Probing primordial non-Gaussianity with SKA galaxy redshift surveys: a fully relativistic analysis

Stefano Camera,[1,2]★ Mário G. Santos[1,2,3] and Roy Maartens[2,4]

[1]*CENTRA, Instituto Superior Técnico, Universidade de Lisboa, P-1049-001 Lisboa, Portugal*
[2]*Physics Department, University of the Western Cape, Cape Town 7535, South Africa*
[3]*SKA SA, The Park, Park Road, Pinelands 7405, South Africa*
[4]*Instiue of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*

ABSTRACT

The Square Kilometre Array (SKA) will produce spectroscopic surveys of tens to hundreds of millions of neutral hydrogen (H I) galaxies, eventually covering 30 000 deg^2 and reaching out to redshift $z \gtrsim 2$. The huge volumes probed by the SKA will allow for some of the best constraints on primordial non-Gaussianity, based on measurements of the large-scale power spectrum. We investigate various observational set-ups for H I galaxy redshift surveys, compatible with the SKA Phase 1 and Phase 2 (full SKA) configurations. We use the corresponding number counts and bias for each survey from realistic simulations and derive the magnification bias and the evolution of source counts directly from these. For the first time, we produce forecasts that fully include the general relativistic effects on the galaxy number counts. These corrections to the standard analysis become important on very large scales, where the signal of primordial non-Gaussianity grows strongest. Our results show that, for the full survey, the non-Gaussianity parameter f_{NL} can be constrained down to $\sigma(f_{NL}) = 1.54$. This improves the current limit set by the *Planck* satellite by a factor of 5, using a completely different approach.

Key words: relativistic processes – cosmological parameters – cosmology: observations – early Universe – large-scale structure of Universe – radio lines: galaxies.

1 INTRODUCTION

The large-scale distribution of matter contains information not only about the current state of the Universe – such as the dark energy – but also about the primordial state of the Universe, at times well before the onset of structure formation. This remarkable fossil record is imprinted in the correlations of galaxies by the primordial fluctuations which evolve to become the seeds of structure, and it contains vital clues as to the inflationary mechanism (or an alternative mechanism) for generating these fluctuations.

One of the most important questions is whether or not the primordial fluctuations are Gaussian. Primordial non-Gaussianity (PNG) imprints a characteristic feature, via the bias b, in the galaxy power spectrum $P_g = b^2 P_m$. (It also imprints features in the bispectrum and higher moments, which we do not consider here.) This feature causes a growth of power $\propto k^{-2}$ on large scales. The excess power is 'frozen' on super-Hubble scales during the evolution of the galaxy overdensity – it is not affected by non-linearity on small scales. By contrast, PNG is continually generated on small scales by structure

formation so that the record of PNG can be washed out on these scales.

The best current constraints on PNG are from cosmic microwave background (CMB) measurements by the *Planck* satellite (Ade et al. 2014). In particular, for the local form of PNG (which has the strongest impact on galaxy bias), we have

$$f_{NL}^{CMB} = 2.7 \pm 5.8 \ (1\sigma). \tag{1}$$

This result already rules out inflationary models with large PNG. In order to discriminate amongst the remaining models, we need to bring down the errors. Ultimately, we need to probe the simplest single-field models that generate negligible PNG. [Note that there is a subtle point here, which is discussed in the next section; see in particular equation (11).]

Galaxy surveys currently lag well behind *Planck*, but they can provide a valuable consistency check by probing PNG at low redshifts (see e.g. Ross et al. 2013; Giannantonio et al. 2014, for the current level of constraints on f_{NL}). Furthermore, future surveys covering a large fraction of the sky, such as those planned by the Square Kilometre Array (SKA),[1] will be able to probe PNG over

---

★ E-mail: stefano.camera@manchester.ac.uk

[1] http://www.skatelescope.org/

a significant range of redshifts, giving a much higher number of modes than the CMB, which is confined effectively to a single redshift. If systematics and complications in the modelling of bias and haloes can be controlled, then the future of PNG constraints lies with huge-volume galaxy surveys.

Here, we investigate the constraining power of future neutral hydrogen (H I) galaxy redshift surveys with the SKA, which are envisaged to cover 3/4 of the sky and reach a depth of $z \sim 2$ in the full SKA. We use the Fisher forecast method, where we include the relevant cosmological parameter σ_8. We carefully include the effects of magnification bias and evolution of source counts, based on simulations of H I galaxy number densities. In addition, we include all relativistic effects that enter into the theoretical modelling of the *observed* power spectra. This removes an important theoretical systematic—present if one neglects relativistic effects—which can bias a reconstruction of f_{NL} (Camera, Santos & Maartens 2014).

We assume a 'concordance' [flat Λ cold dark matter (ΛCDM)] model for the background.

2 PNG WITH RELATIVISTIC EFFECTS

If the distribution of primordial curvature perturbations is not Gaussian, it cannot be fully described by a power spectrum $P_\Phi(k)$, where Φ is Bardeen's gauge-invariant potential. We need higher order moments, starting with the bispectrum $B_\Phi(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$, where $\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = 0$. Different models of inflation give rise to different shapes of the bispectrum. For the effects of PNG on galaxy bias, the 'local' (or 'squeezed') configuration is dominant on large scales (Taruya, Koyama & Matsubara 2008; Fedeli, Moscardini & Matarrese 2009). In this case, the bispectrum is maximized for triangles with one of the three momenta much smaller than the other two. Simple single-field inflation models generate local PNG. There are inflationary models that generate different shapes for the primordial bispectrum, e.g. 'equilateral' (Crociani et al. 2009; Fedeli et al. 2009) and 'folded' (Holman & Tolley 2008; Meerburg, van der Schaar & Corasaniti 2009; Verde & Matarrese 2009). However, the local shape gives the largest effects – particularly on bias – and we confine our analysis to local PNG, i.e. $f_{NL} = f_{NL}^{local}$ from now on.

A convenient way to parametrize the deviation from Gaussianity is (Salopek & Bond 1990; Gangui et al. 1994; Verde et al. 2000; Komatsu & Spergel 2001)

$$\Phi = \varphi + f_{NL} \left(\varphi^2 - \langle \varphi^2 \rangle \right), \tag{2}$$

where φ is a linear Gaussian potential. PNG modulates the formation of haloes in the cold dark matter, inducing a scale (and redshift) dependence in the halo bias (Dalal et al. 2008; Matarrese & Verde 2008). The modification to the Gaussian bias on large scales is given by

$$b(z) \rightarrow b(z) + \Delta b(z, k), \tag{3}$$

$$\Delta b(z, k) = [b(z) - 1] \frac{3\Omega_m H_0^2 q \delta_{cr}}{k^2 T(k) D(z)} f_{NL}, \tag{4}$$

where the factor $q = \mathcal{O}(1)$ is included to obtain agreement with *N*-body simulations (Carbone, Verde & Matarrese 2008; Grossi et al. 2009; Wagner & Verde 2012). Given the uncertainty in modelling q and in extracting it reliably from *N*-body simulations, we follow Giannantonio et al. (2012), and most works that constrain f_{NL} via the bias, in taking $q = 1$. This means that we in fact constrain an effective $f_{NL}^{eff} = q \times f_{NL}$ until such time as the theoretical uncertainty in q is resolved. In equation (4), $\Omega_m = \Omega_b + \Omega_c$ is

the total matter fraction at $z = 0$, H_0 is the Hubble constant, $\delta_{cr} \simeq 1.69$ is the critical matter density contrast for spherical collapse, $T(k)$ is the matter transfer function ($T \rightarrow 1$ as $k \rightarrow 0$) and $D(z)$ is the linear growth function of density perturbations (normalized to unity today). The most important feature in equation (4) is the k^{-2} term, which comes from relating Φ to the overdensity δ via the Poisson equation. This term means that the PNG signal in the bias grows stronger as $k \rightarrow 0$. Note that Δb is strongly suppressed on scales smaller than the equality scale, so that even if we include such scales in computations, there is no false gain in constraining power over f_{NL} (see also Giannantonio et al. 2012).

In equation (4) we have used the large-scale structure (LSS) convention for defining f_{NL}, as opposed to the CMB convention (Afshordi & Tolley 2008; Carbone et al. 2008; Grossi et al. 2009; Pillepich, Porciani & Hahn 2010). The difference arises since Φ is evaluated at decoupling for f_{NL}^{CMB} and at low z for f_{NL}^{LSS}. The relation is

$$f_{NL}^{LSS} = \frac{g(z_{dec})}{g(0)} f_{NL}^{CMB} \simeq \frac{1}{5} \left(3 + 2\Omega_m^{-0.45} \right) f_{NL}^{CMB} \simeq 1.3 f_{NL}^{CMB}, \tag{5}$$

where $g(z) = (1 + z)D(z)$. This means that when we constrain PNG using equation (4), the *Planck* constraint of equation (1), given in the CMB convention, translates in the LSS convention to

$$f_{NL} = 3.5 \pm 7.5 \Rightarrow \sigma(f_{NL})_{Planck} = 7.5. \tag{6}$$

Here and in the rest of this paper, f_{NL} is in the LSS convention, i.e. $f_{NL} = f_{NL}^{LSS}$. Thus, the target for galaxy surveys to beat is $\sigma(f_{NL}) = 7.5$.

The standard approach to constraining PNG via the galaxy bias proceeds from equation (4), and uses power spectrum measurements that include redshift space distortions (RSDs) via the Kaiser term,

$$\delta_g^z = \delta_g - \frac{(1 + z)}{H} (n^i \partial_i)^2 V \quad \text{where } v_i = \partial_i V. \tag{7}$$

Here, n^i is the direction of the galaxy and v^i is its peculiar velocity. In addition, the fiducial value for forecasting constraints on PNG is taken as

$$f_{NL}^{fid} = 0. \tag{8}$$

This approach is reasonable for large f_{NL}, which overwhelms any relativistic corrections. But *Planck* has shown that f_{NL} is not large – and so we need to be more careful and include all relativistic effects.

These effects are of two kinds:

(1) a non-linear primordial correction;
(2) linear projection effects from observing in redshift space on the past light cone.

2.1 Primordial relativistic correction

In the Newtonian approximation, an exactly Gaussian distribution of the primordial curvature perturbation translates into an exactly Gaussian distribution of density perturbations via the Newtonian Poisson equation, $\nabla^2 \Phi = 4\pi G a^2 \rho \delta$. In general relativity however, the Newtonian Poisson equation is not correct at second order. There is a relativistic non-linear correction to the Poisson constraint that links curvature to density perturbations. As a result, an exactly Gaussian distribution of primordial curvature perturbations does not lead to a Gaussian distribution of density perturbation, even on super-Hubble scales.

This non-linearity was first derived in Bartolo, Matarrese & Riotto (2005), and the implications for PNG were then developed in Verde & Matarrese (2009, see also Hidalgo, Christopherson &

Malik 2013; Bruni et al. 2014; Villa, Verde & Matarrese 2014). The effective local PNG parameter that describes this primordial general relativistic correction on large scales is

$$f_{\rm NL}^{\rm GR} = -\frac{5}{3} \quad \text{(CMB convention)}. \tag{9}$$

This correction is derived in the CMB convention because it is based on the primordial Φ. It does not affect CMB measurements of PNG, because these are independent of the Poisson constraint. But for LSS, it must be added to the local PNG parameter. Translating equation (9) to the LSS convention via equation (5), we have

$$f_{\rm NL} \rightarrow f_{\rm NL} + f_{\rm NL}^{\rm GR} \simeq f_{\rm NL} - 2.17. \tag{10}$$

Equation (10) has interesting implications. First, it means that the standard fiducial value of equation (8) should be corrected to

$$f_{\rm NL}^{\rm fid} \simeq -2.17. \tag{11}$$

If the primordial curvature perturbation is Gaussian, which is effectively the case for slow-roll single-field inflation, then this is the local PNG that we expect to measure in galaxy surveys. For the concordance model with slow-roll single-field inflation, galaxy surveys should measure $f_{\rm NL} \simeq -2.17$. This subtle point (see Verde & Matarrese 2009) is missed in most previous work on PNG from single-field inflation (e.g. Pajer, Schmidt & Zaldarriaga 2013).

Finally, equation (11) implies that the necessary accuracy for testing single-field inflation through PNG in galaxy surveys is $\sigma(f_{\rm NL}) \lesssim 2$.

2.2 Relativistic projection effects

The Kaiser RSD term in equation (7) is the dominant term on sub-Hubble scales of a more complicated set of relativistic terms that arise from radial and transverse perturbations along the light ray of observation. (Note that these terms are at linear order, unlike the non-linear correction in the previous subsection.)

The first such term is the contribution of the lensing convergence κ to the observed number density contrast, where

$$\kappa = \int_0^\chi d\tilde{\chi}\,(\chi - \tilde{\chi})\frac{\tilde{\chi}}{\chi}\nabla_\perp^2 \Phi. \tag{12}$$

Here χ is the radial comoving distance, related to the Hubble rate $H(z)$ by $d\chi = dz/H(z)$, and $\nabla_\perp^2$ is the Laplacian on the transverse screen space. Gravitational lensing affects the observed number density in two competing ways – enhancing it by bringing faint galaxies into the observed patch, and decreasing it by broadening the area of the patch. The competition is controlled by the magnification bias $\mathcal{Q}$, i.e. the slope of the (unlensed) galaxy luminosity function at the survey flux limit. It is given by

$$\mathcal{Q} = -\frac{\partial \ln N_{\rm g}}{\partial \ln \mathcal{F}}\bigg|_{\mathcal{F}=\mathcal{F}_*}, \tag{13}$$

where $N_{\rm g}(z, \mathcal{F} > \mathcal{F}_*)$ is the background galaxy number density at redshift z and with flux $\mathcal{F}$ above the flux detection threshold $\mathcal{F}_*$. Then the contribution of lensing to equation (7) is $2(\mathcal{Q} - 1)\kappa$.

The remaining relativistic contributions are both local and integrated terms (Yoo, Fitzpatrick & Zaldarriaga 2009; Yoo 2010; Bonvin & Durrer 2011; Challinor & Lewis 2011; Bertacca et al. 2012; Jeong, Schmidt & Hirata 2012), namely

$$\delta_{\rm g}^{\rm obs} = (b + \Delta b)\delta_{\rm cs} - \frac{(1+z)}{H}(n^i\partial_i)^2 V - 2(1-\mathcal{Q})\kappa + \delta_{\rm loc} + \delta_{\rm int}. \tag{14}$$

In order to define the bias consistently on super-Hubble scales (Challinor & Lewis 2011; Bruni et al. 2012; Jeong et al. 2012), we use the comoving-synchronous matter overdensity

$$\delta_{\rm cs} = \delta - 3aHV, \tag{15}$$

where $a = 1/(1 + z)$ is the scale factor. The local and integrated relativistic terms are

$$\delta_{\rm loc} = (3 - b_{\rm e})\frac{H}{(1+z)}V + An^i\partial_i V + (2\mathcal{Q} - 1 - A)\Phi + \frac{\dot{\Phi}}{H}, \tag{16}$$

$$\delta_{\rm int} = 4\frac{(1-\mathcal{Q})}{\chi}\int_0^\chi d\tilde{\chi}\,\Phi - 2A\int_0^\chi d\tilde{\chi}\frac{\dot{\Phi}}{(1+z)}, \tag{17}$$

where $b_{\rm e}$ is the evolution bias, giving the source count evolution

$$b_{\rm e} = \frac{\partial \ln a^3 N_{\rm g}}{\partial \ln a}, \tag{18}$$

and the factor A is

$$A = b_{\rm e} - 2\mathcal{Q} - 1 - \frac{\dot{H}}{H^2} + \frac{2(\mathcal{Q}-1)(1+z)}{\chi H}. \tag{19}$$

The local term $\delta_{\rm loc}$ has Doppler and Sachs–Wolfe-type contributions. The integrated term $\delta_{\rm int}$ contains time delay and integrated Sachs–Wolfe contributions. These relativistic terms can become significant near and beyond the Hubble scale. (See Yoo et al. 2009; Yoo 2010; Bertacca et al. 2012; Jeong et al. 2012; Raccanelli et al. 2014a for the local terms and Bonvin & Durrer 2011; Challinor & Lewis 2011; Raccanelli et al. 2013 for the integrated terms.) The growth of relativistic effects occurs on the same scales where the effect of PNG is growing through the galaxy bias of equation (4). The relativistic effects are easily confused with the PNG contribution (Bruni et al. 2012; Jeong et al. 2012; Yoo et al. 2012; Raccanelli et al. 2014a). In order to remove this theoretical systematic, it is necessary to include all the relativistic effects in an analysis of PNG in galaxy surveys (see Maartens et al. 2013, for surveys in the radio continuum).

The computation of the observed number density contrast with all relativistic contributions can be done with CAMB_sources (Challinor & Lewis 2011, which we use here) or CLASSgal (Di Dio et al. 2013). These codes compute the angular power spectra $C_\ell(z_i, z_j)$, which correspond to what is observed on the light cone.

3 SKA H I GALAXY REDSHIFT SURVEYS

The SKA project is set to be the largest radio telescope ever built, coordinating efforts from several countries and sharing facilities between South Africa and Australia. It is planned to be developed in two main phases. Phase 1 currently encompasses two mid-frequency facilities ($\sim$1 GHz) operating within South Africa (SKA1-MID) and Australia (SKA1-SUR). A low-frequency array (SKA1-LOW $\sim$100 MHz) will be set in Australia. We refer to Dewdney et al. (2009) for a description of the set-ups. In the second phase, the full SKA, the plan is to extend the array by about a factor of 10 in one of the sites, both in collecting area and primary beam (field of view), thus significantly increasing the survey power of the facility.

One of the key science goals for the SKA will be the detection of a large amount of H I galaxies up to high redshifts using the 21 cm line of neutral hydrogen to measure very accurate redshifts. Cosmological applications will require detecting large numbers of galaxies to beat shot noise and over a large sky area to reduce cosmic variance. With the sensitivities for SKA1 and taking 10,000 h of observing time, the optimal survey area will be around 5000 deg^2 using band
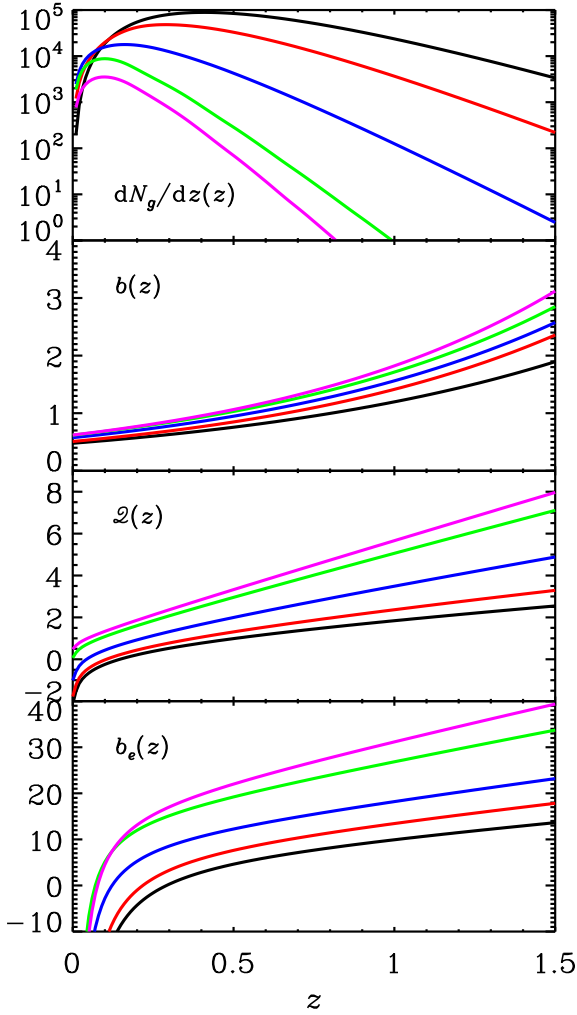
Figure 1. Key functions derived from simulations for SKA H I galaxy redshift surveys, with sensitivities of 3, 7.3, 23, 70 and 100 μJy (black, red, blue, green and magenta curves, respectively). Top: H I galaxy number distribution per square degree, $dN_g(z)/dz$ [deg^{-2}]. Second: H I galaxy bias, $b(z)$, for $f_{NL} = 0$. Third: magnification bias, $\mathcal{Q}(z)$. Bottom: evolution bias, $b_e(z)$.

2 from SKA1-MID ($>$950 MHz) or SKA1-SUR ($>$650 MHz). This will enable the detection of about 5×10^6 galaxies up to $z \sim 0.5$, with a flux cut five times above a flux sensitivity of $\sim$70 μJy. The full SKA, on the other hand, should be capable of detecting about 9×10^8 galaxies over a 30 000 deg^2 area, up to $z \sim 2.0$, with a flux cut 10 times above a flux sensitivity of $\sim$5 μJy, making it the largest galaxy redshift survey ever.

The noise calculations and parameters for these H I galaxy surveys can be found in Yahya et al. (2014). We used the SAX-sky simulation to calculate the H I galaxy number density and bias as a function of flux sensitivity.[2] The results are summarized in Fig. 1. Note that the flux numbers quoted in Fig. 1 refer to noise sensitivities – as explained above, the actual flux threshold is assumed to be five times this number for an SKA1 survey and ten times for the full SKA.

The numbers obtained below (presented in Section 5) show that a galaxy redshift survey with SKA1 will not improve on *Planck* for

PNG constraints. An H I intensity mapping survey has been proposed as a way to surpass this limitation with SKA1 and push the limits on cosmological constraints (see Bull et al. 2014, for details). However, this type of survey will rely on sophisticated foreground cleaning techniques. The advantage of an H I galaxy redshift survey is that it will be much less contaminated by foregrounds or calibration artefacts, making the detection of the signal much cleaner. This is particularly important on very large scales where we are probing PNG. Our results show that as sensitivity is improved towards the full SKA, the number of galaxies detected will allow unmatched constraints on most cosmological parameters, including PNG.

4 FORECASTING CONSTRAINTS ON PNG

The observable that we employ in this analysis is the tomographic angular power spectrum of H I galaxy number density fluctuations (Challinor & Lewis 2011),

$$C_\ell^{ij} = 4\pi \int d\ln k \, \mathcal{W}_\ell^i(k) \mathcal{W}_\ell^j(k) \Delta_\zeta^2(k). \quad (20)$$

Here the index i refers to the ith tomographic redshift bin, which the redshift distribution of H I galaxies, $dN_g(z)/dz$, has been divided into, ℓ is the Legendre multipole of the spherical Bessel–Fourier decomposition, k is the three-dimensional wavenumber and $\Delta_\zeta^2(k)$ is the dimensionless power spectrum of primordial curvature perturbations. Equation (20) describes the same observable as the angular power spectrum $C_\ell(z_i, z_j)$ mentioned in Section 2.2. However, we use from now on the notation of equation (20) to emphasize that we bin the galaxy redshift distribution into $i, j = 1, \ldots, N_z$ slices and then consider the angular power spectrum between the ith and jth bins, rather than picking up two infinitesimal redshift surfaces at z_i and z_j, which is unfeasible for galaxy surveys. (Note that this is not the case for intensity mapping, see e.g. Camera et al. 2013.)

The window functions in equation (20) are derived from binning $\mathcal{W}_\ell(z, k)$, and contain the redshift distribution of the galaxy population and the transfer function for the galaxy number overdensity. In the standard Kaiser analysis, we have

$$\mathcal{W}_\ell^{std}(k) = \int d\chi \, \frac{dN_g(\chi)}{d\chi} \left\{ [b(\chi) + \Delta b(\chi, k)] j_\ell(k\chi) \right.$$

$$\left. + \frac{kv(1+z)}{H(z)} j_\ell''(k\chi) \right\} T_\delta(\chi, k), \quad (21)$$

where j_ℓ is a spherical Bessel function and j_ℓ'' its second derivative, and v is the Newtonian-gauge velocity of the sources (no velocity bias assumed).

We want to perform for the first time a theoretically complete and self-consistent analysis of PNG forecasts, including all the relativistic effects described in Section 2.2, which are known to be important on horizon scales where the PNG signal is also growing. In the fully relativistic analysis, we need to modify equation (21) to include the additional terms in equation (14), beyond the Kaiser term. As mentioned before, to compute the full relativistic angular power spectrum of equation (20), we use the CAMB_sources package (Challinor & Lewis 2011), which we modified in order to include the non-Gaussian bias correction of equation (4).

To forecast the SKA potential for constraining PNG, we perform a Fisher matrix analysis (Fisher 1935; Tegmark, Taylor & Heavens 1997) focused on the PNG parameter f_{NL}. If we assume that the model likelihood surface in parameter space can be well approximated by a multivariate Gaussian (at least in a neighbourhood of

its peak), the Fisher matrix **F** is then a good approximation for the inverse of the parameter covariance. This means that for a set of parameters $\{\vartheta_\alpha\}$, we can write

$$\mathbf{F}_{\alpha\beta} = f_{\text{sky}} \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{2\ell+1}{2} \frac{\partial C_\ell^{ij}}{\partial \vartheta_\alpha} \Sigma_\ell^{jp} \frac{\partial C_\ell^{pq}}{\partial \vartheta_\beta} \Sigma_\ell^{qi}, \tag{22}$$

where we assume summation over repeated indices, and Σ_ℓ^{ij} is the inverse of $C_\ell^{ij} + \mathcal{N}_\ell^{ij}$, due to the fact that measurements of C_ℓ^{ij} are affected by Poisson noise,

$$\mathcal{N}_\ell^{ij} = \frac{\delta^{ij}}{\bar{N}_g^i}. \tag{23}$$

Here $\bar{N}_g^i$ is the total number of galaxies per steradian in the ith redshift bin, and f_{sky} is the fraction of the sky covered by the survey, namely $f_{\text{sky}} \equiv \text{Area}/(4\pi)$, corresponding to 0.73 for 30 000 deg^2 and 0.12 for 5000 deg^2.

If all the model parameters were uncorrelated with each other, the forecast error on parameter ϑ_α would simply be (no equal index summation)

$$\Delta(\vartheta_\alpha) = \frac{1}{\sqrt{\mathbf{F}_{\alpha\alpha}}}. \tag{24}$$

In Bayesian statistics this is known as 'conditional error'. On the other hand, to take parameter degeneracies properly into account, one should always consider 'marginal errors', i.e.

$$\sigma(\vartheta_\alpha) = \sqrt{(\mathbf{F}_{\alpha\alpha})^{-1}}. \tag{25}$$

In the following we shall use equation (25) and only occasionally refer to the conditional errors of equation (24).

For the choice of parameters, one should in principle consider the full ΛCDM parameter set. One possibility to account for our current knowledge on ΛCDM parameters is to add afterwards to the full Fisher matrix a matrix of priors coming from external experiments. Another way, widely adopted in the literature (e.g. Sartoris et al. 2010; Fedeli et al. 2011; Fedeli, Bartelmann & Moscardini 2012; Ade et al. 2014; Raccanelli et al. 2014b; Raccanelli, Dore & Dalal 2014c), is to vary freely only the parameters one is interested in and fix all the others to their fiducial values, as fitted for instance by *Planck*. On the one hand, this possibly leads to an overestimation of a survey constraining power, for by doing so one may as well disregard parameter degeneracies. However, it is known that the inclusion of PNG does not significantly broaden other parameter confidence intervals (Giannantonio et al. 2012).

5 FORECAST RESULTS

In order to forecast constraints for f_{NL}, we proceed as follows. We limit ourselves to a multipole range $2 \leq \ell \leq 700$ (as also done by Ferramacho et al. 2014), where non-linear corrections to galaxy clustering may be safely neglected. On these angular scales, the cosmological parameter that can be most degenerate with f_{NL} is the normalization of the matter power spectrum, σ_8. Hence, we adopt $\vartheta_\alpha = \{f_{\text{NL}}, \sigma_8\}$ and leave the other parameters fixed. (We discuss the impact of this assumption in Section 5.3.)

5.1 Constraints for $z \leq 1.5$

We limit ourselves for now to $z \leq 1.5$. This corresponds to the redshift range that most galaxy surveys will be able to probe. We divide the H I galaxy redshift distribution into 20 equally spaced

Table 1. Summary of forecast constraints on f_{NL} for $z \leq 1.5$. The third and fourth columns refer respectively to conditional errors and errors marginalized over σ_8.

Sensitivity (μJy)	f_{sky}	$\Delta(f_{\text{NL}})$	$\sigma(f_{\text{NL}})$
3	0.73	3.12	3.12
7.3	0.73	3.23	3.23
23	0.73	7.80	8.30
70	0.12	18.5	24.6
100	0.12	53.3	79.3

bins in the range $0 < z \leq 1.5$. Note that we fully account for auto- and cross-correlations amongst all redshift bins.

Table 1 summarizes the constraints that we forecast for various SKA experimental set-ups, which can be easily related to SKA1 or full SKA configurations. It is apparent from the third and fourth columns that our assumption regarding the parameter set appears to be well founded. Indeed, we see that for the highest sensitivities (lowest flux cuts), the conditional and marginal errors on f_{NL} coincide. From equation (6) it follows that SKA H I galaxy redshift surveys will be able to achieve more than twice the accuracy of *Planck* in a measurement of f_{NL}, using the redshift range $0 < z \leq 1.5$, i.e.

$$\sigma(f_{\text{NL}}) = 3.12 \quad (0 < z \leq 1.5). \tag{26}$$

The left-hand panel of Fig. 2 shows the fourth column of Table 1. In the right-hand panel, we illustrate the dependence of $\sigma(f_{\text{NL}})$ on the minimum angular multipole considered in the analysis, $\ell_{\min}$. Again, we present this for all the flux cuts discussed above, as well as for a more conservative value of the maximum angular multipole, namely $\ell_{\max} = 300$. The constraints weaken as expected, when we increase $\ell_{\min}$, since it is on the largest angular scales that PNG has the strongest impact. However, the choice of $\ell_{\max}$ does not influence much the final outcome. Indeed, a more conservative value like $\ell_{\max} = 300$ yields almost the same results.

5.2 Pushing to higher redshifts

It has been shown (see for instance Camera et al. 2013 or Ferramacho et al. 2014) that to probe PNG effectively, one needs to access as big a 'cube' of Universe as possible, in order to include the largest modes and lessen cosmic variance. [Moreover, if possible we can use a multitracer analysis (Seljak 2009) with two or more proxies of the underlying dark matter distribution.]

Large volumes, with redshift information, can in principle be delivered by an SKA1 H I intensity mapping survey (see e.g. Bull et al. 2014), but several technical issues still need to be addressed. SKA galaxy redshift surveys will achieve exquisite accuracy in the measurement of H I emitting galaxies, cataloguing almost one billion such sources with precise redshift estimates. As can be seen in Fig. 1, the redshift distribution of SKA H I galaxies has a non-negligible high-redshift tail, when high sensitivities such as 3 μJy are achieved. This is an advantage of H I surveys compared to other galaxy redshift surveys, given the abundance of H I at high z.

By exploiting this, we can explore the capabilities of SKA H I galaxy redshift surveys in accessing the largest scales at higher redshifts. The outcome of such an analysis is not straightforward to predict. On one hand, the higher the redshift the larger the bias (since we can probe smaller k's in the linear regime), and thus the stronger the PNG signal. On the other hand, the decrease in the number of sources as we go to high z reduces the constraining power because of the Poisson noise term in equation (23).
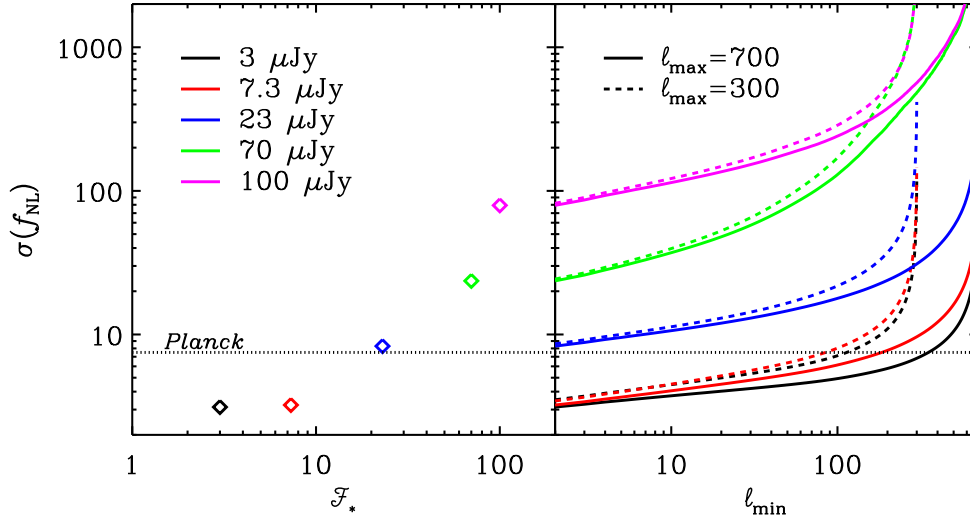
Figure 2. Forecast 68 per cent errors on f_{NL} marginalized over σ_8 for $0 < z \le 1.5$. Left: as a function of flux cut for $\ell_{min} = 2$ and $\ell_{max} = 700$. The dotted horizontal line shows the current *Planck* constraint (in the LSS convention). Right: as a function of ℓ_{min} and for various flux cuts and two values of ℓ_{max}.

We consider the case of a 3 µJy flux sensitivity and perform the same Fisher analysis described before but this time considering the range $1.5 \le z \le 3$, again subdivided into 20 equally spaced bins. Although the noise is indeed larger in this case compared to the case of the previous section, we nevertheless find a tighter constraint

$$\sigma(f_{NL}) = 1.88 \quad (1.5 \le z \le 3). \tag{27}$$

This is a very promising result, and represents the most stringent constraint forecast for a future galaxy redshift survey exploiting one single tracer.

In fact, we can go further, and combine the two redshift chunks. The proper way to do so would be to bin C_ℓ^{ij} in the whole $0 < z \le 3$ interval. However, the calculation of a large number of redshift bins and all their cross-correlations is computationally cumbersome. To overcome this problem, we follow Camera et al. (2013) and construct a 40×40 block-diagonal tomographic C_ℓ^{ij} matrix, and then correct for the disregarded cross-correlations by overlapping a further 20-bin square matrix for the intermediate range $0.75 \le z \le 2.25$. This allows us to put stronger bounds on f_{NL}:

$$\sigma(f_{NL}) = 1.54 \quad (0 < z \le 3). \tag{28}$$

This constraint is five times tighter than the current *Planck* constraint. Furthermore, it is tight enough for testing slow-roll single-field inflation.

Fig. 3 illustrates this major result by showing the marginal 1σ error contours in the σ_8–f_{NL} plane for the three redshift chunks separately (red, blue and green ellipses) and for their combination (magenta contour). We can explain the different orientations and sizes of the ellipses as follows.

First, as we probe higher redshifts, constraints on σ_8 broaden. This is because the matter power spectrum decreases in amplitude as z increases, and so the sensitivity of C_ℓ^{ij} to σ_8 also decreases with redshift.

Secondly, let us focus on the f_{NL}-axis, where there are more subtle effects. It appears as though the most stringent constraint is obtained in the middle redshift interval. Indeed, we get $\sigma(f_{NL}) = 3.12, 1.82$ and 1.88 for the first, second and third chunk, respectively. At first glance, this seems to contradict the fact that PNG deviations from the Gaussian prediction monotonically increase with redshift. However, we must not forget that Poisson noise also grows with z. As a result,
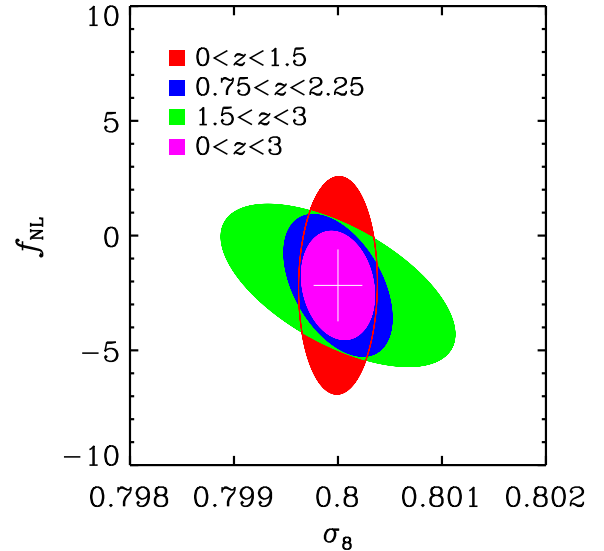

Figure 3. Marginal 1σ error contours in the σ_8–f_{NL} plane at 3 µJy sensitivity for the three redshift chunks and their combination.

the middle redshift interval is the one where the balance between PNG effects and noise is more optimal. A further check can be made by comparing the forecast constraints on f_{NL} in the idealistic case of zero noise. We find that $\sigma(f_{NL}) = 2.82, 1.29$ and 0.75 for the first, second and third chunk, respectively – in agreement with our qualitative expectations. All these major results are summarized in Table 2, where we denote by $\widehat{\sigma}(f_{NL})$ the marginal errors in the noiseless case.

5.3 PNG parameter degeneracies

Our analysis assumed that all parameters were fixed except for f_{NL} and σ_8. This could lead to some concerns about how this assumption might affect the constraints on f_{NL}. However, there are two important points that should be taken into account and which give us some security that the analysis is solid. First, most cosmological parameters do not really have an effect on the large scales (small ℓ) that we are considering. Moreover, small effects on these scales should be

Table 2. Forecast constraints on $f_{\rm NL}$ marginalized over σ_8 for 3 μJy sensitivity and the redshift ranges discussed in Section 5.2, with and without noise (second and third columns, respectively).

Redshift range	$\sigma(f_{\rm NL})$	$\widehat{\sigma}(f_{\rm NL})$
$0 < z \leq 1.5$	3.12	2.82
$0.75 \leq z \leq 2.25$	1.82	1.29
$1.5 \leq z \leq 3$	1.88	0.75
$0 < z \leq 3$	1.54	0.71

Figure 4. Ratio of the angular power spectrum to the fiducial one for the third redshift slice of the 3 μJy case when $f_{\rm NL}$ and $\mathcal{Q}$ are changed.

easy to disentangle from the k^{-2} signature that one expects for PNG. One such example is σ_8, which only shifts the power spectrum by an overall amplitude. Secondly, most of these same cosmological parameters should be very well measured at other scales and in particular using CMB data, which is independent of the $f_{\rm NL}$ effects. Summarizing, not only we do expect most parameters to be very well constrained already, but their effects on large scales should be negligible or easily distinguishable from the k^{-2} signature we are looking for.

Two issues however need a more detailed analysis: the effect of the galaxy bias b and the effect of the extra parameters such as $\mathcal{Q}$ and $b_{\rm e}$ in the relativistic corrections to C_ℓ^{ij}. Changes in these parameters could in principle mimic $f_{\rm NL}$ effects, since they also can affect power on large scales [the bias does this through the $(b-1)f_{\rm NL}$ term in equation (4)].

However, as above, $b(z)$ should be reasonably well constrained by measurements at smaller (but still linear) scales, using a combination of RSDs and the overall amplitude of the power spectrum (where we can safely neglect the possible k^{-2} dependence). We do need to assume that the bias is scale independent so that we can use the same value on large scales, but this should be reasonable in the linear regime. The information from the CMB will further lift any degeneracies. Moreover, even if small uncertainties in the bias still exist, the fact that there should be a reasonable monotonic evolution with redshift while $f_{\rm NL}$ is constant can be used to further disentangle any remaining degeneracies.

A very pessimistic approach would be to consider the bias as a free parameter in each redshift bin. However, we do not expect this to degrade considerably our constraints since the measurement of $f_{\rm NL}$ relies on the ability to pick up the k^{-2} dependence; this should thus not be affected by the uncertainty in the bias as long as it is scale independent on these scales. Note also that we have marginalized over σ_8 which is equivalent to marginalizing over an overall constant bias.

Regarding the relativistic parameters, i.e. the magnification bias $\mathcal{Q}$ and the evolution bias $b_{\rm e}$, the situation is similar. They both depend on the total galaxy number counts which should be very well measured at small scales using deep galaxy surveys. Obviously, we need to be careful with possible cosmic variance effects. Even if there remain small uncertainties in $\mathcal{Q}$ or $b_{\rm e}$, they should not impact on the extraction of $f_{\rm NL}$. We illustrate this in the case of $\mathcal{Q}$ in Fig. 4. This shows that the effect of $\mathcal{Q} \to 2\mathcal{Q}$ is much smaller than the effect of changing $f_{\rm NL}$ within the 1σ limits of our constraints.

6 CONCLUSIONS

In this paper, we made a detailed analysis of the expected PNG constraints from an H I galaxy survey, by using its effect on the large-scale power spectrum and fully taking into account the relativistic corrections expected at these scales. The relativistic corrections
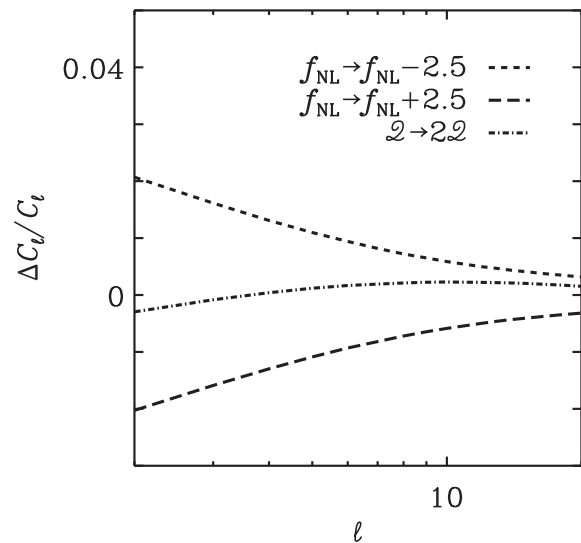
are small compared to the effects of astrophysical modelling, but this is not the relevant comparison: the point is that the relativistic effects are comparable to the PNG signal. By omitting these effects and using the Newtonian Kaiser approximation, one introduces an important theoretical systematic. We avoid this problem by using the correct relativistic analysis.

There are two aspects to the relativistic effects.

(i) The first is an important non-linear correction to the Poisson constraint which induces a correction $f_{\rm NL} \to f_{\rm NL} - 2.17$ when using LSS observations. This changes the fiducial value of $f_{\rm NL}$ from 0 to -2.17, and, more importantly, it shows that the simplest inflation models would give $f_{\rm NL} \simeq -2.2$ in LSS and $f_{\rm NL} \simeq 0$ in the CMB. We have shown that H I galaxy redshift surveys with the full SKA are able, in combination with the CMB, to detect the PNG signal from the simplest inflation models.

(ii) The second is the linear corrections to the observed galaxy number counts and power spectrum. The effect on $\sigma(f_{\rm NL})$ is typically small, but there will be a bias in the best-fitting estimates of $f_{\rm NL}$ from observations, given that the relativistic effects are comparable to the PNG signal.

Measuring the PNG signature on the clustering bias of dark matter tracers has the advantage that it is free from non-linearities that can affect the $f_{\rm NL}$ measurements on smaller scales using the bispectrum estimator. Furthermore, it will not be subject to foreground residuals or calibration artefacts that might contaminate the higher order correlation functions commonly used to estimate the amount of PNG.

In addition to H I galaxy (threshold) surveys, radio telescopes can also conduct H I intensity mapping surveys, which do not detect individual galaxies but the total H I emission, and are therefore ideally suited to map the large-scale distribution relevant for PNG. The intensity mapping surveys cover the same sky area but to greater depth (or cover the same depth in less time). This is shown to be particularly advantageous for an H I intensity mapping survey with SKA1 (Camera et al. 2013). On the other hand, galaxy surveys should give a pristine signal on these scales while foreground cleaning issues and correlated noise effects are still not completely clarified for H I intensity mapping surveys.

Our results show that a survey with Phase 1 of the SKA should reach 1σ constraints on $f_{\rm NL}$ of about 20. This is mostly due to the quick drop with redshift of the expected galaxy number density at this flux levels. On the other hand, as we move towards the full SKA, it will be possible to make a detection below $f_{\rm NL} \sim 2$, thanks to the large number of H I galaxies that will be detected up to high z. This will be five times better than the best current constraints (from *Planck*) and will start to probe into 'standard' inflationary model physics. Other galaxy surveys such as the European Space Agency *Euclid* satellite[3] (Laureijs et al. 2011; Amendola et al. 2013) will also target this type of measurement but only the full SKA should be able to probe such high redshifts ($z \sim 3$) over large areas of the sky, thanks to the abundance of H I galaxies as we move to higher redshifts, compared to other lines.

Finally, we would like to emphasize that the numbers we used for the bias and galaxy redshift distribution for different flux cuts were taken directly from simulations. The other parameters required for the relativistic corrections were in turn directly derived from these parameters, which allowed us to provide a fully consistent analysis of the fluctuations on ultra-large scales taking into account the general relativistic corrections. The obtained constraints should be reasonably insensitive to changes in other parameters due to the specific k^{-2} signature from PNG as well as the fact that most parameters should be already measured from surveys at smaller scales and the CMB itself.

REFERENCES

Ade P. et al., 2014, A&A, 571, A24
Afshordi N., Tolley A. J., 2008, Phys. Rev. D, 78, 123507
Amendola L. et al., 2013, Living Rev. Relativ., 16, 6
Bartolo N., Matarrese S., Riotto A., 2005, J. Cosmol. Astropart. Phys., 0510, 010
Bertacca D., Maartens R., Raccanelli A., Clarkson C., 2012, J. Cosmol. Astropart. Phys., 1210, 025
Bonvin C., Durrer R., 2011, Phys. Rev. D, 84, 063505
Bruni M., Crittenden R., Koyama K., Maartens R., Pitrou C., Wands D., 2012, Phys. Rev. D, 85, 041301
Bruni M., Hidalgo J. C., Meures N., Wands D., 2014, ApJ, 785, 2
Bull P., Ferreira P. G., Patel P., Santos M. G., 2014, preprint (arXiv:1405.1452)
Camera S., Santos M. G., Ferreira P. G., Ferramacho L., 2013, Phys. Rev. Lett., 111, 171302
Camera S., Santos M. G., Maartens R., 2014, preprint (arXiv:1412.4781)
Carbone C., Verde L., Matarrese S., 2008, ApJ, 684, L1
Challinor A., Lewis A., 2011, Phys. Rev. D, 84, 043516
Crociani D., Moscardini L., Viel M., Matarrese S., 2009, MNRAS, 394, 133
Dalal N., Dore O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514
Dewdney P., Hall P., Schilizzi R., Lazio J., 2009, Proc. IEEE, 97, 1482
Di Dio E., Montanari F., Lesgourgues J., Durrer R., 2013, J. Cosmol. Astropart. Phys., 1311, 044
Fedeli C., Moscardini L., Matarrese S., 2009, MNRAS, 397, 1125
Fedeli C., Carbone C., Moscardini L., Cimatti A., 2011, MNRAS, 414, 1545
Fedeli C., Bartelmann M., Moscardini L., 2012, J. Cosmol. Astropart. Phys., 1210, 018
Ferramacho L. D., Santos M. G., Jarvis M. J., Camera S., 2014, MNRAS, 442, 2511
Fisher R. A., 1935, J. R. Stat. Soc., 98, 39
Gangui A., Lucchin F., Matarrese S., Mollerach S., 1994, ApJ, 430, 447
Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, MNRAS, 422, 2854
Giannantonio T., Ross A. J., Percival W. J., Crittenden R., Bacher D., Kilbinger M., Nichol R., Weller J., 2014, Phys. Rev. D, 89, 023511
Grossi M., Verde L., Carbone C., Dolag K., Branchini E., Iannuzzi F., Matarrese S., Moscardini L., 2009, MNRAS, 398, 321
Hidalgo J. C., Christopherson A. J., Malik K. A., 2013, J. Cosmol. Astropart. Phys., 1308, 026
Holman R., Tolley A. J., 2008, J. Cosmol. Astropart. Phys., 0805, 001
Jeong D., Schmidt F., Hirata C. M., 2012, Phys. Rev. D, 85, 023504
Komatsu E., Spergel D. N., 2001, Phys. Rev. D, 63, 063002
Laureijs R. et al., 2011, Euclid Definition Study Report, ESA-SRE-12
Maartens R., Zhao G.-B., Bacon D., Koyama K., Raccanelli A., 2013, J. Cosmol. Astropart. Phys., 1302, 044
Matarrese S., Verde L., 2008, ApJ, 677, L77
Meerburg P. D., van der Schaar J. P., Corasaniti P. S., 2009, J. Cosmol. Astropart. Phys., 0905, 018
Pajer E., Schmidt F., Zaldarriaga M., 2013, Phys. Rev. D, 88, 083502
Pillepich A., Porciani C., Hahn O., 2010, MNRAS, 402, 191
Raccanelli A., Bertacca D., Maartens R., Clarkson C., Dore O., 2013, preprint (arXiv:1311.6813)
Raccanelli A., Bertacca D., Dore O., Maartens R., 2014a, J. Cosmol. Astropart. Phys., 1408, 022
Raccanelli A. et al., 2014b, preprint (arXiv:1406.0010)
Raccanelli A., Dore O., Dalal N., 2014c, preprint (arXiv:1409.1927)
Ross A. J. et al., 2013, MNRAS, 428, 1116
Salopek D., Bond J., 1990, Phys. Rev. D, 42, 3936
Sartoris B., Borgani S., Fedeli C., Matarrese S., Moscardini L., Rosati P., Weller J., 2010, MNRAS, 407, 2339
Seljak U., 2009, Phys. Rev. Lett., 102, 021302
Taruya A., Koyama K., Matsubara T., 2008, Phys. Rev. D, 78, 123534
Tegmark M., Taylor A., Heavens A., 1997, ApJ, 480, 22
Verde L., Matarrese S., 2009, ApJ, 706, L91
Verde L., Wang L.-M., Heavens A., Kamionkowski M., 2000, MNRAS, 313, L141
Villa E., Verde L., Matarrese S., 2014, Class. Quantum Gravity, 31, 234005
Wagner C., Verde L., 2012, J. Cosmol. Astropart. Phys., 1203, 002
Yahya S., Bull P., Silva M., Santos M., Maartens R., Okouma P., Bassett B., 2014, preprint (arXiv:1412.4700)
Yang X., Zhang P., 2011, MNRAS, 415, 3485
Yoo J., 2010, Phys. Rev. D, 82, 083508
Yoo J., Fitzpatrick A. L., Zaldarriaga M., 2009, Phys. Rev. D, 80, 083514
Yoo J., Hamaus N., Seljak U., Zaldarriaga M., 2012, Phys. Rev. D, 86, 063514

APPENDIX A: FITTING FORMULAS FOR MAGNIFICATION BIAS

Yahya et al. (2014) provide fitting formulas for $\mathrm{d}N_{\rm g}(z)/\mathrm{d}z$ as a function of redshift for various flux sensitivities. In this appendix, we give similar fittings for $\ln N_{\rm g}(z, \mathcal{F} > \mathcal{F}_*)$ as a function of flux cut at a given redshift, useful for evaluating $\mathcal{Q}(z, \mathcal{F}_*)$ [see equation (13)].

Table A1. Coefficients of the fitting formula of equation (A1).

z	a_1	a_2	a_3
0.1	9.442 15	$-2.679\ 26$	0.157 762
0.3	11.7533	$-1.173\ 86$	0.183 97
0.5	12.2699	$-0.307\ 84$	0.233 865
0.7	12.2542	0.368 527	0.290 786
0.9	11.9786	0.952 263	0.351 139
1.1	11.5483	1.480 91	0.413 533
1.3	11.0155	1.972 99	0.477 282
1.5	10.4094	2.439 02	0.541 996

We propose the formula

$$\ln N_{\rm g}(\mathcal{F} > \mathcal{F}_*) = a_1 - \ln \mathcal{F}_* - \frac{a_2}{2}\left(\ln \mathcal{F}_*\right)^2 - \frac{a_3}{6}\left(\ln \mathcal{F}_*\right)^3 , \quad \text{(A1)}$$

with values of the fitted coefficients a_1, a_2 and a_3 given in Table A1 for various redshifts in the range $0 < z \leq 1.5$. (Intermediate redshift values can be obtained by interpolating.) As is clear from Fig. A1, our fit is in very good agreement with the tabulated $\mathrm{d}N_{\rm g}(z, \mathcal{F}_*)/\mathrm{d}z$, with errors on average below a few per cent. Note that equation (A1) contains no coefficient for the linear term in $\ln \mathcal{F}_*$. We found that the agreement was as good with a coefficient of -1 as with a free coefficient. Moreover, the correspondent $\mathcal{Q}$ is in good qualitative agreement with the phenomenological fitting formulas provided in the literature by Yang & Zhang (2011).
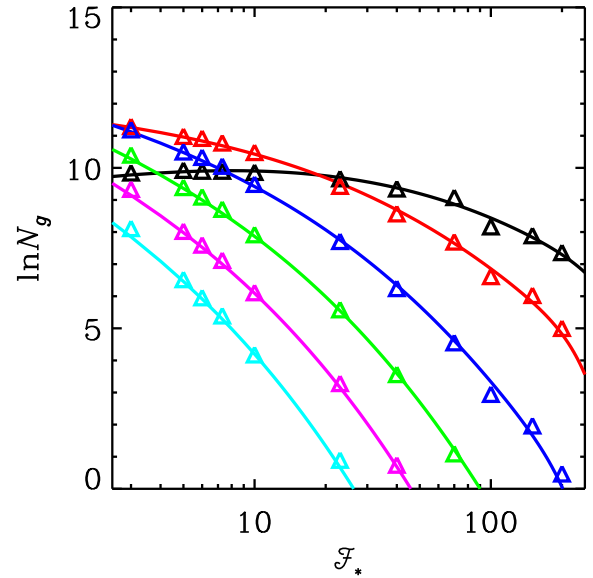

Figure A1. Background galaxy number density with flux larger than the flux threshold for $z = 0.1$, 0.3, 0.6, 0.9, 1.2 and 1.5 from top (right) to bottom. Solid lines are the outcome of equation (A1), and empty triangles come from the tables in Yahya et al. (2014).

This paper has been typeset from a TeX/LaTeX file prepared by the author.