

Speeding up N -body simulations of modified gravity: Chameleon screening models

Sownak Bose,^{1,*} Baojiu Li,^{1,†} Alexandre Barreira,² Jian-hua He,¹
Wojciech A. Hellwing,^{3,4} Kazuya Koyama,³ Claudio Llinares,¹ and Gong-Bo Zhao^{5,3}

¹*Institute for Computational Cosmology, Department of Physics, Durham University, Durham DH1 3LE, UK*

²*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany*

³*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*

⁴*Janusz Gil Institute of Astronomy, University of Zielona Góra, ul.Szafrańska 2, 65-516 Zielona Góra, Poland*

⁵*National Astronomy Observatories, Chinese Academy of Science, Beijing, 100012, P.R. China*

(Dated: March 14, 2017)

We describe and demonstrate the potential of a new and very efficient method for simulating certain classes of modified gravity theories, such as the widely studied $f(R)$ gravity models. High resolution simulations for such models are currently very slow due to the highly nonlinear partial differential equation that needs to be solved exactly to predict the modified gravitational force. This nonlinearity is partly inherent, but is also exacerbated by the specific numerical algorithm used, which employs a variable redefinition to prevent numerical instabilities. The standard Newton-Gauss-Seidel iterative method used to tackle this problem has a poor convergence rate. Our new method not only avoids this, but also allows the discretised equation to be written in a form that is analytically solvable. We show that this new method greatly improves the performance and efficiency of $f(R)$ simulations. For example, a test simulation with 512^3 particles in a box of size 512 Mpc/h is now 5 times faster than before, while a Millennium-resolution simulation for $f(R)$ gravity is estimated to be more than 20 times faster than with the old method. Our new implementation will be particularly useful for running very high resolution, large-sized simulations which, to date, are only possible for the standard model, and also makes it feasible to run large numbers of lower resolution simulations for covariance analyses. We hope that the method will bring us to a new era for precision cosmological tests of gravity.

I. INTRODUCTION

Modified gravity theories [1, 2] are popular alternatives to the cosmological constant and dark energy models [3] to explain the observed accelerating expansion of our Universe [4–9]. Rather than invoking a cosmological constant (Λ), or a new energy component to drive the dynamics of the cosmos, these theories suggest that the Universe contains only normal and dark matter (which is often assumed as cold dark matter, or CDM), but the law of gravitation deviates from that prescribed by Einstein’s General Relativity (GR) on large scales, resulting in an acceleration of the expansion rate.

Since the law of gravity is universal, deviations from GR on large scales are often associated with changes in the behaviour on small scales. Any such small scale changes, however, must be vanishingly small due to the strong constraints placed by numerous local tests of gravity [10]. Consequently, viable modified gravity theories usually have some mechanism by which such modifications are suppressed, recovering GR in dense regions like the Solar System, where those gravity tests have been carried out and their resulting constraints apply. These are commonly referred to as ‘screening mechanisms’ in the literature, and are an inherent (instead of an add-on) property which comes from the dynamics of the theory. The screening effect implies that gravity behaves differently in different environments; this environmental dependence is often reflected in strong nonlinearities in the field equations, which make both analytical and numerical studies of such theories challenging.

In most theories that are currently being investigated, the modification to GR boils down to an extra (so-called fifth) force that is mediated by a new scalar field, and screening in this context means suppression of the fifth force. In one class of such theories, this is achieved by a coupling of the scalar field to matter and a nonlinear self-interaction potential of the scalar field. With appropriate choices of the coupling and potential, the dynamics of the scalar field can ensure that, in high density regions, the fifth force it mediates decays exponentially fast with distance, or becomes extremely small in its amplitude. Chameleon theories [11, 12], with $f(R)$ gravity [13] (see also [14–16]) as a representative example, is an instance of the former case, while the dilaton [20] and symmetron [21] models belong to the latter case. Amongst these models, $f(R)$ gravity is currently the most well-studied case, and there exist numerous works investigating in detail its predictions for large-scale structure formation in the nonlinear regime. This has been made possible by the continuous development of N -body simulation codes [e.g., 22–26, 28–36]. An efficient code amongst these is ECOSMOG [37], based on the publicly available N -body and hydro code RAMSES [38], which makes large simulations for $f(R)$ gravity feasible. Using the generic parametrisation for modified gravity theories [39, 40], ECOSMOG was extended to incorporate chameleon, dilaton and symmetron models [41, 42] in general. ECOSMOG has recently been compared with other codes developed subsequently, including MG-GADGET [43], Isis [44] and MG-ENZO [45] and very good agreement was found between all these codes [46].

There are other modified gravity theories, such as the Dvali-Gabadadze-Porrati [47] (DGP) brane-world model, in which screening is achieved by nonlinear derivative self-couplings of a scalar field. Well-studied examples include the K-mouflage

* Email address: sownak.bose@durham.ac.uk

† Email address: baojiu.li@durham.ac.uk

[48, 49] and Vainshtein [50] mechanisms, the latter being originally studied in massive gravity theories as a means to suppress the extra helicity modes of massive gravitons so that GR is recovered in the massless limit. In addition to the nonlinear massive gravity [51–53] and braneworld models, the Vainshtein mechanism is also employed in general setups, such as the Galileon models [54–59], which have been the subject of various recent studies [e.g., 60–79].

The first two generations of modified gravity simulation codes [e.g., 22, 27–29] were either not parallelised or had a uniform resolution across the whole simulation box, resulting in insufficient resolution and inefficiency. The current generation of codes, such as ECOSMOG, MG-GADGET, Isis and MG-ENZO, are all efficiently parallelised. These codes solve the nonlinear field equations in modified gravity on meshes (or their equivalents), and employ the adaptive mesh refinement (AMR) technique to generate ever finer meshes in high density regions to increase resolution. However, even with these parallelised codes, modified gravity simulations currently are still very slow compared to the fiducial GR case. As we shall discuss below, this is partially due to the nonlinear nature of the equations to be solved, and partly due to the specific numerical algorithms used. The greater computational cost of modified gravity simulations makes it difficult to achieve the resolution and volume attained in state-of-the-art simulations of standard gravity.

The coming decade will see a flood of high-precision observational data from a new generation of cosmological surveys, such as eROSITA [80], the Dark Energy Spectroscopic Instrument (DESI) [81–83], EUCLID [84] and the Large Synoptic Survey Telescope (LSST) [85]. These surveys will provide us with golden opportunities to perform cosmological tests of gravity [see ref. 86, for a recent review] and seek a better understanding of the origin of cosmic acceleration. As things stand now, it is the lack of more powerful simulation methods that limits the accuracy and size that modified gravity simulations can possibly attain, therefore preventing us from fully exploiting future observations. This has led to many attempts to speed up simulations using approximate methods [e.g., 87, 88], or develop alternative methods to predict theoretical quantities [e.g., 89–92]. These alternative methods are fast substitutes of full simulations and powerful when quickly exploring a large parameter space is the primary concern. However, simulations are nevertheless necessary to calibrate these methods or when better (e.g., %-level) accuracy is needed, as well as to study the impact of different theories of gravity on galaxy formation.

In [88], an approximate method to speed up N -body simulations of Vainshtein-type models was presented and shown to reduce the overhead¹ of solving the modified gravity equation to the level of 50 ~ 100%, with the errors in various cosmological quantities being controlled to well under ~ 1% or

smaller (comparable to the discrepancies in the predictions of different modified gravity simulation codes [46]). The same method, however, does not work as accurately in chameleon-type models (see Appendix A), the simulations for which are much more expensive than those for the Vainshtein-type models. Given that chameleon models are a large class of modified gravity models that are of interest to the theoretical and observational community, there is an equally urgent need for fast simulation methods for them – this is precisely the purpose of this paper.

Unlike the truncated simulation method in [88], which artificially suppresses the solver of the modified gravity equation on higher refinement levels of the AMR meshes, and instead interpolates the solution on lower (or coarser) refinement levels to find *approximate* solutions on higher levels, the method proposed here still solves the *full* modified gravity equations on *all* levels. The improved efficiency comes instead from a different way to discretise the equation on meshes, that makes it less nonlinear and greatly enhances the rate of convergence of the solution. The new scheme boosts the performance of the code *by a factor of 5* for a simulation with a periodic box of size 512 Mpc/h and 512^3 particles, and by a factor of *more than 20*, for a higher resolution setup with a box size of 128 Mpc/h and 512^3 particles. The method has its own limitation, namely that the existence of analytical solutions is a particular property of Hu-Sawicki (HS) $f(R)$ gravity – as well as a few other examples of chameleon, symmetron (see Appendix B) and dilaton – models. However, the generic nature of the HS model (in the sense that with varying parameters it covers a wide range of cosmological behaviours predicted by various other classes of models) and the lack of a *preferred* fundamental model make a good argument for using this model as a testbed, given that it is both *impossible* and *unnecessary* to study all chameleon-type models using simulations.

This paper will be arranged as follows. In § II we briefly describe the $f(R)$ gravity model and the chameleon screening mechanism. In § III we recap the method currently employed in $f(R)$ simulations and explain why it is inefficient, before describing the new method. In § IV we perform some tests as validation of this new method. Finally, we discuss and summarise in § V.

Throughout the paper we shall follow the metric convention $(+, -, -, -)$, and set $c = 1$ except in the expressions where c appears explicitly. Greek indices μ, ν, \dots run over 0, 1, 2, 3. A subscript ₀ denotes the present-day value of a quantity.

II. THE HU-SAWICKI $f(R)$ GRAVITY MODEL

In this section, we briefly review the Hu-Sawicki design [15] of $f(R)$ gravity, which is currently the most widely studied one in the literature.

¹ Throughout this paper, the term ‘overhead’ is used to refer to the *extra* computational time (using the same machine and number of cores) involved in running a modified gravity simulation compared to standard gravity. For example, an overhead of 110% means that the modified gravity run requires $2.1 \times$ the CPU time of a Λ CDM simulation.

A. The model

$f(R)$ gravity is obtained by replacing the Ricci scalar R in the standard Einstein-Hilbert action with an algebraic function of R [see, e.g., ref. 13, for a review]:

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{|g|} [R + f(R)] + \int d^4x \mathcal{L}_m, \quad (1)$$

in which G is Newton's gravitational constant, g the determinant of the metric tensor $g_{\mu\nu}$ and \mathcal{L}_m is the Lagrangian density for matter fields (including cold dark matter, baryons, radiation and neutrinos).

The inclusion of $f(R)$ in Eq. (1) changes the Einstein equation from second order to 4th order in derivatives of the metric tensor. However, one can straightforwardly rewrite the equation into two second order ones by defining a new variable, a scalar field, $f_R \equiv df(R)/dR$. These include an equation of motion that governs the dynamics of the scalar field:

$$\square f_R = \frac{1}{3} [R - f_R R + 2f(R) - 8\pi G \rho_m] \equiv \frac{\partial V_{\text{eff}}(f_R)}{\partial f_R}, \quad (2)$$

and a modified Einstein equation:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}^m + X_{\mu\nu}, \quad (3)$$

with

$$X_{\mu\nu} \equiv -f_R R_{\mu\nu} + \left(\frac{1}{2} f - \square f_R \right) g_{\mu\nu} + \nabla_\mu \nabla_\nu f_R, \quad (4)$$

in which $T_{\mu\nu}^m$ is the matter energy-momentum tensor, ρ_m the mass density of non-relativistic matter species, ∇_μ the covariant derivative compatible with metric $g_{\mu\nu}$, \square the Laplacian operator, $R_{\mu\nu}$ the Ricci tensor and $G_{\mu\nu}$ the Einstein tensor.

In the nonlinear regime of structure formation, assuming the quasistatic [93] and weak-field approximations, Eq. (3) simplifies to a modified Poisson equation,

$$\nabla^2 \Phi = \frac{16\pi G}{3} \delta\rho_m - \frac{1}{6} \delta R(f_R), \quad (5)$$

which relates the gravitational potential Φ at a given position to the density ($\delta\rho_m \equiv \rho_m - \bar{\rho}_m$, where a bar denotes the cosmic mean of a quantity) and curvature ($\delta R \equiv R - \bar{R}$) at that position.

Similarly, Eq. (2) reduces to

$$\nabla^2 f_R = \frac{1}{3c^2} [\delta R(f_R) - 8\pi G \delta\rho_m]. \quad (6)$$

Eqs. (5) & (6) need to be solved in cosmological simulations for $f(R)$ gravity to predict the modified gravitational force that is responsible for structure formation. It can be seen that Eq. (6) has a similar form to the Poisson equation, but $\delta R(f_R)$ is generally a nonlinear function of f_R , and this makes it more difficult to numerically solve this equation.

Of course, to fully specify a $f(R)$ model one must fix the functional form $f(R)$. Without the guidance of a fundamental theory, it is not hard to imagine that there is no unique, or even preferred, way to do this. However, there are indeed practical considerations that mean that the functional form cannot be

arbitrary either. This is because the choice of $f(R)$ must serve the purpose that it is originally designed for: namely, to explain the accelerated cosmic expansion. Moreover, as we shall see below, the design of $f(R)$ must ensure that any deviation from GR is suppressed to an insignificant level in places such as the Solar System, where numerous tests have confirmed compatibility with GR to high precision. Indeed, it is known [e.g., refs. 16–19] that for any $f(R)$ model to pass Solar System gravity tests, the background evolution must be close to (practically indistinguishable from) that of Λ CDM.

One functional form of $f(R)$ is proposed by Hu & Sawicki [HS, ref. 15], and has been shown to satisfy these requirements. It is given as:

$$f(R) = -m^2 \frac{c_1 (-R/m^2)^n}{c_2 (-R/m^2)^n + 1}, \quad (7)$$

where n, c_1, c_2 are dimensionless model parameters, and m is another model parameter of mass dimension one. To relate f_R to R , we write:

$$f_R = -n \frac{c_1}{c_2^2} \frac{(-R/m^2)^{(n+1)}}{[1 + (-R/m^2)^n]^2} \approx -n \frac{c_1}{c_2^2} \left(\frac{m^2}{-R} \right)^{(n+1)}, \quad (8)$$

where in the approximation we have used the fact that:

$$-\bar{R} \approx 8\pi G \bar{\rho}_m - 2\bar{f}(R) = 3m^2 \left[a^{-3} + \frac{2}{3} \frac{c_1}{c_2} \right]. \quad (9)$$

Setting $c_1/c_2 = 6\Omega_\Lambda/\Omega_m$, this model reproduces a Λ CDM background expansion history (in which Ω_m, Ω_Λ are, respectively, the present-day fractional density of non-relativistic matter and the cosmological constant for this background). Taking the values of Ω_m, Ω_Λ from any preferred cosmology, we get $-R \gg m^2$. In what follows, we define $\xi \equiv c_1/c_2^2$ for brevity, and the background value of f_R today, f_{R0} , which can be obtained from ξ .

B. Chameleon screening mechanism

In the absence of the chameleon screening mechanism [11, 12], all $f(R)$ gravity models would have been ruled out by local gravity tests due to the enhanced strength of gravity. The chameleon screening acts to suppress this enhancement, thereby allowing $f(R)$ gravity models like HS to pass experimental constraints on deviations from GR.

The chameleon mechanism works as follows: because the fifth force is mediated by a scalar field that has a nonzero mass given by:

$$m_{f_R}^2 \equiv \frac{\partial^2 V_{\text{eff}}(f_R)}{\partial^2 f_R},$$

it is of Yukawa type and proportional to $\exp(-m_{f_R} r)$, where r is the distance between two test particles. In high matter density environments, m_{f_R} is heavy and the suppression of the fifth force becomes very efficient. In reality, this is equivalent to having $|f_R| \ll 1$ in high density regions, which therefore suppresses the fifth force since it is proportional to $\vec{\nabla} f_R$ (this

can be seen most easily by eliminating the δR term in Eq. (5) using Eq. (6), which shows that f_R is essentially the potential of the fifth force).

Hence, the functional form of $f(R)$ is critical in determining if the fifth force can be sufficiently suppressed in dense environments. For the HS model, it was shown by [15] that $|f_{R0}| < 10^{-5}$ is required to screen the Milky Way. Currently, the strongest constraint on the value of $|f_{R0}|$ in the HS model comes from the screening of dwarf galaxies, which requires $|f_{R0}| \lesssim 10^{-7}$ (95% C.L.) [94, 95]. This is a promising way to constrain $f(R)$ gravity, provided astrophysical systematics are well controlled and the environmental impact on screening is modelled accurately (which itself will benefit from high resolution simulations).

In cosmology, there are many constraints on f_{R0} as well, and for recent reviews on this topic the readers are referred to [96, 97]. In [98, 99], X-ray and weak lensing estimates for the mass of the Coma cluster are combined to constrain $|f_{R0}| \lesssim 10^{-4.2}$ (95% C.L.). Two of the strongest constraints to date both come from cluster abundance. In [100] the authors use X-ray cluster abundance while in [101] the counts of high-significance weak lensing convergence peaks are used as a proxy for cluster counts; both studies find that $|f_{R0}| \lesssim 10^{-5.2}$ after carefully accounting for systematics, even though the data and analyses are very different. In [102], it was found that stacked lensing tangential shear of cosmic voids could potentially place constraints at a similar level. More recently, a study by [103], which uses Planck Sunyaev-Zel'dovich cluster counts, constrains $|f_{R0}| \lesssim 10^{-5.8}$, although the result is quite sensitive to the halo mass function used in the analysis. All the constraints are quoted at 95% confidence. There are many other cosmological constraints in the literature but it is beyond the scope of this paper to mention all of them (some of these studies were carried out by using linear perturbation

theory, which underestimates the effectiveness of screening and can therefore overestimate the strength of the constraints on the model – this is why simulations that fully capture the nonlinearity of the theories are useful).

III. N -BODY EQUATIONS AND ALGORITHM

In this section, we describe the N -body equations in appropriate code units and their discretised versions that ECOSMOG solves in simulations.

A. The Newton-Gauss-Seidel relaxation method

Like its base code RAMSES [38], ECOSMOG adopts supercomoving coordinates [104] to express the field equations in terms of dimensionless quantities. The newly defined variables in these code units are summarised as follows (the tilded quantities are those in the code units):

$$\begin{aligned} \tilde{x} &= \frac{x}{B}, & \tilde{\rho} &= \frac{\rho a^3}{\rho_c \Omega_m}, & \tilde{v} &= \frac{av}{BH_0}, \\ \tilde{\Phi} &= \frac{a^2 \Phi}{(BH_0)^2}, & d\tilde{t} &= H_0 \frac{dt}{a^2}, & \tilde{f}_R &= \frac{c^2 a^2 f_R}{(BH_0)^2}, \end{aligned}$$

in which x is the comoving coordinate, ρ_c is the critical density today, and v the particle velocity. In addition, B is the comoving size of the simulation box in units of Mpc/h and H_0 is the Hubble expansion rate today in units of $100h$ km/s/Mpc. Note that with these conventions, the mean matter density is $\tilde{\rho} = 1$. All the tilded quantities are dimensionless.

In terms of these variables, the Poisson and scalar field equations (Eqs. 5 & 6) in the HS model can be rewritten as:

$$\tilde{\nabla}^2 \tilde{\Phi} = 2\Omega_m(\tilde{\rho} - 1) - \frac{1}{6}\Omega_m a^4 \left[\left(-\frac{na^2 \xi}{\tilde{f}_R} \right)^{\frac{1}{n+1}} - 3 \left(a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_m} \right) \right], \quad (10)$$

$$\tilde{\nabla}^2 \tilde{f}_R = -\frac{1}{\tilde{c}^2} \Omega_m a (\tilde{\rho} - 1) + \frac{1}{3\tilde{c}^2} \Omega_m a^4 \left[\left(-\frac{na^2 \xi}{\tilde{f}_R} \right)^{\frac{1}{n+1}} - 3 \left(a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_m} \right) \right], \quad (11)$$

in which we have used the relation $m^2 = \Omega_m H_0^2$, and defined $\tilde{c} = c/(BH_0)$, which is the speed of light in code units.

In principle, Eqs. (10) & (11) can be directly discretised on a mesh and can then be solved numerically. For chameleon-type models, however, there is a further subtlety: namely, the value of $-\tilde{f}_R$ is very small at early times and in high density regions. This property is desirable in order that the model can pass Solar System tests of gravity by virtue of the chameleon mechanism, but it also poses a challenge when numerically solving Eq. (11). In the relaxation method that is employed to solve the discrete version of this equation, $-\tilde{f}_R$ in each mesh cell gets updated until the solution is close enough to

the true value (more details below). This updating procedure is a numerical approximation, and it is possible that $-\tilde{f}_R$ can acquire negative numerical values in some cells as a result. Taking the case of the HS $n = 1$ model as an example: the quantity $(-\tilde{f}_R)^{\frac{1}{n+1}}$ is not physically defined if $-\tilde{f}_R < 0$, and the code then runs into trouble.

To overcome this numerical issue, in [22] Oyaizu proposes to replace $-\tilde{f}_R$ with $\exp(u)$ in Eq. (11). As $\exp(u)$ can only be positive, this guarantees that the nonphysical situation described above will never appear. This change of variable has since then been used in all simulation codes of chameleon models to our knowledge [23–26, 28, 29, 37, 43–45].

In terms of this new variable, Eq. (11) can be discretised as:

$$\begin{aligned}
& \frac{1}{h^2} \left[b_{i+\frac{1}{2},j,k} u_{i+1,j,k} - u_{i,j,k} \left(b_{i+\frac{1}{2},j,k} + b_{i-\frac{1}{2},j,k} \right) + b_{i-\frac{1}{2},j,k} u_{i-1,j,k} \right] + \\
& \frac{1}{h^2} \left[b_{i,j+\frac{1}{2},k} u_{i,j+1,k} - u_{i,j,k} \left(b_{i,j+\frac{1}{2},k} + b_{i,j-\frac{1}{2},k} \right) + b_{i,j-\frac{1}{2},k} u_{i,j-1,k} \right] + \\
& \frac{1}{h^2} \left[b_{i,j,k+\frac{1}{2}} u_{i,j,k+1} - u_{i,j,k} \left(b_{i,j,k+\frac{1}{2}} + b_{i,j,k-\frac{1}{2}} \right) + b_{i,j,k-\frac{1}{2}} u_{i,j,k-1} \right] + \\
& \frac{1}{3\tilde{c}^2} \Omega_m a^4 \left(n a^2 \xi \right)^{\frac{1}{n+1}} \exp \left[-\frac{u_{i,j,k}}{n+1} \right] - \frac{1}{\tilde{c}^2} \Omega_m a (\rho_{i,j,k} - 1) - \frac{1}{\tilde{c}^2} \Omega_m a^4 \left(a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_m} \right) = 0, \tag{12}
\end{aligned}$$

in which we have used the second order finite difference scheme to calculate $\tilde{\nabla}^2 \left(-\tilde{f}_R \right)$. Taking the second order derivative with respect to the x coordinate as an example, this scheme gives:

$$\frac{\partial^2 \phi}{\partial x^2} \rightarrow \frac{1}{h^2} (\phi_{i+1,j,k} - 2\phi_{i,j,k} + \phi_{i-1,j,k}),$$

where h is the size of the mesh cell and the subscript i,j,k refers to the cell that is i -th in the x direction, j -th in the y direction and k -th in the z direction. Note that the discrete Laplacian in Eq. (12) looks slightly more complicated because $\tilde{\nabla}^2 \exp(u) \equiv \tilde{\nabla} \cdot (e^u \tilde{\nabla} u)$, and we have defined $b \equiv \exp(u)$ such that:

$$\begin{aligned}
b_{i+\frac{1}{2},j,k} &\equiv \frac{1}{2} [\exp(u_{i+1,j,k}) + \exp(u_{i,j,k})], \\
b_{i-\frac{1}{2},j,k} &\equiv \frac{1}{2} [\exp(u_{i-1,j,k}) + \exp(u_{i,j,k})].
\end{aligned}$$

Defining the left-hand side of Eq. (12) as the operator \mathcal{L}^h , where a superscript h is used to denote that the equation is discretised on a mesh with cell size h , the equation can be written symbolically as:

$$\mathcal{L}^h(u_{i,j,k}) = 0. \tag{13}$$

This is a nonlinear equation for $u_{i,j,k}$, and the most commonly used method to solve it is relaxation, which begins with some initial guesses of $u_{i,j,k}$ (for all mesh cells) and iteratively improves the old guess to get a new guess according to the Newton-Raphson method (same as the one used for solving nonlinear algebraic equations):

$$u_{i,j,k}^{h,\text{new}} = u_{i,j,k}^{h,\text{old}} - \frac{\mathcal{L}^h(u_{i,j,k}^{h,\text{old}})}{\frac{\partial \mathcal{L}^h(u_{i,j,k}^{h,\text{old}})}{\partial u_{i,j,k}^h}}, \tag{14}$$

until $u_{i,j,k}$ (for all mesh cells) is close enough to the true solution or, equivalently, some all-mesh average of $\mathcal{L}^h(u_{i,j,k})$ gets close enough to zero. A widely used definition of this

all-mesh average (the so-called residual) is given by:

$$\text{Residual} \equiv \left[\sum_{i,j,k} [\mathcal{L}^h(u_{i,j,k})]^2 \right]^{1/2}, \tag{15}$$

where the summation is performed over all mesh cells on a given refinement level.

The implementation of this method is fairly straightforward in principle, but in practice there are a number of subtleties that need to be taken into account. For example, refined meshes usually have irregular shapes and their boundary conditions should be carefully set up by interpolating the values of u from coarser levels. The relaxation method is also notoriously slow to converge (convergence here meaning that the residual becomes smaller than some pre-fixed threshold) if it is only done on a fixed level, and in practice the so-called multigrid method is commonly used to remedy this [105]. This consists of moving the equation to coarser meshes, solving it there, and then using the coarse-mesh solutions to correct the fine-mesh one. These subtleties have been discussed in detail in the literature; as they are not the main concern of this paper, we refer interested readers to, e.g., [37], for a more elaborate description.

Although the multigrid method improves convergence in general, the rate of convergence is still very slow in $f(R)$ simulations, and the relaxation is some times unstable and diverges. One way to improve both the rate of convergence and the stability of the Newton-Gauss-Seidel relaxation method is to impose the condition:

$$\left| \mathcal{L}^h(u_{i,j,k}^{h,\text{new}}) \right| < \left| \mathcal{L}^h(u_{i,j,k}^{h,\text{old}}) \right|,$$

i.e., requiring that the residual after the new iteration gets monotonically smaller than in the previous one. When the condition is not met, we retain the value of the scalar field from the previous step ($u_{i,j,k}^{h,\text{old}}$). While satisfying this condition can be costly on each step, the overall efficiency of the code can be significantly increased by the improved numerical stability and reduced number of iterations required to reach convergence.

Finally, a similar discretisation can be done for the modified Poisson equation:

$$\begin{aligned} & \frac{1}{h^2} \left(\Phi_{i+1,j,k} + \Phi_{i-1,j,k} + \Phi_{i,j+1,k} + \Phi_{i,j-1,k} + \Phi_{i,j,k+1} + \Phi_{i,j,k-1} - 6\Phi_{i,j,k} \right) \\ &= 2\Omega_m a (\rho_{i,j,k} - 1) - \frac{1}{6} \Omega_m a^4 \left[(na^2 \xi)^{\frac{1}{n+1}} \exp\left(-\frac{u_{i,j,k}}{n+1}\right) - 3 \left(a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_m} \right) \right]. \end{aligned} \quad (16)$$

This equation is solved after Eq. (12), by which time $u_{i,j,k}$ is already known. As a result, this is a linear equation for $\Phi_{i,j,k}$ which is easier to solve than Eq. (12), and we shall not discuss it further here. Structurally, Eq. (16) is the same as the Poisson equation for standard gravity (with a modified source term); hence, one may simply use the standard RAMSES implementations for solving the Poisson equation.

B. The new method

The discretisation used in the scalar field equation (Eq. 12) has a number of drawbacks:

- Depending on the value of ξ , the original scalar field equation can be very nonlinear (when ξ is small, the term involving $(-\tilde{f}_R)^{\frac{1}{n+1}}$ is large and non-negligible, c.f. Eq. 8) or close to linear (when ξ is large, that term is small and negligible so that the equation becomes nearly linear in \tilde{f}_R)². In the former case, introducing the new variable $u = \log(-\tilde{f}_R)$ makes the equation even more nonlinear; in the latter case, it non-linearises an almost linear equation. The high nonlinearity makes the relaxation method very slow to converge, which is why simulations of $f(R)$ gravity are generally much more costly than Λ CDM simulations with the same specifications. Indeed, even with parallelised

codes such as ECOSMOG, MG-GADGET, Isis and MG-ENZO (Zhao et al. in prep.), very large-sized and high resolution $f(R)$ simulations are currently still difficult to run, and this situation needs to be improved if we want to compare future survey data to theoretical predictions to perform accurate tests of modified gravity.

- As we have already seen above, the discrete Laplacian $\tilde{\nabla}^2 e^u$ is more complicated than the simple discretisation of $\tilde{\nabla}^2 \tilde{\Phi}$, resulting in a more complex equation that needs to be solved.
- The code ends up with a lot of exp and log operations. This is not optimal from a practical viewpoint, because the cost of these operations is generally much higher than that of simple arithmetic ones, such as summation, subtraction and multiplication.

The method described here alleviates the nonlinearity problem by defining a new variable $u = (-\tilde{f}_R)^{1/2}$, so that the scalar field equation for the HS model with $n = 1$ (the most widely studied $f(R)$ model in the literature) becomes a simple cubic equation in u , which can be solved analytically instead of resorting to the approximation in Eq. (14):

$$u_{i,j,k}^3 + pu_{i,j,k} + q = 0, \quad (17)$$

where:

$$p \equiv \frac{h^2}{6\tilde{c}^2} \Omega_m a \tilde{\rho}_{i,j,k} + \frac{2h^2}{3\tilde{c}^2} \Omega_\Lambda a^4 - \frac{1}{6} (u_{i+1,j,k}^2 + u_{i-1,j,k}^2 + u_{i,j+1,k}^2 + u_{i,j-1,k}^2 + u_{i,j,k+1}^2 + u_{i,j,k-1}^2), \quad (18)$$

$$q \equiv -\frac{h^2}{18\tilde{c}^2} \Omega_m a^4 \xi^{1/2}. \quad (19)$$

Note that here we focus on the case of $n = 1$; other cases will be discussed later.

While Eq. (17) can be solved analytically (and therefore accurately), it has three branches of solutions and, depending on the numerical values of p and q , all these branches can be real. Therefore, extra care has to be taken to make sure that the

correct branch of solutions is chosen. For this, let us define:

$$\begin{aligned} \Delta_0 &\equiv -3p, \\ \Delta_1 &\equiv 27q. \end{aligned} \quad (20)$$

As $q < 0$ is a constant in a given time step of the simulation, we see that $\Delta_1 < 0$.

The case $p > 0$ can occur in high density regions where $u > 0$ is small (and u^2 smaller still) because of the chameleon screening. In these cases, $\Delta_0 < 0$ and thus $\Delta_1^2 - 4\Delta_0^3 > 0$. The cubic equation then admits only one real solution, which must be the one we choose:

$$u_{i,j,k} = -\frac{1}{3} \left(C + \frac{\Delta_0}{C} \right) \quad (21)$$

² Note that, on first glance at Eq. (12), this may appear counter-intuitive. This dependence of the degree of linearity of Eq. (12) on the size of ξ can be explained by the fact that as ξ becomes smaller, the value of \tilde{f}_R also becomes smaller (c.f. Eq. 8), making Eq. (12) on the whole more nonlinear. The converse is true when ξ is large.

with

$$C \equiv \left[\frac{1}{2} \left(\Delta_1 + (\Delta_1^2 - 4\Delta_0^3)^{1/2} \right) \right]^{1/3}. \quad (22)$$

Note that Eq. (22) implies that $C = 0$ only when $\Delta_0 = p = 0$. This ensures that for the $p > 0$ case, $C \neq 0$ in Eq. (21), and the solution is never undefined.

In the case of $p = 0$, the solution is simply:

$$u_{i,j,k} = (-q)^{1/3}. \quad (23)$$

$p < 0$ can occur for density peaks in an overall low density region (where u and hence u^2 can be large). $\Delta_1^2 - 4\Delta_0^3$ can then take either positive or negative values. In the former case, the solution in Eq. (21) still holds, while in the latter case the equation has three real solutions:

$$u_{i,j,k} = -\frac{2}{3}\Delta_0^{1/2} \cos \left[\frac{1}{3}\Theta + \frac{2}{3}j\pi \right], \quad (24)$$

where $j = 0, 1, 2$ and $\cos \Theta \equiv \Delta_1 / (2\Delta_0^{3/2})$. It is straightforward to decide which branch we should take: as $\Delta_1 < 0$, we have $\cos \Theta < 0$ and so $\Theta \in (\pi/2, \pi)$. Given that we require $u_{i,j,k}$ to be positive-definite:

- If $j = 0$, $u_{i,j,k} \sim -\cos(\frac{1}{3}\Theta) < 0$ and is unphysical;
- If $j = 1$, $u_{i,j,k} \sim -\cos(\frac{1}{3}\Theta + \frac{2}{3}\pi) > 0$ and is physical;
- If $j = 2$, $u_{i,j,k} \sim -\cos(\frac{1}{3}\Theta - \frac{2}{3}\pi) < 0$ and is unphysical.

This new method has a few interesting features:

- The discrete equation to be solved is significantly simpler. In particular, q is the same in all cells, so it only needs to be calculated once for a given time step and on a given mesh refinement level.
- There is a substantial reduction of costly computer operations as we get rid of operations. Some \cos and \cos^{-1} operations are introduced, but they will not be executed for all cells (depending on which branch of solutions we take); even for cells in which they need to be executed, they are only executed once. In the old method, \exp is executed on both the cell and its neighbours.
- The cubic equation is solved analytically and a physical solution always exists. The variable redefinition in the old method, $\tilde{f}_R = \exp(u)$, was chosen so as to avoid the unphysical solution $-\tilde{f}_R < 0$; the new method avoids this situation automatically by selecting the physical solution $u = \left(-\tilde{f}_R\right)^{1/2} > 0$ analytically. As a result, we can expect this new method to be both

more stable (i.e., not suffering from catastrophic divergences due to numerics) and more efficient (i.e., the solution to Eq. (17) is exact for each Gauss-Seidel iteration, while Eq. (14) implicitly uses the approximate Newton-Raphson method and may need to be executed many times to arrive at what the new method achieves in one go).

Note that this new method does not really get rid of Gauss-Seidel relaxation, because the quantity p in Eq. (17) depends on the values of the scalar field in (the 6 direct) neighbouring cells, which are not accurate values but temporary guesses. It therefore still needs to do the Gauss-Seidel iterations (we use the standard red-black chessboard scheme here). What it does get rid of is the ‘Newton-Raphson’ part [Eq. (14)] of the Newton-Gauss-Seidel (or nonlinear GS, or NGS) relaxation which updates the old guesses using a linear approximation of the full nonlinear equation. The speedup is also largely assisted by the simplicity of Eq. (18) compared to Eq. (12), which comes about due to the new variable redefinition. Therefore, while Gauss-Seidel iterations are still required, the savings using the new method can still be significant.

IV. TESTS AND SIMULATIONS OF THE NEW METHOD

In this section we present the results of several test runs of the new ECOSMOG code. In what follows, we will only consider the F6 model of $f(R)$ gravity, in which the present-day value of the scalar field is given by $|\tilde{f}_{R0}| = 10^{-6}$. In this model, the chameleon screening is particularly efficient, meaning that deviations from GR are very small. To capture the effects of screening, accurately solving the nonlinear scalar field equations is therefore necessary.

We have simulated the F6 model at three resolution levels: ‘Low res’, ‘Medium res’ and ‘High res’ (the box size and number of particles used in each of these runs are summarised in Table I). In each case, we have also run a Λ CDM simulation starting from the same initial conditions. The mesh refinement criteria used for the ‘High res’ simulation allows us to resolve small scales comparable to those in the Millennium simulation [106]. While the ‘Low’ and ‘High res’ runs use Planck 2015 [107] cosmological parameters (with $\Omega_m = 0.308, \Omega_\Lambda = 0.692, h = 0.6781, \sigma_8 = 0.8149$), the cosmological parameters for the ‘Medium res’ run are obtained from WMAP-7 [108] data (with $\Omega_m = 0.271, \Omega_\Lambda = 0.729, h = 0.704, \sigma_8 = 0.8092$).

In Fig. 1, we compare the nonlinear matter power spectrum, $P_{\delta\delta}(k)$, from the ‘Medium res’ simulations using the old and new methods. $P_{\delta\delta}(k)$ was computed using the publicly-available POWMES code [109]. The solid and dashed curves are $P_{\delta\delta}(k)$ computed at $z = 0$ and $z = 0.5$, respectively, for F6. The results of the two methods are indistinguishable at both redshifts, and this is quantified more clearly in the bottom panel of Fig. 1, which shows the relative difference between the old and new methods. The shaded grey band in this panel represents a 1% error around zero; clearly, the new

Name	Model	B [Mpc/h]	N_p	Speed up	Overhead (new method)
Low res	Λ CDM, F6	512	512^3	$5\times$	110%
Medium res	Λ CDM, F6	250	512^3	$15\times$	180%
High res	Λ CDM, F6	128	512^3	$> 20\times$	190%

TABLE I. Details of the simulations performed in this work. The columns B and N_p , respectively, refer to the comoving box size and number of particles in each of these runs. The starting redshift in all simulations was $z_{\text{ini}} = 49$. The second last column summarises the factor by which the new method is faster than the old one in each case. Note that the $> 20\times$ speedup for the ‘High res’ simulation is an estimate - we have not run an F6 simulation at this resolution using the old method. The last column shows the percentage overhead of the F6 simulations using the new method compared to Λ CDM. The level of speedup that can be achieved in the F6 simulations depends on the convergence criteria used: in all cases, convergence is considered as achieved when the residual is $< 10^{-8}$ on the domain level, and $< 10^{-7}$ on the fine levels.

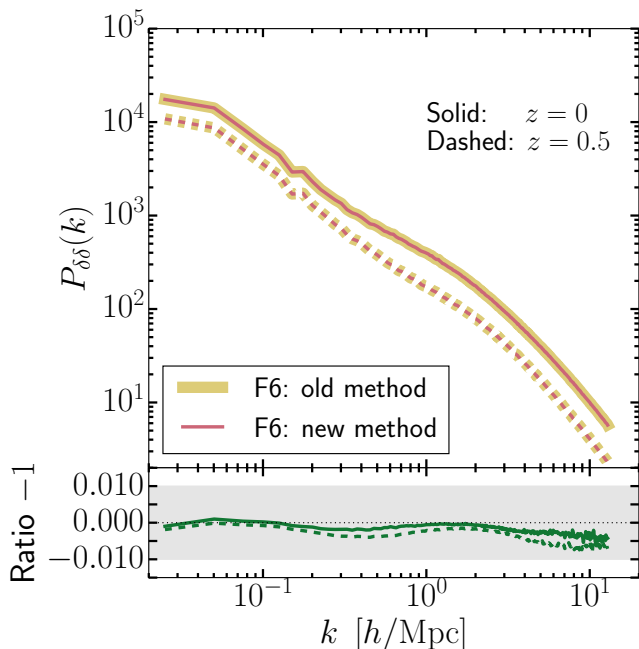


FIG. 1. *Top panel:* Comparison of the nonlinear matter power spectra at $z = 0, 0.5$ for the F6 model using the old method (yellow, § III A) and the new method (red, § III B) for solving the scalar field equations of motion. The results shown are for the ‘Medium res’ simulation. *Bottom panel:* Ratio of the power spectra corresponding to the upper panel. The shaded grey band represents a 1% error region.

and old methods agree to well below 1% at all scales resolved in the simulation. The same is true even at higher redshift ($z = 1, 2$, not shown). We have checked that the agreement also holds in the case of the velocity divergence power spectrum, $P_{\theta\theta}(k)$, which, being just the first integral of the gravitational acceleration, would be more sensitive to differences in the gravitational forces between the two methods. Agreement for $P_{\theta\theta}(k)$, which is calculated in a volume-weighted way, shows that the two methods agree well even in regions of the cosmic web that are not mass-dominated. This is not un-

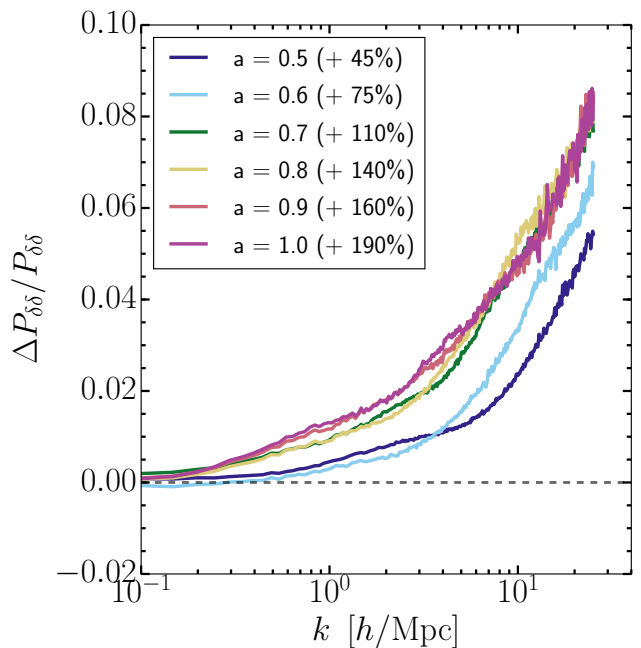


FIG. 2. Enhancement of the F6 matter power spectrum relative to Λ CDM for the ‘High res’ simulation ($B = 128$ Mpc/h, $N_p = 512^3$). The different coloured curves show the relative difference at different scale factors, as indicated in the legend. Alongside the legend labels, we also note the percentage overhead associated with the F6 run compared to the Λ CDM run at the same scale factor.

expected: after all, the new method solves the same equation of motion, without needing to use the approximate and inefficient Newton-Raphson scheme. As a consequence, the simulation is now *significantly* faster than before: the new method boosts the speed of the F6 calculation by a factor of 15 relative to the old implementation in ECOSMOG (see the last column of Table I).

Two-point statistics such as the power spectrum offer a complete description of clustering properties only for Gaussian fields. Gravitational instability theory predicts that the nonlinear evolution induced by gravity drives away the PDF of these fields from Gaussianity at late times and small scales [see e.g., refs. 110, 111]. This is reflected in growing skewness and kurtosis of cosmic density and velocity fields. $f(R)$ theories show systematic deviations from Λ CDM for these statistics, and these can therefore be used as a test of the theory [112]. We have computed PDFs and their higher-order moments for the density and velocity divergence fields to test how well the old and new methods agree beyond simple two-point statistics. We find that the differences are very small and comparable to the differences seen in the $P(k)$. As an additional test, we have also computed the Fourier mode decoherence functions [113, 114], defined as Pearson correlation coefficients for the Fourier modes of the two fields:

$$C(k) \equiv \frac{\langle f_1 f_2^* \rangle}{\langle f_1^2 \rangle^{1/2} \langle f_2^2 \rangle^{1/2}},$$

where f_1 and f_2 are the density or velocity divergence fields for the $f(R)$ runs computed using the two methods. $C(k) = 1$

when both fields being compared have Fourier modes at given k that correspond exactly. The density and velocity divergence fields for the F6 runs using the two methods take $C(k) = 1$ for almost the entire range of k , up until the Nyquist limit of the simulations. These tests reassure us that the density and velocity fields produced by the old and the new method are, for all practical purposes, indistinguishable.

Results from the ‘High res’ simulations are shown in Fig. 2, where we plot the relative difference in $P_{\delta\delta}(k)$ of F6 with Λ CDM – only results using the new method are shown. Curves of different colours represent the relative difference at different scale factors, as labelled in the legend. The legend labels also list the percentage overhead involved in the F6 simulation compared to the Λ CDM run at the same scale factor. With the new method, the F6 simulation is now only $\sim 45\%$ slower than the Λ CDM run at $a = 0.5$ ($z = 1$), and only $\sim 190\%$ slower at the final time. Compared to F6 simulations with comparable resolution using the old method [e.g., ref. 115], the new implementation is estimated to be *more than* $20\times$ faster.

The degree to which the new method improves the efficiency of ECOSMOG over the old one depends on resolution. Indeed, in going from the ‘Low res’ to the ‘High res’ simulations, the gain in performance increases from a factor of 5 to a factor of over 20 (the overhead increases considerably with resolution in the old method). The improved efficiency of the numerical algorithm will enable us to run simulations of chameleon models that would previously have been computationally very expensive to perform. Future applications of the method could include running hydrodynamical simulations (where high resolution is required to follow accurately the hydrodynamics and to resolve spatial scales important for star formation and feedback), and running large numbers of low resolution volumes to estimate the covariance matrix in non-standard gravity.

V. SUMMARY AND DISCUSSIONS

Modified gravity models are an umbrella group of theories seeking to explain the apparent accelerated cosmic expansion by assuming modifications to the Einsteinian gravitational law on cosmological scales. Usually, such modifications must be small in high density environments in which gravity is known to be accurately described by GR, and this can be achieved by screening mechanisms, resulting in highly nonlinear field equations. Studying the cosmological implications of these theories and observational constraints on them is an active research topic in cosmology, but the nonlinear nature of these theories means that one has to resort to numerical simulations, which can be prohibitively slow. This has, up until now, limited the scope of accurately testing gravity using precision observational data.

In this paper, we proposed and demonstrated the power of a new and more efficient method to solve the nonlinear field equation in one of the most popular modified gravity models – the Hu-Sawicki variant of $f(R)$ gravity. The current method used to simulate this model is slow mainly because of a vari-

able redefinition aimed at making the relaxation algorithm numerically stable, but has the negative side effect of making the discrete equation even more nonlinear and, therefore, harder to converge. As a result, modified gravity simulations which match the size *and* resolution of the state-of-the-art Λ CDM N -body or hydrodynamical simulations have thus far been beyond reach [but see refs. 116, 117].

The new method avoids the specific variable redefinition used in the old method, and therefore does not further increase the nonlinearity of the discrete equation to be solved. More importantly, it enables the discrete equation to be written in a form that is *analytically* solvable at each Gauss-Seidel iteration. This is what ultimately makes the method efficient: compare solving a highly nonlinear algebraic equation analytically and solving the same equation using the Newton-Raphson iteration method (Eq. 14), and it is clear that the latter is generally much more inefficient.

We have performed test simulations using the new method, and confirmed that it is indeed very efficient. The working model for the tests is the F6 variant of Hu-Sawicki $f(R)$ gravity. The chameleon screening is very efficient in F6, and it is therefore important that the nonlinear scalar field equations are solved accurately. In Fig. 1, we have confirmed that the new and old methods agree at the sub-percent level when comparing the nonlinear matter power spectrum, $P_{\delta\delta}(k)$. The good agreement continues to hold at higher redshift, as well as for the velocity divergence power spectra, $P_{\theta\theta}(k)$. Next, in Fig. 2, we presented results from our ‘High res’ simulations, which are comparable in resolution to the Millennium simulation. The total overhead in the F6 simulation is $\sim 190\%$ compared to the equivalent Λ CDM run; this represents a boost in performance of $> 20\times$ compared to an F6 simulation of similar resolution using the old method.

The improved performance of the new simulation algorithm compared to the old one serves to highlight the importance of the way in which one discretises partial differential equations for the efficiency of numerically solving them. This is particularly true for highly nonlinear equations, such as those encountered in many modified gravity theories. Our work highlights the following:

(1) There is not a single way of discretisation, and this usually depends on the specific equations to be solved. In general, the discretisation should be chosen to preserve the original degree of nonlinearity of the equation as much as possible, and avoid further nonlinearising the equation.

(2) Where possible, exact solutions to the nonlinear discrete equation should be used instead of the approximate solution in Eq. (14). The latter, despite being commonly used in relaxation solutions to nonlinear differential equation, is a second option only for cases where $\mathcal{L}^h(u_{i,j,k}) = 0$ has no analytical solution in general.

The same observations and conclusions apply to other classes of partial differential equations, such as those involving higher order powers of the derivatives of the scalar field (e.g., $(\nabla^2\varphi)^2$, $\nabla^i\nabla^j\varphi\nabla_i\nabla_j\varphi$, $(\nabla^2\varphi)^3$, $\nabla^i\nabla_j\varphi\nabla^j\nabla_k\varphi\nabla^k\nabla_i\varphi$), which are commonly encountered in Vainshtein-type theories. In fact, in the most popular examples of such models – the DGP, cubic Galileon and quartic Galileon

models – we also found that the discretisation could be done in a way such that $\mathcal{L}^h(u_{i,j,k}) = 0$ is a quadratic or cubic equation that can be solved analytically. This fact has been used in [88, 118–120] to make simulations of these models possible, more efficient and free from numerical instabilities.

Unfortunately, the new method does *not* apply to *all* non-linear partial differential equations, because it relies on $u_{i,j,k}$ being analytically solvable in the discrete equation. In the HS $f(R)$ model with $n = 1$, $u_{i,j,k}$ satisfies a cubic equation, which does have analytical solutions. This neat property does not hold for other models. However, this method will still be very useful for the following reasons:

- At the moment, no specific functional forms of $f(R)$ – or more generally, no specific chameleon models – are known to be fundamental. Different models often share similar qualitative behaviours though the predictions can be quantitatively different. For what it is worth, the HS model serves as a great test case to gain insights into the question ‘How much deviation from GR (in the manner prescribed by the large class of chameleon models) is allowed by cosmological data?’. Indeed, all current observational constraints on modified gravity are to be considered as attempts to answer this question. In this context, the exact functional form of $f(R)$ is not critical, because whatever form we adopt, it is unlikely to be the true theory. Actually, the HS model is capable of reproducing the behaviours of many classes of models, and is therefore a representative example.
- There are other models that this method can be applied to. One example is the HS $f(R)$ model with $n = 2$. In this case Eq. (17) becomes a quartic equation, which also has analytical solutions. A further example is the logarithmic $f(R)$ model studied in the literature [e.g., ref. 16]:

$$f(R) \sim -2\Lambda - \eta \log(R/R_*),$$

where Λ is the cosmological constant, and η and R_* are some model parameters. In this case, $f_R \sim 1/R$, and we could define $u = -\tilde{f}_R$ so that Eq. (17) becomes a quadratic equation. Moreover, looking beyond $f(R)$ gravity, there are also other chameleon models with different coupling strengths from the value of 1/3 for $f(R)$ models, and can be simulated using this method [42]. The method can also be applied to the symmetron model [21], in which the equation $\mathcal{L}^h(\varphi_{i,j,k}) = 0$ is a cubic equation [121] for the symmetron field φ , and certain variants of the dilaton model [20, 41], though our initial tests showed that the improvement of the efficiency is far smaller than in the $f(R)$ case (Appendix B).

Efforts towards generalising the new method to the models mentioned above, and to running large high resolution simulations including baryonic physics, are currently ongoing and will be the subject of future works.

ACKNOWLEDGMENTS

We thank the organisers of the 1st Mexican Numerical Simulations School held in Mexico City (03-06 October 2016), for inspiring discussions that motivated this study. We are also grateful to Fabian Schmidt for many enlightening comments. SB is supported by STFC through grant ST/K501979/1. BL acknowledges support by STFC Consolidated Grant Nos. ST/L00075X/1 and RF040365. WAH acknowledges support from the Polish National Science Center under contract #UMO-2012/07/D/ST9/02785. KK is supported by STFC grant ST/N000668/1. Both WAH & KK are also supported by the European Research Council through grant 646702 (CosTesGrav). CLL acknowledges ICRAR (University of Western Australia) and Queensland University for their warm hospitality and Chris Power and David Parkinson for discussions. CLL also acknowledges support from STFC consolidated grant ST/L00075X/1 as well as funding support from the University of Queensland - University of Western Australia Bilateral Research Collaboration Award. This work was supported by resources provided by The Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. The simulations described in this study used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-Infrastructure Capital grant ST/K00042X/1, STFC Capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. Part of this work was performed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1066293.

Appendix A: The poorer performance of the speed up truncation method in chameleon models

In [88], the authors proposed a method to speed up N -body simulations of modified gravity models with Vainshtein screening. The speed up in this method is achieved by truncating the Gauss-Seidel iterations of the scalar field above a certain refinement level, and then computing the solution on those fine levels by interpolating from coarser levels. This *approximate* method agrees very well with the results of the full N -body calculation (see [88] for details) due to the fact that in Vainshtein screening models, there is a correlation between higher density regions (or, equivalently, higher refined regions in the simulation box) and screening efficiency. Even when the error induced on the fifth force on the refinements is large, it does not propagate to the total gravitational force because the amplitude of the fifth force is small/screened.

In chameleon models, however, the correlation between high density regions and screening efficiency becomes less marked because of the dependence on the environmental density (in Vainshtein models, the screening efficiency depends on the local density only). For example, in $f(R)$ models, a low mass halo in a dark matter void constitutes an example

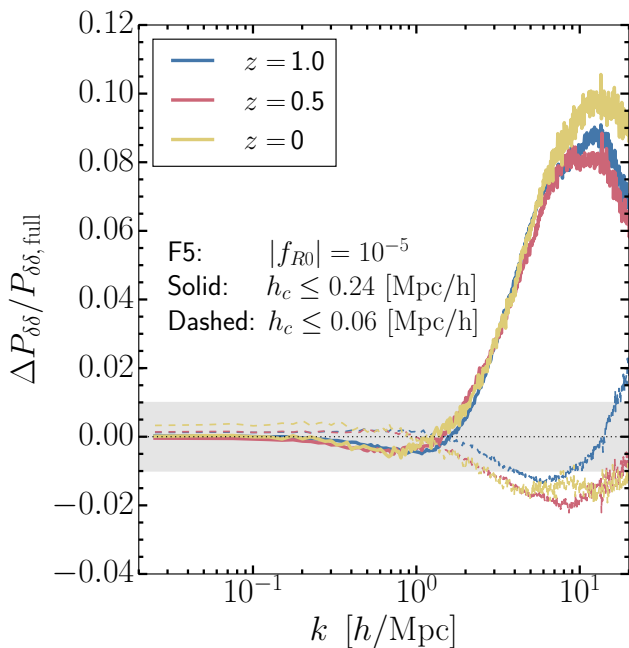


FIG. 3. Relative difference in the matter power spectra at $z = 0, 0.5$ & 1 between a full F5 simulation, and two truncated runs where the Gauss-Seidel iterations of the scalar field have been truncated on finer refinement levels (see the accompanying text). The dashed and solid lines, respectively, correspond to less and more aggressive truncation schemes; h_c is the cell size of the first truncated level in each simulation. The shaded grey band represents the 1% error region around zero.

of a highly-refined region (the centre of the halo can be very concentrated) that may not be screened (either by itself or by the low density environment it lives in). It is therefore interesting to determine whether or not the same truncation method, which works well for Vainshtein models, works equally well in chameleon-type theories.

Fig. 3 shows the relative difference of two truncated $f(R)$ simulations to a full (ie., not truncated) simulation. The simulation box used for this test is the same as the ‘Medium res’ setup in the main text, but with $f_{R0} = -10^{-5}$ (the so-called F5 model). The result is shown at three different redshifts and the two labelled truncation schemes are as follows. The case $h_c \leq 0.24$ Mpc/h indicates that the scalar field was only explicitly solved on the coarse level, with this solution being interpolated to all finer levels. In the case of $h_c \leq 0.06$ Mpc/h, the scalar field was explicitly solved on the coarse, first refinement and second refinement levels, with the solution at the second level being interpolated to all other finer levels. The values 0.24 Mpc/h and 0.06 Mpc/h indicate the cell size of the first truncated level in both these simulations, which ran, respectively, ≈ 10 and ≈ 2 times faster than the full run. For both these truncation criteria, the figure shows that the error can be kept $< 1\%$ for $k \lesssim 2$ h/Mpc, but for higher modes, it grows to unacceptably large values. For example, at $k \approx 5$ h/Mpc, the error is of $\approx 6\%$.

The result shown in Fig. 3 for $f(R)$ should be contrasted

with the corresponding picture in the DGP model (which employs Vainshtein screening), in which for the same truncation criteria, the error is always kept below 1% for $k < 5$ h/Mpc (see e.g., Fig. 5 of [88]). Furthermore, the method described in the main body of this paper results in comparable boosts in performance compared to previous $f(R)$ simulations, but without any loss in accuracy. From this we can conclude that the truncation scheme that works well in Vainshtein screening models is not suitable for chameleon theories.

Appendix B: Performance of the new method for the symmetron model

As a test of the performance of our new method for other classes of screening mechanisms, we implemented our method for the case of the symmetron model. The code used in this case, Isis [44], is a modified version of RAMSES developed independently of ECOSMOG. Details of the symmetron model and its implementation in Isis are described in [44]. Briefly, the equation of motion for the scalar field is given by:

$$\nabla^2 \phi \propto (A\rho - 1)\phi + \phi^3, \quad (\text{B1})$$

where the quantity A is a function of the parameters of the symmetron model. While the equation is formally equivalent to the $f(R)$ in the main text (Eq. 6), the screening mechanism operates differently. In the $f(R)$ model, the scalar field screens itself by becoming very massive. On the other hand, in the symmetron model, the screening occurs when a particular symmetry is restored (i.e., when the factor in front of the linear term of the source of Eq. (B1) becomes positive). Consequently, the model behaves in a different manner to $f(R)$. For instance, negative solutions for the symmetron field, ϕ , are allowed and, thus, the constraints implemented in the $f(R)$ case (§ III B) are not required. We refer the reader to [122] for a summary of the complex phenomenology associated with this property of the symmetron field.

The nonlinear modified gravity solver in Isis is very similar to that of the $f(R)$ model in ECOSMOG. The code uses an implicit multigrid solver with full approximation storage, which means that the code relies on a Newton-Raphson algorithm to evolve the solution in every step of the Gauss-Seidel iterations. As the discretised equation is cubic, the method proposed in this paper can be applied in a straightforward manner. As a check of the accuracy of the new method in solving the symmetron field equations, we have repeated satisfactorily the static test presented in the original Isis paper [Fig. 2 in ref. 44]. However, we find that there is no major difference in the performance of the standard Isis implementation compared to using the new method, either in terms of the run time, or the convergence rate of the iterative solver.

In order to gauge the difference in computing time between the old and new methods for the symmetron model, we have run a set of five different realisations of a box of size 150 Mpc/h on a side, containing 256^3 particles. For each realisation, there are three sets of simulations: Λ CDM and the symmetron model using the old and new methods. Overall, we do not find any improvement in the performance of Isis

using the new method. For both the old and new methods, the overhead compared the Λ CDM simulation is of the order of $\sim 170\%$ and, in fact, the run time using the new method is actually $\sim 1\%$ slower than using the default implementation - this is explained by the fact that $\sim 1\%$ more iterations were required for the whole set of five realisations using the new method. The convergence criterion on the residual was set to 10^{-6} for both symmetron runs; we find that, unlike in the $f(R)$ model, making the convergence criterion even stricter does not impact the run time of the symmetron simulations by a great amount.

The reason why the performance of the code appears to be insensitive to the details of the iteration scheme is seemingly related to the type of screening mechanism used by the symmetron model. The symmetron mechanism is based on a density threshold above which the solution very quickly approaches zero and thus decouples the scalar field from matter. This makes the solutions more stable and, therefore, not strongly dependent on the details of the solver employed. Since the default solver in Isis does not involve a nonlinear change of variables to force a stable, positive solution (as in the $f(R)$ case), the performance is already similar to what ECOSMOG can do for $f(R)$ using the new method.

-
- [1] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Phys. Rept.* **513** (2012) 1.
- [2] A. Joyce, L. Lombriser, & F. Schmidt, *Ann. Rev. of Nuc. and Part. Sc.*, **66**, 95, (2016).
- [3] E. J. Copeland, M. Sami and S. Tsujikawa, *Int. J. Mod. Phys. D*, **15**, 1753 (2006).
- [4] J. Guy, M. Sullivan, A. Conley, N. Regnault, P. Astier, C. Bolland *et al.*, *Astron. Astrophys.* **523**, A7 (2010).
- [5] W. J. Percival, B. A. Reid, D. J. Eisenstein, N. A. Bahcall, T. Budavari, J. A. Frieman *et al.*, *Mon. Not. R. Astron. Soc.*, **401**, 2148 (2009).
- [6] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell *et al.*, *Mon. Not. R. Astron. Soc.*, **416**, 3017 (2011).
- [7] B. A. Reid, L. Samushia, M. White, W. J. Percival, M. Manera, N. Padmanabhan *et al.*, *Mon. Not. R. Astron. Soc.*, **426**, 2719 (2012).
- [8] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley *et al.* (2012), arXiv:1212.5226 [astro-ph.CO].
- [9] A. G. Riess, L. Macri, S. Casertano, M. Sosey, H. Lampeitl, H. C. Ferguson *et al.*, *Astrophys. J.* **699**, 539 (2009).
- [10] C. M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Relativity*, **17**, 4 (2014).
- [11] J. Khoury and A. Weltman, *Phys. Rev. D*, **69**, 044026 (2004).
- [12] D. F. Mota and D. J. Shaw, *Phys. Rev. D*, **75**, 063501 (2007).
- [13] T. P. Sotiriou and V. Faraoni, *Rev. Mod. Phys.*, **82**, 451 (2010).
- [14] B. Li and J. D. Barrow, *Phys. Rev. D* **75**, 084010 (2007).
- [15] W. Hu and I. Sawicki, *Phys. Rev. D*, **76**, 064004 (2007).
- [16] P. Brax, C. van de Bruck, A. C. Davis and D. J. Shaw, *Phys. Rev. D*, **78**, 104021 (2008).
- [17] J. Wang, L. Hui and J. Khoury, *Phys. Rev. Lett.*, **109**, 241301 (2012).
- [18] M. Raveri, B. Hu, N. Frusciante and A. Silverstri, *Phys. Rev. D*, **90**, 043513 (2014).
- [19] J. Ceron-Hurtado, J. He and B. Li, *Phys. Rev. D* **94**, 064502 (2016).
- [20] P. Brax, C. van de Bruck, A. C. Davis and D. J. Shaw, *Phys. Rev. D* **82**, 063519 (2010).
- [21] K. Hinterbichler and J. Khoury, *Phys. Rev. Lett.*, **104**, 231301 (2010).
- [22] H. Oyaizu, *Phys. Rev. D* **78**, 123523 (2008).
- [23] H. Oyaizu, M. Lima and W. Hu, *Phys. Rev. D*, **78**, 123524 (2008).
- [24] F. Schmidt, M. Lima, H. Oyaizu and W. Hu, *Phys. Rev. D*, **79**, 083518 (2009).
- [25] B. Li and H. Zhao, *Phys. Rev. D* **80**, 044027 (2009).
- [26] H. Zhao, A. V. Maccio, B. Li, H. Hoekstra and M. Feix, *Astrophys. J.*, **712L**, 179 (2010).
- [27] C. Llinares, A. Knebe, & H. Zhao, *Mon. Not. R. Astron. Soc.*, **391**, 1778 (2008).
- [28] B. Li and H. Zhao, *Phys. Rev. D* **81**, 104047 (2010).
- [29] G. Zhao, B. Li and K. Koyama, *Phys. Rev. D* **83**, 044007 (2011).
- [30] G. Zhao, B. Li and K. Koyama, *Phys. Rev. Lett.*, **107**, 071303 (2011).
- [31] Y. Li and W. Hu, *Phys. Rev. D* **84**, 084033 (2011).
- [32] B. Li, G. Zhao and K. Koyama, *Mon. Not. R. Astron. Soc.*, **421**, 3481 (2012).
- [33] L. Lombriser, K. Koyama, G. Zhao and B. Li, *Phys. Rev. D* **85**, 124054 (2012).
- [34] J. Lee, G. Zhao, B. Li and K. Koyama (2012), *Astrophys. J.*, **763**, 28 (2013).
- [35] E. Jennings, C. M. Baugh, B. Li, G. Zhao and K. Koyama (2012), *Mon. Not. R. Astron. Soc.*, **425**, 2128 (2012).
- [36] B. Li, W. A. Hellwing, K. Koyama, G. Zhao, E. Jennings and C. M. Baugh, *Mon. Not. R. Astron. Soc.*, **428**, 743 (2013).
- [37] B. Li, G. Zhao, R. Teyssier and K. Koyama, *J. Cosmo. Astropart. Phys.*, **01**, 051 (2012).
- [38] R. Teyssier, *Astron. Astrophys.* **385**, 337-364 (2002).
- [39] P. Brax, A. C. Davis and B. Li (2011), *Phys. Lett. B*, **715**, 38 (2012).
- [40] P. Brax, A. C. Davis, B. Li and H. A. Winther (2012), *Phys. Rev. D* **86**, 044015 (2012).
- [41] P. Brax, A. C. Davis, B. Li, H. A. Winther and G. Zhao (2012), *J. Cosmo. Astropart. Phys.*, **10**, 002 (2012).
- [42] P. Brax, A. C. Davis, B. Li, H. A. Winther and G. Zhao (2013), *J. Cosmo. Astropart. Phys.*, **04**, 029 (2013).
- [43] E. Puchwein, M. Baldi and V. Springel, *Mon. Not. R. Astron. Soc.*, **436**, 348 (2013).
- [44] C. Llinares, D. F. Mota and H. A. Winther, *Astron. & Astrophys.* **562**, 78 (2014).
- [45] H. Wilcox, R. C. Nichol, G. B. Zhao, D. Bacon, K. Koyama and R. A. Kathy, *Mon. Not. R. Astron. Soc.*, **462**, 715 (2016).
- [46] H. A. Winther, F. Schmidt, A. Barreira, C. Arnold, S. Bose, C. Llinares, M. Baldi, B. Falck, W. A. Hellwing, K. Koyama, B. Li, D. F. Mota, E. Puchwein, R. E. Smith and G. B. Zhao, *Mon. Not. R. Astron. Soc.*, **454**, 4208 (2015).
- [47] G. Dvali, G. Gabadadze and M. Porrati, *Phys. Lett.* **B485**, 208 (2000).
- [48] P. Brax and P. Valageas, *Phys. Rev. D* **90**, 023507 (2014).
- [49] P. Brax and P. Valageas, *Phys. Rev. D* **90**, 023508 (2014).
- [50] A. I. Vainshtein, *Phys. Lett.* **B39**, 393 (1972).

- [51] C. de Rham, G. Gabadadze and A. J. Tolley, Phys. Rev. Lett. **106** 231101 (2011).
- [52] F. Sbisà, G. Niz, K. Koyama and G. Tasinato, Phys. Rev. **D86** 024033 (2012).
- [53] G. Chkareuli and D. Pirtskhalava, Phys. Lett. **B713**, 99 (2012).
- [54] A. Nicolis, R. Rattazzi and E. Trincherini, Phys. Rev. **D79**, 064036 (2009).
- [55] C. Deffayet, C. Esposito-Farese and A. Vikman, Phys. Rev. **D79**, 084003 (2009).
- [56] C. Burrage and D. Seery, J. Cosmo. Astropart. Phys., **08** 011 (2010).
- [57] N. Kaloper, A. Padilla and N. Tanahashi, JHEP **10** 148 (2011).
- [58] A. De Felice, R. Kase and S. Tsujikawa, Phys. Rev. **D85** 044059 (2012).
- [59] R. Kimura, T. Kobayashi and K. Yamamoto, Phys. Rev. **D85** 024023 (2012).
- [60] N. Chow and J. Khoury, Phys. Rev. **D80**, 024037 (2009).
- [61] F. P. Silva and K. Koyama, Phys. Rev. **D80**, 121301 (2009).
- [62] R. Gannouji and M. Sami, Phys. Rev. **D82**, 024011 (2010).
- [63] A. De Felice and S. Tsujikawa, Phys. Rev. Lett., **105**, 111301 (2010).
- [64] S. Nesseris, A. De Felice and S. Tsujikawa, Phys. Rev. **D82**, 124054 (2010).
- [65] A. Ali, R. Gannouji and M. Sami, Phys. Rev. **D82**, 103015 (2010).
- [66] P. Brax, C. Burrage and A. C. Davis, J. Cosmo. Astropart. Phys., **09**, 020 (2011).
- [67] S. A. Appleby and E. V. Linder, J. Cosmo. Astropart. Phys., **03**, 043 (2012).
- [68] S. A. Appleby and E. V. Linder, J. Cosmo. Astropart. Phys., **08**, 026 (2012).
- [69] A. Barreira, B. Li, C. M. Baugh and S. Pascoli, Phys. Rev. **D86**, 124016 (2012).
- [70] J. Neveu, V. Ruhlmann-Kleider, A. Conley, N. Palanque-DeLabrouille, P. Astier, J. Guy and E. Babichev, Astron. & Astrophys., **555**, 53 (2013).
- [71] A. Barreira, B. Li, A. Sanchez, C. M. Baugh and S. Pascoli, Phys. Rev. **D83**, 103511 (2013).
- [72] A. Barreira, B. Li, W. A. Hellwing, L. Lombriser, C. M. Baugh and S. Pascoli, J. Cosmo. Astropart. Phys., **04**, 029 (2014).
- [73] B. Falck, K. Koyama, G. B. Zhao and B. Li, J. Cosmo. Astropart. Phys., **07**, 058 (2014).
- [74] B. Falck, K. Koyama and G. B. Zhao, J. Cosmo. Astropart. Phys., **07**, 049 (2015).
- [75] H. Okada, T. Totani and S. Tsujikawa, Phys. Rev. **D87**, 103002 (2013).
- [76] M. Wyman, E. Jennings and M. Lima, Phys. Rev. **D88**, 084029 (2013).
- [77] Y. Zu, D. H. Weinberg, E. Jennings, B. Li and M. Wyman, Mon. Not. R. Astron. Soc., **445**, 1885 (2014).
- [78] A. Barreira, S. Bose, B. Li and C. Llinares, 2016, arXiv:1605.08436 [astro-ph.CO].
- [79] J. Neveu, V. Ruhlmann-Kleider, P. Astier, M. Besancon J. Guy, A. Moller and E. Babichev (2016), arXiv: 1605.02627 [astro-ph.CO].
- [80] A. Merloni, *et al.* (2012), arXiv:1209.3114 [astro-ph.CO] (<http://www.mpe.mpg.de/eROSITA/>).
- [81] R. Laureijs, *et al.*, Euclid Definition Study Report (2011), arXiv:1110.3193 [astro-ph.CO] (<http://desi.lbl.gov/>).
- [82] DESI Collaboration, A. Aghamousa, J. Aguilar, *et al.* (2016), arXiv:1611.00036 [astro-ph.CO].
- [83] DESI Collaboration, A. Aghamousa, J. Aguilar, *et al.* (2016), arXiv:1611.00037 [astro-ph.CO].
- [84] M. Levi, *et al.* (2013), arXiv:1308.0847 [astro-ph.CO] (<http://www.euclid-ec.org/>).
- [85] Z. Ivezić, *et al.* (2008), arXiv:0805.2366 (<http://www.lsst.org/>).
- [86] K. Koyama, Reports on Progress in Physics, **79**, 046902 (2016).
- [87] H. A. Winther and P. G. Ferreira, Phys. Rev. **D91**, 123507 (2015).
- [88] A. Barreira, S. Bose and B. Li, J. Cosmo. Astropart. Phys., **12**, 059 (2015).
- [89] B. Li and G. Efstathiou, Mon. Not. R. Astron. Soc., **421**, 1431 (2012).
- [90] G. B. Zhao, Astrophys. J. Suppl., **211**, 23 (2014).
- [91] A. Mead, J. Peacock, L. Lombriser and B. Li, Mon. Not. R. Astron. Soc., **452**, 4203 (2015).
- [92] M. Cataneo, D. Rapetti, L. Lombriser and B. Li (2016), arXiv:1607.08788 [astro-ph.CO].
- [93] S. Bose, W. A. Hellwing, & B. Li, J. Cosmo. Astropart. Phys., **2**, 034, (2015).
- [94] B. Jain, V. Vikram and J. Sakstein, Astrophys. J., **779**, 39 (2013).
- [95] V. Vikram, A. Cabre, B. Jain and J. T. Van der Plas, J. Cosmo. Astropart. Phys., **08**, 020 (2013).
- [96] L. Lombriser, Annalen der Physik, **526**, 259 (2014).
- [97] C. Burrage, & J. Sakstein (2016), arXiv:1609.01192 [astro-ph.CO].
- [98] A. Terukina, L. Lombriser, K. Yamamoto *et al.*, J. Cosmo. Astropart. Phys., **4**, 013 (2014).
- [99] H. Wilcox, D. Bacon, R. C. Nichol *et al.*, Mon. Not. R. Astron. Soc., **452**, 1171 (2015).
- [100] M. Cataneo, D. Rapetti, F. Schmidt, A. B. Mantz, S. W. Allen, D. E. Applegate, P. L. Kelly, A. von der Linden and R. Glenn Morris, Phys. Rev. **D92**, 044009 (2015).
- [101] X. Liu, B. Li, G. B. Zhao, M. C. Chiu, W. Fang, C. Pan, Q. Wang, W. Du, S. Yuan, L. Fu and Z. Fan, Phys. Rev. Lett., **117**, 051101 (2016).
- [102] Y. Cai, N. Padilla and B. Li, Mon. Not. R. Astron. Soc., **451**, 1036 (2015).
- [103] S. Peirone, M. Raveri, M. Viel, S. Borgani and S. Ansoldi (2016), arXiv:1607.07863 [astro-ph.CO].
- [104] H. Martel and P. R. Shapiro, Mon. Not. R. Astron. Soc., **297**, 467 (1998).
- [105] A. Brandt, Math. of Comp., **31**, 333 (1977).
- [106] V. Springel, S. D. M. White, A. Jenkins *et al.*. Nature (London), **435**, 629 (2005).
- [107] Planck Collaboration, P. A. R. Ade, N. Aghanim *et al.*, Astron. & Astrophys., **594**, A13 (2016).
- [108] E. Komatsu, K. M. Smith, J. Dunkley *et al.*, Astrophys. J. Suppl., **192**, 18 (2011).
- [109] S. Colombi, A. H. Jaffe, D. Novikov and C. Pichon, Mon. Not. R. Astron. Soc., **393**, 511 (2009).
- [110] R. Juszkiewicz, F. R. Bouchet, & S. Colombi, Astrophys. J., **412**, L9 (1993).
- [111] F. Bernardeau, Astrophys. J., **433**, 1 (1994).
- [112] W. A. Hellwing, B. Li, C. S. Frenk, S. Cole, Mon. Not. R. Astron. Soc., **435**, 2806 (2013).
- [113] M. A. Strauss, A. Yahil, M. Davis, J. P. Huchra, & K. Fisher, Astrophys. J., **397**, 395 (1992).
- [114] M. J. Chodorowski, & P. Ciecielag, Mon. Not. R. Astron. Soc., **331**, 133 (2002).
- [115] D. Shi, B. Li, J. Han, L. Gao, & W. A. Hellwing, Mon. Not. R. Astron. Soc., **452**, 3179 (2015).
- [116] A. Hammami, C. Llinares, D. F. Mota, & H. A. Winther, Mon. Not. R. Astron. Soc., **449**, 3635 (2015).
- [117] C. Arnold, V. Springel, & E. Puchwein, Mon. Not. R. As-

- tron. Soc., **462**, 1530 (2016).
- [118] B. Li, G. B. Zhao and K. Koyama, *J. Cosmo. Astropart. Phys.*, **05**, 023 (2013).
- [119] B. Li, A. Barreira, C. M. Baugh, W. A. Hellwing, K. Koyama, S. Pascoli and G. B. Zhao, *J. Cosmo. Astropart. Phys.*, **11**, 012 (2013).
- [120] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh and S. Pascoli, *J. Cosmo. Astropart. Phys.*, **10**, 027 (2013).
- [121] A. C. Davis, B. Li, D. F. Mota and H. A. Winther, *Astrophys. J.* **748**, 61 (2012)
- [122] C. Llinares & L. Pogosian, *Phys. Rev. D* **90**, 124041 (2014).