Model selection for modified gravity

BY T. D. KITCHING*, F. SIMPSON, A. F. HEAVENS AND A. N. TAYLOR
SUPA, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

In this article, we review model selection predictions for modified gravity scenarios as an explanation for the observed acceleration of the expansion history of the Universe. We present analytical procedures for calculating expected Bayesian evidence values in two cases: (i) that modified gravity is a simple parametrized extension of general relativity (GR; two nested models), such that a Bayes' factor can be calculated, and (ii) that we have a class of non-nested models where a rank-ordering of evidence values is required. We show that, in the case of a minimal modified gravity parametrization, we can expect large area photometric and spectroscopic surveys, using three-dimensional cosmic shear and baryonic acoustic oscillations, to ‘decisively’ distinguish modified gravity models over GR (or vice versa), with odds of $\gg 1 : 100$. It is apparent that the potential discovery space for modified gravity models is large, even in a simple extension to gravity models, where Newton’s constant $G$ is allowed to vary as a function of time and length scale. On the time and length scales where dark energy dominates, it is only through large-scale cosmological experiments that we can hope to understand the nature of gravity.

Keywords: cosmology; gravity; gravitational lensing

1. Introduction

Modified gravity scenarios have become popular as a possible explanation of the observed acceleration of the expansion rate of the Universe, attributed to a ‘dark energy’. However, even without this motivation, there is a significant discovery space that remains in the gravitational sector. Figure 1 illustrates a selection of experimental constraints on the effective value of the gravitational constant $G_{\text{eff}}$, a generalization of the constant $G$ that varies as a function of the length and scale factor, which enters into the Poisson equation. In modified theories of gravity, $G_{\text{eff}}$ can be a function of both length scale and redshift. To probe the gravitational sector in the regime where dark energy effects are observed, only gravitational lensing and galaxy clustering probes (for example redshift space distortions), which are large-scale and late time measurements, can be expected to place constraints.

Here, we discuss the need for a model comparison methodology to be used when assessing modified gravity scenarios. Such approaches are necessitated when we have nested models, where deviations from general relativity (GR) are described by additional degrees of freedom. The addition of such parameters is naturally

*Author for correspondence (tdk@roe.ac.uk).

One contribution of 16 to a Theo Murphy Meeting Issue ‘Testing general relativity with cosmology’.
weighted by an evidence calculation that includes an Occam razor-like term. Furthermore, we may have a set of un-nested models which we need to assess in a statistically rigorous manner. Here, we present Bayes’ factor predictions [2] and discuss a multi-model comparison methodology [3] that we propose can be used to distinguish modified gravity models.

In this article, we focus on three-dimensional cosmic shear as a probe of modified gravity scenarios. Three-dimensional cosmic shear [4–7] uses the weak lensing shape distortion induced in galaxy images by dark matter along the line of sight, and the redshift information from every galaxy. It contains information from the matter power spectrum and the expansion history, which allows constraints on both the growth of large-scale structure and the geometry of the Universe to be made.

In §2, we present an analytical expression for the Bayesian evidence and present predicted constraints on a minimal modified gravity scenario. In §3, we discuss rank-ordering of evidence values and apply this to predicted constraints on a redshift-dependent dark energy equation of state. We discuss conclusions in §4.

2. Nested-model selection

Here, we summarize results from Heavens et al. [2]. The aim here is to compute the expected Bayesian evidence ratio for two different models—one that describes GR and one that represents a modified gravity scenario.

*Phil. Trans. R. Soc. A* (2011)
Two competing models are denoted by $M$ and $M'$, where it is assumed that $M'$ is a simpler model that has fewer ($n' < n$) parameters. It is further assumed that the models are nested, i.e. the $n'$ parameters of model $M'$ are common to $M$, which has $p = n - n'$ extra parameters in it. $D$ represents the data vector, and $\theta$ and $\theta'$ the parameter vectors (of length $n$ and $n'$). The posterior probability of each model comes from Bayes’ theorem,

$$p(M|D) = \frac{p(D|M)p(M)}{p(D)} \quad (2.1)$$

and similarly for $M'$. By marginalization $p(D|M)$, known as the evidence, is

$$p(D|M) = \int d\theta p(D|\theta, M)p(\theta|M). \quad (2.2)$$

Hence, the posterior relative probabilities of the two models, regardless of their parameters, are

$$\frac{p(M'|D)}{p(M|D)} = \frac{p(M')}{{p(M)}} \int d\theta' p(D|\theta', M')p(\theta'|M') \quad (2.3)$$

With uniform priors on the models, $p(M') = p(M)$, this ratio simplifies to the ratio of evidences, called the Bayes’ factor,

$$B \equiv \frac{\int d\theta' p(D|\theta', M')p(\theta'|M')}{\int d\theta p(D|\theta, M)p(\theta|M)}. \quad (2.4)$$

Note that the more complicated model $M$ will inevitably lead to a higher likelihood (or at least as high), but the evidence will favour the simpler model if the fit is nearly as good, through the smaller prior volume.

Here, we show how the expected Bayes’ factor from an experiment for two competing models can be calculated under Gaussian assumptions. We assume uniform and separable prior on the two models, and we also assume that the expected value of $B$ is given by the ratio of the expected evidences and we approximate the likelihoods by multi-variate Gaussians,

$$p(D|\theta, M) = L_0 \exp \left[ -\frac{1}{2}(\theta - \theta^0)^t \mathbf{F}_{\alpha\beta}(\theta - \theta^0) \right], \quad (2.5)$$

and similarly for $p(D|\theta', M')$. $\mathbf{F}_{\alpha\beta}$ is the Fisher matrix, given for Gaussian-distributed data by Tegmark et al. [8],

$$\mathbf{F}_{\alpha\beta} = \frac{1}{2} \text{Tr}[C^{-1}C, C^{-1}C, C^{-1}(\mu, \mu^t) + (\mu, \mu^t)], \quad (2.6)$$

where $C$ is the covariance matrix of the data, and $\mu$ is its mean. A comma indicates a partial derivative with respect to a parameter. The peak likelihood $L_0$ is located at $\theta = \theta_0$. In cases where this approximation is not a good one, the Markov chain Monte Carlo approach of ExPO [9] may be a useful alternative, at more computational expense.
To compute the ratio of likelihoods, we need to take into account the fact that, in the incorrect model, the maximum of the expected likelihood will not in general be at the correct values of the parameters. In the incorrect model, some parameters are assumed to be fixed at values which differ by \(d\alpha\) from their true values. The others are shifted on average by an amount that is readily computed under the assumption of the multi-variate Gaussian likelihood [10],

\[
\delta\theta'_\alpha = -(F'^{-1})_{\alpha\beta} G_{\beta\gamma} \delta\psi_{\gamma} \quad \alpha, \beta = 1 \ldots n, \gamma = 1 \ldots p,
\]

where

\[
G_{\beta\gamma} = \frac{1}{2} \text{Tr} \left[ C^{-1} C_{\beta} C^{-1} C_{\gamma} + C^{-1}(\mu_{\gamma}^t \mu_{\beta}^t + \mu_{\beta}^t \mu_{\gamma}^t) \right],
\]

which we recognize as a subset of the Fisher matrix. For clarity, we have given the additional parameters the symbol \(j\); \(g = 1, \ldots, p\) to distinguish them from the parameters in \(M'\). The final expression for the expected Bayes’ factor can be written

\[
B = (2\pi)^{-p/2} \frac{\sqrt{\det F}}{\sqrt{\det F'}} \exp \left( -\frac{1}{2} \delta\theta_\alpha F_{\alpha\beta} \delta\theta_\beta \right) \prod_{q=1}^{p} \Delta\theta_{n'+q}.
\]

with \(\delta\theta_\alpha\) given by equation (2.7). Note that \(F\) and \(F^{-1}\) are \(n \times n\) matrices, \(F'\) is \(n' \times n'\), and \(G\) is an \(n' \times p\) block of the full \(n \times n\) Fisher matrix \(F\).

Note that the ‘Occam’s razor’ term (see Saini et al. [11] for an example), common to evidence calculations, is encapsulated in the \((2\pi)^{-p/2} \sqrt{\det F} / \sqrt{\det F'}\) factor: models with more parameters are penalized in favour of simpler models. Such terms should be treated with caution; as pointed out by Linder & Miquel [12], simpler models do not always result in the most physically realistic conclusions. However, in the example given here, we are only comparing the relative evidence of the same model with different parameter values and not different models containing different numbers of parameters. We use the descriptions defined by Jeffreys [13], where \(1 < \ln B < 2.5\) is referred to as ‘substantial’ evidence in favour of a model, \(2.5 < \ln B < 5\) is ‘strong’ and \(\ln B > 5\) is ‘decisive’.

(b) Minimal modified gravity

To apply these results to cosmological probes of dark energy and modified gravity, we use the convenient minimal modified gravity parametrization introduced by Linder [14] and expanded by Linder & Cahn [15] and Huterer & Linder [16], where the perturbations that parametrize modified gravity are described by a growth factor \(\gamma\). Although this is not the most general modification of gravity by any means, it serves to illustrate how one can approach the problem from a model selection viewpoint. The growth rate of perturbations in the matter density \(\rho_m\), \(\delta \equiv \delta\rho_m/\rho_m\), is accurately parametrized as a function of the scale factor \(a(t)\) by

\[
\frac{\delta}{a} = g(a) = \exp \left\{ \int_0^a \frac{da'}{a'} \left[ \Omega_m(a')^\gamma - 1 \right] \right\},
\]

where \(\Omega_m(a)\) is the density parameter of the matter. The parameter \(\gamma\) has a relatively well-constrained value for standard GR \(\gamma \simeq 0.55\) (however, this is not an equality, as shown by Simpson et al. [17]), whereas, for modified gravity...
theories, it may strongly deviate from this value. As an example, the flat Dvali–Gabadaze–Porrati (DGP) braneworld model \[18\] predicts $\gamma \simeq 0.68$ \[15\], on scales much smaller than those where cosmological acceleration is apparent. Here, we use $\gamma$ as an additional parameter in model $M$, i.e. $M$ represents extensions beyond GR, whereas $M'$ represents GR with dark energy.

The full set of parameters that we explore in $M'$ is $\Omega_m$, $\Omega_b$, $h$, $\sigma_8$, $n_s$, $\alpha_n$, $\tau$, $r$, $w_0$, $w_a$, being the density parameters in matter and baryons, the Hubble constant (in units of $100 \text{km}^{-1} \text{Mpc}^{-1}$), the amplitude of fractional density perturbations, the primordial scalar spectral index of density fluctuations and its running with $k$, the reionization optical depth, and the tensor-to-scalar ratio. We also explore two parameters characterizing the expansion history of the Universe \[15,19\], $w(a) = w_0 + w_a(1 - a)$ \[20\]. Note that $w_0$ and $w_a$ are not necessarily associated with a dark energy component in this case, as outlined in the study of Huterer & Linder \[16\] and expanded upon by Linder & Cahn \[15\]; $\gamma$ is an additional parameter in $M$ (set fixed at 0.55 in $M'$), which parametrizes the growth of structure, and $w_0$ and $w_a$ parameterize the expansion history. For example, for the DGP model, the expansion history is described by $w_0 = -0.78$ and $w_a = 0.32$.

The Fisher matrices are almost unchanged if we take this as the fiducial model, so we present results for $w_0 = -1$ and $w_a = 0$.

\(\text{(c) Experiments}\)

The question we want to address is the following: assume that the model of the Universe is modified gravity, which of the experimental set-ups we consider will have enough statistical power to distinguish this model from a dark energy model with the same expansion history? In this application, we take the parameters of the model to be the flat DGP ones.

The experiments considered are the Planck microwave background survey \[21\], including polarization information, three three-dimensional cosmic shear surveys and proposed supernova (SN) and baryonic acoustic oscillation (BAO) surveys. Note that, as discussed above, we set the cosmic microwave background (CMB) constraint on $\gamma$ to zero. The constraints on $r$ and $\tau$ from the weak lensing is similarly zero—these are assumed to be fixed in the weak lensing-alone experiments; in the weak lensing plus CMB, the constraints on $r$ and $\tau$ come from the CMB.

We consider a number of three-dimensional weak lensing surveys: firstly, a survey covering 5000 square degrees to a median redshift of $z_m = 0.8$ with a source density of 10 galaxies per square arcminute, such as what might be achieved with the dark energy survey (DES, \[22\]); secondly, a survey covering 30000 square degrees of the sky \[23\] to a median depth of $z_m = 0.75$ with five galaxies per square arcminute, as might be achieved with Pan-STARRS 2; thirdly, a survey of 35 sources per square arcminute, $z_m = 0.90$ and an area of 20000 square degrees (next-generation weak lensing survey; WLNG), as might be observed by a space-based survey such as Euclid, which is a candidate for the ESA Cosmic Vision programme \[24\]. Note that the characteristics of the Large Synoptic Survey Telescope (LSST) dataset are not too dissimilar from these, so the reported numbers would be very close to those of LSST. For all surveys, we assume a redshift dependence of source density $n(z) \propto z^2 \exp[-(z/z^*)^{1.5}]$, with $z^* = 1.4 z_m$, and use the three-dimensional cosmic shear power spectrum analysis method studied earlier \[4,5,7\].

\(\textit{Phil. Trans. R. Soc. A} (2011)\)
The survey parameters are summarized in Table 1. The Fisher matrices for the four experiments are available at http://www.roe.ac.uk/~afh. As there is a degeneracy between $w_0$, $w_a$ and $\gamma$ for Planck + WL, better constraints on the Universe expansion history lead to a better determination of $\gamma$ and therefore better model selection power. For probes of the expansion history, we consider supernovae and a sample of 2000 supernova type 1a at $0 < z \leq 1.8$ [25,26] that could be produced by a next-generation space-based experiment. For BAOs, we consider a Euclid-like space-based experiment and a wide-field multi-object spectrograph/Subaru measurement of imaging and redshift (WFMOS/Sumire). From these Fisher matrices (the Fisher matrix of a combination of independent datasets is the sum of the individual Fisher matrices), we compute the expected evidence ratio assuming that the true model is a DGP braneworld. Table 2 shows the expected evidence for the three-dimensional weak lensing surveys with and without a Planck prior.

**Table 1.** The critical survey parameters for the weak lensing experiments considered.

<table>
<thead>
<tr>
<th>survey</th>
<th>area (square degrees)</th>
<th>$z_m$</th>
<th>$n_0/sq''$</th>
<th>$\sigma_z(z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DES</td>
<td>5000</td>
<td>0.80</td>
<td>10</td>
<td>0.05(1 + $z$)</td>
</tr>
<tr>
<td>Pan-STARRS</td>
<td>30000</td>
<td>0.75</td>
<td>5</td>
<td>0.06(1 + $z$)</td>
</tr>
<tr>
<td>WLNG</td>
<td>20000</td>
<td>0.90</td>
<td>35</td>
<td>0.025(1 + $z$)</td>
</tr>
</tbody>
</table>

**Table 2.** The evidence ratio for the three weak lensing experiments considered with and without Planck, SN and BAO priors.

<table>
<thead>
<tr>
<th>survey</th>
<th>$\ln B$</th>
<th>predicted result</th>
</tr>
</thead>
<tbody>
<tr>
<td>DES + Planck + BAO + SN</td>
<td>1.28</td>
<td>substantial</td>
</tr>
<tr>
<td>DES + Planck</td>
<td>0.56</td>
<td>inconclusive</td>
</tr>
<tr>
<td>DES</td>
<td>0.54</td>
<td>inconclusive</td>
</tr>
<tr>
<td>Pan-STARRS + Planck + BAO + SN</td>
<td>3.78</td>
<td>strong</td>
</tr>
<tr>
<td>Pan-STARRS + Planck</td>
<td>2.04</td>
<td>substantial</td>
</tr>
<tr>
<td>Pan-STARRS</td>
<td>0.62</td>
<td>inconclusive</td>
</tr>
<tr>
<td>WLNG + Planck + BAO + SN</td>
<td>63.0</td>
<td>decisive</td>
</tr>
<tr>
<td>WLNG + Planck</td>
<td>52.2</td>
<td>decisive</td>
</tr>
<tr>
<td>WLNG</td>
<td>11.8</td>
<td>decisive</td>
</tr>
</tbody>
</table>

**Results**

We find that $\ln B$ obtained for the standard GR model is only approximately 1 for DES + Planck, whereas, for Pan-STARRS-2 + Planck, we find that $\ln B \sim 2$, for Pan-STARRS-2 + Planck + SN + BAO $\ln B \sim 3.61$ and, for WLNG + Planck, $\ln B$ is a decisive 52.2. Furthermore, a WLNG experiment could still decisively distinguish dark energy from flat DGP-modified gravity without a Planck prior. The expected evidence in this case scales proportionally as the total number of
Figure 2. The expected value of $|\ln B|$ from WL$_{NG}$ (solid line), Pan-STARRS-2 (dot-dashed line) and DES (dashed line), in combination with Planck, as a function of the difference in the growth rate between the modified gravity model and GR. The crossover at small $\delta \gamma$ occurs because Occam’s razor will favour the simpler (GR) model unless the data demand otherwise. The dotted vertical line shows the offset of the growth factor for the DGP model. The horizontal lines mark the boundaries between ‘significant’, ‘strong’ and ‘decisive’ according to Jeffreys’ [13] terminology.

galaxies in the survey. Pan-STARRS and Planck should be able to determine the expansion history, parametrized by $w(a)$ to very high accuracy in the context of the standard GR cosmological model, with an accuracy of 0.03 on $w(z \simeq 0.4)$; it will be able to substantially distinguish between GR and the simplification of the DGP braneworld model considered here, although this does depend on there being a strong CMB prior.

Alternatively, we can ask the question of how different a modified gravity model would have to be for these experiments to be able to distinguish the model from GR. This is shown in figure 2. It shows how the expected evidence ratio changes with progressively greater differences from the GR growth rate. We see that a WL$_{NG}$ survey could even distinguish ‘strongly’ $\delta \gamma = 0.048$, Pan-STARRS-2 $\delta \gamma = 0.137$ and DES $\delta \gamma = 0.179$. A combination of WL$_{NG}$ + Planck + BAO + SN should be able to distinguish $\delta \gamma = 0.041$, at 3.41 sigmas.

3. Multi-model selection

In addition to calculating ratios of evidences to help in favouring one model over another in a pair-wise fashion, we may also hope to distinguish a variety of models, and in modified gravity scenarios this is likely to be the case. In the study of Taylor & Kitching [3], we show how, given a Gaussian assumption for likelihood space, an expression for the absolute evidence for any model given some data can be computed. Here, we summarize these results, show an example for the dark energy equation of state model selection and suggest that a similar approach could be used in modified gravity model selection.
(a) Analytical evidence

If the likelihood for the data is Gaussian and parameters appear in the mean of the statistic under consideration, Taylor & Kitching [3] show that an analytical expression for the evidence can be written as

\[
\mathcal{E}(D|\mathcal{M}) = \Delta D (C^{-1} - C^{-1} \mu_i F_{ij}^{-1} \mu_j C^{-1}) \Delta D^t + \text{Tr} \ln C + 2 \ln(V \sqrt{\det F_{ij}}),
\]

where \( F \) is the Fisher matrix for the parameters in the model considered, \( C \) is the covariance of the data and \( D \) is the data. Note that, by taking differences of expected evidence, this expression can be used to derive equation (2.9).

A common approach to model selection is the use of the Bayes’ factor [27], which we explore in §2, the ratio of pairs of models or its logarithm,

\[
B_{AB} = -2 \ln B_{AB} = \mathcal{E}(D|\mathcal{M}_A) - \mathcal{E}(D|\mathcal{M}_B).
\]

An alternative is to rank-order models by their evidence, with a uniform prior, \( p(\mathcal{M}) = 1/N_M \), where \( N_M \) is the number of models. Even though we do not expect to have a complete set of all possible models, we can still normalize the set we have to estimate the posterior probability for each model, \( \mathcal{M}_A \),

\[
p(\mathcal{M}_A|D) = \frac{p(D|\mathcal{M}_A)p(\mathcal{M}_A)}{\sum_B p(D|\mathcal{M}_B)p(\mathcal{M}_B)},
\]

where we consider independent models to form a countable set. By this definition, we refer to uncountable sets of models that can be distinguished by continuous parameters, which is then just a model with a variable parameter, i.e. we class a model as the set of parameters, and not a set of parameter values. Even though the models may be incomplete, \( p(\mathcal{M}_A|D) \) is an upper limit on the true probability for each model with this dataset. Adding any new model can only reduce the probability. Since the prior is uniform, we expect a new model to appear at random in the distribution.

This scheme not only assesses ‘goodness-of-fit’ to the data, but also the competitiveness of models. If one model does well when compared with other proposed models, we rightly attach more belief to it. However, it does not prevent a new model appearing with a higher evidence which could become the best model. In this scheme, one would not necessarily truncate or throw away models, as they contribute to the normalization of the probabilities—although, if the contribution is negligible, it would seem sensible to drop outliers such that the model space is of a manageable size.

Even though the scheme outlined above puts an upper limit on the absolute model probability, it will still return the following result: that if we have only one model, Bayes’ theorem tells us that we must assign it a 100 per cent probability (as it is the only viable model available). Instead, we could judge a model in relation to the prior we assign it. To do this, we define a significance factor,

\[
S = \frac{p(\mathcal{M}|D)}{p(\mathcal{M})} = \frac{p(D|\mathcal{M})}{p(D)},
\]

1This makes use of the quasi-Newtonian maximum-likelihood method from Taylor & Kitching [3] and the Woodbury matrix identity.

Phil. Trans. R. Soc. A (2011)
where, by definition, \( S \geq 1 \), as we cannot lose information by adding data. The evidence for any model is only significant if the ratio, \( S \), of the evidence to the prior for the model, \( M \), is much larger than unity. For example, if we consider again the situation when we have only one model, the prior probability is \( p(M) = 1 \), so that \( S = 1 \), and we have not learned anything about the absolute validity of the model.

As an aside, we can now estimate the number of models needed for any model to be convincing in an absolute sense. For two models, the uniform prior for each model is \( p(M_A) = 1/2 \), so that the maximum significance is 2. While the Bayes’ factor between the two models could ‘decisively’ favour one model over the other (odds of \( \gg 1:100 \) on Jeffreys’ scale), one could only be at most ‘inconclusive’ (odds of 1:2) that the model is correct in an absolute sense. For absolute confidence, we need at least three models for comparison.\(^2\) This argument can be used to retrospectively understand the history of model selection. For example, when given the choice of a steady-state model over the Big Bang, the latter was clearly favoured owing to a large Bayes’ factor. However, the absolute confidence in the Big Bang could not be high as there were no alternative theories. Indeed, once inflationary cosmologies appeared, this new theory became preferable.

\((b)\) Results

In figure 3, we show an example of how the evidence can be used in practice, for the predicted evidence for a Euclid three-dimensional cosmic shear experiment to measure dark energy (table 1). In this example, we have assumed a dark energy equation of state, \( w(z) \), as a function of redshift, \( z \), which we use to construct mock lensing data. We fit these data using models that assume a cosmology with different \( w(z) \) models. We have chosen some non-nested basis set expansions for our \( w(z) \) models that have a maximum order of 2 (these phenomenological models are described in an earlier study [28]). For each \( w(z) \) realization, we rank-order the evidence for each model. In the first example, the cosine model has the highest probability with 0.4 and the distribution in model space is Gaussian-like. In the second example, the Chebyshev model fits the data very well, creating a spike in model space. In the third example, there is no model that favours the data over any other. These three examples represent the three broad classes of behaviour we can expect for real data, where we hope for example 2 with a spike in model space. The variance in model space is also an interesting quantity, reflecting both the distinguishability of the models and the quality of the data for model selection.

We suggest that, in the scenario where we have multiple, non-nested modified gravity models, such a rank-ordering of absolute evidence values can be used in a consistent way such that we can attach significance to the models in a Bayesian sense.

4. Conclusions

In this article, we have presented a model selection methodology that can be used to assess the evidence for modified gravity models either that are simple

\(^2\)Note the prior on the model is important here. A flat prior of \( 1/N_M \) is only appropriate for equally credible models. Including a vast array of non-credible models can be countered by giving these a low-prior weighting.
parametrized extensions of GR or in the case that we have a set of non-nested models available. We present an analytical expression for the Bayes’ factor for two models, and apply this to a minimal modified gravity scenario. We present an analytical expression for the absolute evidence and apply this to the case where the non-nested dark energy equation of state models are in competition; we find for a Euclid-like all sky three-dimensional cosmic shear experiment that the ability to distinguish a dark energy model will depend on the amplitude and nature of the underlying (true) functional form of the redshift variability.

We find that wide-field photometric surveys, using three-dimensional cosmic shear [4–7], can ‘decisively’ distinguish a minimal modified gravity (flat DGP) scenario over GR. In such a scenario, we can expect the odds for modified gravity over GR to improve from 1 : 20 with near-term experiments such as DES to 1 : 100 with mid-term experiments such as Pan-STARRS-2 and $\gg 1 : 100$ with deep all-sky experiments such as LSST or Euclid.

To illuminate the large remaining discovery space in the modified gravity sector, we show the current constraints on an effective Newton’s constant $G$ as a function of scale factor and length scale. It is clear that there remains a large regime where our knowledge is lacking or non-existent. Furthermore, in the regime where dark energy is observed to dominate, it is only through large-scale cosmological experiments that we can constrain modified gravity models.
T.D.K. was supported by a Royal Astronomical Society Fellowship. We thank the Royal Society for an invitation to present this work at Testing GR with Cosmology. We especially thank Licia Verde, Lance Miller, Richard Massey, Catherine Heymans and Benjamin Joachimi for insightful discussion.

References


