Cosmological constraints on Brans-Dicke theory

A. Avilez and C. Skordis

School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom

We report strong cosmological constraints on the Brans-Dicke (BD) theory of gravity using Cosmic Microwave Background data. We consider two types of models. First, the initial condition of the scalar field is fixed to give the same effective gravitational strength $G_{eff}$ today as the one measured on the Earth, $G_N$. In this case the BD parameter $\omega$ is constrained to $\omega > 177$ at the 95% confidence level, improving by a factor of two over previous constraints. In the second type the initial condition for the scalar is a free parameter leading to a somewhat stronger constrain of $\omega > 288$ while $G_{eff}$ is constrained to $0.97 < G_{eff}/G_N < 1.22$ at the 95% confidence level. We argue that these constraints have greater validity than for the BD theory and are valid for any Horndeski theory which approximates BD on cosmological scales. In this sense, our constraints place strong limits on possible modifications of gravity that might explain cosmic acceleration.

Introduction

The Brans-Dicke theory of gravity (BDT), is one of the simplest extensions of General Relativity (GR) depending on one additional parameter, $\omega$. In addition to the metric, the gravitational field is further mediated by a scalar field $\phi$ whose inverse plays the role of a spacetime varying gravitational strength.

The importance of BDT lies beyond its level of simplicity, in that it is the limit of more sophisticated but also more realistic and physically motivated theories. Its immediate generalizations, the so called scalar-tensor theories, have had strong theoretical support from a variety of perspectives. For example, they manifest in the low-energy effective action for the dilaton-graviton sector in supergravity. More generally, in compactifications of theories with extra dimensions, for instance Kaluza-Klein (KK) type theories or the Dvali-Gabadadze-Porrati theory, the extra-dimensional spacetime metric is decomposed in KK modes acting as effective scalars on our 4-dimensional spacetime in the same way that occurs in scalar-tensor theories. The BDT is a close cousin of the so called Galileon theories, recently proposed to explain cosmic acceleration while evading solar system constraints. In absence of matter fields, the scalar-tensor action also arises as a special sector of the Plebanski action when the trace component of the simplicity constraints is relaxed. Finally as we discuss further below, BDT arises as a particular limit of Horndeski theories, the most general scalar-tensor theory having 2nd-order field equations in four dimensions.

Solar system data put very strong constraints on the BD parameter $\omega$. The measurement of the Parameterized Post-Newtonian parameter $\gamma$ (see from the Cassini mission gives $\omega > 40000$ at the $2\sigma$ level. On cosmological scales, however, the story is somewhat different. Nagata et al. report that $\omega > \{50,100\}$ at $4\sigma$ and $2\sigma$ respectively using Wilkinson Microwave Anisotropy Probe (WMAP) 1st year data. However, as argued in their 2$\sigma$ result is not reliable as the reported $\chi^2$ has a sharp step form, and rather, one should take the 4$\sigma$ result as a more conservative estimate. Better constraints come from Acquaviva et al. who report $\omega > \{80,120\}$ at the 99% and 95% level respectively by using a combination of Cosmic Microwave Background (CMB) data from WMAP 1st year and a set of small scale experiments as well as large scale structure (LSS) data. Finally, Wu et al. reported $\omega > 97.8$ at the 95% level using a combination of CMB data from 5 years of WMAP and other smaller scale CMB experiments and LSS measurements from the Sloan Digital Sky Survey (SDSS) Release 4. Their constraint is weaker than even though newer data are used. As argued in this is the use of flat priors on $\ln(1+\frac{\omega}{3})$ rather than $-\ln \frac{1}{1+\frac{\omega}{3}}$.

Given that solar system data provide a far superior bound on $\omega$, why constrain the BDT with cosmological data? There are two reasons why this is important. Firstly, cosmological constraints on BDT are important as they concern very different spatial and temporal scales. Secondly, as we discuss further below, BDT can be considered as an approximation to a subset of Horndeski-type theories, and thus, cosmological constraints on BDT can be interpreted in a more general setting.

The model

The BDT is described by the action

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ \phi R - 2\Lambda - \omega \phi (\nabla \phi)^2 \right] + S_m$$

where $g$ is the metric determinant, $R$ is the scalar curvature, $\Lambda$ is the cosmological constant, $G$ is the bare gravitational constant and $S_m$ is the matter action. The matter action is independent of $\phi$, hence the weak equivalence principle is satisfied. We have chosen $\phi$ to be dimensionless by convention. In the limit $\omega \to \infty$ the theory tends to GR.

The relevant equations to be solved may be found in and here we quote only the Friedman equation which is

$$3 \left( H + \frac{\dot{\phi}}{2\phi} \right)^2 = \frac{8\pi G}{\phi} \rho + \frac{1}{4} (2\omega + 3) \left( \frac{\dot{\phi}}{\phi} \right)^2$$

where $H$ is the Hubble rate and $\rho$ is the total matter density including $\Lambda$, and the background scalar equation

$$\ddot{\phi} + 3H \dot{\phi} = \frac{8\pi G}{2\omega + 3} (\rho - 3P)$$
where $P$ is the matter pressure including $\Lambda$. We only consider $\omega > -\frac{3}{2}$ since otherwise the scalar is a ghost. As it happens, the field stays constant during the radiation era because $\dot{\phi}$ is sourced by $\rho - 3P$ resulting to $\phi$ behaving like a massless scalar. As the Universe enters the matter era, however, $\phi$ grows but only logarithmically with the scale factor $a$. Thus, the scalar field today, $\phi_0$, is expected to be within a few percent of its initial value in the deep radiation era.

If the scalar is approximately constant then the Friedmann equation becomes $3H^2 \approx 8\pi G_{\text{eff}}\rho$ where the effective cosmological gravitational strength is given by $\xi = G_{\text{eff}}/G = 1/\phi$. For bound states in the quasi-static regime, e.g. our solar system, the effective Newton’s constant is $G_N = G(2\omega + 3)/(2\omega + 4)$, thus, observers in a bound system which formed today will measure the same cosmological and local gravitational strength if

$$\frac{\omega_0}{\omega} = \frac{2\omega + 4}{2\omega + 3} \quad \text{(4)}$$

We call such models **restricted** since to achieve the above condition the initial value of the scalar field, $\phi_i$, must be appropriately fixed. Models for which $\phi_i$ is a free parameter will be called **unrestricted**.

**Analysis and methodology** We numerically solved the FRW and linearized equations in the Jordan frame where the matter equations of motion are unchanged from their GR counterparts. To test the numerical results we implemented the synchronous gauge equations for scalar modes in a modified version of the CAMB package [20] and compared with our own Boltzmann code (derived from CMB-Fast [21] and DASh [22]) in which both the synchronous gauge and the conformal Newtonian gauge was used.

We generated a chain of steps in parameter space by employing the Markov-Chain Monte-Carlo (MCMC) method [23, 24] using the Metropolis-Hastings algorithm [23]. The MCMC method guarantees that after an infinite number of steps the distribution of states visited by the algorithm is identical to the underlying probability distribution. Moreover, the time needed to sample a distribution grows approximately linearly with the dimension of the parameter space. For these reasons, the use of MCMC methods for parameter estimation in cosmology has become common practice and a variety of implementations exist [25–29]. In this work, we used the publicly available MCMC implementation CosmoMC [28].

In practice, the chains obtained have finite length and one must test how closely they have converged to the underlying probability distribution. A variety of convergence tests are provided by CosmoMC. Our chains were long enough not only to pass the convergence diagnostics but also to give very accurate 1D and 2D marginalized posterior distributions.

We used CMB data from WMAP-7 [30], South Pole Telescope (SPT) [31] and a compilation of other small scale CMB experiments, namely the Arcminute Cosmology Barometer Receiver [32], Balloon Observations Of Millimeter Extragalactic Radiation And Geophysics [33-34] and Cosmic Background Imager [35].

The chains were generated for two types of models, the restricted and unrestricted models. Restricted models have 7 parameters which are the dimensionless baryon and dark matter densities $\omega_\text{b}$ and $\omega_\text{c}$, respectively, the ratio of the angular diameter distance to the sound horizon at recombination $\theta$, the reionization redshift $z_\text{re}$, the amplitude and spectral index of the primordial power spectrum $A_\text{s}$ and $n_\text{s}$ respectively and the BDT parameter $\omega$. The (dimensionless) cosmological constant density $\omega_\Lambda$ and the Hubble constant $H_0$ are derived parameters. Unrestricted models have one additional parameter which is the initial condition $\phi_i$. When generating likelihoods for the SPT data, we also use the three parameters $D^C_{\text{5000}}$, $D^P_{\text{3000}}$ and $D^{SZ}_{\text{3000}}$ which respectively normalize the Cluster, Poisson and Sunyaev-Zel’dovich templates [31].

We now turn to the issue of priors. For the non-BDT parameters we assume the same priors as for $\Lambda$CDM since the two types of cosmological evolution are very similar. A prior on $H_0$ from the measurement of the angular diameter distance at redshift $z = 0.04$ [36] is also imposed for some chains. For the BDT parameters we impose flat priors on $\phi_i$ and on $-\log(\omega)$. The later also makes it more convenient to sample the chains, as was also done in [17]. In [18] flat priors on $y = \log(1 + \frac{\omega}{\omega_0})$ were used, in order to accommodate negative $\omega$. However, since $\omega < -\frac{3}{2}$ is a ghost we see little reason for this. Furthermore, since $dy \approx -\frac{d\omega}{\omega_0}$, using $y$ rather than $-\log(\omega)$ penalizes models with large $\omega$ (explains the weaker constraints found in [18]) which we feel is rather artificial.

**Results and discussion** Let us first present the results for the unrestricted models, for reasons to be discussed in more detail further below. Using WMAP7 alone we report $\omega > \{99.4, 55.4\}$ and $\xi = \{0.95_{-0.56}^{+0.55}, 0.95_{-0.94}^{+0.94}\}$ at the 95% and 99% level respectively, while if the HST prior is assumed this is changed to $\omega > \{125.6, 61.9\}$ and $\xi = \{1.07_{-0.43}^{+0.22}, 1.07_{-0.43}^{+0.40}\}$ at the same corresponding levels. Including SPT data along with WMAP7 and HST prior strengthens the constraints to $\omega > \{269.3, 147.6\}$.
and $\xi = \{1.10^{+0.13}_{0.11}, 1.10^{+0.17}_{-0.20}\}$ while further including the "high-$\ell$" data compilation improves to $\omega > \{288.2, 152.3\}$ and $\xi = \{1.09^{+0.13}_{-0.12}, 1.09^{+0.17}_{-0.17}\}$. Constraints on $G$ from CMB can be found in Ref. 37 where $0.74 \leq \xi \leq 1.66$ is found from WMAP-1 alone at the 95% confidence level. Our results improve on Ref. 37 and further put them in context with a realistic theory.

Clearly, including SPT data greatly improves the constraints from just using WMAP. This is mainly due to a better determination of $\theta$ by constraining the location of more CMB peaks but also a better determination of the damping envelope and of the peak heights. As discussed in Ref. 38, the main effect of increasing $\xi$ is to increase the width of the visibility function which in turn increases photon diffusion and damps the CMB temperature anisotropies on small scales. Thus the main constraints on $\xi$ from CMB temperature come from constraining the damping envelope with SPT data. Increasing $\xi$ has a slightly different effect on polarization. The same damping effect occurs on small scales but on large scales we get an enhancement as a thicker last scattering surface increases the amplitude of the local quadrupole which produces larger polarization signal Ref. 38.

The effect of varying $\omega$ is different from varying $\xi$. Varying $\omega$ results to a change in the background expansion history $H$ which results in a shift of the peak locations and peak heights. It makes more sense, however, to consider the changes in the spectrum at fixed $\theta$ as we use it as a parameter. In that case, the dominant effect is on large scales due to the Integrated Sachs-Wolfe effect while small scales are affected due to the perturbations in $\phi$ and are of lesser importance. Thus, the improvement to the constraints on $\omega$ coming from SPT data is indirect and is due to reducing the allowed region by constraining $\xi$ as well as the other parameters.

Let us now turn to the restricted class of models. Using WMAP7 alone we report $\omega > \{89.7, 51.0\}$ at the 95% and 99% level respectively, while if the HST prior is assumed we improve to $\omega > \{177.3, 120.0\}$ at the same corresponding confidence levels. Including SPT data, does not improve the bounds as one would expect but rather we find $\omega > \{156.5, 114.1\}$, although using the "high-$\ell$" data compilation makes things slightly better to $\omega > \{165.7, 122.8\}$. We see a significant improvement over Ref. 17, driven mainly by the inclusion of polarization data which break the degeneracy between $z_{re}$ and $n_s$ and allow the measurements of the damping tail to improve the determination of the other parameters and limit the freedom of $\omega$ to vary.

Why do constraints on the restricted class not improve with including more data but rather become less constrained than the unrestricted class which has one more parameter? In fact, our WMAP+HST results are compatible with the forecasting of Ref. 39 but given that Ref. 39 expect $\omega$ to be constrained close to $\sim 1000$ with Planck, it is reasonable to expect values around 500 – 700 with the inclusion of SPT data, contrary to what we find in practice. The reason we see no improvement with the inclusion of SPT data is uncovered on the right panel of figure 2. The marginalized distribution of $\ln \omega$ exhibits a peak around $\omega \sim 400$, however, the difference in likelihood between the GR limit and this peak is very small which renders this “detection” insignificant. However, its presence makes it difficult to improve the lower bound of $\omega$ unless a different data set is included.

To understand this better recall that in the unrestricted case the preferred value of $\xi$ deviates from unity at around 10%. However, as in the restricted model $\xi \approx 1$, this discrepancy in $\xi$ between the two types of models cuts-off a large portion of good likelihood in parameter space. This is shown on the right panel of figure 2 where for values of $\xi < 1$ there is a small degeneracy between $\ln \omega$ and $\xi$. In order to verify this assertion we consider a fictitious restricted model where $\phi_i$ is fixed so that $\xi$ is the best fit from the unrestricted model. As expected, the best-fit sample for this fictitious model lies in the $\Lambda$CDM large $\omega$ limit and the constraint on omega becomes $\omega > 471$ at the 95% level which is notably stronger than both the realistic models and also in line with the forecasting of Ref. 39. A degeneracy also exists between $\ln \omega$ and $\omega_A$ as shown on the left panel of figure 2 which is lifted in the fictitious model. This suggests that improved constraints on BDT can be obtained by including a different data set to CMB which can determine $\omega_A$ more precisely.
Implications for modified gravity Regardless the simplicity of this model, our constraints are more generally valid as the BDT can be considered an approximation to the most general scalar-tensor theory above some length scale $\ell_*$. The gravitational action is $S[g, \psi] = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \sum_{f=0}^{3} L^{(f)}$ where $\psi$ is by convention dimensionless and $L^{(f)}$ are functions of $\omega > 0$ and letting $\phi = \psi^2$ we recover the BDT. We have ignored the constant terms in $K^{(1)}$ and $K^{(3)}$ as they lead to total derivatives. The constant term in $K^{(2)}$ cannot be ignored in general but would lead to GR coupled to a massless scalar as $\epsilon_i \rightarrow 0$ and is irrelevant to our work. Thus our results hold for any Horndeski theory which can be approximated in the above form on cosmological scales. Choosing $\ell_* \sim 1/16H_0$ (Hubble scale at recombination) and using the BDT solution as $\hat{\phi} = K^{(2)}$ and $\bar{\phi} \sim 1$ we find conservative estimates as $\epsilon_i \sim \{10^{-2}, 10^5, 10^6, 10^7\}$ in order for the additional terms to be smaller than the BDT term.

Conclusions We found strong constraints on the BDT parameter $\omega > 288$ and effective gravitational strength $\xi = 1.99^{+0.13}_{-0.12}$ at the 95% confidence level, significantly improving on previous work. Improvement on these bounds is expected through the inclusion of LSS and weak-lensing data which is left to future work.

We thank Tessa Baker, Jo Dunkley, Pedro Ferreira, Adam Moss, Yong-Seon Song, Tony Padilla and Joe Zuntz for useful discussions and comments. A.A. acknowledges support from CONACYT. C.S. acknowledges support from the Royal Society.

---

* Electronic address: [ppxaa@nottingham.ac.uk]

[1] C. Brans and R. H.Dicke, Phys. Rev. **124**, 925 (1961).
[2] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys.Rept. **513**, 1 (2012).
[3] P. Jordan, Z. Phys. **157**, 112 (1959).
[4] P. Bergmann, Inf. J. Theor. Phys. **1** (1968).
[5] M. B. Green, J. Schwarz, and E. Witten, *Superstring Theory*, vol. 1-2 (1987).
[6] E. Appelquist, T. E. Chodos, A., and E. Freund, P.G.O., *Modern Kaluza-Klein Theories* (Addison-Wesley, 1987).
[7] G. R. Dvali, G. Gabadadze, and M. Porrati, Phys. Lett. **B485**, 208 (2000).
[8] A. Nicolis, R. Rattazzi, and E. Trincherini, Phys.Rev. **D79**, 064036 (2009).
[9] D. Beke (2011), 1111.1139.
[10] D. Beke, J.Phys.Conf.Ser. **360**, 012043 (2012).
[11] G. W. Horndeski, Int.J.Theor.Phys. **10**, 363 (1974).
[12] C. Deffayet, X. Gao, D. Steer, and G. Zahariade, Phys.Rev. **D84**, 064039 (2011).
[13] C. Will, *Theory and experiment in gravitational physics* (Cambridge University Press, Cambridge, 1981).
[14] C. Will, Living Rev. Rel. **9**, 3 (2006).
[15] B. Bertotti, L. Iess, and P. Tortora, Nature **425**, 374 (2003).
[16] R. Nagata, T. Chiba, and N. Sugiyama, Phys.Rev. **D69**, 083512 (2004).
[17] V. Acquaviva, C. Baccigalupi, S. M. Leach, A. R. Liddle, and F. Perrotta, Phys.Rev. **D71**, 104025 (2005).
[18] F. Wu and X. Chen, Phys.Rev. **D82**, 083003 (2010).
[19] M. Tegmark et al. (SDSS Collaboration), Phys.Rev. **D74**, 123507 (2006).
[20] A. M. Lewis, URL: [http:// camb.info](http://camb.info).
[21] U. Seljak and M. Zaldarriaga, Astrophys.J. **469**, 437 (1996).
[22] M. Kaplinghat, L. Knox, and C. Skordis, Astrophys. J. **578**, 665 (2002).
[23] R. M. Neal, Technical Report, Dept. of Computer Science, University of Toronto (1993), eRGT-TR-93-1 (available online).
[24] W. R. Gilks, S. Richardson, and D. J. Speigelhalter, *Markov Chain Monte Carlo in Practice* (Chapman and Hall, London, 1995).
[25] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J.Chem.Phys. **21**, 1087 (1953).
[26] N. Christensen, R. Meyer, L. Knox, and B. Luay, Class.Quant.Grav. **18**, 2677 (2001).
[27] L. Knox, N. Christensen, and C. Skordis, Astrophys.J. **563**, L95 (2000).
[28] A. Lewis and S. Bridle, Phys. Rev. D **66**, 103511 (2002).
[29] J. Dunkley, M. Bucher, P. G. Ferreira, K. Moodley, and C. Skordis, Mon. Not. Roy. Astron. Soc. **356**, 925 (2005).
[30] E. Komatsu et al. (WMAP Collaboration), Astrophys.J.Suppl. **192**, 18 (2011).
[31] K. Schaffer, T. Crawford, K. Aird, B. Benson, L. Bleem, et al., Astrophys.J. **743**, 90 (2011), 1111.7245.
[32] C. Reichardt, P. Ade, J. Bock, J. R. Bond, J. Brevik, et al., Astrophys.J. **694**, 1200 (2009), 0801.1491.
[33] W. C. Jones, P. Ade, J. Bock, J. Bond, J. Borrill, et al., Astrophys.J. **647**, 823 (2006), astro-ph/0507494.
[34] F. Piacentini, P. Ade, J. Bock, J. Bond, J. Borrill, et al., Astrophys.J. **647**, 833 (2006), astro-ph/0507507.
[35] J. L. Sievers, C. Achermann, J. Bond, L. Bronfman, R. Bustos, et al., Astrophys.J. **660**, 976 (2007).
[36] A. G. Riess, L. Macri, S. Casertano, M. Sosey, H. Lampeitl, et al., Astrophys.J. **699**, 539 (2009), 0005.0695.
[37] K.-i. Umezur, K. Ichiki, and M. Yahoio, Phys.Rev. **D72**, 044010 (2005), astro-ph/0503578.
[38] O. Zahn and M. Zaldarriaga, Phys.Rev. **D67**, 063002 (2003).
[39] X.-l. Chen and M. Kamionkowski, Phys.Rev. **D60**, 104036 (1999).