Galaxy-galaxy weak gravitational lensing in $f(R)$ gravity

Baojiu Li$^{1*}$, Masato Shirasaki$^{2†}$

$^1$Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
$^2$Division of Theoretical Astronomy, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

Accepted XXX. Received XXX; in original form XXX

ABSTRACT

We present an analysis of galaxy-galaxy weak gravitational lensing (GGL) in chameleon $f(R)$ gravity – a leading candidate of non-standard gravity models. For the analysis we have created mock galaxy catalogues based on dark matter haloes from two sets of numerical simulations, using a halo occupation distribution (HOD) prescription which allows a redshift dependence of galaxy number density. To make a fairer comparison between the $f(R)$ and $\Lambda$CDM models, their HOD parameters are tuned so that the galaxy two-point correlation functions in real space (and therefore the projected two-point correlation functions) match. While the $f(R)$ model predicts an enhancement of the convergence power spectrum by up to $\sim 30\%$ compared to the standard $\Lambda$CDM model with the same parameters, the maximum enhancement of GGL is only half as large and less than $5\%$ on separations above $\sim 1-2\ h^{-1}{\rm Mpc}$, because the latter is a cross correlation of shear (or matter, which is more strongly affected by modified gravity) and galaxy (which is weakly affected given the good match between galaxy auto correlations in the two models) fields. We also study the possibility of reconstructing the matter power spectrum by combination of GGL and galaxy clustering in $f(R)$ gravity. We find that the galaxy-matter cross correlation coefficient remains at unity down to $\sim 2-3\ h^{-1}{\rm Mpc}$ at relevant redshifts even in $f(R)$ gravity, indicating joint analysis of GGL and galaxy clustering can be a powerful probe of matter density fluctuations in chameleon gravity. The scale dependence of the model differences in their predictions of GGL can potentially allow to break the degeneracy between $f(R)$ gravity and other cosmological parameters such as $\Omega_m$ and $\sigma_8$.

Key words:

1 INTRODUCTION

One of the key, unanswered, questions in modern cosmology is the accelerated expansion rate of our Universe. Since its discovery almost 20 years ago (Riess et al. 1998; Perlmutter et al. 1999), it has been confirmed by various other observations, leading to the establishment of the concordance $\Lambda$ cold dark matter ($\Lambda$CDM) model, in which the Universe is dominated by a small cosmological constant $\Lambda$ which is solely responsible for the accelerated expansion at late times, while the formation of its structures has largely been shaped by dark matter component under the action of gravity, which is assumed to be described by Einstein’s General Relativity (GR). While the simple $\Lambda$CDM model describes almost all cosmological observations very well (WMAP Collaboration 2013; Planck Collaboration 2016), currently there still lacks a satisfactory theoretical explanation for the smallness of $\Lambda$ required by observations, and this has led to significant effort in developing alternative scenarios, such as those which assume that GR is inaccurate and must be replaced by some modified gravity (MG) model on cosmological scales (see, e.g., Joyce et al. 2015; Koyama 2016, for recent reviews). Although a commonly accepted MG model still does not exist, and many of the MG models being proposed indeed introduce a $\Lambda$ through back door, studies in this field have so far led to various interesting possibilities of deviations from GR, which serve as useful testbeds of the validity of GR in cosmology. In recent and coming years, various imaging and spectroscopic galaxy surveys are producing high-quality data for a range of cosmological probes, with which we can hope to improve our understanding of the nature of the cosmic acceleration, along which it is hopeful to push the test of GR to much larger scales than previously attained.

Weak gravitational lensing (Bartelman & Schneider 2001; Refregier 2003; Hoekstra & Jain 2008; Kilbinger 2015) is one of the key cosmological probes that such galaxy surveys offer. It describes the effect that images of distant sources (e.g., galaxies or the CMB itself) are distorted by the intervening large-scale structures, which bend the paths of the photons emitted by the sources. Such bending is caused by visible plus dark matter, so that weak lensing offers a venue to detect the total matter distribution between the source and observer. Weak lensing can be observed in different ways, depending on the size of the lensing objects. At one extreme, the lensing effect by the largest bound objects in the Universe – galaxy clusters – is strong enough that it can be detected around individual clusters. At the other extreme, cosmic

* E-mail: baojiu.li@durham.ac.uk
† E-mail: masato.shirasaki@nao.ac.jp

© 2017 The Authors
shear describes statistically the lensing effect of the entire matter distribution along the lines of sight. Other objects, such as cosmic voids and galaxies, produce lensing signals which are most often not strong or clean enough to allow a clear detection for individual lenses. However, with the stacking of a large number of lenses, a signal can be extracted and it can tell us how matter is distributed around such objects. In this paper we will focus on the lensing of background (source) galaxies by foreground (lens) galaxies, or galaxy-galaxy lensing (GGL; Brainerd et al. 1996; Hudson et al. 1998; Guzik & Seljak 2002; Hoekstra et al. 2004; Mandelbaum et al. 2006a). For some recent works on testing cosmological models with cluster or void lensing, see Cai et al. (2015); Cautun et al. (2017); Barreira et al. (2015a,b, 2017).

As a gravitational effect, weak gravitational lensing can naturally be used to test gravitational physics. On cosmological scales, a deviation from standard GR can leave detectable imprints on lensing observations in various ways. For example, it could change the expansion history through a different mechanism than $\Lambda$ to accelerate the expansion or via different best-fit cosmological parameters such as $\Omega_m$ and $H_0$, leading to different angular diameter distances to lenses and sources at given redshifts. It may also have a different law of gravitational interaction, enhancing or reducing the clustering of matter at cosmological scales. It may also affect the propagation of photons, such as in the Galileon (Nicolis et al. 2009; Deffayet et al. 2009), nonlocal (Maggiore & Mancarella 2014; Dirian et al. 2014) and beyond Horndeski (Glazyez et al. 2015) gravity models. These effects, unfortunately, can have certain degeneracies between each other (and degenerate with the effects of cosmological parameters such as $\Omega_m$ and $\sigma_8$), and in this paper we focus on a simple model – chameleon $f(R)$ gravity – which practically does not modify the expansion history and photon propagation. This is one of the most popular classes of modified gravity models, which has the property of chameleon screening (Khoury & Weltman 2004; Mota & Shaw 2007) to suppress the deviation from GR in regions of deep gravitational potential (such as our Solar System) and ensure that the theory passes local tests of gravity.

The study of GGL in chameleon $f(R)$ gravity presented here is based on numerical simulations. Instead of doing a ray tracing to calculate the stacked lensing signal around galaxies, we will follow an equivalent approach by integrating the cross correlation function $\xi_{gm}(r)$ of galaxies and matter along the line of sight. The calculation is standard, but a significant effort of this study will be devoted to the construction of mock galaxy catalogues used to find $\xi_{gm}(r)$. The reason for this carefullness is twofold. Firstly, galaxies are observable and biased tracers of the underlying matter density field, and different populations of galaxies have different biases and density profiles around them. Therefore, it is important to know which galaxy population in observations should our simulation prediction of GGL signal be confronted to. Secondly (and which has perhaps not been emphasised enough so far), there is only one observable Universe, while there are many theories. For probes such as GGL, which involve the distributions of both source galaxies and the total matter field, the difference between the predictions of two models can come from differences in both. If, say, two models predict very different clustering of foreground galaxies, then one of them may already be incompatible with the spectroscopic observation used to identify these lenses, and the comparison of its GGL prediction to observations no longer makes sense. Of course, given the complex and poorly understood galaxy-mass and galaxy-halo connection, at this point it is premature to rule out the models studied here purely based on their predicted galaxy clustering. For example, commonly adopted frameworks to populate galaxies into simulations, such as halo occupancy distribution (HOD), abundance matching (SHAM) and semi-analytical modelling (SAM), usually have or can accept enough free parameters by tuning which the predicted and observed galaxy clustering can be matched (as we shall show below).

Therefore, in this work we follow a pragmatic approach, by assuming that we have no idea whether GR or $f(R)$ gravity is the correct theory, and that for both of them the free parameters of the HOD (which we use to make galaxy mocks) can be tuned to produce acceptable clustering properties of the resulting galaxy catalogues, which will then be used to make predictions of GGL.

The layout of this paper is as follows. In §2 we present a short description of the $f(R)$ gravity model studied here, a concise summary of weak gravitational lensing and GGL, and a detailed discussion of our numerical simulations and mock HOD galaxy catalogues. In §3 we present the main results of this paper, including the power spectrum of cosmic shear, cross correlation between galaxies and matter, galaxy bias, and GGL in both GR and $f(R)$ gravity. We will also present a forecast of the potential to distinguish the two models using cosmic shear and GGL, for galaxy surveys such as Dark Energy Survey (DES)\(^1\), the Hyper Supreme Camera (HSC)\(^2\) and Large Synoptic Survey Telescope (LSST)\(^3\). Finally, in §4 we sum up and discuss potential ways in which this work can be further improved.

Throughout this paper, we use the unit convention $c = 1$ where $c$ is the speed of light. An overbar denotes the background value, and a subscript 0 denotes the present-day value, of a quantity. Greek indices $\mu, \nu, \cdots$ run over 0, 1, 2, 3 (space-time coordinates) while Latin indices $i, j, k, \cdots$ run over 1, 2, 3 (space coordinates only). The Einstein summation convention is used across the paper unless otherwise stated. Also, we only consider a spatially flat universe.

2 THE THEORETICAL FRAMEWORK

We start with a concise description of $f(R)$ gravity and the model studied in this paper (§2.1), the formulae of weak gravitational lensing (§2.2), and the $N$-body simulations used in the analyses and the catalogues of HOD galaxies which are needed to study their cross correlation with lensing (§2.3).

2.1 The $f(R)$ gravity theory

2.1.1 Generic $f(R)$ gravity

$f(R)$ gravity as an attempt to explain the accelerated late-time expansion of the Universe without invoking a cosmological constant was first proposed in Carroll et al. (2004, 2005). It is constructed by replacing the Ricci scalar $R$ in the standard Einstein-Hilbert action for GR with a function of $R$, $f(R)$:

$$S = \int d^4x\sqrt{-g}\ 1\over 2 M^2_{P} [R + f(R)],$$

in which $M^2_{P}$ is the reduced Planck mass, with $M^2_{P} = 8\pi G$ and $G$ being Newton’s constant, $g$ is the determinant of the metric $g_{\mu\nu}$. The above is the gravitational action; the matter part is assumed to be the same as in standard ΛCDM, and so not explicitly given here.

\(^{1}\) https://www.darkenergysurvey.org

\(^{2}\) http://hsc.mtk.nao.ac.jp/zsp/

\(^{3}\) https://www.lsst.org
A modified version of the Einstein equation can be derived by varying the action Eq. (1) with respect to the metric $g_{\mu \nu}$:

$$G_{\mu \nu} + f_k R_{\mu \nu} + \left[ \Box f_k - \frac{1}{2} f \right] g_{\mu \nu} - \nabla_\mu \nabla_\nu f_k = 8 \pi G T_{\mu \nu},$$  \hspace{1cm} (2)

where $G_{\mu \nu} \equiv R_{\mu \nu} - \frac{1}{2} g_{\mu \nu} R$ is the Einstein tensor, $\nabla_\mu$ the covariant derivative, $f_k \equiv df/\delta R$ is a (new) scalar degree of freedom in this theory, $\Box \equiv \nabla_\mu \nabla^\mu$ the Laplacian and $T_{\mu \nu}$ the energy-momentum tensor for matter.

Eq. (2) is a fourth-order equation in the metric tensor because the Ricci scalar $R$ itself contains second-order derivatives of the latter. It is convenient to cast it into a form involving the usual Einstein equation in GR plus the new scalar field (sometimes called the scalaron) $f_k$. The Klein-Gordon equation for $f_k$ can be derived as the trace of the modified Einstein equation:

$$\Box f_k = \frac{1}{3} [R - f_k R + 2f + 8\pi G \rho_m],$$  \hspace{1cm} (3)

in which $\rho_m$ is the density for non-relativistic matter. We have neglected the contribution of relativistic matter species in this paper because we will focus on late times only.

The analysis in this study will be restricted to lengths scales much smaller than the horizon, in which case we can use the quasi-static approximation to drop all time derivatives of $f_k$ as compared with their spatial derivatives (see, e.g., Bose et al. 2015, for a detailed discussion of this approximation). As a result of this simplification, Eq. (3) becomes

$$\tilde{\Box} f_k = \frac{1}{3} [R - f_k R + 2f + 8\pi G \rho_m],$$  \hspace{1cm} (4)

where $\tilde{\Box}$ is the spatial gradient and $\alpha$ the scale factor ($\alpha = 1$ today).

Also under the quasi-static and weak-field approximation, the Poisson equation that governs the Newtonian potential $\Phi$ is modified:

$$\tilde{\Box}^2 \Phi = 16 \pi G \alpha^2 (\rho_m - \bar{\rho}_m) + \frac{1}{6} (R(f_k) - \bar{R}).$$  \hspace{1cm} (5)

The scalar field is what mediates a fifth force between massive particles in $f(R)$ gravity, and $f_k$ plays the role of its potential. To see this, we can combine Eqs. (4, 5) to obtain the following equation:

$$\tilde{\Box}^2 \Phi = 4 \pi G \alpha^2 (\rho_m - \bar{\rho}_m) + \frac{1}{2} \tilde{\Box} f_k.$$  \hspace{1cm} (6)

Eq. (6) makes explicit the behaviour of the theory in the limit when $|f_k| \ll |\Phi|$: here we can to a good approximation neglect the effect of $f_k$ and recover the usual Poisson equation

$$\tilde{\Box}^2 \Phi = 4 \pi G \delta \rho_m,$$  \hspace{1cm} (7)

in which $\delta \rho_m \equiv \rho_m - \bar{\rho}_m$ is the matter density perturbation. The GR solution $R = -8 \pi G \rho_m$ is also recovered in this regime, where the fifth force is effectively suppressed, as the consequence of the so-called chameleon screening mechanism (Khoury & Weltman 2004; Mota & Shaw 2007).

As the opposite limit, when $|f_k| \geq |\Phi|$, we have $|\delta \rho_m| \ll \delta \rho_m$ with $\delta R \equiv R - \bar{R}$, and so the second term on the right-hand side of the modified Poisson equation, Eq. (5), can be neglected:

$$\tilde{\Box}^2 \Phi \approx \frac{16}{3} \pi G \delta \rho_m.$$  \hspace{1cm} (8)

Compare this with Eq. (6), we can notice a $1/3$ enhancement of the strength of gravit, independent of the functional form of $f(R)$. The exact form of $f(R)$, though, determines the transition between the two limiting regimes. We shall call these two regimes respectively the screened and unscreened regimes.

We can see that the chameleon mechanism works depending on the local Newtonian potential $\Phi$, and it efficiently screens the fifth force in environments where $\Phi$ is deep (i.e., $|\Phi| \gg |f_k|$). This can be qualitatively understood as follows: the fifth force is mediated by the scalar field $f_k$, which has a mass $m_s$ given by

$$m_s^2 = \frac{d^2 V_{\text{eff}}(f_k)}{df_k^2},$$  \hspace{1cm} (9)

where the effective potential $V_{\text{eff}}(f_k)$, due to the self-interaction of the scalar field and its interactions with (non-relativistic) matter, is given by

$$\frac{dV_{\text{eff}}(f_k)}{df_k} = \frac{1}{3} [R - f_k R + 2f + 8\pi G \rho_m].$$  \hspace{1cm} (10)

As a result of $m_s \neq 0$ in general, the fifth force mediated by $f_k$ has the Yukawa form, with a potential $\sim r^{-1} \exp(-m_s r)$ where $r$ is the distance from a massive particle. The complicated form of $V_{\text{eff}}(f_k)$ makes $m_s$ depend on environment, and becomes heavy in deep $\Phi$, where the Yukawa force decays very quickly with distance, such that its effect is not felt beyond $\sim m_s^{-1}$. This is the origin of the chameleon screening, and this property can help the theory to pass stringent solar system tests, since it is expected that screening has effectively suppressed the fifth force to an undetected level at locations where such tests have been performed.

2.1.2 The choice of $f(R)$ model

The requirement of chameleon screening be in place for the model to pass solar system tests does not significantly constrain the possible functional form of $f(R)$. As a result, many different forms have been studied in the literature with differing details. While many of these are interesting, in practice it is both impossible and unnecessary to study all of them with equal detail. Instead, there are reasons why we should focus on a particular example which is to be studied in greater details.

First, there is currently not a fundamental theory to naturally motivate a functional form of $f(R)$ that leads to the cosmic acceleration, and therefore all choices of $f(R)$ used in the literature so far are phenomenological. There is no clear reason why any choice should be preferred over the others, apart from possibly a more apparent simplicity in the functional form of $f(R)$. However, it is known that for general $f(R)$ models that realise an efficient chameleon screening, the background expansion history has to be very close to that of $\Lambda$CDM (e.g., Brax et al. 2008; Wang et al. 2012; Ceron-Hurtado et al. 2016); in other words, a $\Lambda$ is usually introduced to the theory, possibly in an implicit way, regardless of the simplicity of $f(R)$.

Second, as mentioned above, the different $f(R)$ models generally share some common features, such as chameleon screening in deep Newtonian potentials, and differ primarily in how efficient the screening is. However, given a functional form of $f(R)$, the screening efficiency also depends on the parameter used. At least qualitatively, the change of behaviours by varying the form of $f(R)$ can be mimicked by varying the parameters with a chosen $f(R)$. Instead of letting observations determine the form of $f(R)$, it is pragmatically more useful to use observations to determine to what extent deviations from GR as prescribed by $f(R)$ gravity are allowed. The latter task can be carried out by working on a particular model which is to be tested against as many observational data sets as possible, as precisely as possible.

For these reasons, this work is based on the model proposed by Hu & Sawicki (2007, hereafter HS). This is the most well-studied model.
example of $f(R)$ gravity, so that the results of this paper will be built upon various existing tests of this model. The particular functional form of $f(R)$ in this model makes it possible to implement an efficient algorithm to speed up simulations of it (Bose et al. 2017), which we adopt for the simulations used in this work. Additionally, this model is a representative example of classes of scalar-tensor type modified gravity theories, in that by varying its parameters we can have a range of behaviours ranging from strong screening to no screening.

The HS model has the following functional form of $f(R)$,

$$f(R) = -\frac{c_1}{c_2} \left( \frac{R}{M^2} - \frac{1}{2} \right) M^2,$$

(11)

where $M$ is a parameter of mass dimension that is given by $M^2 = 8\pi G\rho_{m0}/3 = H_0^2\Omega_m$, $H$ the Hubble rate, $\Omega_m$ the present-day density parameter for non-relativistic matter, and $c_1, c_2$ are dimensionless model parameters.

In the limit $|R| \gg M^2$, $f(R)$ is approximately a constant $-\frac{c_1}{c_2} M^2$, so that $f_R$ and its derivatives are small. In this case, Eq. (3) can be simplified as

$$-\dot{R} \approx -2f + 8\pi G\rho_m \approx 3M^2 \left[ \frac{2c_1}{3c_2} - a^{-3} \right].$$

(12)

The background expansion rate is therefore close to that of $\Lambda$CDM, if we make the choice $\Omega_m = 0$, where $\Omega_m \equiv 1 + \Omega_m$. For $\Omega_m \sim 0.7$ and $\Omega_m \sim 0.3$, we have $|\dot{R}| \sim 0.7M^2 \gg M^2$ today. As $|\dot{R}|$ increases with redshift, the approximation in Eq. (12) is good all the time. In this approximation, we have the following simplified relation between $f_R$ and $R$,

$$f_R \approx -\frac{c_1}{c_2} \left( \frac{M^2}{-\dot{R}} \right)^{n+1}.$$  

This relation can be inverted to find $R(f_R)$, as the latter appears in field equations. Therefore, with choices of $n$ and $c_1/c_2$, as well as $\Omega_m, \Omega_k, H_0$, the model ($f(R)$) model can be fully specified. In the literature, instead of $c_1/c_2$, people usually use $f_{R0}$, the current value of $f_R$, as the model parameter. We have

$$\frac{c_1}{c_2} = -\frac{n}{f_{R0}} \left[ 1 + 4\Omega_m \right]^{n+1}.$$  

(14)

In this paper, we shall focus on a particular choice of model parameters: $n = 1$ and $f_{R0} = -10^{-5}$, which we refer to as F5. This choice of $f_{R0}$ is almost certainly incompatible with local gravity tests as the Milky Way galaxy is unlikely to be screened. However, the choice is not yet completely ruled out by cosmological observations, and for the GGL analysis we would like to choose a model that can maximise the difference from GR (see, e.g., Cai et al. 2015; Cataneo et al. 2015; Liu et al. 2016; Shirasaki et al. 2016, 2017; Peirone et al. 2017; Cautun et al. 2017, for some recent studies on the current and potential constraints $f(R)$ gravity).

### 2.2 Weak gravitational lensing

Weak lensing is a matured field with a huge body of research works in the literature (see, e.g., Bartelmann & Schneider 2001; Hoekstra & Jain 2008; Kilbinger 2015, for some reviews). Here we only give a very quick catch-up of some essential equation to be used in the discussion below. For simplicity, we assume a flat Universe.

As photons pass through the large-scale structure in the Universe, their paths are bent by the latter, resulting in a change of their apparent angular position $\xi_\theta \equiv \bar{\xi}(\chi = 0)$ as seen by the observer at today (where $\chi$ is the comoving distance, as compared to the true one $\xi_\theta \equiv \bar{\xi}(\chi_s)$, at the source (where $\chi_s$ is the comoving distance to the source). This is given by

$$\xi_\theta = 2\int_0^{\chi_s} g(\chi, \chi_s) \bar{\xi} \, d\chi,$$

(15)

where

$$g(\chi, \chi_s) \equiv \frac{\chi}{\chi_s} (\chi_s - \chi),$$

(16)

is the lensing kernel, and $\Phi \equiv \Phi(\chi, \bar{x})$ is the lensing potential along the line of sight, and $\bar{\nabla}_x$ is the two-dimensional derivative in the plane perpendicular to the l.o.s.. The resulting distortions of source images can be described by a $2 \times 2$ distortion matrix given by

$$A_{ij} = \delta_{ij} - 2 \int_0^{\chi_s} \frac{g(\chi, \chi_s)}{\chi} \nabla_x \bar{\xi} \nabla_x \Phi(\chi, \bar{x}) \, d\chi,$$

$$\approx \delta_{ij} - 2 \int_0^{\chi_s} \frac{g(\chi, \chi_s)}{\chi} \nabla_x \bar{\xi} \nabla_x \Phi(\chi, \bar{x}) \, d\chi,$$

(17)

where $i, j = 1, 2$, $\delta_{ij}$ is the Kronecker delta, $\nabla_x \bar{\xi}$, and $\nabla_x \Phi$ are respectively the derivative with respect to $\bar{x}_i$ at $\chi = 0$ and $\chi$. We have used $\nabla_x \bar{\xi} \approx \nabla_x \bar{\xi}$ in the second step of the above equation. This is known as the Born approximation, and its validity on weak lensing power spectrum and GGL has been studied in, e.g., Hilbert et al. (2009).

The lensing convergence $\kappa$, shear $\gamma_1, \gamma_2$ and rotation $\omega$ are related to the distortion matrix as

$$A = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 - \omega \\ -\gamma_2 + \omega & 1 - \kappa + \gamma_1 \end{pmatrix},$$

(18)

where one can relate the convergence $\kappa$ to the density contrast $\delta \equiv \rho_m/\rho_m - 1$, by using the Poisson equation:

$$\delta = \frac{3}{2} \Omega_m \bar{H}_0^2 \frac{\chi_s}{\chi} g(\chi, \chi_s) \nabla^2 \delta \, d\chi.$$  

(19)

Using the Limber approximation, the convergence power spectrum can be written as

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

(20)

where the lensing weight function $W(\chi)$ is given by

$$W(\chi) = \frac{3}{2} \Omega_m \bar{H}_0^2 g(\chi, \chi_s) |1 + z(\chi)|,$$

(21)

$P_{\delta \delta}$ is the matter power spectrum, and $z(\chi)$ is the redshift corresponding to comoving distance $\chi$.

For GGL, we study the tangential shear profile around galaxies, which is given by

$$\gamma_{\parallel}(r_p) = \frac{\Delta \Sigma(r_p)}{\Sigma_{\text{crit}}},$$

(22)

where $r_p$ is the projected distance from the galaxy, $\Delta \Sigma(r_p) \equiv \Sigma(< r_p) - \Sigma_{\text{crit}}(r_p)$ is the excess surface density defined as the difference between the mean surface (projected) mass density at $r < r_p$, $\Sigma(< r_p)$, and the surface mass density at $r_p$, $\Sigma_{\text{crit}}(r_p)$:

$$\Sigma(< r_p) = \frac{1}{2\pi r_p} \int_0^{r_p} \Sigma(r) 2\pi dr,$$

(23)

and $\Sigma_{\text{crit}}$ is the critical surface mass density given as

$$\Sigma_{\text{crit}} \equiv \frac{1}{4\pi G} \frac{X_s}{|X_s - X_c|} \left( 1 + z(\chi) \right),$$

(24)
in which $\chi_l$ is the comoving distance of the lens galaxy. Note that, given that the background expansion history is in practice indistinguishable in the GR and F5 models studied here, $\Sigma_{\text{crit}}$ is the same for both models.

For individual galaxies, the lensing signal is weak, and so we consider the stacking of the tangential shear profile around many galaxies. In practice, the excess surface density profile can be computed as

$$\Delta \Sigma(\chi_p) = \rho_{\text{crit}} \Omega_m \frac{2}{r_p^2} \int_{-\infty}^{\infty} d\chi \int_0^{r_p} dr \cdot r \xi_{\text{gm}} \left( \sqrt{r^2 + \chi^2} \right)$$

$$- \rho_{\text{crit}} \Omega_m \int_{-\infty}^{\infty} d\chi \xi_{\text{gm}}(r_p, \chi),$$

(25)

where $\xi_{\text{gm}}$ is the cross correlation function between galaxies and the matter density field.

### 2.3 Simulations and halo/galaxy catalogues

To accurately predict lensing effect in the nonlinear regime requires numerical simulations of the matter distribution in the Universe. To cross correlate this with galaxies requires mock galaxy catalogues with mimic the real galaxy distribution as observed by galaxy surveys. In this subsection, we shall describe in detail the simulations performed for this study, and halo and galaxy catalogues used in the analysis.

#### 2.3.1 N-body simulations

The $f(R)$ gravity simulations used in our analysis have been run using the ECOsmog code (Li et al. 2012), which is a modified version of the publicly available simulation code RAMSES (Teyssier 2002), by adding new modules to solve the $f(R)$ and Einstein equations. In this work we have used the optimised version of the code (Bose et al. 2017), which adopts a new algorithm to speed up simulations for HS $f(R)$ model with $n = 1$ by a factor of up to $\sim 20$. RAMSES, and therefore ECOsmog, is efficiently parallelised using MPI, and is an example of the class of so-called adaptive-mesh-refinement (AMR) codes, which hierarchically refine a regular base mesh that covers the whole periodic simulation volume. In this way, it achieves the necessary high resolution in high density regions, without wasting substantial amount of computing resources in low density regions in which the demand for resolution is not as strong. The high resolution is also critical in order to resolve the fifth force effects in high density regions, such as in dark matter haloes, where the chameleon screening makes them weak (but not always negligible).

The cosmological parameters, listed in Table 1, have been chosen from the best-fit (WMAP Collaboration 2013) WMAP9 $\Lambda$CDM model, and their physical meanings are explained in the table caption. The technical parameters for the simulations are given in Table 2. All simulations have 1024$^3$ particles.

### Table 1. The cosmological parameters for the models investigated in this work. $\Omega_m$ and $\Omega_{\Lambda}$ are respectively the present-day fractional density of matter and the cosmological constant (in the case of $f(R)$ gravity it is simply $1 - \Omega_m$), $h = H_0/(100 \text{km s}^{-1} \text{Mpc}^{-1})$ with $H_0$ being the present-day Hubble rate, $n_s$ is the spectral index of the primordial density fluctuations, $A_s$ the amplitude of the primordial power spectrum, and $\sigma_8$ the root-mean-squared (RMS) linear density fluctuation in spheres of radius $8h^{-1}\text{Mpc}$ at $z = 0$ (the value quoted here is for $\Lambda$CDM model only, as F5 has a different value). $\Omega_b$ is the baryon density used for the linear matter power spectrum, to generate the initial conditions only.

| parameter | physical meaning | value |
|-----------|------------------|-------|
| $\Omega_m$ | present fractional matter density | 0.2819 |
| $\Omega_{\Lambda}$ | present fractional baryon density | 0.0461 |
| $\Omega_{\Lambda}$ | $1 - \Omega_m$ | 0.7181 |
| $h$ | $H_0/(100 \text{km s}^{-1} \text{Mpc}^{-1})$ | 0.6970 |
| $n_s$ | primordial power spectral index | 0.9710 |
| $\log_{10} A_s$ | amplitude of the primordial power spectrum | $-8.622$ |
| $\sigma_8$ | RMS density fluctuation at $8h^{-1}\text{Mpc}$ for $\Lambda$CDM | 0.8178 |
| $n$ | HS $f(R)$ parameter | 1.0 |
| $f(0)$ | HS $f(R)$ parameter | $-10^{-5}$ |

### Table 2. The simulation specifications. $L_{\text{box}}$ is the simulation box size, $N_p$ the number of simulation particles, $m_p$ the simulation particle mass, $N_t$ the number of realisations for each box size and $N_e$ the number of particle outputs (which are spaced every $75h^{-1}\text{Mpc}$ in comoving distance from today). $z_{\text{ini}}$ is the starting redshift of the simulations, and the initial condition is generated using the 2LPTIC code. Out of the $N_t$ snapshots, 33 are between $z = [0, 1]$.

| $L_{\text{box}}[h^{-1}\text{Mpc}]$ | $z_{\text{ini}}$ | $N_p$ | $m_p [10^9M_{\odot}]$ | $N_t$ | $N_e$ | IC |
|------------------------|-----------|-------|----------------|-------|-------|----|
| 450                    | 49.0      | 1024$^3$ | 6.64$\times 10^9$ | 1      | 37    | 2LPT |
| 900                    | 36.0      | 1024$^3$ | 5.31$\times 10^9$ | 1      | 37    | 2LPT |

The different between the two models is negligible. We only have one realisation of simulation for each box size.

#### 2.3.2 Halo and galaxy catalogues

The dark matter haloes used in this paper are found using the publicly available phase-space friend-of-friend halo finder ROCKSTAR (Behroozi et al. 2013), and we have chosen $M_{200c}$ as the mass definition, in which the subscript means that the average mass density within halo radius $R_{200c}$ is 200 times the critical density $\rho_c$.

We populate dark matter haloes using halo occupation distribution (HOD; Berlind et al. 2003; Zheng et al. 2005). In the simplest form of this model, the mean number density of galaxies in a host dark matter halo is a function of the halo mass $M$:

$$\langle N_g | M \rangle = \langle N_g | M \rangle_1 \left[ 1 + \langle N_g | M \rangle_1 \right],$$

(26)

where $N_g$, $N_c$ and $N_s$ are respectively the number of all, central and satellite galaxies, given as

$$\langle N_g | M \rangle_1 = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right],$$

(27)

$$\langle N_c | M \rangle = \left( \frac{M - M_{\text{cmin}}}{M_1} \right)^{\alpha} \Theta(M - M_0),$$

(28)

where $\text{erf}(x)$ and $\Theta(x)$ are respectively the error and Heaveside
step functions. The model has five parameters: \( M_{\text{min}}, M_0, M_1, \Omega_{\log M} \) and \( \alpha \), which give it great freedom to tune the galaxy catalogues to match their statistical properties to observables.

To implement the HOD model, we sift through the halo catalogue where subhaloes are eliminated, and for each halo let it host a central galaxy if \( n \leq (N_i/M) \) where \( n \) is a random number generated from a uniform distribution between \([0, 1]\). The number of satellite galaxies is set to a random number generated from the Poisson distribution with mean \( \langle N_i \rangle \), and the satellite galaxies are radially distributed within the dark matter halo following the Navarro-Frenk-White (NFW; Navarro et al. 1997) profile, using the concentration parameters measured by ROCKSTAR. In the rare cases where a halo has no central galaxy but does have satellites, we promote the first satellite galaxy to a central galaxy.

Although this paper does not aim to compare theoretical predictions with real observations, we still want the HOD catalogues to bear a certain degree of reality. To this end, we focus on galaxies at relatively low redshift (where the model difference is expected to be larger) and make the resulting galaxies satisfy a redshift distribution similar to that of the Low Redshift Sample (LOWZ) of BOSS data release 11. Instead of down-sampling generated HOD catalogues, we follow Manera et al. (2015) to allow a redshift dependence of the HOD mass parameters as follows:

\[
\log M_{\text{min}} = \log M_{\text{min}}^* + S n(z)/\left(10^{-4}\right),
\]

\[
\log M_{0,1} = \log M_{0,1}^* + T n(z)/\left(10^{-4}\right),
\]

in which \( n(z) \) is the target galaxy number density for redshift \( z \), \( S = -0.925, T = -0.928 \), and \( M^* \) are the respective mass parameter values at \( n^* = 2.98 \times 10^{-4} \). Note that \( M_1/M_0 \) has no time evolution. We consider galaxies in the redshift range of \( 0.16 \leq z \leq 0.43 \). We use 29 redshift bins in this range, with equal comoving thickness, with 3 bins being taken from each of the 10 snapshots of this box. For the 3 bins from the same simulation snapshot, the evolution of \( n(z) \) is only down to Eqs. (29, 30), while for the bins from different snapshots \( n(z) \) is also affected by the fact that different halo catalogues have been used to build the HOD catalogues.

Should the HOD catalogues in F5 have been constructed with the same HOD parameters as in their \( \Lambda \)CDM counterparts, there would generally be a difference of \( \sim 10-20\% \) in the number density and clustering of galaxies. Because there is only one observed galaxy catalogue in the Universe (assuming an ideal full-sky survey), if we do not know which is the correct cosmological model, we can only demand that both models make predictions that agree with observations. For this reason, we have opted to tune the HOD parameters for F5 so that the resulting galaxy catalogues match the number density and clustering of their \( \Lambda \)CDM counterparts. The assumption that F5 and GR should have different HOD parameters is not unreasonable, given their different dark matter evolutions and galaxy assembly histories. This choice indeed also helps to fix the galaxy clustering and single out the expected difference of the galaxy-galaxy lensing signals in these two models.

The tuning of F5 HOD parameters has been performed using a search with the Nelder-Mead simplex algorithm through the 5-dimensional parameter space. The 3D real-space correlation functions in the F5 and GR HOD catalogues are measured between co-moving separations of 0.6 and 600k-1Mpc, in 40 equally-spaced logarithmic bins. The r.m.s. difference between the two models is calculated with equal weight 1 for all bins. To try to make the two models have similar \( n(z) \), the relative difference in their \( n(z) \) values is also added into the r.m.s., with a weight 8. The code then searches through the 5D parameter space looking for the smallest r.m.s. difference (\( \chi^2 \)). The search stops when \( \chi^2 \leq 0.03 \), meaning that the overall agreement (as defined in the above way) is better than 3% (for some redshift bins better than 2%). We have not attempted to do better than this accuracy, or perform a full parameter search using Markov chain Monte Carlo, because it is not the purpose of this paper to study in great detail the HOD model in the context of modified gravity, and because it is in general difficult to do better than this level of accuracy (for example, the galaxy clustering may change at few percent level if a different set of random seeds is used in generating the HOD catalogues). For the same reason, we shall not list the HOD parameters in this paper.

In the left panel of Figure 1 we compare the measured 3D real-space correlation functions \( \xi(r) \) as a function of galaxy separation \( r \), respectively for F5 (red dashed lines) and GR (blue solid) at 6 of the 29 redshift bins. In the right panel of Fig 1 (the same colour scheme) shows the redshift distribution of the HOD galaxies for the two models – again there is a good agreement (generally better than 4%) after tuning the HOD parameters for F5.

### 3 Weak Gravitational Lensing in \( f(R) \) Gravity

In this section we present the main result of this research, beginning with an analysis of the weak lensing convergence. The convergence power spectrum in \( f(R) \) gravity, amongst various other things, have been studied in a few works (Shirasaki et al. 2015, 2017; Higuchi & Shirasaki 2016; Tasseur et al. 2015; Pratten et al. 2016), using various techniques. We revisit this topic in §3.1 with a detailed decomposition of the convergence power spectrum \( C_{\kappa\kappa}(\ell) \) in terms of the length scales and times where the contributions to the relative difference between \( f(R) \) gravity and GR come from. The analysis not only serves as a sanity check of the simulations, and their corresponding particle snapshots, but also will be used for comparison with results on galaxy-galaxy lensing from §3.2, as they are based on the same set of simulations.

#### 3.1 The convergence field

The weak lensing convergence is an important physical quantity for weak lensing research: it can be calculated directly from a simulation by a projection of the density field, and in the regime of weak lensing its power spectrum coincides with that of the cosmic shear, which is a directly-observable quantity. It has been one of the most widely used quantities in theoretical analyses (see reference above for examples in \( f(R) \) gravity).

##### 3.1.1 The convergence power spectrum in \( f(R) \) gravity

In this work we compute the weak lensing convergence power spectrum \( C_{\kappa\kappa}(\ell) \) by directly integrating the 3D matter power spectrum \( P_{\delta\delta}(k, z) \) measured from the simulations, following the Limber approximation Eq. (20), instead of using full ray tracing. This is partly because the former approach essentially samples all the relevant \( k \)-modes provided by the simulation, while ray tracing only samples those \( k \) modes that fall into the designed lightcone. While the latter more closely mimics the real Universe, the former gives smoother theoretical curves. Eq. (20), together with an accurate prediction of the nonlinear \( P_{\delta\delta} \) such as from HALOFIT (Smith et al. 2003; Takahashi et al. 2012), is a powerful tool to predict \( C_{\kappa\kappa} \).

The nonlinear \( P_{\delta\delta} \) used in this work are measured using the
Figure 1. (Colour Online) Left Panels: Comparison of the GR (black solid) and F5 (red dashed) correlation functions, after tuning the HOD parameters for F5 such that the two models predict similar galaxy number densities and clusterings. For cleanness we only show results for 6 of the 29 redshift bins from the 900h⁻¹ Mpc box. Each bin is taken from HOD catalogues built using a different output snapshot. Right Panel: Comparison of the predicted GR (black solid) and F5 (red dashed) galaxy redshift distributions. See the text for more details on how the galaxy numbers in each redshift bin are determined for GR and F5.

Figure 2. (Colour Online) Left Panels: the effect of choosing a coarser – Δχ = 75h⁻¹ Mpc in blue solid vs Δχ = 25h⁻¹ Mpc in black solid – sampling of P(k, z) (generated using HALOFIT) when using the Limber approximation Eq. (20) to compute the convergence power spectrum C_{κκ}(ℓ). The solid red line shows the result of doing the same integration with Δχ = 75h⁻¹ Mpc, but for the simulated P(k, z). The lower sub-panel shows the relative differences of the other two cases w.r.t. the integration of HALOFIT P(k, z) with Δχ = 25h⁻¹ Mpc. Right Panels: the convergence power spectra for F5 (red) and GR (black), using two source redshifts: z_{source} = 1.0 (solid) and 0.5 (dashed) – all results are from integrating the simulated P(k, z) following Eq. (20). The lower sub-panel shows the relative differences between F5 and GR. See the main text for discussions of this plot.
publicly available POWMES code (Colombi & Novikov 2011) with a Fast Fourier Transform (FFT) grid of size 2048. As mentioned above, our simulations have outputs at 33 snapshots between $z = 1$ and $z = 0$ equally spaced in comoving distance ($\Delta \chi = 75h^{-1}\text{Mpc}$), and therefore this is the finest grid in $z$ (or in $\chi$) that our integration of Eq. (20) can be carried out. We used the 5-point Newton-Cotes formula for the numerical integration over $\chi$, and values of $P_{\delta\delta}(k)$ at a given redshift $z(\chi)$ were obtained by linear interpolation of the corresponding values at $k_i, k_{i+1}$ with $k_i < k < k_{i+1}$, where $k_i$ is the $i$-th grid point in the POWMES output. To check that such a coarse spacing in $\chi$ is not causing substantial numerical errors, we made a test using a finer sampling of $P_{\delta\delta}(k, z(\chi))$ in the $\chi$ direction, with a comoving thickness of $25h^{-1}\text{Mpc}$, generated using HALOFIT. We then integrated this sample (1) using all $\chi$ bins, and (2) using one from every three neighbouring bins (the latter mimicking the simulation binning of $\Delta \chi = 75h^{-1}\text{Mpc}$). The results of these two tests are shown as black and blue solid lines in the upper left panel of Figure 2, and their relative difference is presented as the blue solid line in the lower left panel. The tests showed that the difference is much smaller than 1%, and that $\Delta \chi = 75h^{-1}\text{Mpc}$ is fine enough for our study. We have also checked that the convergence power spectrum (for GR) computed in this way agreed well with the result of CAMB (Lewis & Challinor 2011) for the same cosmology, though note that to match the $k$ range of the simulated $P_{\delta\delta}(k, z)$ we have limited the maximum $k$ in the integration to $k_{\text{max}} = 28.6 h^{-1}\text{Mpc}$, which affects $C_{\delta\delta}$ a little bit at $\ell \sim 10^4$. The red solid line in the upper left panel of Fig. 2 is the result of the same integration of the simulated $P_{\delta\delta}(k, z)$ for GR, and the red solid line in the lower left panel is the relative difference from the integration of HALOFIT $P_{\delta\delta}(k, z)$. The simulation result peels off at $\ell \lesssim 20$ because the lowest $k$ values used in the integration is $k_{\text{min}} \sim 0.014 h^{-1}\text{Mpc}$, and the same feature appeared in the HALOFIT results if the same $k_{\text{min}}$ was applied there. We notice that the simulation and HALOFIT results agree reasonably well, with the former lower by ~2% in a wide range of $\ell$, possibly due to sampling variance (recall that the simulation box is 450$h^{-1}\text{Mpc}$). The difference between the two further changes to 5–6% when $\ell$ approaches $10^5$, due to slight loss of resolution by the simulation. The right panels of Figure 2 compare the convergence power spectra of F5 (red) and GR (black), for two lensing source redshifts, $z_{\text{source}} = 1.0$ (solid) and $z_{\text{source}} = 0.5$ (dashed). For both source redshifts, an enhancement of order 10%–30% between $\ell \sim 100$ and $\ell \sim 10^4$ is found, which increases with $\ell$, and the relative difference is larger in the case of $z_{\text{source}} = 0.5$. 3.1.2 The connection between matter and convergence spectra To gain a clearer understanding of the behaviour shown in the right panels of Figure 2, recall from Eq. (20) that $C_{\delta\delta}(\ell)$ at a given angular scale $\ell$ gets contribution from various $k$ modes of the nonlinear $P_{\delta\delta}$ at different times. It is therefore useful to decompose the contributions from the different times and $k$ modes, onto which we can readily map the relative differences in $P_{\delta\delta}(k, z)$ themselves. Figure 3 shows the results from such an attempt for $z_{\text{source}} = 1.0$ (left panel) and $z_{\text{source}} = 0.5$ (right panel). In both cases, the plot is the fractional contribution (black solid curves) from distances below a specific $\chi$ (indicated with the blue numbers to the right of the right axis; in units of $h^{-1}\text{Mpc}$) against $\ell$. On the top of each black solid curve the red number indicates the redshift of the corresponding $\chi$. The colour-coded lines along the vertical direction denote the values of $\ell$ that contributes at the given $\ell$ (as shown in the horizontal axis) and $\chi$ (as represented by the black solid curves); each of them corresponds to a fixed $k$ with the colour scheme as displayed in the colourbar on the top (running from black for $k \sim 0.004 h^{-1}\text{Mpc}$ to red for $k \sim 50 h^{-1}\text{Mpc}$). From Figure 3 we see the well-known result that most contributions at large (small) angular scales come from large (small) length scales; it also indicates that most contributions come from middle and low redshifts – for example, ~90% contributions in the case of $z_{\text{source}} = 1.0$ come from the redshift range $\lesssim 0.7$ (z $\lesssim 0.38$). Fig. 4 shows how the relative difference in the nonlinear $P_{\delta\delta}$ of F5 and GR evolves in time. We can see the known feature (e.g., Li et al. 2013) that the F5 prediction agrees with GR on large scales due to the finite ranges of the modified gravitational force, and the enhancement of clustering on nonlinear scales ($k \sim 0.1 h^{-1}\text{Mpc}$). We show 33 curves corresponding to the 33 snapshots between $z = 1$ (black) and $z = 0$ (red), which give us a comprehensive picture of not only the $k$- but also the time-dependences of the enhancement. For examples: (i) At $k \sim 0.1 h^{-1}\text{Mpc}$, the relative enhancement of $P_{\delta\delta}$ for F5 is within $1 - 3\%$ between $z = 1$ and $z = 0$. According to Fig. 3 this $k$-mode is most relevant for $\ell \sim 100$, which explains the ~2% enhancement of $C_{\delta\delta}$ there (cf. Fig. 2, the $z_{\text{source}} = 1$ case – the same below). (ii) At $k \sim 1 h^{-1}\text{Mpc}$, $P_{\delta\delta}$ in F5 is enhanced by 10–18% between $z \sim 0.7$ and $z = 0$, leading to a ~15% of $C_{\delta\delta}$ at $\ell \sim 1000$. (iii) Between $k = 1$ and a few $h^{-1}\text{Mpc}$ is the transition regime between 1- and 2-halo terms in the halo model, where the enhancement in F5 $P_{\delta\delta}$ increases more slowly with $k$ than at lower $k$ values. This is reflected as a slight flattening of $\Delta C_{\delta\delta}/C_{\delta\delta}$ at $\ell$ just above 1000 (Fig. 2). (iv) At $k \sim 5 h^{-1}\text{Mpc}$, the relative enhancement of F5 $P_{\delta\delta}$ remains nearly a constant at ~20% within $z = 1$ and $z = 0$, which translates into a ~20% enhancement of F5 $C_{\delta\delta}$ at $\ell \sim 4000 - 5000$ (Fig. 2). (v) At $k > 5 h^{-1}\text{Mpc}$, there is another steep increase of the enhancement in F5 $P_{\delta\delta}$ with $k$, indicating that dark matter halos (at least the small ones which dominate in number) in F5 are more concentrated. This further increases $\Delta C_{\delta\delta}/C_{\delta\delta}$ above $\ell \sim 4000 - 5000$. (vi) At all $\ell$, $\Delta C_{\delta\delta}/C_{\delta\delta}$ is larger in the case of $z_{\text{source}} = 0.5$ than for $z_{\text{source}} = 1.0$, because in the former case the result is dominated by lower redshifts in the integration Eq. (20), at which the enhancement of the nonlinear matter power spectrum in F5 relative to GR is stronger. The results here agree with those found in (Tessore et al. 2015; Shirasaki et al. 2017), and so serve as a sanity check of the simulations. Moreover, the decomposition of Figure 3 provides a way to qualitatively understand the behaviour of the convergence power spectra based on knowledge about the scale- and time-dependences of the nonlinear matter power spectra. It can also prove useful when decomposing the degeneracies of modified gravity with other physical effects that can modify $P_{\delta\delta}$ on intermediate and small scales, such as Active Galactic Nuclei feedback and massive neutrinos (see Sembolini et al. 2011; Osato et al. 2015; Mummery et al. 2017, for some general examples); (and see also e.g., Arnold et al. 2014; Harmois-Deraps et al. 2015, for some examples in the framework of modified gravity). 3.2 Galaxy galaxy weak lensing Having looked at the behaviour of the lensing power spectra, now we move to the theoretical predictions of galaxy-galaxy weak lensing.
Figure 3. (Colour Online) A detailed decomposition showing how much the different time (redshift) intervals and \( k \)-modes in the nonlinear matter power spectrum \( P_{\delta \delta}(k, z) \) contribute to the convergence power spectrum \( C_\kappa\kappa(\ell) \) at a given \( \ell \). The decomposition is done for two source redshifts: \( z_{\text{source}} = 1 \) (left panel) and 0.5 (right panel). The vertical axis is the fractional contribution from distances smaller than a given \( \chi \), defined as \( C_\kappa\kappa(<\chi)/C_\kappa\kappa(<\chi_s) \), and the black solid lines are the values for a selection of \( \chi \) values (indicated in blue; the corresponding redshift values are indicated in red) in step of 100 \( h^{-1} \) Mpc (for \( z_{\text{source}} = 1 \)) and 50 \( h^{-1} \) Mpc (for \( z_{\text{source}} = 0.5 \)). The colour-coded solid lines (roughly) in the vertical direction are curves with constant \( k \) values that contribute to a given \( \ell \) at a given time \( z \) (or \( \chi \)), with \( k = \ell/r(\chi) \). The value of \( k \) for each of these curves can be found using the colour-bar on the top, which runs from black (\( k = 0.004 h \) Mpc\(^{-1} \)) to red (\( k = 50 h \) Mpc\(^{-1} \)). Note that each \( \ell \) in \( C_\kappa\kappa(\ell) \) receives contributions from a limited range of \( k \) (which shifts to larger values for decreasing \( z_{\text{source}} \)), and that most contributions are from low and middle redshifts.

3.2.1 The galaxy matter cross correlation

The correlation of shearing of source galaxies behind lens galaxies is given by the tangential shear profile around the lenses, which in turn is related to the excess surface density profile \( \Delta \Sigma(r) \). This can be calculated by integrating \( \xi_{\delta \delta} \) along the l.o.s., and means that information about \( \xi_{\delta \delta} \) can be obtained by studying galaxy galaxy lensing. For this reason in this subsection we shall compare the theoretical predictions of \( \xi_{\delta \delta} \) by GR and F5.

We measure \( \xi_{\delta \delta} \) directly by cross correlating the mock galaxy catalogues with the particle data from which they (and the corresponding halo catalogues) are constructed. This is done by a modified version of the publicly available CUTEBOX code (Alonso 2012). We have not down-sampled the dark matter particles, and for a galaxy number of \( N_g \sim 3 \times 10^5 \) and dark matter particle number \( N_p = 1024^3 \) the code generally finishes for a snapshot in a few hours using 12 threads. For the auto-correlation function of matter, \( \xi_{\delta \delta} \), the cost is prohibitive without down-sampling particles, and we instead calculate it by directly transforming the matter power spectra \( P_{\delta \delta} \) (see below).

In order to estimate the uncertainties caused by the larger-scale environment, we use the internal jackknife resampling method with \( N_{\text{resample}} \) resamples by dividing the simulation box into \( N_{\text{resample}} \) subboxes of equal volume. For each resample, we discard the galaxies in one of the subboxes, but still include matter from that subbox when cross correlating with galaxies in other subboxes. The cross correlation

Figure 4. (Colour Online) The relative enhancement of the nonlinear matter power spectrum \( P_{\delta \delta} \) in F5 relative to GR, as a function of \( k \) (the horizontal axis) and redshift \( \chi \) (equally spaced in \( \chi \) with a \( \Delta \chi = 75 h^{-1} \) Mpc; colour-coded according to the colour-bar on the top).
function of the $i$-th jackknife resample, $\xi_{\text{GM}}^{(i)}$, is estimated as

$$1 + \xi_{\text{GM}}^{(i)}(r, r + \Delta r) = \frac{N_{\text{DD}}^{(i)}(r, r + \Delta r) - N_{\bar{p}} N_{\bar{g}}^{(i)} \Delta V/V}{N_{\bar{p}} N_{\bar{g}}^{(i)} \Delta V/V},$$

where $V$ is the whole volume of the simulation box, $\Delta V$ is the volume of the radius bin $[r, r + \Delta r]$, $N_{\bar{p}}$ the total number of dark matter particles in the simulation volume, $n_{\bar{p}} = N_{\bar{p}}/V$, and $N_{\bar{g}}^{(i)}$ the number of galaxies and $N_{\text{DD}}^{(i)}(r, r + \Delta r)$ the number of galaxy-particle pairs in radius bin $[r, r + \Delta r]$ when the $i$-th jackknife subbox is excluded. Note that we only discard the galaxies (and not the dark matter particles) in the $i$-th jackknife subbox when counting the pairs. We use $N_{\bar{g}} = 5$ in this work. Note that the jackknife error estimates do not include shape noises of source galaxies and the noises due to finite number of sources. Also, by directly integrating the galaxy-matter cross correlation function $\xi_{\text{GM}}$, we essentially perform the integration along all possible lines of sight and so the statistical error is smaller than that obtained using a sample of the same volume and number density of lens objects in real observations, where one has only a single line of sight for each object. For a more detailed analysis of errors using mock lensing data, see Shirasaki et al. (2016).

Figure 5 presents the galaxy matter cross correlation functions measured from the $L_{\text{box}} = 900$ h$^{-1}$Mpc (left) and 450 h$^{-1}$Mpc (right) simulations, and for each case we show, in the top panel, the GR results for all 29 snapshots of HOD catalogues (the coloured solid curves, with the colours denoting the redshifts according to the colour bar). In the lower panels we show the relative differences between F5 and GR using the same redshift-colour scheme, and the grey shaded region marks the 1-$\sigma$ Jackknife error. Notice that to improve visibility we have plotted $r^2 \xi_{\text{GM}}(r)$ instead of $\xi_{\text{GM}}(r)$ itself in the top panels. The results from the two boxes agree with each other well.

The GR curves in Fig. 5 show a scatter of up to $\sim 10\%$, which is because the different HOD catalogues do not correspond to the same set of tracers of the density field. To see this more explicitly, let us remark that the galaxy number density in our HOD catalogues peaks at $z \sim 0.16$, and as we go to higher redshifts it first decreases until $z \sim 0.27$, when it begins to increase, and then from $z \sim 0.33$ it starts to decrease again. The same trend can be seen in the $\xi_{\text{GM}}$ curves of Figure 5, namely the curves become higher from $z \sim 0.16$ (red) to 0.27 (green), then lower between $z \sim 0.27$ and $z \sim 0.33$ (blue), and then higher again until $z \sim 0.43$ (black). The trend is the same from both boxes. Physically, according to our HOD model, a lower galaxy number density indicates that only the more massive haloes are populated with galaxies, and these haloes have stronger correlations with matter around and inside. The behaviour is hence not surprising: when we consider the cross correlation between the matter density field and its tracers, the choice of the latter is critical (we will revisit this point later).

What is more interesting is the relative difference between F5 and GR, which is shown in the lower panels of Figure 5. There we see that the two models agree with each other (within the Jackknife error) at $r > 10$ h$^{-1}$Mpc, that F5 $\xi_{\text{GM}}$ starts to get enhanced – with a modest enhancement factor of $3-5\%$ – relative to the GR result at $r < 10$ h$^{-1}$Mpc, and finally near the transition scale between 1- and 2-halo terms ($r \sim 1-2$ h$^{-1}$Mpc) the enhancement factor increases to $\sim 10\%$. The stronger enhancement in the regime where the 1-halo term dominates is related to the more concentrated matter distribution in F5 compared to GR. Note that the model difference, unlike the cross correlation function $\xi_{\text{GM}}$ itself, is less dependent on red-
shift and therefore less sensitive to the choice of tracers. This is also as expected, because the selection of tracers in our HOD modelling affects both models in a similar way.

3.2.2 The excess surface density profiles

The averaged excess surface density profile, $\Delta \Sigma(r_p)$, is calculated by directly integrating the galaxy-mass cross correlation functions, following the prescription of Eq. (25).

To calculate this, we first use cubic spline to interpolate the $\xi_{gm}(r = \sqrt{r^2 + z^2})$ measured from the simulations onto a grid in $(\log(r_p), z)$, where $r_p$ is the projected separation and $z$ is the l.o.s. coordinate (with $z = 0$ at the position of the galaxy)$^4$, and then perform the integrations using direction summation. We have chosen a fine interpolation grid, as well as a large enough $z_{\text{max}}$ for the los integration, so that the percentage error for the numerical integrations is within $\sim 0.1\%$. For the $L_{\text{box}} = 900$ h$^{-1}$Mpc (450 h$^{-1}$Mpc) simulations, this allows us to have $\Delta \Sigma(r_p)$ up to $r_p = 30$ h$^{-1}$Mpc (15 h$^{-1}$Mpc), with $z_{\text{max}} = 90$ h$^{-1}$Mpc ($\sim 65$ h$^{-1}$Mpc).

The left (right) panel of Figure 6 shows the results of $\Delta \Sigma(r_p)$ from the $L_{\text{box}} = 900$ (450) h$^{-1}$Mpc simulation. A comparison of the upper panels indicates that the $900$ h$^{-1}$Mpc box starts to lose resolution at $r_p \sim 1$ h$^{-1}$Mpc, but the relative difference between F5 and GR agrees down to $r_p \sim 0.5$ h$^{-1}$Mpc. In both cases, we have split the lens galaxies into two separate redshift ranges (a high-$z$ bin with $0.30 \lesssim z \lesssim 0.43$ and a low-$z$ bin with $0.16 \lesssim z \lesssim 0.29$), but the results – dotted and dashed lines respectively – display barely any difference from using the whole lens galaxy sample, suggesting that the conclusion does not depend sensitively on redshift or the source number density, at least in the redshift range covered here.

The stacked excess surface density profiles in F5 and GR differ most significantly at $r_p \lesssim 2$ h$^{-1}$Mpc, which reflects the bigger difference in the 1-halo term of the matter-galaxy correlation function as shown in Fig. 5. The model difference between F5 and GR has a distinct dependence on $r_p$ from that of $\xi_{gm}(r)$, which is $\sim 5$-10% at $1 \lesssim r_p \lesssim 10$ h$^{-1}$Mpc, because $\Delta \Sigma(r_p)$ depends on the average $\xi_{gm}$ from $R = 0$ to $R = r_p$, where $R$ is the projected distance from the lens. Although this means that one can see a strong signal up to larger values of $r_p$, it also implies that uncertainties that affect the prediction of the 1-halo term in $\xi_{gm}$, such as baryonic physics, can have an impact on $\Delta \Sigma$ to larger $r_p$, something which we should bear in mind when considering observational constraints.

The 10-15% relative difference in $\Delta \Sigma$ at $r_p \lesssim 1$ h$^{-1}$Mpc between the two models is roughly the same as the difference in $C_{kk}$ at $\ell \lesssim 2000$, but smaller than the difference in $C_{kk}$ at larger $\ell$ (see Fig. 2), and therefore we are interested in whether there is any difference in the sign-to-noise (S/N) of these two probes. The S/N quantifies the distinguishability of the two models and is defined as

$$S/N = \left[ \frac{d_{GR}(x_i) - d_{GGL}(x_i)}{\text{C}} \right]^T \text{C}^{-1} \left[ \frac{d_{GR}(x_i) - d_{GGL}(x_i)}{\text{C}} \right],$$

in which $x_i$ is the data in the $i$th bin ($x = R_p$ for GGL and $x = \ell$ for lensing convergence), and $\text{C}$ is the covariance matrix.

To estimate the S/N values, we consider three fiducial surveys with roughly the DES, HSC and LSST specifications, as summarised in Table 3. We expect that Euclid will have a similar performance as LSST (see, e.g., Cautun et al. 2017, for some relevant S/N analysis but for void lensing using Euclid and LSST). For simplicity, let us assume that there are overlapping spectroscopic surveys which can provide galaxy catalogues with a number density at least as high as the ones used in our lensing galaxies, between $0.16 \leq z \leq 0.43$. As shown above, the model difference in $\Delta \Sigma(r_p)$ is insensitive the lens redshift, so we assume a single lens redshift $z_{\text{sour}} = 0.3$ and a single source redshift $z_{\text{source}} = 1$ in the simplified study. The covariance matrix $\text{C}$ is calculated based on a halo model prescription following Jeong et al. (2009), which accounts for contributions from cosmic variance, the Poisson noise of lensing galaxies and the shape noise of source galaxies, and it adopts single source and lens redshifts as was mentioned above. The GR HOD and cosmological parameters for the snapshot at $z \sim 0.3$ are used to generate $\text{C}$. We also adopt the halo-model approach developed in Sato et al. (2009) to estimate the covariance matrix for $C_{kk}$. In this matrix, we properly include the non-Gaussian term induced by convergence tri-spectrum and halo sample variance.

Figure 7 shows the S/N based on the our analysis which takes into account statistical uncertainties only, for the GGL (left panel) and $C_{kk}$ (right panel). The black, blue and red curve are respectively for a HSC, DES and LSST-like survey. In the case of GGL we show $S/N$ as a function of the minimum $r_p$ used in the calculation, while for the case of $C_{kk}$ the S/N is displayed as a function of the maximum $\ell$ included in the analysis. These are because using a smaller (larger) cutoff in $r_p$ ($\ell$) means more data are included in the model test, which can increase the S/N; but on the other hand, by doing these we are moving to smaller scales, where we would need to worry about other theoretical uncertainties such as the impact of baryons. Eventually, a comprise between these two considerations will have to be made; but for this study we are mainly interested in how the S/N values vary with the $r_p$ and $\ell$ cuts.

From Fig. 7 we can see that LSST, being a stage-4 experiment, has significantly more power to distinguish the two models, thanks to its larger sky coverage and higher source galaxy number density; this is true for both GGL and $C_{kk}$, HSC and DES, both of which are stage-3 experiments, show similar performances. When comparing the two probes, we find that $C_{kk}$ gives larger S/N values for all 3 surveys – this is not surprising as both can be symbolically written as $\langle AB \rangle$ where $A, B$ are two statistics of the underlying density field and $\langle \cdot \rangle$ denotes ensemble average. For $C_{kk}, A = B = k \times$ for GGL we have $A = \kappa$ and $B = \delta_n$, where $\delta_n$ is the galaxy number density fluctuation, the model difference in which has been greatly reduced by the tuning of HOD parameters. Therefore, we expect the model difference in GGL to be smaller than that in $C_{kk}$.

3.2.3 Effect of tracers

Because the galaxy-matter cross correlation quantifies the distribution of matter around galaxies, the choice of galaxies is important,
which is why we made effort to tune the HOD parameters in the F5 case to match the 3D galaxy two point correlation function $\xi_{gg}$ (and the galaxy number density) in the two models. To gain a feeling of the effect of choosing different tracers of the matter field, we have, for the snapshot at $z = 0.16$ of the $L_{\text{box}} = 900 h^{-1}\text{Mpc}$ simulation, tested how $\xi_{gm}$ changes by using other tracers. The result is shown in Figure 8 (see the figure caption for a more detailed description of the different curves). There are several interesting features in this figure.

First of all, we see that at small separations $\xi_{gm}$ is indeed very sensitive to the choice of tracers, as expected. The difference caused by this is substantially larger than that between F5 and GR as we have show above. Evidently, for the one-halo term of $\xi_{gm}$, changing the tracer is equivalent to placing the tracer to a different location in its host halo, or even outside the halo, which can impact the matter distribution around. This sensitivity is reduced at large separations, at which the effect of slightly relocating the tracer is smaller.

Second, for the two cases of cross correlating dark matter distribution from the F5 simulation with haloes (green dot-dashed and red dot-dot-dot-dashed), the lowest points in the curve shift toward smaller separations, and $\xi_{gm}$ is overall lower than the cross correlations by using HOD galaxies. Using haloes as tracers is similar to using only the central galaxies, and misses the contribution from satellite-matter cross correlation, which may have caused this feature. To check this, we show the result of cross correlating the dark matter field with only the central HOD galaxies, in the case of F5, as the orange long dashed curve in Figure 8. The behaviour is very similar to the cross correlation with haloes, and the slight difference is because the two tracer classes do not correspond to exactly the same halo population since central galaxies are populated in haloes in a random way.

Third, for the two cases in which we cross correlate the dark matter field from the F5 simulation with tracers from the GR simulation (cyan dashed and red dot-dot-dot-dashed), $\xi_{gm}$ decreases toward very small separations ($r \lesssim 0.3 h^{-1}\text{Mpc}$), and this is because GR tracers usually do not coincide with the highest density peaks in the dark matter field of the F5 simulation so that the highest values of $\xi_{gm}$ appear to be away from the GR tracers themselves.

All in all, the test highlights the importance of using carefully constructed mock galaxy catalogues when we compare the predictions of galaxy-galaxy lensing in different models. As an example, the 3D real-space two point correlation functions and the projected 2D correlation functions in our GR and F5 HOD catalogues agree with each other within 2-3% in the range of separations considered here, but such agreements are in a statistical, rather than an object-by-object, sense. When cross correlated with the same (F5) matter field, the $\xi_{gm}$ results by using F5 and GR galaxies are indeed different (cf. blue dotted and cyan dashed curves in Fig. 8), in particular at small separations. This indicates that high-resolution simulations enable to accurately resolve the internal structures of haloes and to allow realistic galaxy mocks to be constructed, are necessary in or-

---

5 We only use one snapshot because, as we have seen above, the model difference between F5 and GR is shows little redshift evolution in the redshift range considered in this paper.
under to make galaxy-galaxy lensing more useful for distinguishing $f(R)$ gravity and GR.

3.3 Galaxy bias

Finally, we compare the galaxy biases in GR and $f(R)$ gravity. We are interested in this comparison for two reasons.

Firstly, in standard ΛCDM, the linear bias factor, defined as

$$b_{\text{lin}}(r) = \frac{\xi_{\text{gm}}(r)}{\xi_{\text{mm}}(r)} = \frac{\xi_{gg}(r)}{\xi_{mm}(r)} = \sqrt{\frac{\xi_{gg}(r)}{\xi_{mm}(r)}}.$$  \hspace{1cm} (33)

is scale-independent on large scales. On the other hand, it is known that $f(R)$ gravity, and in general chameleon models, predict a scale-dependent linear growth rate of matter density perturbations, which may result in scale-dependent $b_{\text{lin}}$, because we have tuned the HOD parameters to match $\xi_{gg}(r)$ in the two models, while their $\xi_{mm}(r)$ can have a scale-dependent difference. We want to check whether, at least for F5, such a scale dependence of $b_{\text{lin}}$ is significant enough to make it of interest in observations.

Secondly, assuming that the linear bias factor $b_{\text{lin}}$ gives a correct quantification of galaxy bias at some given scale $r$, the galaxy-matter cross correlation coefficient $R_{\text{gm}}$, defined as

$$R_{\text{gm}}(r) = \xi_{\text{gm}}(r)/\sqrt{\xi_{gg}(r)\xi_{mm}(r)},$$  \hspace{1cm} (34)

is equal to 1. This offers an opportunity to derive the matter correlation function $\xi_{mm}$ from observational determinations of $\xi_{gg}$ and $\xi_{gm}$. We want to see whether the minimum length scale $R_{\min}$ above which this can be performed without worrying about the nonlinear effects in galaxy bias is significantly different in GR and in F5.

Figure 9 shows the results of our tests to answer these questions. The top bunch of curves in each panel (left for GR and right for F5) show the linear bias factors calculated in different ways (the different line styles) at a selection of redshifts (shown by different colours; see the legend for more details). We see that the different ways of calculating $b_{\text{lin}}$ agree well with each other, and that $b_{\text{lin}}$ remains roughly constant down to $r \sim 4\ h^{-1}\Mpc$, in agreement with previous results (e.g., Yoo et al. 2006). This holds true for all redshifts and both models – the latter is perhaps not surprising given that $\xi_{gg}$ has been tuned to match in GR and F5, while $\xi_{gm}$ in these models differ by at most a few percent above $r \sim 4\ h^{-1}\Mpc$, so that any scale dependence of $b_{\text{lin}}$ should be very weak. Note that $b_{\text{lin}}$ is larger at higher redshifts since the number densities of galaxies are lower for those redshifts and so we are looking at more biased tracers of the underlying matter field.

The results of $R_{\text{gm}}$ are shown as dashed lines near the bottom of the left (right) panel for GR (F5), for the same redshifts and using the same colour scheme. We can see that for $r$ down to $1-2\ h^{-1}\Mpc$ $R_{\text{gm}}$ is equal to unity to a good accuracy, and there is no significant difference between the different redshifts or between GR and F5. This suggests that the reconstruction of $\xi_{mm}$ from $\xi_{gg}$ and $\xi_{gm}$ can be done with the same reliability whether the underlying model is GR or F5 – this is even with the consideration that $b_{\text{lin}}$ may vary by a few percent (in theory) from vary large scales to $r \sim 4\ h^{-1}\Mpc$.

4 SUMMARY AND DISCUSSION

We have studied the possibility of using galaxy-galaxy weak lensing to distinguish between the standard ΛCDM paradigm and a lead
Before looking at GGL, we have first inspected the weak lensing convergence power spectrum in the selected $f(R)$ model. This has been studied before, but here we focus on the connection to the matter power spectrum and its time evolution. Figure 3 is a useful plot to understand how much contribution the convergence power spectrum has received from the matter clustering at different scales and epochs. We have checked that the behaviour of the former can be qualitatively explained using the latter.

Compared to the convergence and matter power spectra, which quantify the auto-correlation of the matter density field, we find that the cross correlation between galaxies and matter, $\xi_{gm}(r)$, shows a $\sim 50\%$ smaller difference between the two model studied, because the galaxy clustering in these models has been tuned to agree well with each other. This is further verified by an analysis of the signal-to-noise (S/N), which quantifies the distinguishability of the models when statistical-only errors are used, for three imaging surveys similar respectively to the DES, HSC and LSST specifications, with synergy data from a spectroscopic survey that can at least match the BOSS LOWZ galaxy number density at $0.16 \leq z \leq 0.43$. We find that the S/N is smaller than $2 \sim 3$ for DES and HSC, if we exclude GGL data within $2 \sim 3h^{-1}$Mpc, while for the LSST-like survey we have $S/N \sim 10$. The S/N is $2 \sim 4$ times higher if the cosmic shear power spectrum, with a conservative cut at $\ell_{\text{min}} = 1000 \sim 2000$, is used for constraints. In the redshift range covered in this study, we do not find any strong dependence of the model difference in GGL on the lens redshift – indeed, Fig. 6 shows that the low-$z$ and high-$z$ samples give almost identical results. Note that in the above forecast we have included only statistical and not systematic uncertainties. Although the cosmic shear power spectrum can have a higher S/N than GGL for distinguishing $f(R)$ gravity, it still remains useful and complementary to consider GGL. Current two-point correlation analysis of cosmic shear is subject to observational systematics induced by imperfect measurement of distant galaxies. The multiplicative bias is one of the most important systematics in cosmic shear analysis and can make the amplitude of lensing power spectrum uncertain (e.g. Huterer et al. 2006). In recent cosmological analyses, the impact has been controlled by image simulations (e.g., Hildebrandt et al. 2017) or appropriate parametrisation (e.g. DES Collaboration 2017). On the other hand, GGL analysis can separate the multiplicative bias and lensing effects from observed signal if one can use single source population and multiple foreground objects at different redshifts (Oguri & Takada 2011). We also note that the intrinsic alignments of galaxy ellipticities, which are a significant source of systematic errors for cosmic shear measurements (e.g. Hirata & Seljak 2004), are not important for GGL with single source population. This is because the stacked lensing is linear in shear, and does not include any correlation between the shapes of different galaxies.

We have investigated galaxy bias in the $\Lambda$CDM and $f(R)$ models. It is well known that $f(R)$ gravity models have scale-dependent linear growth rate, that goes back to the $\Lambda$CDM prediction on large scales. Figure 4 is an example of this scale dependence, although it shows the nonlinear growth of matter density perturbations. Hence, we expect the linear galaxy bias to be different in these two models. However, from Fig. 9 we do not find a clear difference: the different values at the different redshifts are more likely due to the different corresponding galaxy number densities, as we know that the more luminous galaxies (for which $n_g$ is smaller) are more strongly biased tracers of the dark matter field. The F5 galaxy bias, at a given redshift, only shows a rather weak scale dependence and is slightly smaller than the GR value, the latter being because $\xi_{gg}$ is tuned to agree in the two models while $\xi_{gm}$ is larger in $f(R)$ gravity. On the
other hand, we find that, as in GR, the galaxy-mass cross correlation coefficient $\xi_{gm}$ remains at 1.0 down to $\sim 3h^{-1}\text{Mpc}$, suggesting that one can infer $\xi_{mm}$ from measurements of galaxy auto correlation and GGL at scales larger than $\sim 3h^{-1}\text{Mpc}$.

As a first study of GGL in $f(R)$ gravity, this work can be further improved in several ways, all of which require more (and more advanced) simulations than used here.

First, as shown in Fig. 5, the galaxy-mass cross correlation $\xi_{gm}$ (and therefore GGL) depends sensitively on the number density of lenses. Our galaxy catalogues are made to mimic the BOSS LOWZ sample, with a number density of order $10^{-4} \langle h^{-1}\text{Mpc} \rangle^{-3}$, and this number will be further increased with ongoing and future galaxy surveys. It remains interesting to see how effectively the increased lens number density can help improve the S/N, given that many of the newly added galaxies will be fainter. In order to check this, we will need simulations with higher resolutions to populate fainter galaxies, and such simulations may also enable mock galaxy catalogues constructed using other techniques such as SHAM and SAM, to understand the impact of different methods of galaxy-halo connection.

Second, the simulations used here are for a choice of particular cosmological parameters (WMAP9). It is important to understand whether other choice of cosmological parameters can have a significant impact on the model difference in GGL. It is also important to check if there is a degeneracy between modified gravity and cosmology, e.g., whether the effect of $f(R)$ gravity can be mimicked by a different value of $\Omega_m$ or $\sigma_8$, as such a degeneracy can downgrade the potential of using GGL to constrain chameleon $f(R)$ models. For example, it was shown (Yoo et al. 2006) that the GGL signal $\Delta\Sigma(r_p) \propto \Omega_m \sigma_8^2 \delta^{\alpha/b}$, with $\alpha$ and $b$ depending weakly on $r_p$, and therefore it is possible that the $r_p$-dependence of $\Delta(\Delta\Sigma)/\Delta\Sigma_{GR}$ is not completely degenerate with the effects of varying $\Omega_m$ and $\sigma_8$, allowing these different parameters to be constrained simultaneously using GGL alone, or possible synergies of GGL with other probes. This will require simulations with different cosmologies.

Third, in the forecast above we have deliberately not used signals on smaller scales ($R_c \lesssim 2 \sim 3h^{-1}\text{Mpc}$ and $\ell > 1000 \sim 2000$), not just for resolution consideration, but also to be conservative as on those scales the theoretical predictions might be significantly affected by poorly-understood baryonic physics (which is not included in the simulations used in this work). The modified gravitational force can change the halo density profiles, which is reflected in the GGL signals. However, to reliably include this in the model test, we have to fully understand how baryonic processes affect the (re)distribution of matter inside dark matter haloes for $f(R)$ gravity, and for this hydrodynamical simulations with suitably adjusted subgrid physics are essential. This is an almost entirely unexplored regime to date.

**ACKNOWLEDGEMENTS**

We would like to thank Christian Arnold, Alex Barreira, Carlton Baugh, Yan-Chuan Cai, Marius Cautun, Jianhua He, Peder Norborg for useful discussions. The authors thank the kind host by Naoki Yoshida at IPMU, the University of Tokyo, when part of this work was carried out. BL is supported by an European Research Council Starting Grant (ERC-StG-716532-PUNCA) and UK Science and Technology Facilities Council (STFC) Consolidated Grants (ST/P000541/1 and ST/L00075X/1). This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC Data Centre at Durham, which is funded by STFC and the large clusters of the DiRAC service at the Science and Technology Facilities Council’s HPC facility at RAL. The authors thank Naoki Yoshida at IPMU, the University of Tokyo, for useful discussions.

---

**Figure 9.** (Colour Online) Left panel: The bunch of curves on the top of the panel show the linear galaxy bias defined in three ways -- $b_{lin} \equiv \xi_{gm}/\xi_{mm}$ (solid), $b_{lin} \equiv \sqrt{\xi_{gg}/\xi_{mm}}$ (dashed) and $b_{lin} \equiv \xi_{gg}/\xi_{mm}$ (dotted) -- from the GR $L_{box} = 900 h^{-1}\text{Mpc}$ simulation. The redshifts of the curves are given in the legend, using the same colour scheme and order (from top to bottom) as the curves themselves. The dashed lines at around 1.0 are the galaxy matter correlation coefficient defined as $R_{gm} \equiv \xi_{gm}/\sqrt{\xi_{gg}\xi_{mm}}$ for the same redshifts. Right panel: The same as the left panel, but for F5.
