorrect predictions on these scales to be misinterpreted as violations of GR [48–51].

To compute the weak lensing Fisher matrix, we need to compute the covariances between the observable spectra. The total observed power in a given redshift bin will be a combination of signal and noise:

$$C_{l;ij}^{\mathrm{obs}} \equiv C_{l;ij} + \delta_{ij}\frac{\gamma_{\mathrm{rms}}^2}{n^{\mathrm{gal}}{}_i} \tag{B15}$$

where γ_{rms} is the intrinsic scatter of one polarization of the galaxy shears. The covariance between the observables is

$$\mathrm{Cov}[C_{l;ij}^{\mathrm{obs}}, C_{l';mn}^{\mathrm{obs}}] = \frac{\delta_{ll'}}{(2l+1)\Delta l f^{\mathrm{sky}}}\left(C_{l;im}^{\mathrm{obs}}C_{l;jn}^{\mathrm{obs}} + C_{l;in}^{\mathrm{obs}}C_{l;jm}^{\mathrm{obs}}\right) \tag{B16}$$

where Δl is the l-bin width [52]. Our results are insensitive to Δl since the cosmic shear power spectra are relatively featureless. For DES, $f^{\mathrm{sky}} = 0.12$ and $\gamma_{\mathrm{rms}} = 0.16$. Figure 2 shows the DES projections for lensing. If the true model were MG (right panel), then the lensing constraint would shift to more early dark energy (higher w) to accommodate the slower growth of structure.

---

[1] The formula was only calibrated for flat ΛCDM models and open models. It has been extrapolated to closed models and assumed to approximately hold for $w \neq -1$.
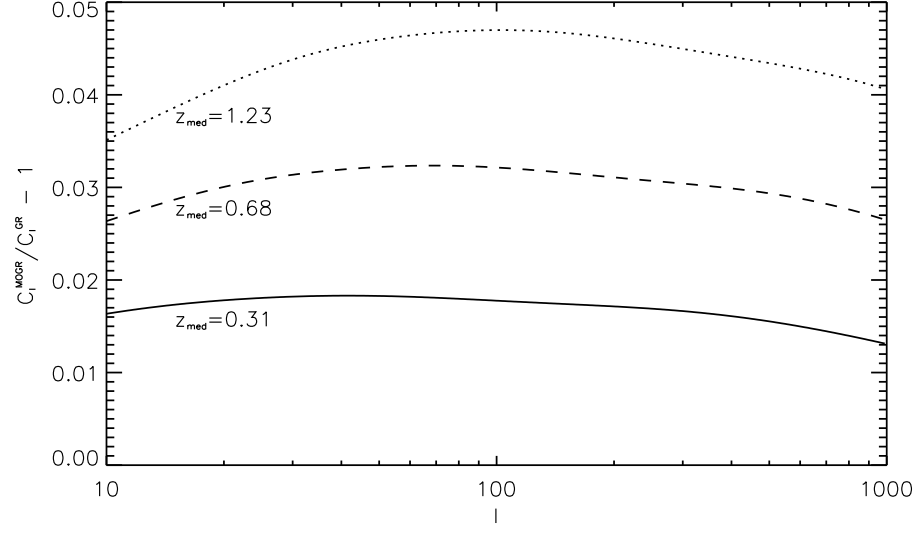
FIG. 4: Fractional difference between weak lensing power spectra predictions for GR and our MG model with $\gamma = 0.68$. Since we fix σ_8, the models diverge with increasing redshift. Agreement is enforced towards large l by the interpolation in (B13).