

International Journal of Modern Physics D
© World Scientific Publishing Company

Approximation Methods in Modified Gravity Models

Baojiu Li

*Institute for Computational Cosmology, Department of Physics, Durham University, Durham
DH1 3LE, UK
baojiu.li@durham.ac.uk*

Received Day Month Year

Revised Day Month Year

We review some of the commonly used approximation methods to predict large-scale structure formation in modified gravity (MG) models for the cosmic acceleration. These methods are developed to speed up the often slow N -body simulations in these models, or directly make approximate predictions of relevant physical quantities. In both cases, they are orders of magnitude more efficient than full simulations, making it possible to explore and delineate the large cosmological parameter space. On the other hand, there is a wide variation of their accuracies and ranges of validity, and these are usually not known a priori and must be validated against simulations. Therefore, a combination of full simulations and approximation methods will offer both efficiency and reliability. The approximation methods are also important from a theoretical point of view, since they can often offer useful insight into the nonlinear physics in MG models and inspire new algorithms for simulations.

Keywords: Keyword1; keyword2; keyword3.

PACS numbers:

1. Overview

Up until today, N -body simulations are the only tool to be able to predict the large-scale structure formation in the nonlinear regime with any desired accuracy. The need for full simulations in modified gravity (MG) models is even stronger because of the difficulty to model the effects of nonlinear physics using other approaches, which is why a great deal of effort has been devoted to the development of simulation algorithms and codes in the past decade or so. Our ability of simulating MG models has improved by orders of magnitude during this period (e.g., Refs. 1, 2, 3, 4, 5, see also the previous article⁶ in this Special Issue for review). It is with little doubt that advancements in supercomputing infrastructure and paradigm shift in programming design will allow further improvements in the coming years.

However, numerical simulations have their own limitations in practice. A major drawback, from this point of view, is their lack of efficiency, as is evident in the effort poured into making existing codes faster to cope with the ever increasing demand for numbers of independent realisations, volume and resolution, in order to match

2 *Baojiu Li*

those of current and future observational surveys. Unlike Λ CDM, which is unique in being widely accepted as a standard paradigm, there are many MG models, which makes it more difficult to allocate much effort to simulating individual models or performing a continuous parameter search. Furthermore, simulations are often used as a black box, making the underlying physics intractable, which does not help in developing reliable theoretical templates used in model constraints. Finally, although a simulation can predict structure formation down to very small scales, it may be that information from such small scales is not strictly necessary and trustable (e.g., due to the uncertainties in modelling baryonic physics).

This is certainly not to downplay the importance of simulations, which are not replaceable in many aspects, e.g., when we need accurate model predictions deep in the nonlinear regime and when the physics governing galaxy formation and evolution needs to be taken into account. However, if approximation methods can be developed for the MG models, and demonstrated to be valid in certain regimes, then it would be realistic to use observations in those regimes to make model constraints much more efficiently.

In this article we review some of the main approximation methods that have been developed in recent years to make predictions of the MG effect on large-scale structure formation. We will follow a logic of introducing methods that are closer to full simulations and then describing ones that can directly predict observables. The former cover fast, approximate simulations and rescaled simulations, which produce particle snapshots like full simulations, while the latter predict quantities based on simplified treatments of the underlying physics, such as perturbation theory and spherical collapse. It is worthwhile to point out, before we get into the details, that these approaches (especially the latter) usually have unknown validity regimes *a priori* or free parameters describing physical effects that are not part of the modelling, and these must be validated or calibrated using full simulations.

2. Approximate simulation techniques

In this section we briefly review some recent algorithm developments to allow approximate particle snapshots to be obtained in much more efficient ways than full MG simulations. We will describe three classes of methods in the order of moving further away from the full simulation technique and thus with increasing efficiency (and less well-controlled accuracy).

2.1. *Truncated refinement-level simulations*

This class of method is built upon the observation that many of the MG models of interest in cosmology have the screening property, which means that their deviation from GR is small in high-density regions. A major source of cost of modern adaptive-mesh-refinement (AMR) simulations is the use of refinements where the Poisson equation (in GR) and the MG partial differential equations (PDE)s have to be solved. Since screening happens exactly in those regions where AMR is needed, one

may ask whether a high resolution for the MG PDE is necessary: after all, if the fifth force is weak in high-density regions, then a small inaccuracy in it will probably have little impact on the simulation as a whole.

In Ref. 3 the authors propose a truncated simulations, where the standard Newtonian gravity is solved on all levels of the simulation mesh, while the scalar field that determines the fifth force is only solved on the domain grid (the lowest-resolution mesh that covers the whole simulation domain). The fifth force from the domain grid is then interpolated to the finer refinement levels. In this manner, both the fifth and the Newtonian forces are solved ‘accurately’, but to different accuracies as dictated by the fineness of the domain and most refined mesh, respectively. More explicitly, let G_{eff} be Newton’s constant that includes the effects of both the Newtonian (F_N) and the fifth (F_5) forces, then in the full and truncated simulations one has, respectively,

$$G_{\text{eff}}^{\text{full}} = 1 + \frac{F_5^{\text{full}}}{F_N^{\text{full}}}, \quad G_{\text{eff}}^{\text{trunc}} = 1 + \frac{F_5^{\text{trunc}}}{F_N^{\text{full}}}.$$

In dense regions, where the truncated fifth force has a fractional error ϵ and $F_5^{\text{full}} = \eta F_N^{\text{full}}$, with $\epsilon, \eta \ll 1$, one has $F_5^{\text{trunc}} = F_5^{\text{full}}(1 + \epsilon) = F_N^{\text{full}}\eta(1 + \epsilon) \approx \eta F_N^{\text{full}} = F_5^{\text{full}}$, and so $G_{\text{eff}}^{\text{trunc}} \approx G_{\text{eff}}^{\text{full}}$. The left panels of Fig. 1 show how the truncations at different refinement levels result in $G_{\text{eff}}^{\text{trunc}} \approx G_{\text{eff}}^{\text{full}}$ even for the most aggressive truncation (i.e., only solving the scalar field on the domain level). The right panel shows that the truncation has little impact on the matter power spectrum ($< 1\%$ at $k < 5h\text{Mpc}^{-1}$). The same has been found in most other observables (e.g., $< 3\%$ for halo abundance, mass and density profiles, $< 0.05\%$ for halo positions and velocities). The gain in simulation speed is over an order of magnitude, marking a significant improvement.

Although the idea of allowing lower accuracy in the fifth-force calculation in high-density regions sounds physical, and is shown to work well for Vainshtein-type screened scenarios such as the DGP model, one should be cautious in generalising it to other types of screened theories. Indeed, it was found⁴ that this approach works less well for chameleon models, e.g., $f(R)$ gravity, possibly because the Vainshtein mechanism is very efficient at screening the fifth force inside a matter clump, while the chameleon screening depends also on environments and works less well for matter clumps in underdense or intermediate-density regions. The symmetron mechanism has an explicit dependence on the local matter density and thus this approach may work well for it, but to date this has not been checked explicitly.

Another approach that follows a similar logic is that proposed earlier in Ref. 7. There, again the standard Newtonian force is solved accurately using AMR, but the fifth force is obtained in an approximate manner on all refinement levels including the domain one. More explicitly, because for chameleon and Vainshtein models the scalar field value can be calculated analytically for spherical tophat density profiles, they use this analytical approximation (with matter density estimated locally) to calculate the fifth force during the simulation. The agreements with full simulations are good at $k \leq 1h\text{Mpc}^{-1}$, with details depending on model and redshift.

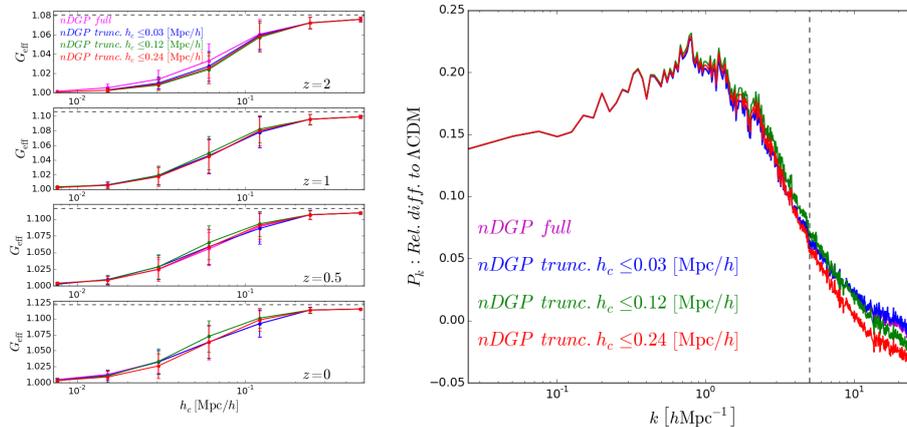
4 *Baojiu Li*


Fig. 1. *Left*: Total-force-to-normal-gravity ratio G_{eff} at particle positions as a function of the cell size of the finest refinement level the particles are in, for four redshifts ($z = 2, 1, 0.5, 0$) and for nDGP simulations truncated at four different refinement levels, as shown in the legends (e.g., $h_c \leq 0.03$ means that the scalar field is not directly solved but is interpolated from coarser levels on all refinement levels with cell size $\leq 0.03h^{-1}$ Mpc). The full curve (magenta) means there is no truncation. *Right*: The enhancements of the matter power spectrum wrt Λ CDM in the 4 truncated simulations as in the left panels. The vertical dashed line indicates $k = 5h\text{Mpc}^{-1}$, which is the k value below which the different N-body simulations codes compared by Winther et al. (2015)¹ have agreement better than 1%. The figure is adapted from Barreira et al. (2015).³

2.2. COLA

The truncated approach above boils down to sacrificing the accuracy of the fifth force calculation to gain improvement in efficiency. One can certainly take this logic one step further, and ask whether it is possible to give up the high accuracy of the Newtonian force calculation in return of further increases in performance.

Indeed, a majority of computer time in any AMR simulation is spent on the gravity solver and particle movements on refinements, which is particularly true for high-resolution simulations. Refinements are crucial for accurately resolving structures and substructures on small, highly-nonlinear, scales. But if one is only interested in mildly-nonlinear scales, e.g., the matter power spectrum down to $k \sim 0.3 h/\text{Mpc}$, then not using refinements at all may not be too bad an approximation.

In Λ CDM AMR simulations, the relaxation method is often used to solve the standard Poisson equation on refinements, partially because it is hard to make fast Fourier transform (FFT) work for the irregular-shaped refinement domains. If there is no refinement, the Poisson equation can be solved straightforwardly using FFT, which is efficient as the simulation becomes a pure particle-mesh (PM) one.

The situation is slightly worse for MG, because even for a cubic periodic box the nonlinear PDEs of these equations cannot be solved using FFT (or at least the

simple version of it), but instead has to be solved using relaxation. Nevertheless, by completely abandoning refinements one basically follows a more drastic approach than the truncated simulations described above, by calculating both the Newtonian and fifth forces accurately but with a lower resolution.

An alternative method in addition to pure PM simulations is the so-called Co-moving Lagrangian Acceleration (COLA) approach,⁸ which is also an approximation scheme. This method takes advantage of the fact that the second-order Lagrangian perturbation theory (2LPT) prediction of the evolution of particle trajectories, $\mathbf{x}_{2\text{LPT}}$, can be obtained with little computational effort. As a result, one can write the full particle position as $\mathbf{x} \equiv \mathbf{x}_{2\text{LPT}} + \delta\mathbf{x}$, where $\delta\mathbf{x}$ denotes the difference. The standard acceleration equation $\ddot{\mathbf{x}} = -\nabla\Phi$ then can be written as

$$\ddot{\delta\mathbf{x}} = -\nabla\Phi - \ddot{\mathbf{x}}_{2\text{LPT}}, \quad (1)$$

where now the known $\ddot{\mathbf{x}}_{2\text{LPT}}$ serves as an effective friction. This method is flexible and can be very efficient – 2LPT accuracy is guaranteed even if the simulation consists of a single time step – if one is interested in large scales only the simulation can afford to have much fewer time steps, but with increasing number of time steps small-scale structures can be recovered as well. If one is interested in mildly nonlinear scales and requires a large number of simulations, say, to estimate the covariance matrices of observables, then COLA can be an efficient substitute for full N -body simulations.

The COLA approach is extended to MG and massive neutrino scenarios in Refs. 9, 10, 11. The basic framework requires two major generalisations: (i) the 2LPT solution $\mathbf{x}_{2\text{LPT}}$ needs to be generalised to MG models, some of which involve scale-dependent linear growth factors (or scale-dependent Newton’s constant) and additional source terms in the Poisson equation; (ii) at nonlinear scales the screening effects should be included, which is achieved using spherical approximations to the fifth force as in Ref. 7.

Figure 2 shows that the MG-COLA approach is able to reproduce the MG effects in $f(R)$ and nDGP models to a few percent on length scales down to $k \sim \mathcal{O}(1)h\text{Mpc}^{-1}$ but with much fewer time steps and less computational time. Due to the approximate modelling of screening it is not straightforward to see how an increase of time steps improves the agreement with full simulations on all scales – one will perhaps need to use both more steps and accurate computation of the screening.

2.3. Rescaled simulations

All the fast simulation approaches described so far involve tracking of particles throughout a simulation, and only differ in how the forces are calculated or how the particles are followed. The method discussed in this subsection is different in that it does not follow the movements of particles during a period for all models. Instead, it only needs a single simulation – a ΛCDM one which is designed to have specific cosmological parameters and output times. These outputs of particle snapshots can

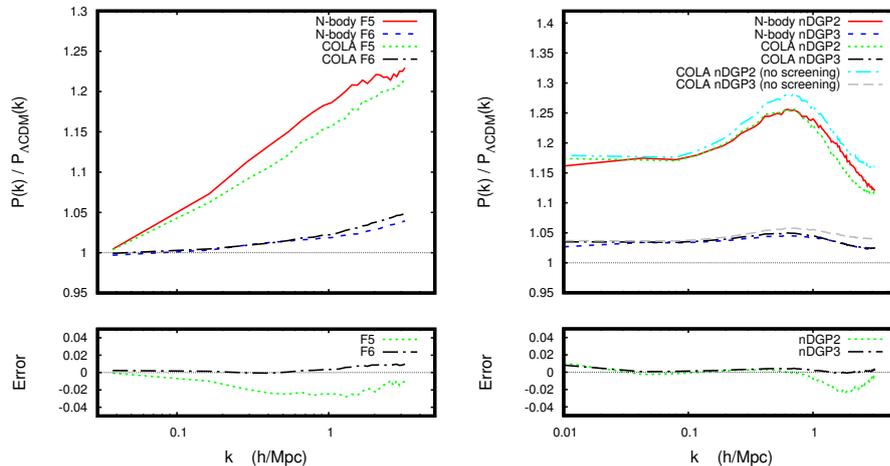


Fig. 2. *Left*: Comparisons between the predictions of the relative enhancement of the matter power spectra wrt Λ CDM, by MG-COLA and full N-body simulations for the Hu-Sawicki $f(R)$ model F5 and F6 (at $z = 0$). *Right*: The same as the left but for two nDGP models. The figure is adapted from Winther et al. (2017).¹⁰

then be ‘rescaled’ following some physically-motivated recipe to get an *approximate particle distribution* for a target MG model.

The basic idea is first described in Ref. 12, which shows that one can rescale an N -body particle snapshot to approximate simulation result of a Λ CDM model with different cosmological parameters, by changing the box size, particle mass and redshift of the original box to match the halo mass functions (or HMFs for short) in the two cosmologies, and then correcting the linear modes using the Zel’dovich approximation.

Consider a ‘target’ particle snapshot at redshift z' in a comoving box size of L' , which one hopes to obtain by rescaling an ‘original’ simulation particle snapshot at redshift z with box size $L = L'/s$ (we follow the convention that all quantities in the ‘target’ cosmology are primed, and s is the rescaling factor of box size and particle coordinates). The first step is to find the correct z, s , since the original simulation is a full simulation that needs to be run with these specifications. These are determined by minimising

$$\delta_{\text{rms}}^2(s, z; z') = \frac{1}{\ln R_2 - \ln R_1} \int_{R_1}^{R_2} \frac{dR'}{R'} \left[1 - \frac{\sigma(s^{-1}R', z)}{\sigma'(R', z')} \right]^2, \quad (2)$$

where $\sigma(R, z)$ is the linear density field variance at smoothing scale R at z :

$$\sigma^2(R, z) = \frac{1}{4\pi} \int_0^\infty dk k^2 P_{\text{lin}}(k) \tilde{W}^2(kR), \quad (3)$$

in which P_{lin} is the linear matter power spectrum of the corresponding cosmology and $\tilde{W}(kR)$ is the Fourier transform of the window function (taken as a real-space spherical tophat window of radius R). The minimisation of Eq. (2) can ensure that the linear fluctuation amplitudes in the original and target cosmologies are as close as possible over a scale range $[R_1, R_2] = [s^{-1}R'_1, s^{-1}R'_2]$ at redshift $z = z(z')$. In excursion set theory (see below), the halo mass function (HMF) depends on $\sigma(R, z)$, and so this is equivalent to minimising the differences in the HMFs of the target and original cosmologies in the mass range $[M(R_1), M(R_2)]$.¹²

The rescaling consists of a scaling of box size according to

$$L \rightarrow L' = sL, \quad (4)$$

and accordingly particle coordinates, a redshift mapping

$$z \rightarrow z', \quad (5)$$

and a particle mass rescaling

$$m_p \rightarrow m'_p = \frac{\Omega'_m H'^2 L'^3}{\Omega_m H^2 L^3} m_p, \quad (6)$$

where m_p is the simulation particle mass, and $H = H(z)$ (and $H' = H'(z')$) is the Hubble expansion rate at z (and z') for the given cosmology.

While the minimisation of Eq. (2) is designed so that the target and original simulations have matched power spectrum at scales where the dimensionless linear power spectrum $\Delta_{\text{lin}}^2(k) \equiv k^3 P(k)/(2\pi^2) \sim 1$, the matching of matter power spectra is not automatically guaranteed at much larger scales. To overcome this, one can add additional displacements to the particles in the scaled simulation, which accounts for corrections for the difference on large scales (defined as wavenumbers $k < R_{\text{nl}}^{-1}$ with R_{nl} given by $\sigma'(R_{\text{nl}}, z') = 1$), resulting in (where a subscript \mathbf{k}' means that the operation is in Fourier space):

$$\mathbf{x} \rightarrow \mathbf{x}' = s\mathbf{q} + D'_+(a')s\mathbf{d} + \delta\mathbf{x}'_{\mathbf{k}'}, \quad \delta\mathbf{x}'_{\mathbf{k}'} = \left[\left[\frac{\Delta_{\text{lin}}'^2(k', z')}{\Delta_{\text{lin}}^2(k = sk', z)} \right]^{1/2} - 1 \right] D'_+(a')s\mathbf{d}_{\mathbf{k}'}, \quad (7)$$

in which $D_+(a)\mathbf{d}(\mathbf{q})$ is the displacement field – the difference between the Eulerian ($\mathbf{x}(a)$ at time a) and the initial Lagrangian (\mathbf{q}) coordinates of a particle: $\mathbf{x}(a) = \mathbf{q} + D_+(a)\mathbf{d}(\mathbf{q})$ with $D_+(a)$ being the linear growth factor at a . In this approach, the phase of each Fourier mode of the density field $\delta(a) = -D_+(a)\nabla \cdot \mathbf{d}$ is preserved, while its amplitude is changed to match the target power spectrum. The peculiar velocity (again only on scales $k < R_{\text{nl}}^{-1}$) is corrected using the linear growth rate

$$\mathbf{v}_{\text{lin}, \mathbf{k}} \rightarrow \mathbf{v}'_{\text{lin}, \mathbf{k}'} = s \frac{a' \dot{D}'_+(a')}{a \dot{D}_+(a)} \frac{f'(k', a')}{f(k = sk', z)} \left[\frac{\Delta_{\text{lin}}'^2(k', z')}{\Delta_{\text{lin}}^2(k = sk', z)} \right]^{1/2} \mathbf{v}_{\text{lin}, \mathbf{k}'}, \quad (8)$$

8 *Baojiu Li*

where $\mathbf{v}_{\text{lin}} = a\dot{D}_+(a)\mathbf{d} = aD_+(a)H(a)f(a)\mathbf{d}$, and $f(a) = d \ln D_+(a)/d \ln a$ is the linear growth rate. The nonlinear velocities are rescaled as

$$\mathbf{v}_{\text{nl}} \rightarrow \mathbf{v}'_{\text{nl}} = \left[\frac{aL}{a'L'} \frac{m'_p}{m_p} \right]^{1/2} \mathbf{v}_{\text{nl}} = \left[\frac{\Omega'_m H'^2 L'^2 (1+z')}{\Omega_m H^2 L^2 (1+z)} \right]^{1/2} \mathbf{v}_{\text{nl}}, \quad (9)$$

which comes from a spherical analogy $v^2(r) \propto M(r)/r$.

In Ref. 12 it is found that the rescaling has excellent performance, able to reproduce the matter power spectrum of the target cosmology to better than 0.5% on scales $k < 0.1h\text{Mpc}^{-1}$ and 3% on $k < 1h\text{Mpc}^{-1}$, halo masses and concentrations to below 10%, and halo positions and velocities to better than $0.09h^{-1}\text{Mpc}$ and 5%.

The method can be improved to include a rescaling of halo internal structures¹³ and generalised to rescaling halo rather than particle catalogues.^{14,15} Its extension to modified gravity scenarios is also straightforward,¹⁶ with the linear power spectrum, growth rate and growth factor replaced by their modified gravity counterparts. Figure 3 shows that for ΛCDM and Hu-Sawicki $f(R)$ gravity (F4, F5, F6) it is capable of reproducing the nonlinear matter power spectrum and redshift monopole with 5% accuracy down to $k = 1h\text{Mpc}^{-1}$.

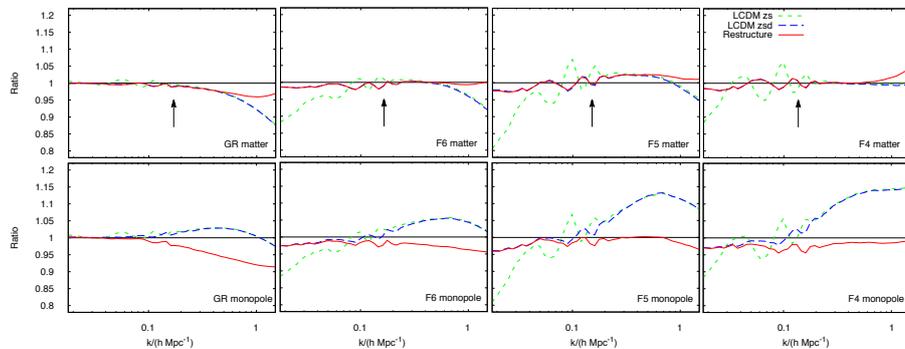


Fig. 3. *Upper panels:* The ratio between the predicted matter power spectra from rescaled and full simulations (at $z = 0$) for four models – from left to right ΛCDM , and Hu-Sawicki $f(R)$ model F6, F5, F4. The curves with different colours and line styles in each panel show the result of different stages of the rescaled simulation: rescaling redshift z and length s only (green shorted dashed), adding large-scale displacement field correction (blue long dashed) and full results by including also the rescaling of halo internal structures (red solid). *Lower panels:* The same as the upper panels for the the monopoles of the redshift space power spectra. This plot is adapted from Mead et al. (2015).¹⁶

3. Semianalytical methods for cosmological observables

The fast and approximate simulation approaches described in the previous subsection have one thing in common – they produce actual particle snapshots. This is particularly useful if mock catalogues of galaxies and clusters are needed, for which one can obtain halo catalogues from the particle distributions and follow recipes such as halo occupation distribution^{17,18} (HOD) to populate galaxies. Such catalogues are of great use when connecting theoretical models with observational data, the latter usually containing catalogues of galaxies.

In other situations, however, people are more interested in certain statistics that can be extracted from the large-scale distribution of galaxies, haloes or matter, the theoretical predictions for which do not necessarily require direct measurement from simulation. For example, when using cluster number counts to test a model, it is handy to have theoretical predictions of the counts (the simulation counterpart of which is the number of dark matter haloes as a function of halo mass, i.e., halo mass function, or HMF for short) which could be easily generated for arbitrary model parameter values; the theoretical HMF does not have to be measured from a simulation if it can be predicted in an alternative way.

Apparently, without tracking the full nonlinear physics of structure formation, these semi-analytical methods to calculate observables are approximate by nature. Their validity must be checked against full simulations and the physics unaccounted for in them are usually lumped together and described by free parameters which must then be calibrated by simulations. In this section we will discuss some of these methods. We will focus on three physical quantities that are fundamental to many observations – the HMF, matter power spectrum and two-point correlation function. For the latter two, there are approximate methods to predict them in both real and redshift spaces. We shall mention void abundance briefly in the end.

3.1. Halo mass functions

The Universe is known to host a ‘cosmic web’ of large-scale matter distribution, in which the majority of space has little matter and is occupied by the ‘cosmic voids’, whose boundaries are defined by rapid increases of matter and galaxy densities near filaments and sheets. Large nodes exist where filaments join, which are sites where large galaxy clusters are usually found. Such structures, seeded by the physics that took place at the primordial Universe and shaped by the late-time evolution under the action of gravity, are complicated and often require full simulations to model.

Yet, if we are interested in certain simple statistics of the cosmic web, it is not always necessary to go a great length to run simulations. One primary example is the abundance of dark matter haloes in the Universe, which can be analytically predicted by the simple Press-Schechter (PS) approach.¹⁹ The PS method was extended later into the excursion set theory²⁰ and has motivated various fitting functions of the HMF that are calibrated by simulations^{21–25} and used in theoretical studies.

A key premise, which is a reasonable physical assumption, in this approach is the

connection between the dark matter haloes found in a late-time simulation snapshot and the small density peaks produced in the early Universe as initial conditions. In the standard picture of cosmic structure formation, these initial density peaks grow by attracting more matter toward them through gravity, and thereby form the late-time matter clumps. To simplify the modelling, the density peaks are usually assumed to be spherical or ellipsoidal, so that their evolution toward singularities can be (semi)analytically followed. A specific point of this collapsing process, usually taken as the end of it, is associated to the time at which a dark matter halo is formed. Apparently, higher initial density peaks collapse earlier and so correspond to haloes which form at higher redshifts. Using this connection, the number counts of the haloes more massive than a given mass threshold can be obtained from the number densities of the density peaks higher than some threshold. Let's formulate this idea more quantitatively now, following the prescription of excursion set theory.²⁰

3.1.1. Spherical collapse and excursion set theory

Let's consider a spherical tophat overdensity described by its density contrast $\delta_i \equiv \rho_i/\bar{\rho}(a_i) - 1$, where a subscript i denotes the initial time and $\bar{\rho}(a_i)$ is the mean matter density at a_i . The initial comoving size of this spherical region is R . Since $\delta_i \ll 1$, ρ_i is approximately the mean density at a_i , so that the enclosed mass is $M(R) \approx \bar{\rho}(a_i)(a_i R)^3 = \bar{\rho}_{m0}R^3$ with $\bar{\rho}_{m0}$ the background matter density today. We assume that during the evolution of this region there is no shell crossing and the density profile remains a tophat, a consequence of which is that the enclosed mass remains constant. Defining $y(t) = r(t)/a(t)R$ where $r(t)$ is the physical size of this patch at time t , the nonlinear evolution in Λ CDM is described by

$$y'' + \left[2 - \frac{3}{2}\Omega_m(N)\right]y' + \frac{1}{2}\Omega_m(N)(y^{-3} - 1)y = 0, \quad (10)$$

where a prime is the derivative wrt $N = \ln(a)$ and $\Omega_m(N)$ is the matter density parameter at N . The initial condition is taken as $y(a_i) = 1 - \delta_i/3$ and $y'(a_i) = -\delta_i/3$.

A dark matter halo of mass $M(R)$ is said to have formed at z_f if $r(z = z_f) = 0$ or $y(z_f) = 0$, and z_f is the halo formation redshift. For this to happen, the initial density contrast δ_i has to be tuned to the correct value $\delta_{i,c}$ (which often involves trial-and-error). In the literature, $\delta_{i,c}$ is usually extrapolated to today according to linear perturbation theory: $\delta_c \equiv \delta_{i,c}D_+(a = 1)/D_+(a_i)$ with $D_+(a)$ the linear growth factor at a . We shall follow the same convention by defining δ_c as the linearly extrapolated value of $\delta_{i,c}$. Throughout this section, we always use the extrapolated density contrast in equations and figures, unless otherwise stated. Apparently, δ_c is in general a function of R and z_f , $\delta_c(R, z_f)$, but this can be equivalently written as $\delta_c(M, z_f)$ thanks to the relation between R and M mentioned above. In Λ CDM, δ_c is independent of R or M – this can be most straightforwardly seen from Eq. (10), which does not depend on R ; this does not hold true, for example, in the chameleon model, as we will see below.

To check if a spherical region with initial radius R or mass M has collapsed to form a dark matter halo by z_f , we only need to see if the initial density contrast inside it, $\delta(R)$, extrapolated to today (again using the linear growth rate), satisfies $\delta(R) \geq \delta_c(R, z_f)$. To calculate the abundance of these haloes, we need the probability distribution of $\delta(\mathbf{x}, R)$, which is the density contrast within a sphere of radius R centred at an arbitrary spatial location \mathbf{x} , given by

$$\delta(\mathbf{x}, R) = \int W(|\mathbf{x} - \mathbf{x}'|; R) \delta(\mathbf{x}') d^3\mathbf{x}' = \int \tilde{W}(k; R) \delta_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{x}} d^3\mathbf{k}, \quad (11)$$

where $W(r; R)$ is a filter or window function of radius R and $\tilde{W}(k; R)$ its Fourier transform, and $\delta_{\mathbf{k}}$ is the Fourier transform of $\delta(\mathbf{x})$. Assuming $\delta_{\mathbf{k}}$ satisfies a Gaussian distribution, statistically it can be fully specified by its power spectrum $P(k)$ given by $(2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') P(|\mathbf{k}|) = \langle \delta_{\mathbf{k}}(\mathbf{k}) \delta_{\mathbf{k}}^*(\mathbf{k}') \rangle$ where $\delta^{(3)}$ is the Dirac δ -function in 3D and $\langle \dots \rangle$ denotes ensemble average. In this case we have

$$\langle \delta^2(\mathbf{x}, R) \rangle \equiv \sigma^2(R) \equiv S(R) = \int 4\pi k^2 P(k) \tilde{W}(k; R) dk, \quad (12)$$

in which the variance $S(R) = \sigma^2(R)$ is defined, with $\sigma(R)$ the root-mean-squared (rms) density fluctuation within spherical window of size R . We still have the freedom to choose the smoothing window function. If W is a sharp k -space window,

$$\tilde{W}(k; R) = \Theta(1 - kR), \quad (13)$$

where $\Theta(x)$ is the Heaviside function with $\Theta(x \geq 0) = 1$ and $\Theta(x < 0) = 0$, then in Eq. (11) any change of $\delta(\mathbf{x}, R)$ due to change of $R \rightarrow R - dR$ comes from extra k -modes that are newly brought into the integration. Because different k -modes in the linear spectrum are independent, the increment of $\delta(\mathbf{x}, R)$ due to an increase in R does not depend on the value of $\delta(\mathbf{x}, R)$ at the previous value of R (the Markov property). Note that, given initial power spectrum $P(k)$, $S(R)$ is a fixed monotonic decreasing function of R , and can be used interchangeably with R, M .

Because of its above properties, one can regard the change of $\delta(\mathbf{x}, R)$ with R or S as a Brownian motion, as schematically shown in Figure 4. $S = 0$ corresponds to $R \rightarrow \infty$ so that $\delta(\mathbf{x}, R) \rightarrow 0$ is the overdensity of the whole Universe, which means that the Brownian motion always starts at the origin (not shown in Figure 4). As S increases, we are looking at smaller and smaller R , and $\delta(\mathbf{x}, R)$ fluctuate more and more strongly. S is effectively a ‘time’ variable for the Brownian motion.

The Gaussian smoothed density field $\delta(\mathbf{x}, R)$, or equally $\delta(\mathbf{x}, S)$ satisfies a Gaussian probability distribution

$$P(\delta, S) d\delta = \frac{1}{\sqrt{2\pi S}} \exp\left[-\frac{\delta^2}{2S}\right] d\delta. \quad (14)$$

At a given \mathbf{x} , a virialised dark matter halo within the mass range $[M - dM, M]$ forms at z_f *if and only if* the Brownian motion makes its *first* crossing of the barrier $\delta_c(M, z_f)$ (which is constant in M or S for Λ CDM; see the horizontal dashed line in Figure. 4) within $[S, S + dS]$. By its nature, the Brownian motion will cross this

12 *Baojiu Li*

barrier infinitely many times at larger S which correspond to collapsed structures of smaller masses, but those are considered as the internal structures rather than main haloes. Similarly, if the Brownian motion crosses the barrier at smaller S , then the structure we look at is a substructure of some more massive halo. The probability density for this first crossing to take place inside $[S, S + dS]$ is given by

$$f(S, z_f) dS = \frac{\delta_c(z_f)}{\sqrt{2\pi} S^{3/2}} \exp\left[-\frac{\delta_c^2(z_f)}{2S}\right] dS, \quad (15)$$

and so the number density of haloes in the mass range $[M - dM, M]$ at z_f is

$$\frac{dn(M)}{dM} dM = \frac{\bar{\rho}_m(z_f)}{M} f(S, z_f) dS, \quad (16)$$

where $\bar{\rho}_m(z_f)$ is the mean matter density at z_f .

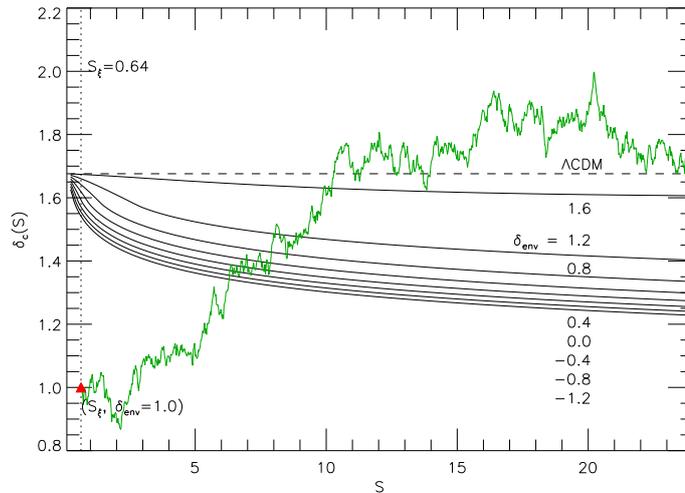


Fig. 4. An illustration of how the excursion set theory works. See the main text for a description. This figure is adapted from Li & Efstathiou (2012).²⁶

While Eq. (16) is generic, Eq. (15) has been derived under the assumption that the critical density for collapse, $\delta_c(M, z_f)$, only depends on z_f and not on M , R or S . In other words, the barrier that the considered Brownian motion has to cross is flat. Also, thanks to the spherical symmetry used in the derivation, $\delta_c(M, z_f) = \delta_c(z_f)$ does not depend on the large scale environment. Both assumptions are model specific and might not be valid in MG models, as we shall see shortly.

3.1.2. Halo bias

Dark matter haloes are biased tracers of the underlying matter density field, which means that the halo number density contrast $\delta_h = n_h/\bar{n}_h - 1$, where n_h is the local halo number density and \bar{n}_h its mean, is generally different from the matter density contrast δ . The two are related by the halo bias b , which depends on the halo mass (high mass haloes are more biased with $b > 1$), redshift and length scales considered.

Consider a halo of mass M which forms at z_f , and use the picture of a Brownian motion which crosses $\delta_c(z_f)$ between $[S, S + dS]$ with $S = S(M)$. This time let us assume that the halo forms in a large scale environment with initial density contrast δ_{env} and an initial comoving size ξ . In Figure 4, this is equivalent to saying that the Brownian motion has passed through the point $(S_\xi, \delta_{\text{env}})$, the red triangle, before crossing the barrier $\delta_c(z_f)$, in which S_ξ is the value of S corresponding to $R = \xi$ and δ_{env} is the initial density contrast enclosed in ξ extrapolated to today using the linear growth factor. Note that this necessarily means that $S_\xi < S$ and $\delta_{\text{env}} < \delta_c(z_f)$ as otherwise the environment itself would have collapsed to form a bigger halo.

The halo number density contrast satisfies²⁷

$$(1 + \delta_h)f(S, z_f)dS = (1 + \delta_{\text{env}}^{z_f})f(S, z_f; S_\xi, \delta_{\text{env}})dS, \quad (17)$$

where $\delta_{\text{env}}^{z_f}$ is the linearly-extrapolated density contrast of the environment at z_f , and $f(S, z_f; S_\xi, \delta_{\text{env}})$ is the *conditional* probability density for the Brownian motion that has first passed through $(S_\xi, \delta_{\text{env}})$ to cross $\delta_c(z_f)$ at S , which for a flat δ_c is:

$$f(S, z_f; S_\xi, \delta_{\text{env}}) = \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z_f) - \delta_{\text{env}}}{(S - S_\xi)^{3/2}} \exp\left[-\frac{(\delta_c(z_f) - \delta_{\text{env}})^2}{2(S - S_\xi)}\right]. \quad (18)$$

Let us take the limit of large environmental regions, with $S_\xi \ll 1$ and $|\delta_{\text{env}}| \ll 1$, as a concrete example. With a Taylor expansion²⁸

$$\delta_h = \sum_{n=0}^{\infty} \delta^n \frac{b_n}{n!} = b_0 + b_1 \delta_{\text{env}}^{z_f} + \dots \quad (19)$$

where ellipsis represents higher-order terms neglected as a result of $\delta \ll 1$, we find b_1 as

$$b_1 = \left. \frac{d\delta_h}{d\delta_{\text{env}}^{z_f}} \right|_{\delta_{\text{env}}^{z_f}=0} = 1 + \frac{d\delta_{\text{env}}}{d\delta_{\text{env}}^{z_f}} \frac{f(S, z_f)dS}{d\delta_{\text{env}}} = 1 + g(z_f) \frac{\delta_c^2/S - 1}{\delta_c}, \quad (20)$$

in which we have used Eqs. (15, 17, 18), neglected the z_f -dependence of $\delta_c(z_f)$ and defined $g(z) \equiv d\delta_{\text{env}}/d\delta_{\text{env}}^{z_f} = D_+(z=0)/D_+(z_f)$. As the coefficient of the linear term of the Taylor expansion, b_1 is called linear halo bias. Eq. (20) makes explicit that b_1 depends on z_f and the halo mass M (or equivalently S). Note that the zeroth order term b_0 in the Taylor expansion vanishes as $\delta_h \rightarrow 0$ when $S_\xi \rightarrow 0$.

3.1.3. Ellipsoidal collapse

The prediction of excursion set theory, Eqs. (15, 16), provides a good starting point to model HMFs and generally shows qualitative agreement with simulations. However, it is not surprising that such a simple approach does not fully match simulation HMFs, especially at the low-mass end (where it overestimates halo abundance).

Refs. 29, 30 suggest that the spherical collapse model used in excursion set theory is far from realistic, and should be replaced with a model in which the initial regions that form haloes at late times have ellipsoidal shapes. Unlike spherical collapse, the collapse of an ellipsoidal region depends on the surrounding shear field and the resulting critical density for collapsing at z_f now depends on the size of the region (or the mass enclosed in it), $\delta_{ec}(M, z_f)$ or $\delta_{ec}(S, z_f)$:

$$\delta_{ec}(S, z_f) = \sqrt{q}\delta_c(z_f) \left[1 + \beta \left(\frac{\delta_c(z_f)}{S} \right)^{-\alpha} \right], \quad (21)$$

in which $\alpha \approx 0.615$, $\beta \approx 0.485$ are derived from ellipsoidal collapse dynamics ($\alpha = \beta = 0$ in the spherical collapse limit), and $q \approx 0.7$ from normalising to simulations. It can be seen that as S increases, the barrier becomes higher, which means that it is harder for the Brownian motion to cross, leading to fewer low-mass haloes.

The first crossing probability density $f(S, z_f)$ in this case can be written as

$$f(S, z_f) = \frac{\sqrt{q}}{\sqrt{2\pi}} A \left[1 + q \frac{\delta_c^2}{S} \right]^{-p} \frac{\delta_c}{S^{3/2}} \exp \left[-\frac{q}{2} \frac{\delta_c^2}{S} \right], \quad (22)$$

where $p = 0.3$ and A is a normalisation factor which is determined by requiring that the integration of $f(S, z_f)$ over $S \in [0, \infty)$ is 1. In this equation we have again not explicitly written the z_f -dependence of $\delta_c(z_f)$ for simplicity.

Eq. (16) can still be used to predict the halo abundance, with now the use of the $f(S, z_f)$ from Eq. (22). Meanwhile, the linear halo bias becomes

$$b_1 = 1 + g(z_f) \left[\frac{q\delta_c^2/S - 1}{\delta_c} + \frac{2p/\delta_c}{1 + (q\delta_c^2/S)^p} \right], \quad (23)$$

where $g(z_f)$ is the same as before and we again omit the z_f -dependence of δ_c .

3.1.4. Excursion set theory in modified gravity

The physical picture of excursion set theory and its generalisations to ellipsoidal collapse hold for MG models. However, due to the fifth force, the nonlinear evolution of a spherical or ellipsoidal tophat overdensity is in general different from the Λ CDM result discussed above. This means that the critical initial density required for such an overdensity to collapse at z_f would be different from the Λ CDM predictions. Indeed, although for Λ CDM spherical collapse δ_c depends only on z_f , in a MG model it can depend on z_f, M and δ_{env} , as happens for chameleon-type theories. Because the analysis for Vainshtein-type models is simpler, in this subsection we shall focus on chameleons and only briefly mention some results for Vainshtein.

Before moving to the details, let us remark that even in the study of MG models, we shall again extrapolate initial density contrasts to today using the linear growth factor of Λ CDM. This might sound counterintuitive, but one has to bear in mind that the linear extrapolation does not change anything in the initial density field apart from rescaling it by *some* constant. Strictly speaking, the extrapolation is not necessary as in the spherical or ellipsoidal collapse model what are needed are the distribution of *initial* density peaks and how they evolve *nonlinearly* under the action of gravity: linear theory has no role to play here apart from setting the initial condition. From a practical viewpoint, the extrapolation is convenient (for example, with it we will be dealing with values of $\delta_c(z_f)$ of order unity, and different people will agree on its value even though they would generally have different values of the linear density contrast at their *own, different*, initial redshifts, e.g., $z_i = 50, 100$) but not necessary. Unless otherwise stated, when talking about linear density contrast, we always mean the one linearly extrapolated to $z = 0$ using Λ CDM cosmology.

For chameleon models, the evolution equation Eq. (10) has to be modified to include the fifth force. Taking $f(R)$ gravity as an example, we have²⁶

$$y_h'' + \left[2 - \frac{3}{2}\Omega_m(N) \right] y_h' + \frac{1}{2}\Omega_m(N) (y_h^{-3} - 1) y_h \left[1 + \frac{1}{3} \min \left\{ 1, 3 \frac{\Delta r}{r} \right\} \right] = 0, \quad (24)$$

in which we have now used y_h instead of y , to denote that this is $r(t)/a(t)R$ for a halo under consideration. The expression in the squared brackets multiplied to the last term incorporates the effects of the fifth force. In $f(R)$ gravity, the fifth force can be between 0 and 1/3 the strength of the Newtonian gravity, but its maximum strength relative to standard gravity can be different in general chameleon models, for which the term in these square brackets should be changed accordingly. Here the fifth force effect has been modelled under the thin-shell approximation,^{31,32} with $\Delta r/r$ to be defined shortly. A more accurate expression for the fifth force may also be used:^{33,34} $\frac{1}{3} \min \left\{ 1, 3 \frac{\Delta r}{r} - 3 \left(\frac{\Delta r}{r} \right)^2 + \left(\frac{\Delta r}{r} \right)^3 \right\}$.

To be more concrete, we assume that the initial region that forms a virialised dark matter halo at z_f is spherical with an initial comoving radius R and constant initial density δ_c (extrapolated to today using Λ CDM linear growth factor) inside. The spherical collapse takes place inside another, much larger, spherical region with initial comoving radius ξ and constant initial density δ_{env} . The latter is called the environment, and is used to model the environmental screening of the fifth force.

As the inner spherical overdensity evolves, the environment itself also evolves (either expands or contracts, depending on the sign of δ_{env}). We have to follow the co-evolution of these two systems. For the former we use Eq. (24) and for the latter we assume Λ CDM spherical evolution, similar to Eq. (10):

$$y_{\text{env}}'' + \left[2 - \frac{3}{2}\Omega_m(N) \right] y_{\text{env}}' + \frac{1}{2}\Omega_m(N) (y_{\text{env}}^{-3} - 1) y_{\text{env}} = 0, \quad (25)$$

where a subscript _{env} is used to distinguish from y_h in Eq. (24). The use of Λ CDM equation for the evolution of the environment is justified if the initial comoving

radius ξ is much larger than the range of the fifth force (or the Compton wavelength of the scalar field mediating the force). Meanwhile, ξ cannot be too big for it to be a faithful environmental definition (if ξ is too large then the environmental density will simply be the cosmic mean). In Ref. 26 $\xi = 8h^{-1}\text{Mpc}$ is adopted, though later we will briefly mention other possibilities explored in the literature. Again, $\delta_{\text{env}} < \delta_c$ so that the environment itself has not collapsed to form virialised halo by z_f .

At any late redshift $z \geq z_f$, the co-evolving system consists of a small spherical tophat with uniform density $\rho_h(z) = \rho_{m0}y_h^{-3}(1+z)^3$, embedded in a larger one with density $\rho_{\text{env}}(z) = \rho_{m0}y_{\text{env}}^{-3}(1+z)^3$. Let $f_{R,\text{env}}$ and $f_{R,h}$ be the values of the scalar field f_R inside constant density fields with $\rho = \rho_{\text{env}}$ and $\rho = \rho_h$ respectively, and Ψ_N be the Newtonian potential of the inner spherical patch at its edge r , then as described in the first article³⁵ of this Special Issue, if $|f_{R,\text{env}} - f_{R,h}| \ll |\Psi_N|$, the scalar field can go from its external environmental value $f_{R,\text{env}}$ to its internal halo value $f_{R,h}$ within a distance Δr inside the edge of inner patch. This is known as the thin-shell^{31,32} which for $f(R)$ gravity can be expressed as:

$$\frac{\Delta r}{r} = \frac{f_{R,\text{env}} - f_{R,h}}{2\Psi_N}, \quad (26)$$

Using $\Psi_N = -\frac{4\pi G}{3}\rho_h(z)r^2 = -\frac{4\pi G}{3}\rho_{m0}R^2y_h^{-1}(1+z)$ and

$$f_{R,h} = f_{R0} \left[\frac{\rho_h(z) + 4\frac{\Omega_\Lambda}{\Omega_m}}{\rho_{m0} + 4\frac{\Omega_\Lambda}{\Omega_m}} \right]^{-2}; \quad f_{R,\text{env}} = f_{R0} \left[\frac{\rho_{\text{env}}(z) + 4\frac{\Omega_\Lambda}{\Omega_m}}{\rho_{m0} + 4\frac{\Omega_\Lambda}{\Omega_m}} \right]^{-2} \quad (27)$$

for Hu-Sawicki $f(R)$ model with $n = 1$ in Eq. (26), we can express $\Delta r/r$ in terms of $z, R, y_h, y_{\text{env}}$ and cosmological parameters such as $H_0, \Omega_m, \Omega_\Lambda$. Hence, Eqs. (24, 25) are coupled differential equations which need to be solved together. As in the standard case, through trial-and-error we can find the correct value of the initial density contrast for the inner patch to collapse at z_f , extrapolated to today using linear growth factor of ΛCDM . The result is written as $\delta_c(S, z_f, \delta_{\text{env}})$, where the dependence on S, δ_{env} is because the thin-shell condition above depends on R, y_{env} . Recall that δ_{env} is the initial density contrast in the environment that is extrapolated to today using the linear growth factor of ΛCDM – it is a single number rather than a function of time such as y_{env} , and so more convenient to use (as the initial size of the environment ξ is fixed, given δ_{env} we can calculate y_{env} at arbitrary times).

The solid lines in Figure 4 are $\delta_c(S, z_f = 0, \delta_{\text{env}})$ for a selection of δ_{env} values as indicated in the legend. Two physical features are observed: (1) the barrier δ_c is lower at larger S (corresponding to smaller mass), meaning that the first crossing of Brownian motions is more likely to take place at smaller S , therefore larger haloes are more likely to form; (2) in higher-density environments (larger δ_{env}), the barrier is closer to the ΛCDM result, a feature of environmental screening.

The dependence of $\delta_c(S, z_f, \delta_{\text{env}})$ on S makes it impossible to find analytical solutions to the first-crossing probability $f(S, z_f, \delta_{\text{env}})$ for a given δ_{env} . Furthermore, the dependence on δ_{env} means that, unlike in ΛCDM , there is no unique barrier but

infinitely many. These both add complications to the solution. In practice, one is interested in the overall, or environment-averaged, first-crossing probability density:

$$f(S, z_f) = \int_{-\infty}^{\delta_c^{\Lambda\text{CDM}}} q(\delta_{\text{env}}, \delta_c^{\Lambda\text{CDM}}, S_\xi) \times f(S, z_f; S_\xi, \delta_{\text{env}}) d\delta_{\text{env}}, \quad (28)$$

in which $f(S, z_f; S_\xi, \delta_{\text{env}})dS$ is the conditional probability that a Brownian motion that has passed through $(S_\xi, \delta_{\text{env}})$ will cross the barrier $\delta_c(S, z, \delta_{\text{env}})$ between $[S, S + dS]$, which can be calculated numerically.³⁶ $q(\delta_{\text{env}}, \delta_c^{\Lambda\text{CDM}}, S_\xi)$ is the distribution of δ_{env} , or the probability density that a Brownian random motion passes through $(S_\xi, \delta_{\text{env}})$ and never exceeds $\delta_c^{\Lambda\text{CDM}}$ at $S \leq S_\xi$ (as otherwise the environment itself would have collapsed). As the environment follows a ΛCDM evolution, we have²⁰

$$q(\delta_{\text{env}}, \delta_c^{\Lambda\text{CDM}}, S_\xi) = \frac{1}{\sqrt{2\pi S_\xi}} \left\{ \exp\left[-\frac{\delta_{\text{env}}^2}{2S_\xi}\right] - \exp\left[-\frac{(\delta_{\text{env}} - 2\delta_c^{\Lambda\text{CDM}})^2}{2S_\xi}\right] \right\}, \quad (29)$$

for $\delta_{\text{env}} \leq \delta_c^{\Lambda\text{CDM}}$ and 0 otherwise. The integration in Eq. (28) can be performed numerically. Note that halo bias can also be calculated in this approach, though the calculation will be more involved and has not been done so far.

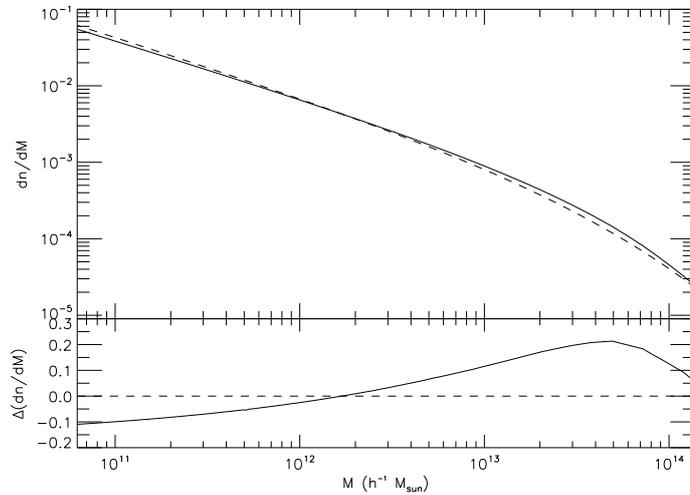


Fig. 5. An example showing the prediction of the HMF in chameleon model from the extended excursion set theory (see text for more details). This figure is adapted from Li & Efstathiou (2012).²⁶

Once $f(S, z_f)$ is obtained, the calculation of the HMF follows the same expression as Eq. (16). Figure 5 shows an example of the derived HMF for a chameleon model (solid) and a ΛCDM model which only differs in the absence of a fifth force

(dashed). The bottom panel shows the relative difference between the two models, which shows the expected result that they predict the same number of massive haloes, but the chameleon model predicts more intermediate-mass haloes due to the enhancement of gravity. Because more small haloes have merged to form intermediate-mass haloes in the chameleon model, fewer are left at the low-mass end.

The method above has gone a great length in modelling the fifth force effect on the spherical collapse, especially its environmental dependence. Given that spherical collapse is at best an approximation, this part can be simplified. For example, one can eliminate the need for integrating over an environment distribution in Eq. (28) by using an average environment $\bar{\delta}_{\text{env}}$ given by $\int \delta_{\text{env}} q(\delta_{\text{env}}, \delta_c^{\Lambda\text{CDM}}, S_\xi) d\delta_{\text{env}}$ to compute a mean barrier^{33,34} $\bar{\delta}_c(S, z_f, \bar{\delta}_{\text{env}})$, which will be a single barrier to calculate the first-crossing probability. Alternatively, one can calculate the mean of $\delta_c(S, z_f, \delta_{\text{env}})$ as $\bar{\delta}_c(S, z_f) = \int \delta_c(S, z, \delta_{\text{env}}) q(\delta_{\text{env}}, \delta_c^{\Lambda\text{CDM}}, S_\xi) d\delta_{\text{env}}$ in which one again ends up with a single barrier. An even simpler treatment is to consider two limiting cases of the fifth force – zero and 1/3 of the strength of standard gravity – which will lead to band of possible values of the HMF.³⁷

On the other hand, the modelling of environment above can also be improved. The environments were taken as spherical regions whose initial comoving (or Lagrangian) radius is $\xi = 8h^{-1}\text{Mpc}$, and thus are called *Lagrangian environments*. However, depending on the initial density inside an environment, its size can become much larger or smaller than $8h^{-1}\text{Mpc}$ at late times, while a good definition of environment should be neither (if the size of the environment is too small it is dominated by the halo under consideration; if the size is too large is simply approaches the cosmological background; neither is a good definition for our modelling). In the mean time, it is known³⁸ that the thin shell model works very well if the environment has a fixed Eulerian size of order $5 \sim 8h^{-1}\text{Mpc}$, and therefore it is physically sensible to define environments as regions whose Eulerian size is similar to the Compton wavelength of the scalar field. Refs. 39, 40 define the environments to have a physical (Eulerian) size of $\zeta = 5h^{-1}\text{Mpc}$ at z_f , and the calculation then proceeds as in the case with Lagrangian environment. The main practical difference is in the way the physical density in the environment is calculated: instead of solving Eq. (25), the nonlinear density contrast in the environment, at arbitrary redshift z , $\Delta_{\text{NL}}(z)$, is given approximately as^{41,42}

$$\Delta_{\text{NL}}(z) = \left[1 - \frac{\delta_{\text{lin}}(z)}{\delta_c^{\Lambda\text{CDM}}} \right]^{-\delta_c^{\Lambda\text{CDM}}}, \quad (30)$$

in which $\delta_{\text{lin}}(z)$ is the initial density contrast extrapolated to redshift z using the linear growth rate of ΛCDM . The probability distribution of δ_{env} is also different because all environments now have the same final Eulerian radius $\zeta = 5h^{-1}\text{Mpc}$, and some approximate analytical expressions can be found in Ref. 43. For details of the implementation and result of this method see Refs. 39, 40.

The fifth force effect can be more easily modelled in other MG (e.g., Vainshtein)

models, in which case the spherical collapse equation can be written as

$$y_h'' + \left[2 - \frac{3}{2}\Omega_m(N) \right] y_h' + \frac{1}{2}\Omega_m(N) (y_h^{-3} - 1) y_h [1 + \Gamma(N, y_h)] = 0, \quad (31)$$

in which $\Gamma(N, y_h) = \Gamma(z, \rho_h(z))$ incorporate the fifth force effect,^{44–46} which depends only on the redshift z and the density inside the spherical tophat at z . There is no dependence on δ_{env} or R (or equivalently halo mass M), and the resulting $\delta_c(z_f)$ is a single flat barrier just like in Λ CDM. The first crossing probability can be obtained analytically and so the HMF (and halo bias⁴⁵) is straightforward to find.

An approximation used in the above method to calculate $\delta_c(S, z_f, \delta_{\text{env}})$ is that haloes have a tophat density profile throughout their evolution history. It is known, however, that an initial tophat overdensity generally develops nontrivial inner profiles during spherical collapse in $f(R)$ gravity,⁴⁷ which increases with radius and has a spike at the surface. Neither of this and a tophat reflects the true shapes of dark matter haloes in simulations, which is why in the references mentioned above a tophat profile is assumed for its simplicity. An alternative method to calculate the collapse threshold was proposed in Ref. 48, where haloes are assumed to take mean initial density profiles by smoothing an initial Gaussian random field, and the density profiles are numerically evolved until collapse allowing the shape to evolve as well. The resulting collapse threshold, for a given model, depends on z and the mass enclosed in the halo, M ; the effect of the environment is not explicitly included in the calculation, and it is suggested that the use of mean density profiles implicitly accounts for that. The method predicts the same behaviour that δ_c decreases as the halo mass M decreases, and once δ_c is available the calculation of the first-crossing probability and the HMF is the same.

3.1.5. Simulation calibrations

The modified excursion set theory with spherical collapse represents a greatly simplified picture of the complicated nonlinear physics of MG, with assumptions such as sphericity and tophat density profile. Deviations from these assumptions affect the accuracy of predictions of the HMF already in Λ CDM, and they further affect the reliability of the thin shell calculation of the fifth force and chameleon screening.^{49–51} One way to improve on this, instead of making the semi-analytical modelling more complicated and intractable, is to include parameterised corrections for which the parameters are calibrated by numerical simulations for a selected few models, and expect these calibrated fitting formulae to work for general models.

Ref. 52 proposes to incorporate all the unaccounted-for physics in the spherical model to obtain a ‘corrected’ critical density (or barrier) for collapse, δ_c^{cor} :

$$\delta_c^{\text{cor}} = \epsilon(S, z_f; M_{\text{th}}^{(1)}, M_{\text{th}}^{(2)}, \eta, \vartheta, \chi) \delta_c(S, z_f, \delta_{\text{env}}^{\text{peak}}), \quad (32)$$

where $\delta_c(S, z_f, \delta_{\text{env}})$ is the environment-dependent barrier calculated using the most

probable value of $\delta_{\text{env}}, \delta_{\text{env}}^{\text{peak}}$, for simplicity. ϵ is a correction function given by

$$\epsilon(S, z_f; M_{\text{th}}^{(1)}, M_{\text{th}}^{(2)}, \eta, \vartheta, \chi) = \frac{1 + \left[\frac{M}{M_{\text{th}}^{(1)}} \right]^{\eta} \left[\frac{\delta_c^{\Lambda\text{CDM}}(S, z_f)}{\delta_c(S, z_f, \delta_{\text{env}}^{\text{peak}})} \right]^{\chi} + \left[\frac{M}{M_{\text{th}}^{(2)}} \right]^{\vartheta} \frac{\delta_c^{\Lambda\text{CDM}}(S, z_f)}{\delta_c(S, z_f, \delta_{\text{env}}^{\text{peak}})}}{1 + \left[\frac{M}{M_{\text{th}}^{(1)}} \right]^{\eta} + \left[\frac{M}{M_{\text{th}}^{(2)}} \right]^{\vartheta}}, \quad (33)$$

where $M = M(S)$ is the halo mass, and $\eta, \vartheta, \chi, M_{\text{th}}^{(1)}, M_{\text{th}}^{(2)}$ are free parameters. ϵ goes to 1 for strongly screened theories where $\delta_c(S, z_f, \delta_{\text{env}}^{\text{peak}}) \rightarrow \delta_c^{\Lambda\text{CDM}}(S, z_f)$.

Using Hu-Sawicki $f(R)$ gravity as an example, Ref. 52 discovers an empirical expression for χ , $\chi = 0.5 - 0.2 \log_{10}(-10^5 f_{R0})$, and fits $\eta, \vartheta, M_{\text{th}}^{(1)}, M_{\text{th}}^{(2)}$ as functions of f_{R0} and z_f using a suite of Gpc-box $f(R)$ simulations. The resulting δ_c^{cor} is then used in the Sheth-Tormen fitting formula of the HMF.

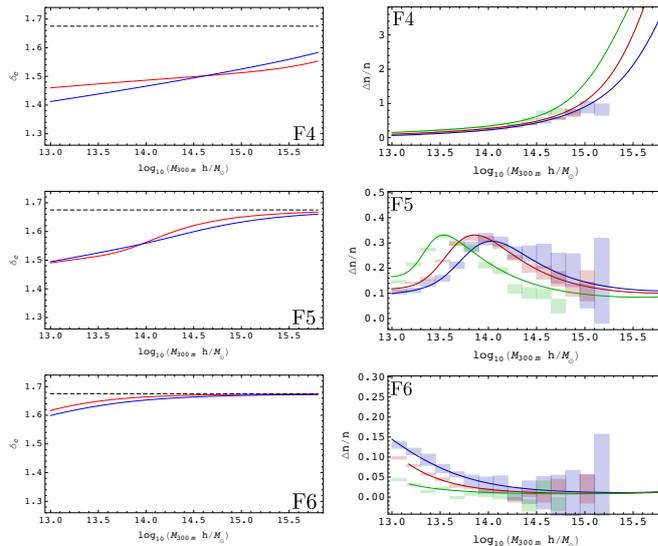


Fig. 6. An exempling showing how the simulation calibration improves the performance in predicting the HMFs for Hu-Sawicki $f(R)$ models. This figure is adapted from Cataneo et al. (2016).⁵²

The left panels of Fig. 6 compare the spherical collapse threshold $\delta_c(S, z_f = 0, \delta_{\text{env}}^{\text{peak}})$ (blue) with the corrected threshold Eq. (32)(red). This demonstrates that the spherical collapse calculation captures the qualitative features of the corrected threshold, while differing quantitatively in fine details. The right panels of Figure 6 show the HMF enhancements relative to ΛCDM for three values of f_{R0} in Hu-Sawicki $f(R)$ gravity, at three redshifts $z_f = 0$ (blue), 0.2 (red) and 0.5 (green),

where the agreement between simulation results (shaded squares) and excursion set predictions using δ_c^{cor} (lines) is found to be within 5% level, which is currently the best theoretical model for halo mass function in chameleon-type models. Fitting formulae for Vainshtein models can be found in Ref. 53.

Calibrated HMFs are a crucial ingredient to test gravity using cluster observations such as abundance, as will be discussed in the article by Cataneo & Rapetti (Ref. 54) in this Special Issue.

3.2. Matter power spectra

The matter power spectrum, $P(k)$, which is the Fourier-space counterpart of the two-point matter correlation function, $\xi(r)$, is an important statistic in cosmology. It measures the clustering power of matter at some given scale k (or particle pair separation r), and is sensitive to the presence of a MG force, with different theories affecting different scales (chameleon models modify scales below the Compton wavelength of the scalar field, while in the DGP model (see Ref. 35) the modifications are above the Vainshtein radius). It is closely related to various other quantities, such as the weak lensing convergence or shear power spectrum and the redshift space matter or galaxy clustering, which are of importance in cosmological tests of models, making it one of the main focuses of studies of theoretical model predictions.

While linear perturbation theory can be used to efficiently and accurately compute $P(k)$ on large linear scales (e.g., $k < 0.1h\text{Mpc}^{-1}$), on small scales it loses accuracy due to the nonlinear evolution of structures. At mildly nonlinear scales, bulk flow causes damping of linear signal and coupling of different Fourier modes (which are decoupled in linear theory), and these effects can be modelled by using higher-order perturbation theories⁵⁵ or removed using reconstruction techniques.^{56–59} On even smaller scales where shell crossing takes place, the perturbation theory breaks down and has to be replaced by numerical simulations. The nonlinear effect is often worse in MG models due to the inherently nonlinear nature of screening mechanisms, making it possible that linear theory fails whenever it predicts a MG effect.⁶⁰

Ideally, one would rely on full simulations to provide theoretical templates for $P(k)$ in MG models, but as mentioned above, their cost is too high for continuous parameter space searches or covariance matrix estimation. (Semi)analytical models, with acceptable accuracy in certain regimes, can be more economic alternatives.

3.2.1. Perturbation theory

The growth of matter density contrast, $\delta = \delta(\mathbf{x}, t)$, is governed by the continuity and Euler equations,

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla \cdot [(1 + \delta)\mathbf{u}] = 0, \quad (34)$$

$$\frac{\partial \mathbf{u}}{\partial t} + H\mathbf{u} + \frac{1}{a} \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{a} \nabla \Phi, \quad (35)$$

in which $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ is the velocity field and Φ is the Newtonian potential given by the Poisson equation

$$\nabla^2 \Phi = 4\pi G \delta \rho \quad (36)$$

in which the density perturbation is defined as $\delta\rho(\mathbf{x}, t) = \rho(\mathbf{x}, t) - \bar{\rho}(t) = \bar{\rho}(t)\delta$ and $\bar{\rho}(t)$ is the mean matter density at time t . Note that for simplicity we have assumed that vorticity (the curl part of the peculiar velocity field) does not exist in the initial conditions and is not generated during the late-time evolution in perturbation theory (we do not consider the vorticity generated when structure formation enters the highly nonlinear regime, which requires higher order perturbation theory or full simulations to study). One can take the divergence of Eq. (35) to rewrite it as a scalar equation for $\theta \equiv \nabla \cdot \mathbf{u}$.

In Fourier space, Eqs. (34, 35) become

$$\tilde{\delta}' + \tilde{\theta} + \frac{1}{(2\pi)^3} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{12}) \alpha(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, a) \tilde{\delta}(\mathbf{k}_2, a) = 0, \quad (37)$$

$$\tilde{\theta}' + \left[2 + \frac{H'}{H}\right] \tilde{\theta} + \frac{1}{2(2\pi)^3} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{12}) \beta(\mathbf{k}_1, \mathbf{k}_2) \tilde{\theta}(\mathbf{k}_1, a) \tilde{\theta}(\mathbf{k}_2, a) = \left[\frac{k}{aH}\right]^2 \tilde{\Phi}, \quad (38)$$

in which a prime is the derivative wrt $\ln(a)$, $k = |\mathbf{k}|$, $\mathbf{k}_{12} \equiv \mathbf{k}_1 + \mathbf{k}_2$, $\delta_D^{(3)}$ is the 3D Dirac- δ function, and a tilder denotes the Fourier transform of a quantity, e.g.,

$$\tilde{\delta}(\mathbf{k}) = \int d^3\mathbf{x} \exp(-i\mathbf{k} \cdot \mathbf{x}) \delta(\mathbf{x}), \quad \delta(\mathbf{x}) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \exp(i\mathbf{k} \cdot \mathbf{x}) \tilde{\delta}(\mathbf{k}), \quad (39)$$

and we use shorthand $\tilde{\delta} = \tilde{\delta}(\mathbf{k}, a)$ for simplicity. α and β are functions encoding the mixing of different Fourier modes:

$$\alpha(\mathbf{k}_1, \mathbf{k}_2) = 1 + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1^2}, \quad \beta(\mathbf{k}_1, \mathbf{k}_2) = \frac{|\mathbf{k}_1 + \mathbf{k}_2|^2 (\mathbf{k}_1 \cdot \mathbf{k}_2)}{k_1^2 k_2^2}. \quad (40)$$

In the regime of smaller field, $\tilde{\delta} \ll 1$, a power series expansion can be done:

$$\tilde{\delta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} \tilde{\delta}^{(n)}(\mathbf{k}, a), \quad \tilde{\theta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} \tilde{\theta}^{(n)}(\mathbf{k}, a), \quad (41)$$

where $\tilde{\delta}^{(n)}$ and $\tilde{\theta}^{(n)}$ are the n -th order term. From Eq. (37) at first order we have:

$$\tilde{\delta}^{(1)}(\mathbf{k}, a) = D_+(a) \delta_0(\mathbf{k}), \quad \tilde{\theta}^{(1)}(\mathbf{k}, a) = -f(a) D_+ \delta_0(\mathbf{k}) = -f(a) \tilde{\delta}^{(1)}(\mathbf{k}, a), \quad (42)$$

where D_+ is the (growing mode of the) linear growth factor normalised to $D_+ = 1$ today, $\delta_0(\mathbf{k})$ the initial density contrast extrapolated to the present day using linear theory, and $f = d \ln D_+ / d \ln a$ is the linear growth rate.

In an Einstein-de-Sitter (EdS) universe, one has $\Omega(a) = 1$ and $H'/H = -3/2$, and $D_+ = a$, $f = 1$. In this case, the time and spatial dependences of the solutions to Eqs. (37, 38) can be separated, as can be easily checked:

$$\tilde{\delta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} a^n \tilde{\delta}^{(n)}(\mathbf{k}), \quad \tilde{\theta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} a^n \tilde{\theta}^{(n)}(\mathbf{k}). \quad (43)$$

The higher-order perturbation terms can then be read recursively from Eqs. (37, 38, 42). For example, at second order we have

$$\tilde{\delta}^{(2)}(\mathbf{k}, a) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{12}) F_2(\mathbf{k}_1, \mathbf{k}_2; a) \tilde{\delta}_0(\mathbf{k}_1) \tilde{\delta}_0(\mathbf{k}_2), \quad (44)$$

$$\tilde{\theta}^{(2)}(\mathbf{k}, a) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{12}) G_2(\mathbf{k}_1, \mathbf{k}_2; a) \tilde{\delta}_0(\mathbf{k}_1) \tilde{\delta}_0(\mathbf{k}_2), \quad (45)$$

with

$$F_2(\mathbf{k}_1, \mathbf{k}_2; a) \equiv a^2 \left[\frac{5}{14} [\alpha(\mathbf{k}_1, \mathbf{k}_2) + \alpha(\mathbf{k}_2, \mathbf{k}_1)] + \frac{1}{7} \beta(\mathbf{k}_1, \mathbf{k}_2) \right], \quad (46)$$

$$G_2(\mathbf{k}_1, \mathbf{k}_2; a) \equiv -a^2 \left[\frac{3}{14} [\alpha(\mathbf{k}_1, \mathbf{k}_2) + \alpha(\mathbf{k}_2, \mathbf{k}_1)] + \frac{2}{7} \beta(\mathbf{k}_1, \mathbf{k}_2) \right]. \quad (47)$$

Even higher-order terms can also be derived straightforwardly, but we shall not go to the details here.

In non-EdS models, the separability of time and spatial dependencies is generally not exact, but for realistic Λ CDM models as a good approximation one can write

$$\tilde{\delta}(\mathbf{k}, a) = \sum_{n=1}^{\infty} F_1^n(a) \tilde{\delta}^{(n)}(\mathbf{k}), \quad \tilde{\theta}(\mathbf{k}, a) = -f(a) \sum_{n=1}^{\infty} F_1^n(a) \tilde{\theta}^{(n)}(\mathbf{k}), \quad (48)$$

in which $F_1(a) = D_+(a)$ is still scale independent. For MG models, however, this is not necessarily true, and one has to refer to the more general expressions:

$$\begin{aligned} \tilde{\delta}^{(n)}(\mathbf{k}, a) &= \frac{1}{(8\pi^3)^{n-1}} \int d^3\mathbf{k}_1 \cdots d^3\mathbf{k}_n \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{1\dots n}) F_n(\mathbf{k}_1, \dots, \mathbf{k}_n; a) \tilde{\delta}_0(\mathbf{k}_1) \cdots \tilde{\delta}_0(\mathbf{k}_n), \\ \tilde{\theta}^{(n)}(\mathbf{k}, a) &= \frac{1}{(8\pi^3)^{n-1}} \int d^3\mathbf{k}_1 \cdots d^3\mathbf{k}_n \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{1\dots n}) G_n(\mathbf{k}_1, \dots, \mathbf{k}_n; a) \tilde{\delta}_0(\mathbf{k}_1) \cdots \tilde{\delta}_0(\mathbf{k}_n). \end{aligned} \quad (49)$$

In addition, the MG effects on matter clustering has to be taken into account, often through the modified Poisson equation, which becomes⁶¹ now

$$-\frac{k^2}{(aH)^2} \tilde{\Phi}(\mathbf{k}, a) = \frac{3}{2} \Omega_m(a) \mu(k, a) \tilde{\delta}(\mathbf{k}, a) + S(\mathbf{k}, a), \quad (50)$$

where $\mu(k, a)$ parameterises the dependences of Newton's constant on space and time, while $S(\mathbf{k}, a)$ incorporates nonlinear source terms, which up to third order in the perturbations can be written as

$$\begin{aligned} S(\mathbf{k}, a) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{12}) \gamma_2(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2; a) \tilde{\delta}_0(\mathbf{k}_1) \tilde{\delta}_0(\mathbf{k}_2) + \\ &\quad \frac{1}{(2\pi)^6} \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 d^3\mathbf{k}_3 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_{123}) \gamma_3(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; a) \tilde{\delta}_0(\mathbf{k}_1) \tilde{\delta}_0(\mathbf{k}_2) \tilde{\delta}_0(\mathbf{k}_3). \end{aligned} \quad (51)$$

In Λ CDM, one has $\mu(k, a) = 1$ and $\gamma_2 = \gamma_3 = 0$. Expressions for these quantities in chameleon and Vainshtein models can be found in, e.g., Refs. 62, 63, 64.

In perturbation theory, the matter and velocity (divergence θ) power spectra can be calculated from the above higher-order perturbation solutions to $\tilde{\delta}$ and $\tilde{\theta}$:

$$P_{ab}(k)\delta_{\text{D}}^{(3)}(\mathbf{k} + \mathbf{k}') = (2\pi)^3 \langle \tilde{\delta}_a(\mathbf{k}) \tilde{\delta}_b(\mathbf{k}') \rangle, \quad (52)$$

where $\{a, b\} = \{\delta, \theta\}$. Up to quadratic order in the linear matter power spectrum, the result is given by

$$P_{ab}(k) = P_{ab}^{(1,1)}(k) + P_{ab}^{(1,3)}(k) + P_{ab}^{(2,2)}(k), \quad (53)$$

where $P_{ab}^{(1,1)}(k)$ is the linear spectrum. The terms are given respectively by

$$\begin{aligned} P_{\delta\delta}^{(1,1)}(k, a) &= F_1^2(k, a)P_0(k), \\ P_{\delta\theta}^{(1,1)}(k, a) &= F_1(k, a)F_1(k, a)P_0(k), \\ P_{\theta\theta}^{(1,1)}(k, a) &= G_1^2(k, a)P_0(k), \end{aligned} \quad (54)$$

$$\begin{aligned} P_{\delta\delta}^{(1,3)}(k, a) &= \frac{6}{(2\pi)^3} \int d^3\mathbf{k}' P_0(k)P_0(k')F_1(k; a)F_3(\mathbf{k}, \mathbf{k}', -\mathbf{k}'; a), \\ P_{\delta\theta}^{(1,3)}(k, a) &= \frac{3}{(2\pi)^3} \int d^3\mathbf{k}' P_0(k)P_0(k')G_1(k; a)F_3(\mathbf{k}, \mathbf{k}', -\mathbf{k}'; a) \\ &\quad + \frac{3}{(2\pi)^3} \int d^3\mathbf{k}' P_0(k)P_0(k')F_1(k; a)G_3(\mathbf{k}, \mathbf{k}', -\mathbf{k}'; a), \\ P_{\theta\theta}^{(1,3)}(k, a) &= \frac{6}{(2\pi)^3} \int d^3\mathbf{k}' P_0(k)P_0(k')G_1(k; a)G_3(\mathbf{k}, \mathbf{k}', -\mathbf{k}'; a), \end{aligned} \quad (55)$$

and

$$\begin{aligned} P_{\delta\delta}^{(2,2)}(k, a) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k}' P_0(|\mathbf{k} - \mathbf{k}'|)P_0(k')F_2^2(\mathbf{k} - \mathbf{k}', \mathbf{k}'; a), \\ P_{\delta\theta}^{(2,2)}(k, a) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k}' P_0(|\mathbf{k} - \mathbf{k}'|)P_0(k')F_2(\mathbf{k} - \mathbf{k}', \mathbf{k}'; a)G_2(\mathbf{k} - \mathbf{k}', \mathbf{k}'; a), \\ P_{\theta\theta}^{(2,2)}(k, a) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k}' P_0(|\mathbf{k} - \mathbf{k}'|)P_0(k')G_2^2(\mathbf{k} - \mathbf{k}', \mathbf{k}'; a). \end{aligned} \quad (56)$$

Following a lengthier calculation, the bispectrum can also be computed within the framework of perturbation theory using its definition,

$$B(\mathbf{k}_1, \mathbf{k}_2)\delta_{\text{D}}^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) = (2\pi)^3 \langle \tilde{\delta}(\mathbf{k}_1)\tilde{\delta}(\mathbf{k}_2)\tilde{\delta}(\mathbf{k}_3) \rangle, \quad (57)$$

though little work has been done on this (e.g., Refs. 65, 66).

In principle, higher-order corrections can be included in this framework straightforwardly, but this approach is known to have divergence problems on both small and large scales, the former being a consequence of perturbations becoming progressively nonlinear and the latter because oscillations on large scales at a given order can affect small-scale modes. Resummation techniques, such as the regularised perturbation theory,^{67–69} overcomes the latter problem by including small- k contributions from all orders in perturbation theory, effectively introducing a damping on the oscillations of the power spectrum. This makes it possible to Fourier transform

$P(k)$ to calculate the correlation function reliably, as well as alleviating small-scale divergence of the standard perturbation theory. This has been used in recent works to predict the nonlinear matter and velocity power spectrum in $f(R)$ and DGP models.^{64,70–73} Another approach is the Lagrangian perturbation theory,^{74–78} which has been extended to MG models recently.⁷⁹

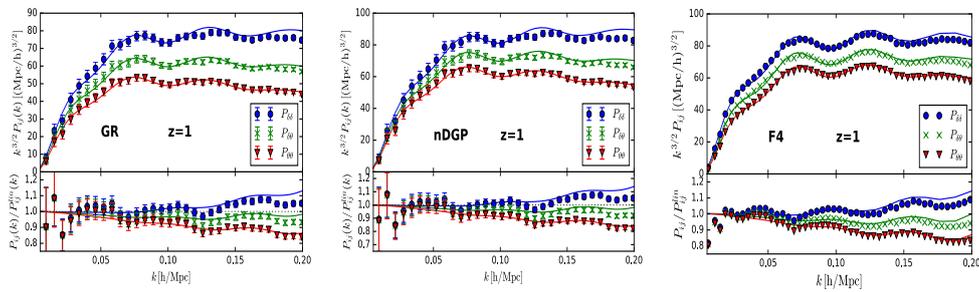


Fig. 7. Perturbation theory predictions of the matter and velocity power spectra (lines), plotted against full simulation results (symbols). The lower panels show the ratios of them. This figure is adapted from Bose et al. (2016)⁷¹ (the left and right panels) and Bose et al. (2017)⁷² (the middle panel).

Figure 7 shows a few examples of perturbation theory predictions for $P_{\delta\delta}$, $P_{\delta\theta}$ and $P_{\theta\theta}$ at $z = 1$, for three models (GR, nDGP N1 and Hu-Sawicki $f(R)$ gravity F4). The agreement with full simulation results is good down to scales $k \sim 0.1-0.2h \text{ Mpc}^{-1}$, below which perturbation theory fails to fully reproduce simulation results.

3.2.2. Halo model and HALOFIT

Moving to the fully nonlinear regime of structure formation, the frequent crossing of particle trajectories and virialised motions make it impossible to track the evolution of density perturbations perturbatively and analytically. One of the commonly used methods to predict the nonlinear matter power spectrum is the halo model.^{80,81}

The main premise of the halo model is that all matter in the universe is in bound structures (haloes). The matter 2-point correlation function becomes the sum of the contributions from the correlations between masses which belong to different haloes (the 2-halo term) and those in the same halo (the 1-halo term). The matter power

26 *Baojiu Li*

spectrum can be written as (see Ref. 82 for a review)

$$P(k) = P_{1h}(k) + P_{2h}(k), \quad (58)$$

with

$$\begin{aligned} P_{1h}(k) &= \int dM \frac{M}{\bar{\rho}_{m0}^2} \frac{dn(M)}{d \ln M} |y(k, M)|^2, \\ P_{2h}(k) &= \left[\int dM \frac{1}{\bar{\rho}_{m0}} \frac{dn(M)}{d \ln M} b_{\text{lin}}(M) |y(k, M)| \right]^2 P_{\text{lin}}(k), \end{aligned} \quad (59)$$

in which $dn/d \ln M$ is the HMF, $b_{\text{lin}}(M)$ the linear halo bias, $P_{\text{lin}}(k)$ the linear matter power spectrum and $y(k, M)$ the Fourier transform of the halo density profile truncated at halo radius and normalised such that $y(k \rightarrow 0, M) \rightarrow 1$.

Of the three main building blocks of the halo model, the halo mass function and linear halo bias have been discussed above in the frameworks of both Λ CDM and modified gravity. The halo density profiles in CDM cosmologies are known to follow the universal Navarro-Frenk-White (NFW) formula,⁸³

$$\rho_{\text{NFW}}(r) = \frac{\rho_s}{(r/R_s)(1+r/R_s)^2}, \quad (60)$$

where haloes are assumed to be spherical and r is the distance from the halo centre. ρ_s and R_s are the characteristic density and scale radius, which vary from halo to halo. The halo mass, is given by integrating Eq. (60) between $r = 0$ and the halo radius R_h :

$$M = 4\pi\rho_s \frac{R_h^3}{c^3} \left[\ln(1+c) - \frac{c}{1+c} \right] \quad (61)$$

where $c = R_h/R_s$ is called the halo concentration. In practice, the halo radius R_h has no unique definition, and some common choices are the radius in which matter is virialised, or the radius enclosing an average density which is Δ times the critical or mean matter density at the halo redshift. In these latter cases, R_h is often written as R_Δ with mass $M_\Delta = \frac{4}{3}\pi\Delta\bar{\rho}R_\Delta^3$ and concentration $c_\Delta = R_\Delta/R_s$. c_Δ, M_Δ can be more conveniently used instead of ρ_s, R_s to characterise a halo's density profile, and a concentration-mass relation, $c_\Delta(M_\Delta)$, can be either fitted using simulations⁸⁴⁻⁸⁷ or computed from physical modelling.⁸⁸ $y(k, M)$ for an NFW profile is given by

$$\begin{aligned} y(k, M) &= 4\pi\rho_s R_s^3 \left[\frac{\sin(kR_s)}{M} (\text{Si}[(1+c)kR_s] - \text{Si}(kR_s)) \right] \\ &+ 4\pi\rho_s R_s^3 \left[\frac{\cos(kR_s)}{M} (\text{Ci}[(1+c)kR_s] - \text{Ci}(kR_s)) - \frac{\sin(ckR_s)}{M(1+c)kR_s} \right], \end{aligned} \quad (62)$$

where c is the concentration, $\text{Si}(x) \equiv \int_0^x dx' \sin(x')/x'$, $\text{Ci}(x) \equiv -\int_x^\infty dx' \cos(x')/x'$.

To a good approximation, the c_Δ - M_Δ relation in Λ CDM cosmologies follows a power law, which depends on cosmological parameters and redshift. In MG models, the prediction further depends on the behaviour of gravity. In models with

Vainshtein mechanisms, such as Galileons and the nDGP model (with realistic parameters), the screening strongly suppresses the fifth force inside haloes, so that the halo density profile is not appreciably affected.^{53,89} In other models, such as chameleons, the screening efficiency depends on various factors including model parameters, redshift, halo mass and environment, leaving nontrivial imprints to the density profiles of haloes. While earlier works indicate that the NFW profile of Eq. (60) is generally still a good description for these models,^{90,91} recent higher-resolution simulations show that the c_Δ - M_Δ relation is no longer a simple power law.² Note that the deviation from a simple power-law in the c_Δ - M_Δ relation has been considered in Ref. 91 as a way to incorporate the chameleon screening effect, by promoting the coefficient of the power to a function of halo mass.

The halo model approach can be straightforwardly applied to MG scenarios with the use of the relevant physical quantities, such as $c_\Delta(M_\Delta)$, for the latter. Therefore, at the basic level its application in modified gravity does not require any extension (unlike, for example, halo mass function or the perturbation theory prediction of the power spectra). Some of its applications in different modified gravity models can be found in Refs. 34, 53, 92. But as a simple model, the quantitative agreement of its prediction with simulations is relatively poor, in particular at intermediate scales where the transition between the 2-halo and 1-halo terms takes place, which is true even when looking at the relative difference between MG and Λ CDM,⁵³ though the inaccuracy of the halo model affects both MG and Λ CDM so that its effects cancel to certain extent when looking at the relative difference between the models.

There are different ways to improve on this, including simple phenomenological fixes such as using an interpolated P_{lin} between the MG and Λ CDM $P_{\text{lin}}(k)$ in Eq. (59) to partially compensate the underestimation of the nonlinear screening effect on relatively large scales (Figure 8). This is similar in philosophy to the methods used in Ref. 93, 94, which use the predictions of perturbation theory to replace $P_{\text{lin}}(k)$ in the 2-halo term. A more elaborated approach is to include new parameters to the default halo model to account for missing physical effects such as halo exclusion and nonlinear damping of small-scale power spectrum, as well as promoting some physical quantities to adjustable parameters to increase the freedom and flexibility of the model prediction,⁹⁵ where the transition between the 1- and 2-halo terms is also modified so that the total $P(k)$ is no longer a simple sum of the two. The approach has been shown to work well for a range of models including chameleons and nDGP.⁹⁶ The most popular extension to the halo model is perhaps HALOFIT,^{97,98} which takes a similar approach by using free parameters mimicking the missing physical effects to gain more flexibility, and tuning these free parameters using simulations. HALOFIT and its variations for non-standard cosmologies (e.g., massive neutrinos⁹⁹) is implemented in standard Boltzmann codes such as CAMB,¹⁰⁰ but it is not tested for modified gravity models and a naive application to the latter can lead to results very different from simulations;⁶⁰ later, in Ref. 101 the author extends it to the case of chameleon $f(R)$ gravity, and the resulting fitting function

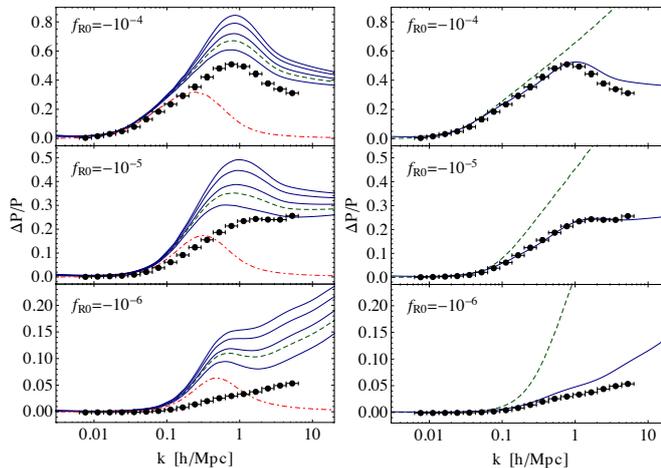


Fig. 8. Halo model predictions of the matter power spectrum enhancements wrt Λ CDM, for three variants of the Hu-Sawicki $f(R)$ model as shown in the legends. *Left panels:* the blue curves come from the use of HMFs predicted by the extended excursion set theory predictions for $f(R)$ gravity, with different assumed values of δ_{env} ; for the green dashed δ_{env} is taken as the mean $\langle \delta_{\text{env}} \rangle$. *Right panels:* The halo model is improved by using an interpolated $P_{\text{lin}}(k)$ between the modified gravity and Λ CDM linear power spectra in the two-halo term; the green dashed line is linear theory prediction here. In all panels black symbols are simulation results at $z = 0$. This figures is adapted from Lombriser et al. (2014)³⁴

is accurate to 6% on scales up to $k = 1h\text{Mpc}^{-1}$ and 12% in $1 < k < 10h\text{Mpc}^{-1}$.

The nonlinear matter power spectrum predicted by the halo model or HALOFIT can be a good starting point in the prediction of other observables. An example is the weak lensing convergence or shear angular power spectrum, which is a weighted integration of $P(k)$ along the line of sight:

$$C_{\kappa\kappa}(\ell) = \int_0^{\chi_s} d\chi \frac{W(\chi)^2}{\chi^2} P_{\delta\delta} \left(k = \frac{\ell}{\chi}, z(\chi) \right), \quad (63)$$

where χ is the comoving distance from the observer, χ_s for the lensing source, $z(\chi)$ the redshift corresponding to comoving distance χ , and $W(\chi) = \frac{3}{2} \Omega_m H_0^2 \frac{\chi}{\chi_s} (\chi_s - \chi) [1 + z(\chi)]$ is a weak lensing kernel for a spatially flat Universe; the integration accumulates the lensing effects of structures between the source and observer. This has been used in, e.g., Refs. 102, 103, using the nonlinear $P_{\delta\delta}(k)$ from either simulations or (generalised) HALOFIT. Other weak lensing statistics have been studied in, e.g., Refs. 104, and lensing tests of gravity will be covered in more detail in the article by Heymans & Zhao¹⁰⁵ in this Special Issue.

3.3. Redshift space distortions (RSD)

In real observations, apart from the nonlinear evolution of structures and biasing of tracers with respect to the underlying dark matter field, there is also another complication in the confrontation with theories – redshift space distortions, the fact that we directly measure the redshifts of galaxies, which are not equivalent to their radial distance due to peculiar velocities. However, if modelled accurately, RSD can be a useful probe which is sensitive to the velocity field and the law of gravity.

Let \mathbf{u} be the peculiar velocity of a particle, then its position in redshift space, \mathbf{s} , is relate to its true position in configuration space \mathbf{r} , as

$$\mathbf{s} = \mathbf{r} + \frac{1}{aH}(\hat{\mathbf{z}} \cdot \mathbf{u})\hat{\mathbf{z}} = \mathbf{r} + \frac{u_z}{aH}\hat{\mathbf{z}}, \quad (64)$$

where $\hat{\mathbf{z}}$ is the unit vector in the line-of-sight (LOS) direction, chosen as the z axis for simplicity. The redshift-space density contrast $\delta^z(\mathbf{s})$ is related to its configuration-space counterpart, $\delta^r(\mathbf{r})$, by mass conservation,

$$[1 + \delta^z(\mathbf{s})] d^3\mathbf{s} = [1 + \delta^r(\mathbf{r})] d^3\mathbf{r} \Rightarrow \delta^z(\mathbf{s}) = \left| \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \right|^{-1} [1 + \delta^r(\mathbf{r})] - 1, \quad (65)$$

where $d^3\mathbf{s}$ and $d^3\mathbf{r}$ are volume elements. The Fourier transform of $\delta^z(\mathbf{s})$ is

$$\tilde{\delta}^z(\mathbf{k}) = \int d^3\mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{s}} \delta^z(\mathbf{s}) = \int d^3\mathbf{r} \left[\delta^r(\mathbf{r}) - \frac{\nabla_z u_z}{aH} \right] e^{-i(\mathbf{k}\cdot\mathbf{r} + \frac{k_z u_z}{aH})}, \quad (66)$$

where $|\partial \mathbf{s} / \partial \mathbf{r}| = 1 + \nabla_z u_z / aH$ is used. The redshift-space power spectrum reads

$$P^z(\mathbf{k}) = \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \left\langle e^{-ik_z[u_z(\mathbf{r}) - u_z(\mathbf{r}')] } [\delta^r(\mathbf{r}) - \nabla_z u_z(\mathbf{r})] [\delta^r(\mathbf{r}') - \nabla_z u_z(\mathbf{r}')] \right\rangle, \quad (67)$$

where $\mathbf{x} = \mathbf{r} - \mathbf{r}'$.

By rewriting Eq. (67) in terms of cumulants, treating the resulting expression perturbatively and neglecting terms with third or higher powers of the linear matter power spectrum, one gets the 1-loop result for the redshift-space power spectrum by Taruya, Nishimichi & Saito (TNS; Ref. 106)

$$P_{\text{TNS}}^z(\mathbf{k}) = D_{\text{FoG}}(k_z \sigma_v) [P_{\delta\delta}(k) - 2\mu^2 P_{\delta\theta} + \mu^4 P_{\theta\theta} + A(\mathbf{k}) + B(\mathbf{k}) + C(\mathbf{k})], \quad (68)$$

where D_{FoG} is a phenomenological factor describing the exponential damping of P^z due to virial motions of particles at small scales, $\mu = k_z/h$ and A, B, C are correction terms whose detailed expressions are not presented here. These terms represent higher-order interactions between the density and velocity fields, neglecting which Eq. (68) reduces to the previous results of Ref. 107. If one further neglects the random-motion-induced Finger-of-God effect (the factor D_{FoG}) and adopts linear theory predictions for P_{ab} 's with $a, b = \delta, \theta$, then Kaiser's formula¹⁰⁸ is recovered:

$$P_{\text{Kaiser}}^z(\mathbf{k}, a) = [F_1^2(k, a) - 2\mu^2 F_1(k, a) G_1(k, a) + \mu^4 G_1^2(k, a)] P_0(k), \quad (69)$$

where F_1, G_1 are the same as used in the previous section.

The redshift-space two-point correlation function can be obtained as Fourier transform of $P^z(\mathbf{k})$, or directly in the configuration space, following the streaming model.^{109,110} Ref. 73 applies the streaming model to MG models, and find their generalised streaming model (GSM) prediction agrees with the Fourier transform of the TNS power spectrum reasonably well, see Figure 9.

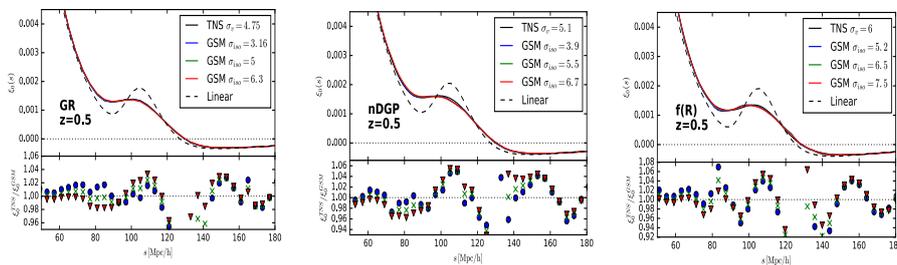


Fig. 9. The comparison of the redshift space correlation function monopole for three models (from left to right: GR, nDGP and $f(R)$ gravity F5, all at $z = 0.5$) predicted by the GSM model and by Fourier transforming the power spectrum from the TNS model. Agreements between both approaches are good, as are the agreements with simulation results. This figure is adapted from Bose & Koyama (2017).⁷³

3.4. Void abundances

Cosmic voids¹¹¹ are a relatively new cosmological probe, believed to be particularly useful for testing MG models with screening mechanisms, due to such mechanisms not working in low-density regions. Unlike large galaxy clusters, cosmic voids primarily develop at sites where the primordial density is low, where matter is evacuated and attracted to their surroundings. Voids and haloes, however, do share an important similarity, namely both form from extreme regimes of the nearly Gaussian distribution of the initial density field ($\delta < 0$ and $\delta > 0$ respectively), so that some of their statistical properties are analytically tractable. Following the same principle of using excursion set theory to predict halo abundance, one can connect the observed abundance of voids to the counts of initial density field with $\delta < \delta_v < 0$, where δ_v is some threshold (similar to δ_c for haloes) to define void formation.

In Ref. 112, in the framework of standard gravity, the authors model voids as spherical tophat underdensities (similar to the case of haloes, which are modelled as spherical overdensities in the simple treatments) which expand over time. The matter evacuated from this region accumulates at its edge, forming an overdensity

ridge which can be a defining feature of voids. The assumption of spherical tophat may sound like an oversimplified assumption, but Ref. 112 shows that different underdensity profiles evolve towards it over time. The threshold δ_v which (like in the halo case) is often taken as the value of $\delta_{v,\text{ini}}$ extrapolated to the present day using linear theory, can be chosen arbitrarily as a free parameter for void definition, but a natural choice is $\delta_v = -2.81$ which in an EdS universe is a scale-independent value at which shell crossing (the event that different layers of the void profile cross each other) happens today. The corresponding value of the nonlinear overdensity inside the spherical tophat at today is $\Delta \approx -0.8$, and the latter is often (but not always) used as the defining criterion for numerical algorithms for void finding. The evolution of the underdensity from $\delta_{v,\text{ini}}$ at initial time to $\Delta = -0.8$ at late times is governed by the growth equation, Eq. (10), for Λ CDM, again.

An important feature of the model used in Ref. 112 is the proper accounting for voids-in-voids (VIV) and voids-in-clouds (VIC) effects. The former says that subregions in a void can satisfy a chosen void definition but should not be counted as independent voids (they are subvoids), and the latter says that a region of space satisfying the void definition must not be considered as viable voids at late times should they happen to live in larger overdense regions which would have collapsed to form haloes (and so squashed the underdense region). In the language of excursion set theory, these are respectively equivalent to saying a void is formed if a random walk crosses δ_v at S_* , with initial Lagrangian radius of $R(S_*)$ and final radius $R_v \approx 1.7R(S_*)$ ($1.7 \approx (1 + \Delta)^{-\frac{1}{3}}$ with $\Delta = -0.8$), but two situations should be excluded: (i) VIV – if the random walk cross δ_v at different values of S then only the first crossing (at the smallest S) counts, and (ii) VIC – if the random walk also crosses δ_c (the halo formation threshold) at smaller S then it does not correspond to a viable void. The void abundance predicted is given by

$$n_v(M) = \frac{\bar{\rho}_m}{M} \mathcal{F}(S, \delta_v, \delta_c) \frac{dS}{dM}, \quad (70)$$

where M is the mass enclosed in the void radius R (recall that R, S, M can be used interchangeably), and $\mathcal{F}(S, \delta_v, \delta_c)$ is the probability density for a random walk to first cross δ_v at S and not cross δ_c at smaller S (where $\mathcal{D} \equiv \delta_v/(\delta_c - \delta_v)$, for a derivation see Ref. 112):

$$\mathcal{F}(S, \delta_v, \delta_c) = \sum_{j=1}^{\infty} \frac{j^2 \pi^2 \mathcal{D}^2}{\delta_v^2} \frac{\sin(j\pi \mathcal{D})}{j\pi} \exp\left[-\frac{j^2 \pi^2 \mathcal{D}^2}{2\delta_v^2/S}\right]. \quad (71)$$

The treatment of VIC in the model described in Eqs. (70, 71) is approximate: it eliminates the possibility of forming voids of any size inside regions which are destined to form haloes, but does not properly account for the effects that voids residing in slightly overdense regions survive being crushed but nevertheless their sizes get squeezed (or grow less) because of the contraction of their surroundings. An indication of the unphysical-ness of this approximation is that in this model the predicted void volume fraction (void volume as a fraction of the total volume of the

universe) is ~ 2 . Ref. 113 shows that this problem can be alleviated by replacing the single second barrier for the random walk, δ_c , with ones corresponding to the Eulerian volumes that the surroundings of the voids-to-be would attain: as the voids cannot grow larger than their surroundings, their sizes should be set to be equal to these Eulerian volumes rather than by the S value where the random walk crosses δ_v . The voids predicted in this way are generally smaller, so that a smaller volume fraction (~ 1.2 though still unphysical) can be obtained. It is also noticed that improvements can be achieved by using correlated random walk steps.

The extension of the semi-analytical models for void formation to MG models can be done similarly as in the halo case. In Ref. 114 the model in Eqs. (70, 71) is extended to chameleon-type models in a conceptually straightforward way: instead of having two constant (scale-independent) barriers δ_v, δ_c , the authors (i) introduce scale- and environment-dependences in δ_v following the method for δ_c (see above), (ii) numerically compute the modified $\mathcal{F}(S, \delta_v, \delta_c)$ – now becoming $\mathcal{F}[S, \delta_v(S, \delta_{\text{env}}), \delta_c(S, \delta_{\text{env}}), \delta_{\text{env}}]$, (iii) average over the distribution of δ_{env} to find $\mathcal{F}_{\text{ave}}(S)$ and (iv) apply Eq. (70) to calculate void abundance. In Ref. 115 this work is further generalised to use the Eulerian model of Ref. 113 to predict void abundance.

Unlike for halo abundance, semi-analytical models for void abundance, despite their successes in qualitatively predicting the behaviour, are still quite far from being used in precision cosmological tests. Theoretically, the assumptions of spherical evolution an excursion set are perhaps an oversimplified description of the evolution and mergers of voids embedded in the complicated cosmic web. Observationally, voids are found to have different and mostly irregular shapes rather than being spherical, adding another layer of complexity as how to define voids and their boundaries, and different void-finding algorithms can return very different void abundances even when applied to the same density field.^{116,117} Furthermore, the above theoretical models only apply to voids found in the dark matter density field, while in reality voids are usually found from a distribution of tracers (such as galaxies), and galaxy bias offers a freedom which can affect the void abundance; this can be translated to a degeneracy – if the two-point galaxy correlation functions of two different models (e.g., Λ CDM and MG) are tuned to agree with each other, this more or less fixes the abundance and distribution of voids identified from the galaxy fields of the two models to be very similar to each other, even though their underlying dark matter density fields can be different.¹¹⁷ For these reasons, the models for void abundance are less widely used than the other techniques mentioned in this review.

The article by Cai¹¹⁸ in this Special Issue will contain a more depth review of gravity tests using void observables.

4. Summary and conclusions

In this **article** we have reviewed a range of approximate simulation techniques and (semi)analytical methods to predict various cosmological observables. Most of the methods described here have previously been used in the studies of Λ CDM, but their

application to MG models can involve nontrivial extensions depending on specific properties of the model being studied. This is one of the main points underlying the developments of these methods: while N-body simulations can appear to be a black box to users, analytical methods can help us track the highly nonlinear physics underlying (modified) gravity; their successes not only offer physical insights into the interpretation of simulation predictions, but are also potentially useful in assisting the development of (model specific) simulation algorithms.

From a practical point of view, the main advantage of approximate and analytical methods is their efficiency. At moderate and controlled loss of accuracy, these methods are usually orders of magnitude faster than full simulations, therefore allowing quick delineation of the model parameter space and serving as guidances for full simulation explorations. This is particularly important considering that we will need to study a large number of theoretical models and compare their predictions with a wide variety of cosmological observables.

Of course, these methods still cannot completely replace full N-body simulations, because they are approximate in nature. Simulations will be needed to assess their performance and range of validity and applicability, as well as calibrate their free parameters. An optimal combination of simulations and fast methods will be an effective tool to reliably test cosmological models and gravity in cosmology.

Acknowledgments

The author is supported by the European Research Council (ERC-StG-716532-PUNCA) and UK STFC Consolidated Grants ST/P000541/1, ST/L00075X/1.

References

1. H. A. Winther *et al.*, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 4208, [arXiv:1506.06384 \[astro-ph.CO\]](#).
2. D. Shi, B. Li, J. Han, L. Gao and W. A. Hellwing, *Mon. Not. Roy. Astron. Soc.* **452** (2015) 3179, [arXiv:1503.01109 \[astro-ph.CO\]](#).
3. A. Barreira, S. Bose and B. Li, *JCAP* **1512** (2015) 059, [arXiv:1511.08200 \[astro-ph.CO\]](#).
4. S. Bose, B. Li, A. Barreira, J.-h. He, W. A. Hellwing, K. Koyama, C. Llinares and G.-B. Zhao, *JCAP* **1702** (2017) 050, [arXiv:1611.09375 \[astro-ph.CO\]](#).
5. C. Arnold, P. Fosalba, V. Springel, E. Puchwein and L. Blot (2018) [arXiv:1805.09824 \[astro-ph.CO\]](#).
6. C. Llinares, *IJMPD* (2018).
7. H. A. Winther and P. G. Ferreira, *Phys. Rev.* **D91** (2015) 123507, [arXiv:1403.6492 \[astro-ph.CO\]](#).
8. S. Tassev, M. Zaldarriaga and D. Eisenstein, *JCAP* **1306** (2013) 036, [arXiv:1301.0322 \[astro-ph.CO\]](#).
9. G. Valogiannis and R. Bean, *Phys. Rev.* **D95** (2017) 103515, [arXiv:1612.06469 \[astro-ph.CO\]](#).
10. H. A. Winther, K. Koyama, M. Manera, B. S. Wright and G.-B. Zhao, *JCAP* **1708** (2017) 006, [arXiv:1703.00879 \[astro-ph.CO\]](#).

11. B. S. Wright, H. A. Winther and K. Koyama, *JCAP* **1710** (2017) 054, [arXiv:1705.08165 \[astro-ph.CO\]](#).
12. R. E. Angulo and S. D. M. White, *Mon. Not. Roy. Astron. Soc.* **405** (2010) 143, [arXiv:0912.4277 \[astro-ph\]](#).
13. R. E. Angulo and S. Hilbert, *Mon. Not. Roy. Astron. Soc.* **448** (2015) 364, [arXiv:1405.5888 \[astro-ph.CO\]](#).
14. A. Mead and J. Peacock, *Mon. Not. Roy. Astron. Soc.* **440** (2014) 1233, [arXiv:1308.5183 \[astro-ph.CO\]](#).
15. A. Mead and J. Peacock, *Mon. Not. Roy. Astron. Soc.* **445** (2014) 3453, [arXiv:1408.1047 \[astro-ph.CO\]](#).
16. A. J. Mead, J. A. Peacock, L. Lombriser and B. Li, *Mon. Not. Roy. Astron. Soc.* **452** (2015) 4203, [arXiv:1412.5195 \[astro-ph.CO\]](#).
17. A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Dave, C. S. Frenk, A. Jenkins, N. Katz and C. G. Lacey, *Astrophys. J.* **593** (2003) 1, [arXiv:astro-ph/0212357 \[astro-ph\]](#).
18. Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Dave, C. S. Frenk, N. Katz and C. G. Lacey, *Astrophys. J.* **633** (2005) 791, [arXiv:astro-ph/0408564 \[astro-ph\]](#).
19. W. H. Press and P. Schechter, *Astrophys. J.* **187** (1974) 425.
20. J. R. Bond, S. Cole, G. Efstathiou and N. Kaiser, *Astrophys. J.* **379** (1991) 440.
21. R. K. Sheth and G. Tormen, *Mon. Not. Roy. Astron. Soc.* **308** (1999) 119, [arXiv:astro-ph/9901122 \[astro-ph\]](#).
22. A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman and N. Yoshida, *Mon. Not. Roy. Astron. Soc.* **321** (2001) 372, [arXiv:astro-ph/0005260 \[astro-ph\]](#).
23. M. S. Warren, K. Abazajian, D. E. Holz and L. Teodoro, *Astrophys. J.* **646** (2006) 881, [arXiv:astro-ph/0506395 \[astro-ph\]](#).
24. D. Reed, R. Bower, C. Frenk, A. Jenkins and T. Theuns, *Mon. Not. Roy. Astron. Soc.* **374** (2007) 2, [arXiv:astro-ph/0607150 \[astro-ph\]](#).
25. J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes, S. Gottlober and D. E. Holz, *Astrophys. J.* **688** (2008) 709, [arXiv:0803.2706 \[astro-ph\]](#).
26. G. Li, Baojiu; Efstathiou, *Mon. Not. Roy. Astron. Soc.* **421** (2012) 1431, [arXiv:1110.6440 \[astro-ph\]](#).
27. H. J. Mo and S. D. M. White, *Mon. Not. Roy. Astron. Soc.* **282** (1996) 347, [arXiv:astro-ph/9512127 \[astro-ph\]](#).
28. J. N. Fry and E. Gaztanaga, *Astrophys. J.* **413** (1993) 447, [arXiv:astro-ph/9302009 \[astro-ph\]](#).
29. R. K. Sheth, H. J. Mo and G. Tormen, *Mon. Not. Roy. Astron. Soc.* **323** (2001) 1, [arXiv:astro-ph/9907024 \[astro-ph\]](#).
30. R. K. Sheth and G. Tormen, *Mon. Not. Roy. Astron. Soc.* **329** (2002) 61, [arXiv:astro-ph/0105113 \[astro-ph\]](#).
31. J. Khoury and A. Weltman, *Phys. Rev. Lett.* **93** (2004) 171104, [arXiv:astro-ph/0309300 \[astro-ph\]](#).
32. J. Khoury and A. Weltman, *Phys. Rev.* **D69** (2004) 044026, [arXiv:astro-ph/0309411 \[astro-ph\]](#).
33. L. Lombriser, B. Li, K. Koyama and G.-B. Zhao, *Phys. Rev.* **D87** (2013) 123511, [arXiv:1304.6395 \[astro-ph.CO\]](#).
34. L. Lombriser, K. Koyama and B. Li, *JCAP* **1403** (2014) 021, [arXiv:1312.1292 \[astro-ph.CO\]](#).

35. K. Koyama, *IJMPD* (2018).
36. J. Zhang and L. Hui, *Astrophys. J.* **641** (2006) 641, [arXiv:astro-ph/0508384](#) [astro-ph].
37. F. Schmidt, M. V. Lima, H. Oyaizu and W. Hu, *Phys. Rev.* **D79** (2009) 083518, [arXiv:0812.0545](#) [astro-ph].
38. B. Li, G. Zhao and K. Kazuya, *Mon. Not. Roy. Astron. Soc.* **421** (2012) 3481, [arXiv:1111.2602](#) [astro-ph].
39. B. Li and T. Lam, *Mon. Not. Roy. Astron. Soc.* **425** (2012) 730, [arXiv:1205.0058](#) [astro-ph].
40. T. Lam and B. Li, *Mon. Not. Roy. Astron. Soc.* **426** (2012) 3260, [arXiv:1205.0059](#) [astro-ph].
41. F. Bernardeau, *Astrophys. J.* **427** (1994) 51, [arXiv:astro-ph/9311066](#) [astro-ph].
42. R. K. Sheth, *Mon. Not. Roy. Astron. Soc.* **300** (1998) 1057, [arXiv:astro-ph/9805319](#) [astro-ph].
43. T. Y. Lam and R. K. Sheth, *Mon. Not. Roy. Astron. Soc.* **386** (2008) 407, [arXiv:0711.5029](#) [astro-ph].
44. F. Schmidt, W. Hu and M. Lima.
45. A. Barreira, B. Li, C. M. Baugh and S. Pascoli, *JCAP* **1311** (2013) 056, [arXiv:1308.3699](#) [astro-ph.CO].
46. L. Lombriser, *JCAP* **1611** (2016) 039, [arXiv:1608.00522](#) [astro-ph.CO].
47. A. Borisov, B. Jain and P. Zhang, *Phys. Rev.* **D85** (2012) 063518, [arXiv:1102.4839](#) [astro-ph.CO].
48. M. Kopp, S. A. Appleby, I. Achitouv and J. Weller, *Phys. Rev.* **D88** (2013) 084015, [arXiv:1306.3233](#) [astro-ph.CO].
49. K. Jones-Smith and F. Ferrer, *Phys. Rev. Lett.* **108** (2012) 221101, [arXiv:1105.6085](#) [astro-ph.CO].
50. C. Burrage, E. J. Copeland and J. Stevenson, *Phys. Rev.* **D91** (2015) 065030, [arXiv:1412.6373](#) [hep-th].
51. R. Pourhasan, N. Afshordi, R. B. Mann and A. C. Davis, *JCAP* **1112** (2011) 005, [arXiv:1109.0538](#) [astro-ph.CO].
52. M. Cataneo, D. Rapetti, L. Lombriser and B. Li, *JCAP* **1612** (2016) 024, [arXiv:1607.08788](#) [astro-ph.CO].
53. A. Barreira, B. Li, W. A. Hellwing, L. Lombriser, C. M. Baugh and S. Pascoli, *JCAP* **1404** (2014) 029, [arXiv:1401.1497](#) [astro-ph.CO].
54. M. Cataneo and R. D., *IJMPD* (2018).
55. F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Phys. Rept.* **367** (2002) 1, [arXiv:astro-ph/0112551](#) [astro-ph].
56. D. J. Eisenstein, H.-j. Seo, E. Sirko and D. Spergel, *Astrophys. J.* **664** (2007) 675, [arXiv:astro-ph/0604362](#) [astro-ph].
57. H.-M. Zhu, Y. Yu, U.-L. Pen, X. Chen and H.-R. Yu, *Phys. Rev.* **D96** (2017) 123502, [arXiv:1611.09638](#) [astro-ph.CO].
58. M. Schmittfull, T. Baldauf and M. Zaldarriaga, *Phys. Rev.* **D96** (2017) 023505, [arXiv:1704.06634](#) [astro-ph.CO].
59. Y. Shi, M. Cautun and B. Li, *Phys. Rev.* **D97** (2018) 023505, [arXiv:1709.06350](#) [astro-ph.CO].
60. B. Li, W. A. Hellwing, K. Koyama, G.-B. Zhao, E. Jennings and C. M. Baugh, *Mon. Not. Roy. Astron. Soc.* **428** (2013) 743, [arXiv:1206.4317](#) [astro-ph].
61. K. Koyama, A. Taruya and T. Hiramatsu, *Phys. Rev.* **D79** (2009) 123512, [arXiv:0902.0618](#) [astro-ph.CO].
62. K. Koyama, A. Taruya and T. Hiramatsu, *Phys. Rev.* **D79** (2009) 123512,

- arXiv:0902.0618 [astro-ph.CO].
63. P. Brax and P. Valageas, *Phys. Rev.* **D88** (2013) 023527, arXiv:1305.5647 [astro-ph.CO].
 64. B. Bose, Cosmological Tests of Gravity, PhD thesis (2018).
 65. N. Bartolo, E. Bellini, D. Bertacca and S. Matarrese, *JCAP* **1303** (2013) 034, arXiv:1301.4831 [astro-ph.CO].
 66. E. Bellini, R. Jimenez and L. Verde, *JCAP* **1505** (2015) 057, arXiv:1504.04341 [astro-ph.CO].
 67. F. Bernardeau, M. Crocce and R. Scoccimarro, *Phys. Rev.* **D78** (2008) 103521, arXiv:0806.2334 [astro-ph].
 68. A. Taruya, F. Bernardeau, T. Nishimichi and S. Codis, *Phys. Rev.* **D86** (2012) 103528, arXiv:1208.1191 [astro-ph.CO].
 69. A. Taruya, T. Nishimichi and F. Bernardeau, *Phys. Rev.* **D87** (2013) 083509, arXiv:1301.3624 [astro-ph.CO].
 70. A. Taruya, T. Nishimichi, F. Bernardeau, T. Hiramatsu and K. Koyama, *Phys. Rev.* **D90** (2014) 123515, arXiv:1408.4232 [astro-ph.CO].
 71. B. Bose and K. Koyama, *JCAP* **1608** (2016) 032, arXiv:1606.02520 [astro-ph.CO].
 72. B. Bose, K. Koyama, W. A. Hellwing, G.-B. Zhao and H. A. Winther, *Phys. Rev.* **D96** (2017) 023519, arXiv:1702.02348 [astro-ph.CO].
 73. B. Bose and K. Koyama, *JCAP* **1708** (2017) 029, arXiv:1705.09181 [astro-ph.CO].
 74. B. T., *A. A.* **223** (1989) 9.
 75. P. Catelan, *Mon. Not. Roy. Astron. Soc.* **276** (1995) 115, arXiv:astro-ph/9406016 [astro-ph].
 76. F. R. Bouchet, S. Colombi, E. Hivon and R. Juszkiewicz, *Astron. Astrophys.* **296** (1995) 575, arXiv:astro-ph/9406013 [astro-ph].
 77. A. N. Taylor and A. J. S. Hamilton, *Mon. Not. Roy. Astron. Soc.* **282** (1996) 767, arXiv:astro-ph/9604020 [astro-ph].
 78. T. Matsubara, *Phys. Rev.* **D77** (2008) 063530, arXiv:0711.2521 [astro-ph].
 79. A. Aviles and J. L. Cervantes-Cota, *Phys. Rev.* **D96** (2017) 123526, arXiv:1705.10719 [astro-ph.CO].
 80. U. Seljak, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 203, arXiv:astro-ph/0001493 [astro-ph].
 81. J. A. Peacock and R. E. Smith, *Mon. Not. Roy. Astron. Soc.* **318** (2000) 1144, arXiv:astro-ph/0005010 [astro-ph].
 82. A. Cooray and R. K. Sheth, *Phys. Rept.* **372** (2002) 1, arXiv:astro-ph/0206508 [astro-ph].
 83. J. F. Navarro, C. S. Frenk and S. D. M. White, *Astrophys. J.* **490** (1997) 493, arXiv:astro-ph/9611107 [astro-ph].
 84. J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack and A. Dekel, *Mon. Not. Roy. Astron. Soc.* **321** (2001) 559, arXiv:astro-ph/9908159 [astro-ph].
 85. A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel and A. Jenkins, *Mon. Not. Roy. Astron. Soc.* **381** (2007) 1450, arXiv:0706.2919 [astro-ph].
 86. A. V. Maccio', A. A. Dutton, F. C. van den Bosch, B. Moore, D. Potter and J. Stadel, *Mon. Not. Roy. Astron. Soc.* **378** (2007) 55, arXiv:astro-ph/0608157 [astro-ph].
 87. A. R. Duffy, J. Schaye, S. T. Kay and C. Dalla Vecchia, *Mon. Not. Roy. Astron. Soc.* **390** (2008) L64, arXiv:0804.2486 [astro-ph], [Erratum: *Mon. Not. Roy. Astron.*

- Soc.415,L85(2011)].
88. A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk and S. D. M. White, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 378, [arXiv:1312.0945 \[astro-ph.CO\]](#).
 89. A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh and S. Pascoli, *JCAP* **1310** (2013) 027, [arXiv:1306.3219 \[astro-ph.CO\]](#).
 90. G.-B. Zhao, B. Li and K. Koyama, *Phys. Rev.* **D83** (2011) 044007, [arXiv:1011.1257 \[astro-ph.CO\]](#).
 91. L. Lombriser, K. Koyama, G.-B. Zhao and B. Li, *Phys. Rev.* **D85** (2012) 124054, [arXiv:1203.5125 \[astro-ph.CO\]](#).
 92. B. Hu, X.-W. Liu and R.-G. Cai (2017) [arXiv:1712.09017 \[astro-ph.CO\]](#).
 93. P. Valageas and T. Nishimichi, *A. & A.* **527** (2015) 87, [arXiv:1009.0597 \[astro-ph.CO\]](#).
 94. U. Seljak and Z. Vlah, *Phys. Rev.* **D91** (2015) 123516, [arXiv:1501.07512 \[astro-ph.CO\]](#).
 95. A. Mead, J. Peacock, C. Heymans, S. Joudaki and A. Heavens, *Mon. Not. Roy. Astron. Soc.* **454** (2015) 1958, [arXiv:1505.07833 \[astro-ph.CO\]](#).
 96. A. Mead, C. Heymans, L. Lombriser, J. Peacock, O. Steele and H. Winther, *Mon. Not. Roy. Astron. Soc.* **459** (2016) 1468, [arXiv:1602.02154 \[astro-ph.CO\]](#).
 97. VIRGO Consortium Collaboration (R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou and H. M. P. Couchmann), *Mon. Not. Roy. Astron. Soc.* **341** (2003) 1311, [arXiv:astro-ph/0207664 \[astro-ph\]](#).
 98. R. Takahashi, M. Sato, T. Nishimichi, A. Taruya and M. Oguri, *Astrophys. J.* **761** (2012) 152, [arXiv:1208.2701 \[astro-ph.CO\]](#).
 99. S. Bird, M. Viel and M. G. Haehnelt, *Mon. Not. Roy. Astron. Soc.* **420** (2012) 2551, [arXiv:1109.4416 \[astro-ph.CO\]](#).
 100. A. Lewis, A. Challinor and A. Lasenby, *Astrophys. J.* **538** (2000) 473, [arXiv:astro-ph/9911177 \[astro-ph\]](#).
 101. G.-B. Zhao, *Astrophys. J. Suppl.* **211** (2014) 23, [arXiv:1312.1291 \[astro-ph.CO\]](#).
 102. M. Shirasaki, T. Hamana and N. Yoshida, *Publ. Astron. Soc. Jap.* **68** (2016) 4, [arXiv:1508.02104 \[astro-ph.CO\]](#).
 103. B. Li and M. Shirasaki, *Mon. Not. Roy. Astron. Soc.* **474** (2018) 3599, [arXiv:1710.07291 \[astro-ph.CO\]](#).
 104. L. Lombriser, F. Schmidt, T. Baldauf, R. Mandelbaum, U. Seljak and R. E. Smith, *Phys. Rev.* **D85** (2012) 102001, [arXiv:1111.2020 \[astro-ph.CO\]](#).
 105. C. Heymans and G. Zhao, *IJMPD* (2018).
 106. A. Taruya, T. Nishimichi and S. Saito, *Phys. Rev.* **D82** (2010) 063522, [arXiv:1006.0699 \[astro-ph.CO\]](#).
 107. R. Scoccimarro, *Phys. Rev.* **D70** (2004) 083007, [arXiv:astro-ph/0407214 \[astro-ph\]](#).
 108. N. Kaiser, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1.
 109. K. B. Fisher, *Astrophys. J.* **448** (1995) 494, [arXiv:astro-ph/9412081 \[astro-ph\]](#).
 110. B. A. Reid and M. White, *Mon. Not. Roy. Astron. Soc.* **417** (2011) 1913.
 111. B. A. Reid and M. White, *International Journal of Modern Physics: Conference Series* **1** (2011) 41.
 112. R. K. Sheth and R. van de Weygaert, *Mon. Not. Roy. Astron. Soc.* **350** (2004) 517, [arXiv:astro-ph/0311260 \[astro-ph\]](#).
 113. A. Paranjape, T. Y. Lam and R. K. Sheth, *Mon. Not. Roy. Astron. Soc.* **420** (2011) 1648, [arXiv:1106.2041 \[astro-ph\]](#).

38 *Baojiu Li*

114. J. Clampitt, Y.-C. Cai and B. Li, *Mon. Not. Roy. Astron. Soc.* **431** (2013) 749, [arXiv:1212.2216](#) [astro-ph].
115. T. Y. Lam, J. Clampitt, Y.-C. Cai and B. Li, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 3319, [arXiv:1408.5338](#) [astro-ph.CO].
116. J. M. Colberg *et al.*, *Mon. Not. Roy. Astron. Soc.* **387** (2008) 933, [arXiv:0803.0918](#) [astro-ph].
117. M. Cautun, E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, B. Li and N. Padilla (2017) [arXiv:1710.01730](#) [astro-ph.CO].
118. Y. Cai, *IJMPD* (2018).