

## ULTRA-LARGE-SCALE COSMOLOGY IN NEXT-GENERATION EXPERIMENTS WITH SINGLE TRACERS

DAVID ALONSO<sup>1</sup>, PHILIP BULL<sup>2</sup>, PEDRO G. FERREIRA<sup>1</sup>, ROY MAARTENS<sup>3,4</sup>, AND MÁRIO G. SANTOS<sup>3,5</sup><sup>1</sup> Astrophysics, University of Oxford, DWB, Keble Road, Oxford OX1 3RH, UK<sup>2</sup> Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, NO-0315 Oslo, Norway<sup>3</sup> Department of Physics, University of the Western Cape, Cape Town 7535, South Africa<sup>4</sup> Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK<sup>5</sup> SKA SA, The Park, Park Road, Pinelands 7405, South Africa

Received 2015 June 9; accepted 2015 October 26; published 2015 November 30

## ABSTRACT

Future surveys of large-scale structure will be able to measure perturbations on the scale of the cosmological horizon, and so could potentially probe a number of novel relativistic effects that are negligibly small on sub-horizon scales. These effects leave distinctive signatures in the power spectra of clustering observables and, if measurable, would open a new window on relativistic cosmology. We quantify the size and detectability of the effects for the most relevant future large-scale structure experiments: spectroscopic and photometric galaxy redshift surveys, intensity mapping surveys of neutral hydrogen, and radio continuum surveys. Our forecasts show that next-generation experiments, reaching out to redshifts  $z \simeq 4$ , will not be able to detect previously undetected general-relativistic effects by using individual tracers of the density field, although the contribution of weak lensing magnification on large scales should be clearly detectable. We also perform a rigorous joint forecast for the detection of primordial non-Gaussianity through the excess power it produces in the clustering of biased tracers on large scales, finding that uncertainties of  $\sigma(f_{\text{NL}}) \sim 1\text{--}2$  should be achievable. We study the level of degeneracy of these large-scale effects with several tracer-dependent nuisance parameters, quantifying the minimal priors on the latter that are needed for an optimal measurement of the former. Finally, we discuss the systematic effects that must be mitigated to achieve this level of sensitivity, and some alternative approaches that should help to improve the constraints. The computational tools developed to carry out this study, which requires the full-sky computation of the theoretical angular power spectra for  $\mathcal{O}(100)$  redshift bins, as well as realistic models of the luminosity function, are publicly available at <http://intensitymapping.physics.ox.ac.uk/codes.html>.

*Key words:* large-scale structure of universe

## 1. INTRODUCTION

The current success of relativistic cosmology is primarily based on the use of observations of large-scale structure (LSS) to infer the properties of the universe. The statistics of temperature and mass fluctuations, from maps of the cosmic microwave background (CMB) and galaxy surveys respectively, have been used to place remarkably tight constraints on the key cosmological parameters (Planck Collaboration 2015a), and we have now measured the spectral index of primordial fluctuations, the Hubble constant, and the present-day densities of dark matter, baryons, and relativistic species to exquisite precision. While many of these parameters can have similar effects on cosmological observations, by combining different observables at different times and length scales, we have been able to break the degeneracies between them.

While a variety of upcoming surveys will certainly improve the existing cosmological constraints, they will also allow us to probe an altogether new regime of large-scale structure—perturbations that span the cosmological horizon. While such scales are routinely studied in the CMB, these measurements consist of a single two-dimensional projection of the radiation density field and gravitational potentials at a fixed time. As such, the amount of information we can obtain from them is fundamentally limited by projection effects and cosmic variance. With three-dimensional maps of the matter density field, on the other hand, it should be possible to greatly refine our measurements of horizon-scale perturbations and, in

doing so, explore a variety of new relativistic effects in cosmology.

As pointed out in Bonvin et al. (2006), Yoo et al. (2009), Yoo (2010), Bonvin & Durrer (2011), and Challinor & Lewis (2011), relativistic effects come into play through apparent distortions of the projected survey volume by lensing, the propagation of light through inhomogeneous potentials, and the large-scale effect of peculiar velocities at the source. While they are strongly sub-dominant on scales of order  $100 h^{-1}$  Mpc, where current galaxy surveys are focused, they can substantially modify the power spectrum of, e.g., the number density of galaxies from the usual Newtonian predictions on extremely large scales.

The familiar Kaiser redshift-space distortion (RSD) is a sub-horizon approximation to a general-relativistic redshift-space distortion, with the post-Kaiser terms becoming non-negligible only around the horizon scale. Terms accounting for lensing, time delays, and the Sachs–Wolfe and integrated Sachs–Wolfe (ISW) effects are also present. (For analysis of each of these effects in large-scale structure observables, see Bertacca et al. 2012; Bruni et al. 2012; Jeong et al. 2012; Yoo et al. 2012; Hall et al. 2013; Lombriser et al. 2013; Yoo & Desjacques 2013; Raccanelli et al. 2014; Yoo & Seljak 2015.)

Most of these additional effects can safely be ignored in standard analyses of galaxy clustering, simply because they are negligible for current surveys of limited volume. A partial exception is the lensing term, which contributes to the observed overdensity (modulated by the magnification bias), and which has occasionally been incorporated into clustering analysis—it

can become significant on sub-horizon scales, but only at high redshift. The remaining terms<sup>6</sup> will be referred to as the “GR terms” for the purposes of this paper, which focuses on very large-scale effects. We will define the GR terms more precisely in due course. The hope, then, is that the GR effects can be teased out of cosmological data on ultra-large scales, and used to test the standard relativistic model of cosmology.

Another effect that can come into play on large scales is a scale-dependent bias due to primordial non-Gaussianity (Dalal et al. 2008; Matarrese & Verde 2008). If the primordial fluctuations are non-Gaussian, there will be a coupling between short- and long-wavelength modes such that the clustering of galaxies with respect to the underlying density field is enhanced on large scales. Specifically, the galaxy bias gains a scale dependence proportional to  $\sim k^{-2}$ , where  $k$  is the wavenumber of the mode being observed. The scale on which this effect comes into play is again of the order the cosmological horizon if the parameter quantifying the non-Gaussianity is  $f_{\text{NL}} \sim 1$ . Hence, in addition to detecting relativistic effects in structure formation, measurements of horizon-scale modes can also be used to constrain the statistics of primordial fluctuations, and thus the mechanism that seeded structure in the early universe.

As we can see, the scientific returns from measuring horizon-size fluctuations are legion. In this paper, we take a comprehensive view of future observations, and attempt to quantify how well both relativistic effects and primordial non-Gaussianity can be constrained with upcoming surveys. To do so, we examine four different types of survey: spectroscopic and photometric galaxy redshift surveys, continuum surveys of radio galaxies, and intensity mapping surveys of neutral hydrogen (H I). Each of these techniques will probe different redshifts with different sensitivities, but all will (in principle) be able to access horizon-scale modes in the next decade. The surveys will also measure different combinations of the relativistic corrections, and will be sensitive to different systematic effects. We will therefore pay close attention to identifying the different obstacles for detecting large-scale modes and discussing possible methods for mitigating their effects.

The paper is structured as follows. In Section 2 we present the various large-scale effects that we are targeting, focusing on the relativistic effects and primordial non-Gaussianity, and discuss how they might show up in galaxy surveys and intensity mapping experiments. We also discuss the nuisance parameters that determine the amplitudes of these effects, and their dependence on astrophysical uncertainties. In Section 3, we briefly present the Fisher forecasting formalism we will be working with. Then, in Section 4, we systematically present the forecasted uncertainties in these observables for the different types of survey, discussing the specific experiments that we consider, their sensitivity to ultra-large-scale observables, possible degeneracies with nuisance parameters, and their main sources of systematic uncertainties. The models used to describe the signal and noise for each experiment are summarized in Appendix B. Finally, in Section 5 we discuss our findings and draw conclusions for the prospects for detecting the ultra-large-scale effects.

<sup>6</sup> While excluded from our list of “GR terms,” we will also pay some attention to the lensing term, as it is a “hybrid” term that is only observable via clustering in very high volume surveys.

## 2. ULTRA-LARGE-SCALE COSMOLOGICAL OBSERVABLES

### 2.1. Relativistic Effects in Large-scale Structure Observables

The aim of this section is to compute all the terms that give rise to fluctuations in the number of light-emitting sources measured by an observer in a given redshift interval  $dz$  and observed solid angle  $d\Omega_o$ . The main result from this calculation is the following: neglecting all perturbations in the trajectory of the photons emitted by these sources, the observed perturbation in the number counts is simply given by the perturbation in the comoving number density of sources. The comoving 4-volume that we ascribe to the patch defined by  $dz$  and  $d\Omega_o$  depends on the direction and redshift of the photons received from these sources, however, and therefore any perturbation in their trajectory will induce additional contributions to the total fluctuation of source number counts.

RSDs are a perfect example of one of these contributions, where the observed redshifts are perturbed by the peculiar velocity of the source. While RSDs have been well understood for decades, several other terms have only recently been rigorously quantified (Bonvin & Durrer 2011; Challinor & Lewis 2011). These terms are relativistic in nature, and affect the clustering spectrum only on extremely large scales, approaching the horizon size. Since these relativistic effects have been thoroughly discussed in the literature, the aim of this section is not to provide a derivation from first principles, but rather to give some physical intuition for the interpretation of each of these terms, as well as to establish the notation that will be used in what follows.

#### 2.1.1. Relativistic Lightcone Effects: Number Counts

Consider a set of sources with comoving number density  $n_s(\eta, \mathbf{x})$  (as measured in their own rest frame) and 4-velocity  $u_s^\mu$ . These sources emit photons with a wave vector  $k^\mu \equiv dx^\mu/d\lambda$  ( $\lambda$  is an affine parameter of the photon geodesic) and rest-frame energy  $k_\mu u_s^\mu$ . During an interval  $d\lambda$  of the affine parameter, the photons cover a volume  $dA_e(k_\mu u_s^\mu)d\lambda$ , where  $dA_e$  is the invariant area of the wavefront corresponding to the observed solid angle  $d\Omega_o$ . Throughout, we have labelled quantities measured in the emitter’s and observer’s frames with subscripts  $e$  and  $o$  respectively.

The total number count in a redshift interval  $dz$  corresponding to  $d\lambda$  is therefore

$$\frac{dN}{dz d\Omega_o} = n_s \frac{dA_e}{d\Omega_o} (k_\mu u_s^\mu) \frac{d\lambda}{dz}. \quad (1)$$

Each of these terms is straightforward to compute in the absence of perturbations:

$$\begin{aligned} n_s(\eta, \mathbf{x}) &= \bar{n}_s(\eta(z)), \\ \frac{dA_e}{d\Omega_o} &= a^2(\eta(z)) r^2(z), \\ (k_\mu u_s^\mu) \frac{d\lambda}{dz} &= \frac{a(\eta(z))}{H(\eta(z))}. \end{aligned}$$

Here  $\eta(z)$  is the background conformal time at redshift  $z$ ,  $r(z)$  is the background comoving angular diameter distance,  $a(\eta(z)) = 1/(1+z)$  is the scale factor, and  $H \equiv \dot{a}/a$  is the expansion rate. For the rest of this work we will assume a flat

background cosmology, so that radial ( $\chi$ ) and angular distances are the same.

In the presence of inhomogeneities, all of these quantities are perturbed with respect to their background values at redshift  $z$ , and in general we can write

$$\eta(z, \hat{n}) \equiv \eta(z) + \delta\eta, \quad (2)$$

$$n_s(z, \hat{n}) = \bar{n}_s(\eta(z)) \left[ 1 + \delta_n + \frac{\partial \ln \bar{n}_s}{\partial \eta} \delta\eta \right], \quad (3)$$

$$\frac{dA_e}{d\Omega_o} \equiv a^2(\eta(z)) \chi^2(z) [1 + 2\delta_\perp], \quad (4)$$

$$(k_\mu u_s^\mu) \frac{d\lambda}{dz} \equiv \frac{a(\eta(z))}{H(\eta(z))} [1 + \delta_\parallel], \quad (5)$$

where  $\delta_n$  is the perturbation to  $n_s$ , and we have defined the perturbations to the conformal time,  $\delta\eta$ , transverse distance,  $\delta_\perp$ , and radial distance,  $\delta_\parallel$ .

One extra detail must be taken into account: not all sources are equally bright, and sources will generally be distributed according to a particular luminosity function,  $n_s(\eta, \ln L, \mathbf{x})$ , which we define as the density of sources in a logarithmic interval of luminosity:

$$n_s \equiv \frac{d(\# \text{ sources})}{dV d \ln L}. \quad (6)$$

Only sources with a flux (observed power per unit detector area) above a given detection cut,  $F_{\text{cut}}$ , will be detected. Flux and luminosity are related by an inverse-square law in angular distance, so perturbations to the angular diameter distance will affect the measured flux. Linearizing with respect to these perturbations gives

$$F(z, \hat{n}) = \frac{L}{4\pi(1+z)^4 a^2(\eta(z)) \chi^2(z)} [1 - 2\delta_\perp]. \quad (7)$$

At a given redshift and flux cut, we will only observe sources with luminosities above a threshold  $L_{\text{cut}}$ , related to  $F_{\text{cut}}$  by the previous equation. In order to take this into account, we must replace  $n_s$  by the cumulative luminosity function,

$$\mathcal{N}(\eta, \mathbf{x}, > \ln L) \equiv \int_{\ln L}^{\infty} d \ln L' n_s(\eta, \mathbf{x}, \ln L'), \quad (8)$$

so that Equation (3) becomes

$$\mathcal{N}(z, \mathbf{n}, F_{\text{cut}}) = \bar{\mathcal{N}} \left[ 1 + \delta_{\mathcal{N}} + \frac{\partial \ln \bar{\mathcal{N}}}{\partial \eta} \delta\eta - 2 \frac{\bar{n}_s}{\bar{\mathcal{N}}} \delta_\perp \right].$$

We have shortened our notation such that

$$\bar{\mathcal{N}} \equiv \bar{\mathcal{N}}(\eta(z), > \ln \bar{L}_{\text{cut}}(z, F_{\text{cut}})) \quad (9)$$

(and likewise for  $\bar{n}_s$ ), and have overlined all quantities evaluated in the background.

The full linear expression for the source number counts can finally be written as

$$\frac{dN}{dz d\Omega_o} = \bar{\mathcal{N}} \frac{a^3(z)}{H(z)} \chi^2(z) [1 + \Delta_N(z, \hat{n})], \quad (10)$$

where the perturbation is given by

$$\Delta_N(z, \hat{n}) = \delta_{\mathcal{N}} + \frac{\partial \ln \bar{\mathcal{N}}}{\partial \eta} \delta\eta + \delta_\parallel + 2\delta_\perp \left[ 1 - \frac{\bar{n}_s}{\bar{\mathcal{N}}} \right]. \quad (11)$$

In order to simplify the notation, from now on we will refer to the observed background number of sources found per unit redshift and solid angle simply as  $\bar{N}(z)$ , i.e.,

$$\bar{N}(z) \equiv \bar{\mathcal{N}} \frac{a^3(z)}{H(z)} \chi^2(z). \quad (12)$$

The terms  $\delta\eta$ ,  $\delta_\parallel$ , and  $\delta_\perp$  can be related to the metric, density, and velocity perturbations by solving the geodesic equation for photons in any gauge. In the conformal Newtonian gauge, defined by the line element

$$ds^2 = -a^2(\eta) [(1 + 2\psi)d\eta^2 - (1 - 2\phi)\delta_{ij}dx^i dx^j], \quad (13)$$

these perturbations read

$$\mathcal{H}\delta\eta = -\psi + \int (\phi' + \psi') d\eta + v_r \quad (14)$$

$$\delta_\parallel = \left[ 1 - \frac{\mathcal{H}'}{\mathcal{H}^2} \right] \mathcal{H}\delta\eta + \psi + v_r \quad (15)$$

$$+ \frac{1}{\mathcal{H}} \left[ -\frac{d\psi}{d\eta} + \phi' + \psi' + \frac{dv_r}{d\eta} \right] \quad (16)$$

$$\delta_\perp = \mathcal{H}\delta\eta - \frac{1}{\chi} [\delta\eta + \int (\phi + \psi) d\eta] - \phi - \kappa \quad (17)$$

$$\kappa \equiv \frac{1}{2} \int_0^x \frac{\chi - \chi'}{\chi \chi'} \nabla_\Omega^2 (\phi + \psi) d\chi', \quad (18)$$

where  $v_r \equiv \hat{n} \cdot \mathbf{v}_s$  is the radial peculiar velocity of the sources,  $\nabla_\Omega^2$  is the Laplacian on the unit sphere, and  $\kappa$  is the lensing convergence. Note that we have denoted all partial derivatives with respect to conformal time as  $\partial_\eta b \equiv b'$  (and  $\mathcal{H} \equiv a'/a$ ), and that the operator  $d/d\eta$  denotes a total lightcone derivative along the unperturbed photon trajectory,

$$\frac{db}{d\eta} \equiv \frac{d}{d\eta} [b(\eta, \mathbf{x} = (\eta_0 - \eta)\hat{n})], \quad (18)$$

where  $\eta_0$  is the age of the universe. Likewise, all integrals shown in the equations above must be understood as lightcone integrals along the same trajectory.

### 2.1.2. Relativistic Lightcone Effects: Intensity Mapping

Besides source number counts, another promising observational tool for studying large-scale structure is a technique known as *intensity mapping*. The technical details of this method are discussed in Section 4.1, but we will describe the relevant relativistic effects here (see also Hall et al. 2013).

In intensity mapping, the observable used to trace the matter density is the intensity received from a line-emitting medium integrated over a patch of the sky (i.e., the total power measured in a frequency interval per unit detector area and observed solid angle). We assume that this line emission is caused by some well-defined transition line, and can therefore be used to recover the redshift of the source by comparing the observed frequency with the known rest-frame one. In the rest

frame of a set of line-emitting sources, the emissivity is

$$\frac{dE_e}{dt_e d\nu_e d\Omega_e dV_e} = \frac{\hbar}{2} A_{21} \nu_e \varphi(\nu_e) \frac{x_2 \rho_a}{m_a}, \quad (20)$$

where  $A_{21}$  is the Einstein coefficient for the transition,  $\rho_a$  is the comoving density of the emitting gas,  $m_a$  is its atomic mass,  $x_2$  is the (number) fraction of the gas in the excited state, and  $\varphi(\nu)$  is the line profile (normalized to unity when integrated over all frequencies).

As shown in Section 2.1.1, the volume covered by the emitted photons in an affine parameter interval  $d\lambda$  is

$$dV_e = dA_e (k_\mu u_s^\mu) d\lambda. \quad (21)$$

Assuming that no absorption or scattering of the emitted photons occurs, the emissivity can be related to the measured intensity by

$$\begin{aligned} I(\nu_o, \hat{n}) &\equiv \frac{dE_o}{dt_o dA_o d\Omega_o d\nu_o} \\ &= \frac{\hbar \nu_o A_{21} x_2 \rho_a}{2 m_a} \varphi(\nu_e) (k_\mu u_s^\mu) d\lambda \frac{dA_e d\Omega_e}{dA_o d\Omega_o} \frac{d\nu_e dt_e}{d\nu_o dt_o}. \end{aligned}$$

The frequencies and time intervals in both frames are directly related through the redshift  $z$ , as are angles and invariant areas (from Etherington's reciprocity relation),

$$\frac{dt_e d\nu_e}{dt_o d\nu_o} = 1, \quad \frac{dA_e d\Omega_e}{dA_o d\Omega_o} = \frac{1}{(1+z)^2}. \quad (22)$$

Using these relations, and assuming that observations will take place on frequency intervals  $\Delta\nu_o$  much larger than the line width, we finally obtain the relation

$$I(\nu_o, \hat{n}) = \frac{\hbar A_{21} \nu_{21} x_2}{2 m_a (1+z)^2} \rho_a (k_\mu u_s^\mu) \frac{d\lambda}{dz}, \quad (23)$$

where  $\nu_{21}$  is the rest-frame line frequency. We can see that this is equivalent to Equation (1) for number counts, except for the factor of the angular diameter distance,  $dA_e/d\Omega_o$ . This is because the observable in intensity mapping is not the total number of objects in a given patch of sky, but the combined emitted light from the same patch. Since luminosities and angular distances are affected in the same way by lightcone effects, these cancel exactly for intensity mapping.

Expanding both  $\rho_a$  and  $(k_\mu u_s^\mu) \frac{d\lambda}{dz}$  to linear order, we can therefore compute all of the linear perturbations to the observed intensity mapping signal,

$$\begin{aligned} I(\nu, \hat{n}) &= \bar{I}(\nu) [1 + \Delta_I] \\ &\equiv \bar{I}(\nu) \left[ 1 + \delta_a + \frac{\partial \ln \bar{\rho}_a}{\partial \eta} \delta\eta + \delta_{\parallel} \right], \end{aligned} \quad (24)$$

where  $\delta_a$  is the intrinsic perturbation of the emitting gas density. By comparing this with Equation (11), we can see that the linear perturbation for intensity mapping is equivalent to the perturbation of the number counts for a population of sources with a particular form of the luminosity function, such that the number of sources observed above a given flux is proportional to the luminosity associated with that flux,

$$\tilde{N}_{\text{IM}}(>L) \propto L. \quad (25)$$

The somewhat unfortunate consequence of this result is that there are no linear perturbations to angular distances for intensity mapping,<sup>7</sup> which could potentially reduce the amount of cosmological information that can be extracted from this probe.

The background term in Equation (24) is commonly expressed in terms of antenna temperature, defined through the Rayleigh–Jeans relation for a blackbody emitter  $T(\nu) = I(\nu) c^2 / (2 k_B \nu^2)$ , where  $k_B$  is the Boltzmann constant. In terms of background quantities, the homogeneous intensity mapping signal is

$$\bar{T}(z) = \frac{3 \hbar A_{21} x_2 c^2}{32 \pi G k_B m_a \nu_{21}^2} \frac{H_0^2 \Omega_{b,0} x_a(z) (1+z)^2}{H(z)}, \quad (26)$$

where  $x_a(z)$  is the fraction of baryons made up by the line-emitting species under study.

### 2.1.3. Clustering Bias

Until now, we have not related the intrinsic perturbation in the number density of sources to the perturbations of the energy–momentum tensor. Assuming that galaxies form in dark matter haloes, which themselves form preferentially in high-density regions, one would expect the halo (or galaxy) number overdensity to trace the fluctuations in the overall matter density on large (linear) scales with a simple linear bias factor,  $\delta_{\text{halo}} \simeq b_{\text{halo}} \delta_M$ . This bias is a central piece of the halo model of structure formation (Mo & White 1996; Peacock & Smith 2000), the validity of which has been extensively tested against numerical simulations (Cole et al. 2008). Although the linear bias is expected to be scale-dependent on nonlinear and mildly nonlinear scales, where nonlinear and stochastic bias terms could also be important, we are mainly interested in large-scale observables in this paper, where the approximation of a linear, scale-independent bias should be valid (although see Section 2.2). This bias will depend on redshift and luminosity, however (e.g., more luminous, and therefore rarer, objects are expected to be more highly biased).

Since it is not possible to unambiguously define the matter overdensity  $\delta_M$  in a gauge-invariant way in a general-relativistic context, a subtle point is the choice of overdensity field on which the bias relation is applied. In this work we take the point of view of Challinor & Lewis (2011), Baldauf et al. (2011), Jeong et al. (2012), and Bruni et al. (2012), and argue that, since the process of galaxy formation is due to local physics, and since we expect our sources to follow the same velocity field as the dark matter, the bias relation should be applied in the synchronous comoving gauge. Note that it is also the comoving gauge perturbation that appears in the Poisson equation. A more complete discussion of this argument, which can also be extended to the case of primordial non-Gaussianity, can be found in Baldauf et al. (2011).

It follows that the intrinsic perturbation to the number density of sources,  $\delta_N$ , in the Newtonian gauge—our choice for this work—is related to the matter overdensity in the synchronous comoving gauge,  $\delta_{M,\text{syn}}$ , through

$$\delta_N = b(L, z, k) \delta_{M,\text{syn}} + \frac{\partial \ln \tilde{N}}{\partial \eta} \frac{v}{k}, \quad (27)$$

<sup>7</sup> Note that this is equivalent to the result that lensing of the CMB is a second-order effect.



where  $v$  is the peculiar velocity in the Newtonian gauge, and we have allowed the bias,  $b$ , to be scale-dependent, in anticipation of the discussion in Section 2.2.

#### 2.1.4. Magnification and Evolution Bias

The amplitudes of the perturbations to the conformal time and transverse distances depend explicitly on the derivatives of the luminosity function of the source population with respect to luminosity and time (see Equation (11)). It has become common to express these derivatives in terms of the so-called *magnification bias*,  $s(\eta)$ , and *evolution bias*,  $f_{\text{evo}}(\eta)$ , defined as

$$s(\eta) \equiv \frac{5}{2} \frac{\bar{n}_s(\eta, \ln \bar{L}_{\text{cut}})}{\bar{N}(\eta, > \ln \bar{L}_{\text{cut}})}, \quad (28)$$

$$f_{\text{evo}}(\eta) \equiv \frac{\partial \ln[a^3 \bar{N}(\eta, > \ln \bar{L}_{\text{cut}})]}{\partial \ln a}. \quad (29)$$

As with the clustering bias, the values of  $s$  and  $f_{\text{evo}}$  depend on the source population under study, so must be modeled correctly in order to maximize the information that can be extracted from any clustering analysis. While  $b(z)$  must be determined directly from clustering statistics, it is possible to estimate  $s$  and  $f_{\text{evo}}$  directly from the overall number counts of sources as a function of redshift and magnitude. Let  $\bar{N}(z, < m_*)$  be the cumulative number of sources with magnitude  $m$  brighter than  $m_*$ , per unit solid angle and redshift interval.  $\bar{N}$  is related to the luminosity function,  $\bar{n}_s$ , through

$$\bar{N}(z, < m_*) = \frac{c \chi^2(z)}{(1+z)^3 H(z)} \int_{\ln L_*}^{\infty} \bar{n}_s(\eta, \ln L) d \ln L, \quad (30)$$

where the threshold luminosity  $L_*$  is  $L_* = 4\pi(1+z)^2 \chi^2(z) F_*$ , and fluxes and magnitudes are related through

$$m = -\frac{5}{2} \log_{10} \left[ \frac{F}{F_0} \right]. \quad (31)$$

Note that we have neglected evolution and  $k$ -corrections.

Using the definitions of  $f_{\text{evo}}$  and  $s$  in Equations (28) and (29), these quantities can be related to the derivatives of  $\bar{N}$  with respect to  $z$  and  $m_*$  by

$$\frac{\partial \log_{10} \bar{N}}{\partial m_*} = s, \quad (32)$$

$$\frac{\partial \log_{10} \bar{N}}{\partial \log_{10}(1+z)} = \frac{(2-5)s}{\chi aH} - 5s + \frac{H'}{aH^2} - f_{\text{evo}}. \quad (33)$$

Note that, in order to use these relations to estimate  $s$  and  $f_{\text{evo}}$ , it is necessary to have full redshift information about the source distribution. While this is available by default for spectroscopic surveys, determining the redshift distribution becomes more involved for photometric and radio continuum surveys. This is nevertheless a necessary task if these probes are to be usable for cosmological studies, where, e.g., the redshift and photometric redshift distributions must be correctly modeled. In any case, the uncertainties on  $s$  and  $f_{\text{evo}}$  will tend to grow toward large  $z$ , and must therefore be taken into account in any cosmological analysis.

As we described in Section 2.1.2, the case of intensity mapping is slightly different. In this case, perturbations to the angular distance cancel, which is equivalent to setting the magnification bias to the critical value  $s_{\text{IM}} = 2/5$ .  $f_{\text{evo}}$  can be determined directly from the redshift dependence of the background brightness temperature,  $T(z)$  (Equation (26)).

#### 2.1.5. Power Spectra

The most informative observable regarding the clustering of astrophysical sources is their two-point correlation,  $\langle \Delta_N(z_1, \hat{n}_1) \Delta_N(z_2, \hat{n}_2) \rangle$ . The perturbation  $\Delta_N$  can be expressed in terms of spherical harmonic coefficients,

$$a_{\ell m}(z) \equiv \int d\hat{n} \Delta_N(z, \hat{n}) Y_{\ell m}(\hat{n}), \quad (34)$$

where  $Y_{\ell m}(\hat{n})$  are the spherical harmonics. The clustering of number counts can then be studied through the angular power spectrum, defined by the correlation

$$\langle a_{\ell m}(z_1) a_{\ell' m'}^*(z_2) \rangle \equiv \delta_{\ell\ell'} \delta_{mm'} C_{\ell}(z_1, z_2), \quad (35)$$

where angle brackets denote an ensemble average.

In practice,  $\Delta(z, \hat{n})$  is not measured in infinitesimal intervals of  $z$ , but by averaging over a set of finite radial bins, which we will label here by a Latin index,  $i$ . The observed anisotropy in bin  $i$  is

$$a_{\ell m}^i \equiv \int dz W_i(z) \Delta(z, \hat{n}), \quad (36)$$

where the window function  $W_i$  is normalized to 1 when integrated over redshift. The shape of  $W_i$  is determined by both the background redshift distribution of observed sources,  $\bar{N}(z)$ , and the probability that a source at redshift  $z$  will be included in the  $i$ th bin  $p_i(z)$ , so that

$$W_i(z) \propto \bar{N}(z) p_i(z). \quad (37)$$

Using this, one can show that the cross-spectrum between two bins can be written as (Di Dio et al. 2013)

$$C_{\ell}^{ij} = 4\pi \int_0^{\infty} \frac{dk}{k} \mathcal{P}(k) \Delta_{\ell}^i(k) \Delta_{\ell}^j(k), \quad (38)$$

where  $\mathcal{P}(k)$  is the dimensionless primordial power spectrum, which is assumed to take the form  $\mathcal{P}(k) = A_s (k/k_0)^{n_s-1}$ , and  $\Delta_{\ell}^i(k)$  contains the transfer functions of the terms that contribute to the anisotropy in bin  $i$ , in Fourier space and projected on the sky. Expanding the various contributions to Equation (11), the functions  $\Delta_{\ell}^i(k)$  can be written as a sum of 10 terms corresponding to different physical effects (Di Dio et al. 2013):

$$\begin{aligned} \Delta_{\ell}^{\text{D},i}(k) &\equiv \int d\eta b \tilde{W}_i \delta_{M,\text{syn}}(k, \eta) j_{\ell}(k\chi(\eta)), \\ \Delta_{\ell}^{\text{RSD},i}(k) &\equiv \int d\eta (aH)^{-1} \tilde{W}_i(\eta) \theta(k, \eta) j_{\ell}''(k\chi(\eta)), \end{aligned} \quad (39)$$

$$\begin{aligned} \Delta_{\ell}^{\text{L},i}(k) &\equiv \ell(\ell+1) \int d\eta \tilde{W}_i^{\text{L}}(\eta) (\phi + \psi)(k, \eta) j_{\ell}(k\chi(\eta)), \\ \Delta_{\ell}^{\text{V1},i}(k) &\equiv \int d\eta (f_{\text{evo}} - 3) aH \tilde{W}_i(\eta) \frac{\theta(k, \eta)}{k^2} j_{\ell}(k\chi(\eta)), \end{aligned} \quad (40)$$

$$\Delta_{\ell}^{V2,i}(k) \equiv \int d\eta \left( 1 + \frac{H'}{aH^2} + \frac{2-5s}{\chi aH} + 5s - f_{\text{evo}} \right) \times \tilde{W}_i(\eta) \frac{\theta(k, \eta)}{k} j_{\ell}'(k\chi(\eta)), \quad (41)$$

$$\Delta_{\ell}^{P1,i}(k) \equiv \int d\eta \left( 2 + \frac{H'}{aH^2} + \frac{2-5s}{\chi aH} + 5s - f_{\text{evo}} \right) \times \tilde{W}_i(\eta) \psi(k, \eta) j_{\ell}(k\chi(\eta)), \quad (42)$$

$$\Delta_{\ell}^{P2,i}(k) \equiv \int d\eta (5s - 2) \tilde{W}_i(\eta) \phi(k, \eta) j_{\ell}(k\chi(\eta)),$$

$$\Delta_{\ell}^{P3,i}(k) \equiv \int d\eta (aH)^{-1} \tilde{W}_i(\eta) \phi'(k, \eta) j_{\ell}(k\chi(\eta)), \quad (43)$$

$$\Delta_{\ell}^{P4,i}(k) \equiv \int d\eta \tilde{W}_i^{P4}(\eta) (\phi + \psi)(k, \eta) j_{\ell}(k\chi(\eta)),$$

$$\Delta_{\ell}^{\text{ISW},i}(k) \equiv \int d\eta \tilde{W}_i^{\text{ISW}}(\eta) (\phi + \psi)'(k, \eta) j_{\ell}(k\chi(\eta)), \quad (44)$$

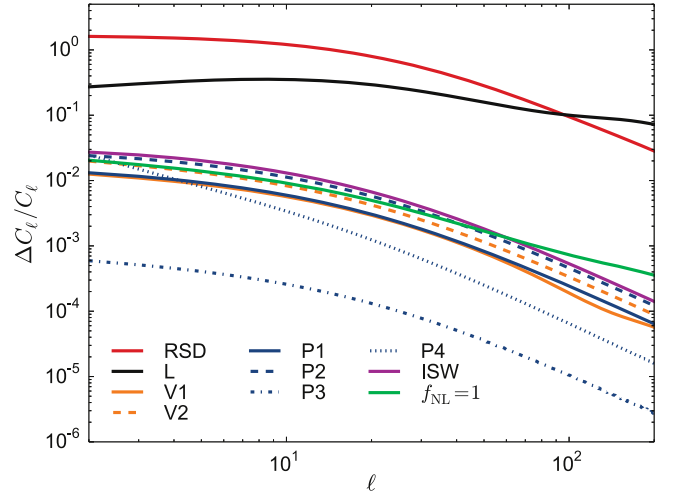
where we have defined the window functions

$$\begin{aligned} \tilde{W}_i(\eta(z)) &\equiv W_i(z) \left( \frac{d\eta}{dz} \right)^{-1}, \\ \tilde{W}_i^L(\eta) &\equiv \int_0^\eta d\eta' \tilde{W}_i(\eta') \frac{2-5s(\eta')}{2} \frac{\chi(\eta) - \chi(\eta')}{\chi(\eta)\chi(\eta')}, \quad (45) \\ \tilde{W}_i^{P4}(\eta) &\equiv \int_0^\eta d\eta' \tilde{W}_i(\eta') \frac{2-5s}{\chi}, \\ \tilde{W}_i^{\text{ISW}}(\eta) &\equiv \int_0^\eta d\eta' \tilde{W}_i(\eta') \\ &\times \left( 1 + \frac{H'}{aH^2} + \frac{2-5s}{\chi aH} + 5s - f_{\text{evo}} \right). \quad (46) \end{aligned}$$

In these equations, the quantities  $\delta_{M,\text{syn}}(k, \eta)$ ,  $\theta(k, \eta)$ ,  $\psi(k, \eta)$ , and  $\phi(k, \eta)$  are the transfer functions for the perturbation in the synchronous comoving gauge matter density, the divergence of the peculiar velocity, and the two metric potentials respectively.

Each term is sourced by a different physical effect.  $\Delta_{\ell}^L$  corresponds to the intrinsic perturbation in the comoving number density of sources, which is the dominant contribution in most cases, and is the only term that has traditionally been taken into account when RSDs and lensing can be neglected.  $\Delta_{\ell}^{\text{RSD}}$  is the usual RSD term corresponding to the Kaiser effect, due to the deformation of the Lagrangian volume in redshift space.  $\Delta_{\ell}^L$  is the lensing convergence term, caused by the deformation of the Lagrangian volume in the transverse directions due to weak lensing. The terms  $\Delta_{\ell}^{V1}$  and  $\Delta_{\ell}^{V2}$  are extra RSD contributions that come from evaluating the background terms at a redshift perturbed by the Doppler effect. The remaining terms correspond to the same effect, but for redshift perturbations caused by gravitational redshifting instead of peculiar velocities. In particular,  $\Delta_{\ell}^{\text{ISW}}$  is the analogue of the integrated Sachs-Wolfe (ISW) effect (Sachs & Wolfe 1967) for number counts.

Of these terms, the first three give the largest contribution to the total clustering anisotropy, so only these have traditionally been included in clustering analyses. The remaining terms are mainly relevant on super-horizon scales at the position of the sources, and even on those scales their amplitude is significantly smaller than the first three (see Figure 1). Nevertheless, these terms contain useful information that could



**Figure 1.** Amplitude of the different terms listed in Equations (39)–(44) relative to the amplitude of  $\Delta_{\ell}^L$  in the power spectrum. The  $C_{\ell}$ s were calculated for a redshift bin at  $z = 2$  with width  $\Delta z = 0.12$  and for constant bias functions ( $b(z) = 1.5$ ,  $s(z) = 1$ ,  $f_{\text{evo}}(z) = 1$ ). We have also included in green the contribution of primordial non-Gaussianity for  $f_{\text{NL}} = 1$ .

potentially be used, for example, to constrain different theories of gravity (Lombriser et al. 2013; T. Baker & P. Bull 2015, in preparation). One of the aims of this paper is to forecast the detectability of these terms by future experiments. In order to do so, we have defined an effective parameter,  $\epsilon_{\text{GR}}$ , which multiplies the terms  $\Delta_{\ell}^{V1,2}$ ,  $\Delta_{\ell}^{P1-4}$ , and  $\Delta_{\ell}^{\text{ISW}}$  and has a fiducial value of 1.  $\epsilon_{\text{GR}}$  therefore parameterizes the amplitude of the relativistic corrections to the clustering of sources.

Even though the origin of the lensing convergence term ( $\Delta_{\ell}^L$ ) is clearly general-relativistic, we have not included it under the umbrella of  $\epsilon_{\text{GR}}$  for two main reasons: first, we would like to focus on yet-undetected effects, and lensing magnification has already been detected by cross-correlating pairs of distant tracers (Scranton et al. 2005; Hildebrandt et al. 2009). Second, in this work we aim to identify possibly detectable observables on ultra-large scales, but the lensing term has a non-negligible effect on small angular scales. Nevertheless, for completeness we have also forecasted for the detectability of lensing magnification by defining an effective amplitude,  $\epsilon_{\text{WL}}$ , multiplying  $\Delta_{\ell}^L$ . In keeping with the main aim of this paper, however, we will only produce forecasts for this parameter based on its effects on the largest angular scales (lowest multipoles).

Thus, to clarify the terminology used here, we will refer to the terms  $\Delta_{\ell}^{V1,2}$ ,  $\Delta_{\ell}^{P1-4}$ , and  $\Delta_{\ell}^{\text{ISW}}$  as “GR effects” or “GR terms” and to  $\Delta_{\ell}^L$  as the “lensing term,” even if the nature of the latter is clearly relativistic.

## 2.2. Primordial Non-Gaussianity

A fundamental assumption of current theories of large-scale structure is that the primordial seed fluctuations can be described as a multivariate Gaussian random process, uniquely characterized by the primordial power spectrum (Baumann 2009). For many years, the Gaussianity of primordial fluctuations was one of the main predictions of inflationary theories of the early universe. More recently, the possibility of non-Gaussian primordial fluctuations has been revisited for two main reasons. On the one hand, a battery of statistical techniques have been developed to quantify primordial non-

Gaussianity, primarily from the CMB, but also adapted to large-scale structure data (Yadav & Wandelt 2008; Fergusson & Shellard 2009; Planck Collaboration 2014c). There is a hope that these techniques will uncover something that will enrich our understanding of the early universe, above simple one- and two-point statistics. On the other hand, the Effective Field Theory approach to inflation (Cheung et al. 2008; Weinberg 2008) can be used to systematically quantify all possible deviations from the quadratic action of linear perturbation theory around quasi-de Sitter space. These deviations, in the form of higher-order terms, will lead to non-trivial Gaussian signatures that are directly related to the fundamental parameters (and more importantly, the fundamental structure) of the theory of the early universe.

A useful (though not universal) way of describing deviations from primordial Gaussianity (Komatsu & Spergel 2001) is to assume that the gravitational potential  $\Phi$  can be decomposed into a quadratic polynomial in a Gaussian random field  $\phi$ , of the form

$$\Phi = \phi + f_{\text{NL}}(\phi^2 - \langle \phi^2 \rangle),$$

where  $f_{\text{NL}}$  is, in the simplest scenarios, a constant.  $f_{\text{NL}}$  has been calculated for a family of inflationary models. In the local (“squeezed”) configuration, it is expected to be of order  $|n_s - 1| \sim \mathcal{O}(10^{-2})$ , where  $n_s$  is the spectral scalar index of primordial fluctuations (although some non-canonical models can lead to  $f_{\text{NL}} \sim 1$ ; Verde et al. 2000; Liguori et al. 2006; Smith & Zaldarriaga 2011). Current constraints from higher order statistics of CMB maps place the constraint at  $|f_{\text{NL}}| \lesssim 7$  (Planck Collaboration 2015b).

A novel proposal is to look for the effects of non-Gaussian initial fluctuations at the level of the power spectrum. It has been shown that primordial non-Gaussianity induces a scale- and redshift-dependence for any biased tracer,  $X$ , of the overall density field (Dalal et al. 2008; Matarrese & Verde 2008). If the density contrast of  $X$  has a linear, Gaussian bias,  $b_X^G$ , then the non-Gaussianity of the distribution will induce a correction of the form

$$\Delta b_X(z, k) = 3f_{\text{NL}} \frac{[b_X^G(z) - 1] \Omega_M H_0^2 \delta_c}{(T(k)D(z)k^2)}. \quad (47)$$

Here,  $\Omega_M = \Omega_b + \Omega_C$  is the fractional energy density in matter (i.e., baryons plus dark matter),  $H_0$  is the Hubble constant,  $\delta_c \simeq 1.686$  is the critical density contrast of matter from the spherical collapse model,  $T(k)$  is the matter transfer function, and  $D(z)$  is the linear growth function for density perturbations.

As can be seen from Equation (47), there will be a substantial enhancement of the tracer power spectrum on large scales, with a  $\sim 1/k^2$  scale dependence (since  $T \sim 1$  on large scales). A rough estimate is that the transition scale is of order  $k_{\text{NG}} \sim f_{\text{NL}} H_0$ , i.e., we expect the scale-dependent signature to kick in close to the horizon scale for  $f_{\text{NL}} \sim 1$ . There have already been some attempts to constrain  $f_{\text{NL}}$  from the scale-dependent galaxy bias, although they have been severely hampered by systematic effects on extremely large scales (Giannantonio et al. 2014). In parallel, there have also been attempts to forecast the possibility of measuring  $f_{\text{NL}} \sim 1$  with future surveys (e.g., Namikawa et al. 2011; Giannantonio et al. 2012; Camera et al. 2013, 2015b; Ferramacho et al. 2014; Raccanelli et al. 2015a).

Given the nature of this signature—the fact that it arises on large scales and has a  $1/k^2$  scale dependence—it has been argued that non-Gaussianity may be degenerate with the relativistic effects we are studying in this paper (Bertacca et al. 2012; Bruni et al. 2012; Jeong et al. 2012). We will therefore include the effect of  $f_{\text{NL}}$  throughout our analysis and, in the process, present the most up-to-date and conservative forecasts for its detectability with future surveys. As illustrated in Figure 1, the extra power induced by a value of  $f_{\text{NL}} \sim 1$  on large scales is typically similar to the amplitude of the relativistic corrections presented in the previous section.

Including our three main observables ( $f_{\text{NL}}$ ,  $\epsilon_{\text{GR}}$ , and  $\epsilon_{\text{WL}}$ ) the total perturbation to the number counts is

$$\begin{aligned} \Delta_\ell^i &= \Delta_\ell^{\text{D},i}(f_{\text{NL}}) + \Delta_\ell^{\text{RSD},i} + \epsilon_{\text{WL}} \Delta_\ell^{\text{L},i} \\ &+ \epsilon_{\text{GR}} [\Delta_\ell^{\text{V}1,i} + \Delta_\ell^{\text{V}2,i} + \Delta_\ell^{\text{P}1,i} + \Delta_\ell^{\text{P}2,i} \\ &+ \Delta_\ell^{\text{P}3,i} + \Delta_\ell^{\text{P}4,i} + \Delta_\ell^{\text{ISW},i}], \end{aligned} \quad (48)$$

where there is an implied scale- and time-dependence in all of these terms.

### 3. FORECASTING FORMALISM

The spherical harmonic coefficients of the fluctuation in the observed number counts in the  $i$ th redshift bin,  $a_{\ell m}^i$ , contain most of the information about the clustering of sources. Assuming that they are Gaussian-distributed with a variance given by the cross-power spectrum  $C_\ell^{ij}$  (Equation (38)), it is straightforward to show that the log-likelihood for a given realization of the harmonic coefficients is

$$\ln \mathcal{L} = -\frac{1}{2} \left[ \sum_{\ell, m} \mathbf{a}_{\ell m}^T \mathbf{C}_\ell^{-1} \mathbf{a}_{\ell m} - \ln(\det[\mathbf{C}_\ell]) \right] + \text{const.}, \quad (49)$$

where we have written  $a_{\ell m}^i$  for each  $\ell$  and  $m$  as a vector  $[\mathbf{a}_{\ell m}]_i \equiv a_{\ell m}^i$ , and the set of cross-spectra  $C_\ell^{ij}$  as a matrix  $[\mathbf{C}_\ell]_{ij} \equiv C_\ell^{ij}$ . Our aim here is to forecast the precision with which different experiments will be able to measure a certain set of parameters,  $\{\theta_\alpha\}$ , which boils down to predicting the parameter covariance matrix. An efficient way of doing this is to use the Fisher matrix formalism, wherein the likelihood is approximated by a Gaussian expansion of Equation (49) around a fiducial set of parameters,

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{\alpha, \beta} (\theta_\alpha - \bar{\theta}_\alpha) F_{\alpha\beta} (\theta_\beta - \bar{\theta}_\beta) + \mathcal{O}(\theta^3), \quad (50)$$

where we have defined the Fisher matrix  $F_{\alpha\beta} \equiv \langle \partial^2 \ln \mathcal{L} / \partial \theta_\alpha \partial \theta_\beta \rangle$ . The covariance matrix  $C_{\alpha\beta} \equiv \langle (\theta_\alpha - \bar{\theta}_\alpha)(\theta_\beta - \bar{\theta}_\beta) \rangle$  can then be approximated as the inverse of  $F_{\alpha\beta}$ . For our likelihood, one can show that

$$F_{\alpha\beta} = f_{\text{sky}} \sum_{\ell=2}^{\ell_{\text{max}}} \frac{(2\ell+1)}{2} \text{Tr}[(\partial_\alpha \mathbf{C}_\ell) \mathbf{C}_\ell^{-1} (\partial_\beta \mathbf{C}_\ell) \mathbf{C}_\ell^{-1}], \quad (51)$$

where  $\partial_\alpha \equiv \partial / \partial \theta_\alpha$ .

We model the observable  $\mathbf{a}_{\ell m}$  as the sum of two contributions:  $\mathbf{a}_{\ell m} = \mathbf{a}_{\ell m}^S + \mathbf{a}_{\ell m}^N$ , corresponding to signal (i.e., cosmological anisotropies) and noise (non-cosmological fluctuations due to instrumental or shot noise). We will also assume that the two contributions are uncorrelated, so that  $\mathbf{C}_\ell = \mathbf{C}_\ell^S + \mathbf{N}_\ell$ . Here,  $\mathbf{C}_\ell^S$  is given by Equation (38), and  $\mathbf{N}_\ell$  is

the noise power spectrum, the exact form of which will depend on the particular type of experiment.

The theoretical power spectra  $C_\ell^S$  were computed using a modified version of the public CLASS code (Lesgourgues 2011; Di Dio et al. 2013). Our modifications are documented in detail in Appendix A. For our fiducial cosmology, we adopted a model consistent with the best-fit flat  $\Lambda$ CDM parameters from Planck (Planck Collaboration 2014b), given by  $(\Omega_M, f_b, h, w, A_s, n_s) = (0.315, 0.156, 0.67, -1, 2.46 \times 10^{-9}, 0.96)$ , where  $f_b \equiv \Omega_b/\Omega_M$  is the baryon fraction. We further set the fiducial value of  $f_{\text{NL}}$  to 0, the value for Gaussian initial conditions, and  $\epsilon_{\text{GR}} = \epsilon_{\text{WL}} = 1$ .

While we are primarily interested in forecasting for  $f_{\text{NL}}$  and  $\epsilon_{\text{GR}}$ , we must also marginalize over other parameters that could be correlated with them, which includes the six other cosmological parameters listed above, as well as the bias nuisance parameters described below. When forecasting for the uncertainty on  $f_{\text{NL}}$ , we do not consider  $\epsilon_{\text{GR}}$  as an extra free parameter, and fix it to its fiducial value of 1. Conversely, for the  $\epsilon_{\text{GR}}$  forecasts we fix all but  $f_{\text{NL}}$  and  $\epsilon_{\text{GR}}$ , assuming that only a possible degeneracy with  $f_{\text{NL}}$  could hamper a detection of the GR effects (all other parameters would simply change the shape of the GR correction “template”).

The derivatives required by Equation (51) were computed using central finite differences,

$$\partial_\alpha f = \frac{f(\theta_\alpha + \delta\theta_\alpha) - f(\theta_\alpha - \delta\theta_\alpha)}{2\delta\theta_\alpha} + \mathcal{O}(\delta\theta^3),$$

where we chose intervals  $\delta\theta_\alpha$  such that the estimated derivatives converged to the required numerical accuracy.

We also need to impose priors on certain parameters (e.g., the bias functions) to mitigate degeneracies. These are straightforward to incorporate into the Fisher matrix formalism: a Gaussian prior on  $\{\theta_\alpha\}$  with a covariance  $C_{\alpha\beta}^p$  can be added directly to the Fisher matrix as

$$F_{\alpha\beta} \longrightarrow F_{\alpha\beta} + (C^p)^{-1}_{\alpha\beta}. \quad (52)$$

For the cosmological parameters in particular, it is useful to add a Planck CMB prior, which we construct by estimating their covariance matrix directly from the corresponding Planck Markov Chain Monte Carlo (MCMC) chains (Planck Collaboration 2014a).

Finally, the constraints that any experiment will be able to yield will depend crucially upon the smallest and largest scales that can be used. For angular scales, this is explicitly taken into account in Equation (51) as the maximum multipole,  $\ell_{\text{max}}$ , that we sum up to. This cutoff is determined either by the angular resolution of the experiment (e.g., the beam size for intensity mapping) or by the nonlinear scale, beyond which the theoretical predictions become unreliable and the modes must be discarded from the analysis. The smallest radial scale corresponds to the comoving width of the redshift bins used, and is also determined by either instrumental effects (e.g., the redshift resolution in continuum and photometric surveys) or the nonlinear scale. In any case, since the effects we aim to study in this work are dominant on large scales, most of the information about them is concentrated on scales well inside the linear regime, and our final results are fairly insensitive to the choice of a minimum scale. By default we assume  $\ell_{\text{max}} = 500$  for all of the probes considered here, and address the redshift binning for each case individually. The maximum

**Table 1**  
Redshift Bins Used for the Bias Function Nuisance Parameters,  
for Each Survey

Survey	# bins	Bin Edges
Intensity mapping (SKA1-MID)	5	0.0, 0.7, 1.3, 2.0, 3.0, 4.0
Continuum survey (SKA)	5	0.0, 0.5, 1.0, 2.0, 3.0, 5.0
Spectroscopic (H $\alpha$ survey)	4	0.5, 0.75, 1.0, 1.3, 2.1
Photometric (LSST), red	4	0.0, 0.35, 0.7, 1.05, 1.4
Photometric (LSST), all	5	0.0, 0.5, 1.0, 1.5, 2.2, 3.0

angular scale is set by the sky area surveyed by each telescope. In our analysis, all surveys can probe down to  $\ell = 2$ . In later sections we analyse the effect of changing this minimum  $\ell$ . The largest radial scale is set by the available redshift range, or the maximum scale that is free from foreground contamination in the case of intensity mapping.

As described in Section 2.1, the amplitude of the number count spectra also depends on three redshift-dependent parameters: the clustering bias,  $b(z)$ , the magnification bias,  $s(z)$ , and the evolution bias,  $f_{\text{evo}}(z)$ . We will refer to these three parameters collectively as the “bias functions” in what follows. The value of  $b(z)$  can only be determined from the clustering statistics of the sample under study, and must therefore be marginalized over in the cosmological analysis. As discussed previously,  $s(z)$  and  $f_{\text{evo}}(z)$  can be estimated from the redshift-magnitude distribution of the sources, although these estimates will inevitably be uncertain and, to some extent, model-dependent. Properly accounting for this uncertainty is vital, as the behavior of the bias functions can strongly affect the detectability of the signal. In the absence of strong prior measurements from observations or simulations, these parameters must therefore also be marginalized over.

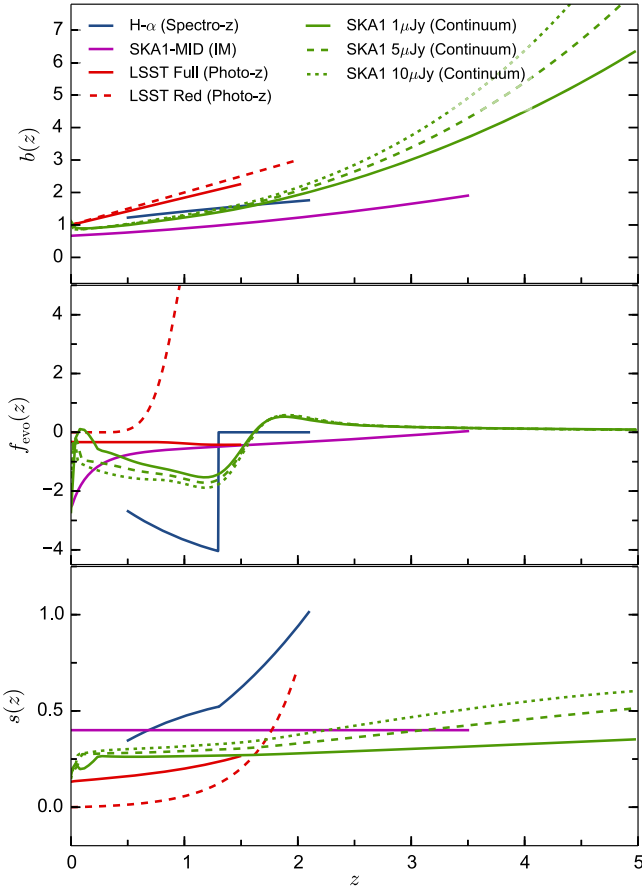
In order to do this, we defined a small number of redshift bins that sample the bias functions well enough for each survey. The mean values of the bias functions were computed in each bin, and a linear interpolation between those values was used to define each fiducial bias function. We then marginalized over the functions by treating each of the mean values as an additional free parameter that was included in the computation of the Fisher matrix (see Figure 3). We also explored other strategies, such as higher order interpolation and local modifications to the fiducial functions in each bin, but the method described above was found to be the best compromise between simplicity and stability to variations in each parameter. We also confirmed that the final results do not change significantly for the different methods. The redshift bins used for the bias functions for each of the four survey types are given in Table 1, and the input bias functions for the surveys<sup>8</sup> are shown in Figure 2.

#### 4. FORECASTS

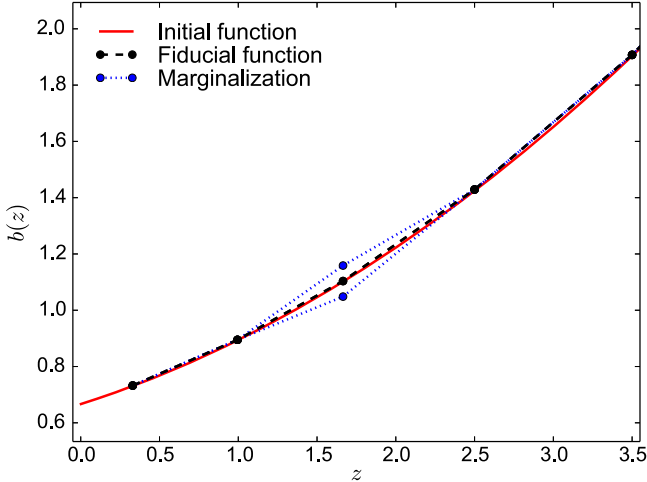
The aim of this section is to present the forecasted uncertainties on our main observables ( $f_{\text{NL}}$ ,  $\epsilon_{\text{WL}}$ , and  $\epsilon_{\text{GR}}$ ) for the main four types of cosmological surveys that will be used to measure them in the future: intensity mapping, radio continuum surveys, spectroscopic redshift surveys, and photometric redshift surveys. For each of them we will also discuss the main sources of systematic uncertainties that could prevent these

<sup>8</sup> The codes used to estimate the bias functions for the models detailed in Sections 4.1–4.4 can be found at <http://intensitymapping.physics.ox.ac.uk/codes.html>





**Figure 2.** Linear clustering bias (top panel), evolution bias (middle panel), and magnification bias (bottom panel) for the different experiments considered here.



**Figure 3.** Illustration of the procedure used in this work to marginalize over the bias functions in the case of the H I clustering bias. First the theoretical function is calculated (solid red line). Then, the mean values of the function in the redshift bins listed in Table 1 are computed (black circles), and the linear interpolation between these values (black dashed line) is used as the fiducial function in the computation of the power spectra. Finally, in order to marginalize with respect to this bias function, the mean values mentioned above are used as free parameters and varied to compute the numerical derivatives in Equation (51) (blue circles and dotted blue lines).

measurements. The signal and noise models assumed for each of these experiments are presented in detail in Appendix B and summarized in Table 2.

#### 4.1. H I Intensity Mapping

Intensity mapping (IM) is a relatively new technique, but has a number of potential advantages for the study of ultra-large scales (Battye et al. 2004; Chang et al. 2008; Wyithe & Loeb 2008). The idea is to observe the unresolved emission integrated over many line-emitting sources that are assumed to trace the large-scale matter distribution, sacrificing angular resolution for survey speed. For source populations with sufficiently narrow, isolated emission lines, high redshift resolution can nevertheless be obtained with a suitable spectroscopic instrument, allowing the 3D redshift-space matter distribution to be reconstructed. Thus, for a given pointing on the sky, one measures the combined emission from all the sources in it essentially as a continuous field, much like the CMB, rather than as a set of separately identifiable objects. There is therefore no need to set a flux threshold that rejects most of the peaks in the signal for being insufficiently far above the noise level; the whole of the signal can be used, but different modes will be recovered with more or less noise. This leads to significant gains in survey speed, making it possible to cover extremely large volumes—and thus constrain ultra-large scales—in a relatively short time.

The most developed technique to date is H I intensity mapping, which uses the redshifted 21 cm radio emission line of neutral hydrogen (H I) as its tracer. H I is ubiquitous even in the post-reionization universe, where it is mostly confined to dense, self-shielded damped Ly $\alpha$  systems. Forthcoming large, high-sensitivity, wide-bandwidth radio arrays such as the Square Kilometre Array (SKA), and purpose-built IM experiments like CHIME, are expected to be able to detect fluctuations in the cosmological H I signal over  $\sim 75\%$  of the sky, for a wide range of redshifts (Bull et al. 2015). We have focused our analysis on Phase I of the SKA, which has the large survey area and extremely wide frequency/redshift coverage needed to access ultra-large scales (although we have also computed the constraints for a cosmic variance-limited experiment).

The SKA is a proposed and partially funded multi-science radio facility that will be able to survey a large fraction of the sky in the frequency range from  $\sim 50$  MHz to  $\sim 20$  GHz. It will comprise two different instruments, built at separate sites in the South African Karoo region and Western Australia’s Murchison region:

1. SKA1-MID: an array of  $\sim 200$  single-pixel, 15 m dishes to be installed in South Africa. It will cover the frequency range 350–1760 MHz ( $z \lesssim 3$ ) in two separate bands.
2. SKA1-LOW: a set of about 455 aperture array stations each with 35 m diameter. It will cover the frequency range 50–350 MHz ( $3 \lesssim z \lesssim 20$ ).

Further information regarding the baseline design for the SKA can be found in Dewdney et al. (2009) and Braun et al. (2015).

As discussed in Bull et al. (2015) and Santos et al. (2015), the most efficient use of SKA1 for intensity mapping at late times ( $z \lesssim 3$ ) would be to use SKA1-MID as an autocorrelation experiment. If used in interferometric mode, the number of short baselines—needed to resolve the scale of baryon acoustic oscillations and larger—is not large enough. This fact is all the more important for ultra-large scales, and so we have focused on the SKA1-MID, single-dish case.

**Table 2**  
Properties of the Experiments Under Consideration

Experiment Type	Experiment	$[z_0, z_f]$	$\langle S/N \rangle$	$\langle b \rangle$	$\langle s \rangle$	$\langle f_{\text{evo}} \rangle$
Intensity mapping	SKA1-MID	[0.1, 3.5]	6.7	1.2	0.4	-0.4
Continuum survey	$S_{\text{cut}} = 10 \mu\text{Jy}$	[0, 3]	8.3	1.9	0.4	-0.6
	$S_{\text{cut}} = 5 \mu\text{Jy}$	[0, 4]	13	2.5	0.3	-0.3
	$S_{\text{cut}} = 1 \mu\text{Jy}$	[0, 5]	32	2.9	0.3	-0.2
Spectroscopic survey	H $\alpha$ survey	[0.65, 2.05]	3.6	1.5	0.55	-2.0
Photometric survey	LSST-red	[0, 1.4]	25	1.75	0.04	3.5
	LSST-full	[0, 2.5]	210	2.1	0.35	-0.7

**Note.** The quantities shown in columns 3–7 are, in order: the approximate redshift range probed by each experiment, the average signal-to-noise ratio ( $S/N$ , computed by averaging the ratio of the signal and noise power spectra over redshift and angular scale), and the clustering, magnification, and evolution biases averaged over redshift.

#### 4.1.1. Constraints on Relativistic Effects

As we have argued above, the two main sources of extra power on ultra-large scales are primordial non-Gaussianity and relativistic corrections. In order to study the detectability of the latter, we can therefore treat only  $f_{\text{NL}}$  and  $\epsilon_{\text{GR}}$  as free parameters, and fix the rest to their fiducial values. While this procedure would clearly yield an optimistic prediction of the actual constraint on  $\epsilon_{\text{GR}}$ , it mimics what a survey attempting a first detection of any new effect would do: fix all non-degenerate parameters to their best-fit values, and fit for the amplitude of the terms related to the new effect. If, in doing this, the  $S/N$  on the amplitude of the effect is smaller than unity, then there is no point in even considering the covariance with other parameters.

We applied this procedure for both SKA1-MID and a cosmic variance-limited survey ( $f_{\text{sky}} = 1$ ,  $N_{\ell}^{ij} = 0$ ), obtaining the following result:

$$\text{SKA1-MID} \longrightarrow \sigma(\epsilon_{\text{GR}}) = 2.75, \quad (53)$$

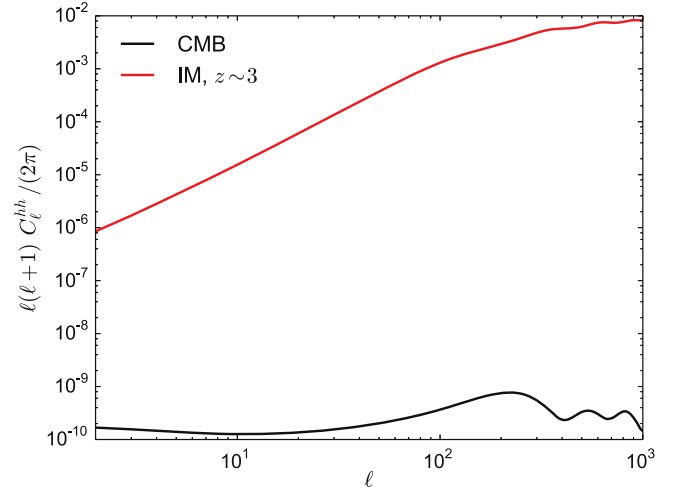
$$\text{CV-limited} \longrightarrow \sigma(\epsilon_{\text{GR}}) = 1.97. \quad (54)$$

Thus, even in the best-case scenario, it is not possible to detect the effect of the relativistic corrections on the brightness temperature fluctuations.

This may seem like a striking result, as one of the relativistic effects is the equivalent of the CMB ISW effect for large-scale structure ( $\Delta^{\text{ISW}}$  in Equation (44)), and the CMB ISW has already been detected above  $4\sigma$  by cross-correlating CMB maps with different LSS data sets (Giannantonio et al. 2008). In simplistic terms, intensity mapping surveys can be thought of as a set of uncorrelated “CMB” maps at different redshifts, so it is legitimate to ask why a similarly significant measurement is not possible in this case.

This can be explained in terms of the clustering variance of both data sets. Consider an attempt to measure the ISW effect by cross-correlating two data sets, one at high redshift (e.g., the CMB or a high- $z$  H I intensity map), which we label here by a superscript  $h$ , and another at low redshift (e.g., a galaxy survey or low- $z$  intensity map), which we label by  $g$ . Assuming that the ISW is the only term that could give rise to a significant cross-correlation between the two data sets, the signal would be given by the cross-power spectrum,

$$S_{\text{ISW}} = C_{\ell}^{hg}. \quad (55)$$

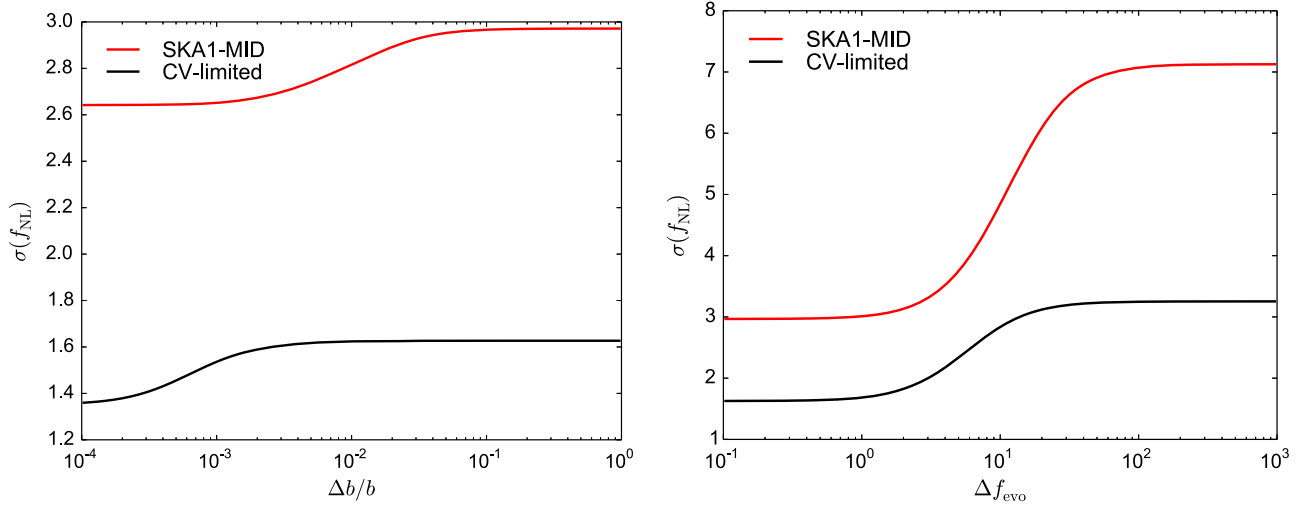


**Figure 4.** Dimensionless angular power spectrum for a H I intensity map at redshift  $z \sim 3$  with width  $\Delta z \sim 0.5$  (orange) and dimensionless temperature angular power spectrum of the CMB (blue). The much larger amplitude of perturbations at lower redshifts for intensity mapping can explain the difficulty in detecting the effect of relativistic corrections in LSS.

Neglecting any instrumental or shot noise, and assuming full-sky coverage and Gaussian statistics, the noise is purely given by the sample variance,

$$N_{\text{ISW}} \simeq \sqrt{\frac{2}{2\ell+1} C_{\ell}^{gg} C_{\ell}^{hh}}, \quad (56)$$

where we have assumed that  $C_{\ell}^{hg} \ll \sqrt{C_{\ell}^{gg} C_{\ell}^{hh}}$ . Except for factors of order unity, the amplitude of the signal depends only on the value of  $\phi' + \psi'$  at the redshift of  $g$ , and not on the nature of the high-redshift sample, so it will be roughly the same for both a high- $z$  intensity mapping bin and the CMB. The difference in  $S/N$  between the two cases must therefore depend primarily on the amplitude of the noise, which differs by the ratio of  $C_{\ell}^{hh}$  for the two cases. Since perturbations have grown significantly since  $z_{\text{CMB}} \sim 1100$ , we can expect the power spectrum of the intensity mapping to have a much larger amplitude,  $C_{\ell}^{hh, \text{IM}} \gg C_{\ell}^{hh, \text{CMB}}$ , which would explain the difficulty of achieving a good  $S/N$  in this case. This is shown explicitly in Figure 4: even at the highest redshift we



**Figure 5.** *Left panel:* uncertainty in  $f_{\text{NL}}$  as a function of the prior on  $b(z)$ . A constant Gaussian relative prior was imposed over the whole redshift range. *Right panel:* assuming a prior of 10% on  $b(z)$  (e.g., 10% uncertainty). This plot shows the dependence of  $\sigma(f_{\text{NL}})$  on the prior for  $f_{\text{evo}}$ .

considered, the power spectrum of the intensity mapping is four orders of magnitude larger than that of the CMB.

Another effect conspiring against the detection of the relativistic terms in an intensity mapping survey is the fact that the perturbations on transverse scales cancel exactly (i.e.,  $s(z) \equiv 2/5$ ), as previously discussed. This further suppresses the overall amplitude of the relativistic effects, and is the reason why we do not present forecasts for  $\epsilon_{\text{WL}}$  in this case.

#### 4.1.2. Constraints on Primordial Non-Gaussianity

The uncertainty on the level of primordial non-Gaussianity measured by an intensity mapping experiment will depend on our prior knowledge of the free parameters of the model. Here we have imposed a (non-diagonal) Gaussian prior on the cosmological parameters  $\{\Omega_M, f_b, h, w_0, n_s, A_s\}$  using a prior covariance matrix estimated from the appropriate Planck 2015 MCMC chains for our set of parameters (Planck Collaboration 2015a). As we have argued, it is reasonable to assume that, by the time the SKA attempts to measure  $f_{\text{NL}}$ , prior information will be available regarding the H I clustering bias (e.g., from experiments on smaller scales) and evolution bias (e.g., from external measurements of  $T_{\text{HI}}(z)$ ).

Before we assume any specific priors for these parameters, it is worth studying their effect on  $\sigma(f_{\text{NL}})$  in order to quantify how good prior measurements will need to be in order to optimize the constraints on  $f_{\text{NL}}$ . We first studied the level of degeneracy between  $f_{\text{NL}}$  and  $b(z)$  by estimating the value of  $\sigma(f_{\text{NL}})$  assuming a relative Gaussian prior on  $b(z)$ , constant across the whole redshift range. The result is shown in the left panel of Figure 5. We observe that a mild improvement on  $\sigma(f_{\text{NL}})$  ( $\sim 10\%$  for SKA1-MID and  $\sim 20\%$  for a CV-limited survey) can be achieved only for extremely accurate prior measurements of the clustering bias ( $\Delta b/b \lesssim 10^{-3}$ ). Since it would not be realistic to expect such a tremendous accuracy, we adopted a fiducial prior on  $b$  of 10%, more compatible with current measurements of the bias of neutral hydrogen (Masui et al. 2013).

Using this fiducial prior, we then explored the degeneracy between  $f_{\text{NL}}$  and  $f_{\text{evo}}$  by studying the dependence of  $\sigma(f_{\text{NL}})$  on the prior uncertainty,  $\Delta f_{\text{evo}}$ . The result is shown in the right panel of Figure 5. While  $\sigma(f_{\text{NL}})$  increases only slightly when

factoring in the uncertainties on the clustering bias, we observe a much larger increase (e.g., by a factor 2.4 for SKA1-MID) when we assume no knowledge about the evolution bias of the sample at all. This suggests a much stronger degeneracy between  $f_{\text{NL}}$  and  $f_{\text{evo}}$ , which could affect any attempt at measuring  $f_{\text{NL}}$  with LSS probes (note that a similar issue was reported by Camera et al. 2015b).

The source of the degeneracy can be understood by comparing the scale dependence of the terms in Equations (39)–(44) that are proportional to  $f_{\text{evo}}$  and  $f_{\text{NL}}$  respectively. Primordial non-Gaussianity introduces a term, included in  $\Delta_\ell^{\text{D}}$ , with a  $k$ -dependence of the form  $\propto \delta(\mathbf{k})j_\ell(k\chi)/(k^2 T(k)) \simeq \delta(\mathbf{k})j_\ell(k\chi)/k^2$ , where the second equality holds on ultra-large scales. The evolution bias, on the other hand, multiplies four different terms:

1.  $\Delta_\ell^{\text{PI}}$ , proportional to  $\psi(\mathbf{k})j_\ell(k\chi) \propto \delta(\mathbf{k})j_\ell(k\chi)/k^2$ .
2.  $\Delta_\ell^{\text{V1}}$ , which is proportional to  $\theta(\mathbf{k})j_\ell(k\chi)/k^2 \propto \delta(\mathbf{k})j_\ell(k\chi)/k^2$ .
3.  $\Delta_\ell^{\text{V2}}$ , proportional to  $\theta(\mathbf{k})j'_\ell(k\chi)/k$ . For sufficiently large  $\ell$  ( $\gtrsim 5$ ), this term is also proportional to  $\delta(\mathbf{k})j_\ell(k\chi)/k^2$ .
4.  $\Delta_\ell^{\text{ISW}}$ , the ISW term, which involves a much wider window function covering the full photon path from the source.

Thus, on large scales, three out of the four terms involving  $f_{\text{evo}}$  have the same scale dependence as the  $f_{\text{NL}}$  term, which explains the result found above. Fortunately, as can be seen in Figure 5, the effect of this degeneracy disappears if we can assume a relatively loose prior on  $f_{\text{evo}}$  of  $\Delta f_{\text{evo}} \lesssim 1$ . That is, if we parameterize the evolution of background density of H I in the universe as  $\rho_{\text{HI}}(a) \propto a^\alpha$ , the slope  $\alpha$  should then be measured with an error  $\Delta\alpha \lesssim 1$ . In what follows we assume that such an accuracy will be available from external measurements of  $T_{\text{HI}}(z)$ , and impose a Gaussian prior of  $\Delta f_{\text{evo}} = 1$ .

For our fiducial set of priors (Planck CMB priors for the cosmological parameters,  $\Delta b/b = 0.1$ , and  $\Delta f_{\text{evo}} = 1$ ), the final constraints on  $f_{\text{NL}}$  for SKA1-MID and for a cosmic variance-limited survey are given in Table 3. We also include results for an IM survey with a higher fiducial bias, which we discuss in Section 5.

**Table 3**  
Forecasted Constraints on  $f_{\text{NL}}$ ,  $\epsilon_{\text{GR}}$ , and  $\epsilon_{\text{WL}}$ .

Experiment Type	Experiment	$\sigma(f_{\text{NL}})$	$\sigma(\epsilon_{\text{GR}})$	$\sigma(\epsilon_{\text{WL}})$
Intensity mapping	SKA1-MID	3.01	2.75	...
	(w. $1.5\times$ bias)	0.90	1.90	...
	CV-limited	1.68	1.97	...
Continuum survey	$S_{\text{cut}} = 10 \mu\text{Jy}$	18.5	26.7	1.90
	$S_{\text{cut}} = 5 \mu\text{Jy}$	16.0	24.6	0.57
	$S_{\text{cut}} = 1 \mu\text{Jy}$	11.8	17.1	0.25
Spectroscopic survey	H $\alpha$ survey	6.64	2.57	0.19
	CV-limited	3.02	1.35	0.10
Photometric survey	LSST-red	4.32	1.41	0.14
	LSST-full	1.71	2.33	0.04

#### 4.1.3. Systematic Uncertainties

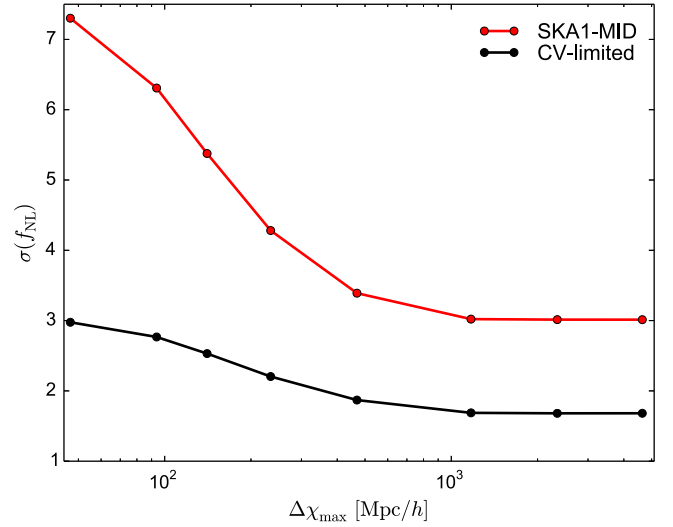
One of the most important observational challenges for intensity mapping is the presence of galactic and extragalactic radio foregrounds (e.g., galactic synchrotron emission and extragalactic continuum radio sources) with amplitudes several orders of magnitude larger than the cosmological H I signal. The potential bias and extra variance induced in the measured signal by the process of foreground removal must be correctly taken into account in any analysis.

The strategy underlying most foreground cleaning methods is to exploit the very different frequency structures of signal and foregrounds. Most foreground signals have a very smooth frequency dependence, while the cosmological signal traces the stochastic fluctuations in the matter density, and is therefore much “noisier” in the radial (frequency) direction. Broadly speaking, most cleaning methods try to remove the foregrounds by fitting and subtracting a set of smooth functions of the frequency from the combined foreground + cosmological signal. See Alonso et al. (2015) for a description and comparison of different methods.

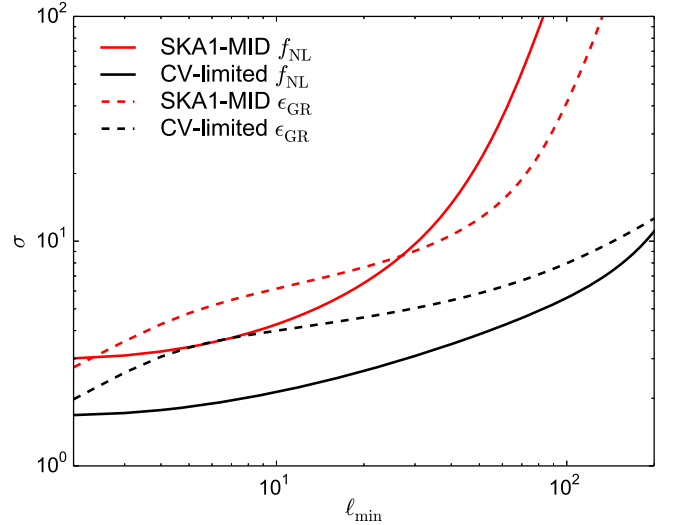
Since foregrounds are smooth in frequency, and frequency is a proxy for radial distances for an IM experiment, we can expect the foreground-cleaned maps to be dominated by systematics on large radial scales. These scales must then be omitted from the analysis, which reduces the sensitivity of an experiment to  $f_{\text{NL}}$ . In order to understand the importance of this effect, we have studied the dependence of  $\sigma(f_{\text{NL}})$  on the maximum radial separation between redshift bins included in the computation of the Fisher matrix, i.e., we set to zero the off-diagonal elements of  $C_\ell$  corresponding to pairs of bins separated by more than some radial separation  $\Delta\chi$ .

The results are shown in Figure 6 for SKA1-MID and for a cosmic variance-limited survey. Reducing the range of the cross-correlations included in the analysis can degrade the sensitivity to  $f_{\text{NL}}$  significantly, enlarging the errors by up to a factor  $\sim 3.3$  in the case of SKA1-MID.

Given that measurements of ultra-large-scale modes will typically need to be done in single-dish (autocorrelation) mode, one also needs to worry about the stability of the instrument and the observational strategy. Much as in CMB experiments, long-term noise drifts (the “ $1/f$ ” noise) will lead to striping in the maps, i.e., a coherent set of large-angle features that have been artificially projected on the sky. A sensible choice of scan strategy that leads to appreciable cross-linking between the scans can mitigate the effect, but there will always be a residual



**Figure 6.** Constraint on  $f_{\text{NL}}$  as a function of the maximum comoving separation between pairs of redshift bins included in the analysis (at  $z \sim 1.7$ , 1 Gpc corresponds to  $\sim 100$  Mpc).



**Figure 7.** Uncertainty on  $f_{\text{NL}}$  as a function of the largest angular scale (minimum  $\ell$ ) included in the Fisher matrix calculation. Systematic effects and incomplete sky coverage can prevent recovery of the lowest multipoles (in this plot, sky coverage,  $f_{\text{sky}}$ , is fixed to the survey specifications).

large-angle contaminant. Again, we can model this effect by not including the very large angular modes in the analysis (tantamount to assuming that they are filtered out by the destriping process). Figure 7 gives some idea of the impact of this effect on the constraints: a severe cut in the large-angle data significantly degrades any attempt to detect large-scale features.

#### 4.2. Radio Continuum Surveys

Radio continuum surveys measure the integrated emission of radio sources in one broad frequency band (or a small number of them). At radio wavelengths, the spectral energy distributions (SEDs) of most sources are generally smooth and featureless, except for a few radio lines such as the aforementioned 21 cm signal (which can only be resolved for individual sources after a long integration time). For this reason, there is little distance information to be gained by



integrating the flux density over more than a single, wide band. In turn, using a wide band significantly increases sensitivity, allowing much fainter sources to be observed than it would be possible to detect otherwise. Continuum surveys can therefore cover extremely large volumes, with the caveat that essentially all information on radial scales (even the ultra-large ones) is completely inaccessible.

That continuum surveys have the potential to constrain the level of primordial non-Gaussianity has already been shown in the literature (Ferramacho et al. 2014; Raccanelli et al. 2015a), especially if the survey can be split into several different populations so that the multi-tracer technique (Seljak 2009) can be used. We attempt to reproduce this result here, as well as exploring the possible degeneracies of  $f_{\text{NL}}$  with the relativistic corrections, and the possibility of detecting the latter. As with the other probes, we treat the continuum survey as being single-tracer only, and will not study the potential of the multi-tracer technique here. (This is left for future work, in which all the possible cross-correlations will be considered systematically.)

Our forecasts focus on a continuum survey with Phase 1 of the SKA, since this would correspond to the widest proposed survey area to date in the  $\sim \mu\text{Jy}$  regime. SKA1 should be able to detect radio sources out to  $z \sim 5$  over an area of about  $3\pi$  steradians. This survey would be carried out using the SKA1-MID facility, integrating the source flux in the band 350–1050 MHz with an rms noise of  $S_{\text{rms}} \simeq 1 \mu\text{Jy}$  (see Jarvis et al. 2015 for details).

It is common to define the source detection limit  $S_{\text{cut}}$  to be several times higher than the noise level of the experiment (usually by a factor of 5 or 10, depending on the intended use of the sample). Assuming an rms instrumental noise for SKA1-MID of  $S_{\text{rms}} = 1 \mu\text{Jy}$ , we have defined a fiducial  $5\sigma$  detection limit (i.e.,  $S_{\text{cut}} = 5 \mu\text{Jy}$ ). In order to explore the dependence on the survey depth, we have also produced forecasts for  $S_{\text{cut}} = 10 \mu\text{Jy}$  and  $S_{\text{cut}} = 1 \mu\text{Jy}$ .

#### 4.2.1. Constraints on Relativistic Effects

Relativistic effects in the continuum angular power spectrum were considered in Maartens et al. (2013) and Chen & Schwarz (2015). As in the previous section, we first present the most optimistic forecasts for  $\epsilon_{\text{GR}}$  by marginalizing only over  $f_{\text{NL}}$ . If the uncertainty on  $\epsilon_{\text{GR}}$  is larger than unity (i.e., no detection of GR effects) in this optimistic case, there is no point in exploring more realistic scenarios.

There are two main differences between intensity mapping and continuum surveys in terms of the quantities that can affect these forecasts. First of all, we can expect the lack of radial information in continuum surveys to considerably degrade the constraints on most parameters. Perturbations to transverse scales will affect the observed clustering of radio sources in this case, however (i.e.,  $s(z) \neq 2/5$ , unlike for intensity mapping), which will enhance the amplitude of the relativistic terms.

The constraints on  $\epsilon_{\text{GR}}$  found for the three flux limits that we considered are shown in Table 3. Due to the lack of radial information, there is no hope of detecting GR effects using only a single-tracer analysis, in spite of the enormous volume probed;  $\sigma(\epsilon_{\text{GR}}) = 17.1$  even for the deepest ( $1 \mu\text{Jy}$ ) survey.

We have also produced forecasts for the detectability of the weak lensing term on large scales for continuum surveys, following the same logic used in the case of  $\epsilon_{\text{GR}}$  (i.e., we keep all other parameters fixed, except for  $f_{\text{NL}}$ ). In order to pick up only the large-scale lensing contribution we also used a more

stringent value of  $\ell_{\text{max}} = 100$ . The results are summarized in Table 3: a continuum survey with a flux limit of  $1 \mu\text{Jy}$  would be able to clearly detect the large-scale lensing effect above  $\sim 4\sigma$ , although the level of this detection would be below  $2\sigma$  for  $S_{\text{cut}} = 5 \mu\text{Jy}$ . No detection would be possible for a flux cut of  $10 \mu\text{Jy}$ .

#### 4.2.2. Constraints on Primordial Non-Gaussianity

In a continuum survey, the available information is compressed into only a small amount of data—the angular clustering statistics of radio sources—due to the lack of any sensitivity to radial modes. We can therefore expect an even larger degree of degeneracy between the various cosmological and nuisance parameters than for intensity mapping. As before, we assume Planck CMB priors on all cosmological parameters (except  $f_{\text{NL}}$ ), and start our discussion of the primordial non-Gaussianity forecasts by exploring the effect of prior information about the bias functions on the  $f_{\text{NL}}$  constraints. The results of this analysis are shown in Figure 8.

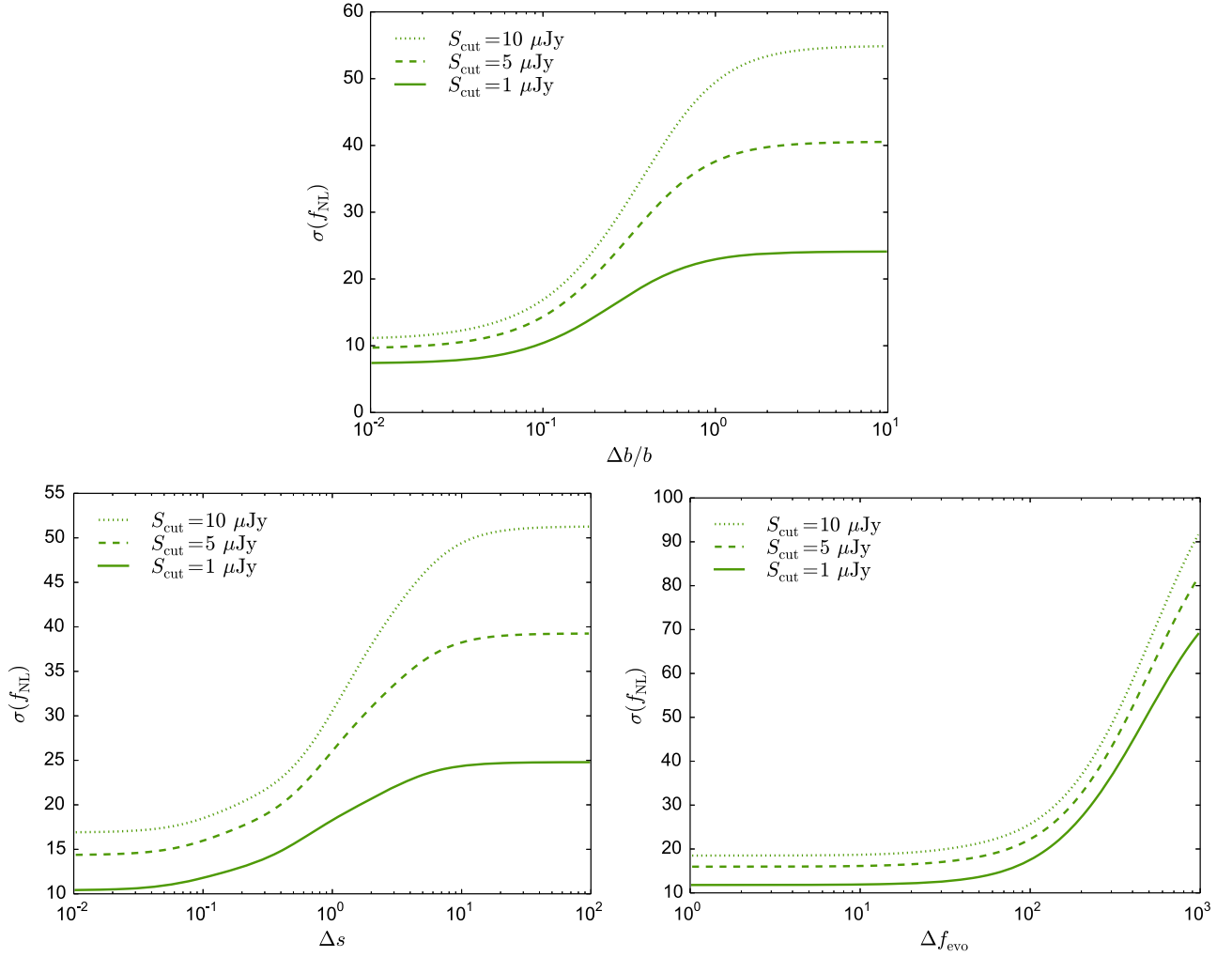
The top panel of Figure 8 shows the forecast uncertainty on  $f_{\text{NL}}$  as a function of a constant relative Gaussian prior for  $b$ . In contrast with the situation for intensity mapping, the constraints are much more sensitive to the prior on the clustering bias. The main reason for this is that, for IM experiments, the availability of redshift information helps to break the degeneracy between the parameters through the scale dependence of the  $f_{\text{NL}}$  term along the line of sight. We find that a  $\sim 10\%$  error on  $b(z)$  would be sufficient to minimize the uncertainty on  $f_{\text{NL}}$ . Since this is compatible with previous smaller-scale observations (see Lindsay et al. 2014), we chose this value as our fiducial prior on the clustering bias.

The effect of a prior on the magnification bias is shown in the bottom left panel of Figure 8. We observe a similar degeneracy with  $f_{\text{NL}}$ , again mainly due to the lack of redshift information, which can only be mitigated by prior measurements of  $s(z)$  with an error better than  $\Delta s \approx \pm 0.1$ . We have assumed that such an accuracy would be achievable using the magnitude–redshift distribution of sources, although this could be an optimistic assumption at the highest redshifts. Finally, as in the case of intensity mapping, we observe a significant degradation in the uncertainty on  $f_{\text{NL}}$  if we assume no knowledge about the evolution bias of the sample (bottom right panel in Figure 8). This is again due to the degeneracy in the scale dependence of the terms corresponding to both quantities (see Section 4.1.2), and can be mitigated by measuring  $f_{\text{evo}}$  to an accuracy of better than  $\Delta f_{\text{evo}} \lesssim 10$ .

In view of this analysis, our final fiducial set of priors on the bias functions is  $\Delta b/b = 0.1$ ,  $\Delta s = 0.1$ , and  $\Delta f_{\text{evo}} = 1$ . Forecasts for  $f_{\text{NL}}$  with these priors are listed in Table 3 and, with  $\sigma(f_{\text{NL}}) = 16$  for the  $5 \mu\text{Jy}$  sample, are compatible with the results of Ferramacho et al. (2014) for their combined sample. This is far worse than the constraints possible with multiple tracers, however.

#### 4.2.3. Systematic Uncertainties

The different frequency range and observational techniques involved in radio astronomy give rise to potentially very different sources of systematics for continuum surveys compared with optical and near-infrared surveys. To begin with, the diffuse nature of galactic synchrotron emission (the largest galactic foreground at radio frequencies) makes it

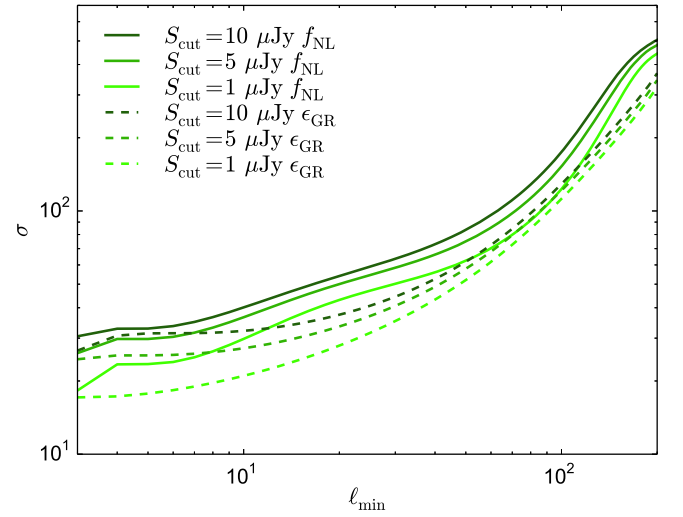


**Figure 8.** Dependence of the uncertainty on  $f_{\text{NL}}$  on the prior assumed for the clustering bias (top panel, in terms of a relative prior) and the magnification and evolution biases (bottom left and right panels, shown in terms of an absolute prior).

virtually transparent to the long interferometer baselines needed for a survey aiming to resolve individual sources, and hence the problem of galactic foregrounds is greatly ameliorated. On the other hand, in order to produce a full-sky catalog, mosaicing of the individual pointings must be implemented. If the mosaicing pattern and correlations in the noise properties between pointings are not fully understood, they could introduce systematic deviations on large angular scales. Ionospheric effects will also be relevant at low frequencies, although this should not be a problem in the SKA1-MID frequency range. Bright point sources would also need to be masked in a non-trivial way, due to dynamical range issues causing increased noise in the far beam sidelobes. The extent to which this would affect ultra-large scales is instrument-dependent, however. Avoiding these systematics might again entail removing the smallest multipoles of the power spectrum from the analysis; we show how the constraints on  $f_{\text{NL}}$  depend on the minimum multipole  $\ell_{\text{min}}$  in Figure 9.

#### 4.3. Spectroscopic Redshift Surveys

Spectroscopic galaxy redshift surveys in the optical and near-infrared represent the current state-of-the-art in large-scale structure observations (Percival 2014). The premise is simple: to detect redshifted emission lines from as many resolved



**Figure 9.** Dependence of the uncertainty on  $f_{\text{NL}}$  (solid lines) and  $\epsilon_{\text{GR}}$  (dashed lines) on the minimum scale probed by the survey for different flux limits.

sources as possible, over as large a spatial volume as possible, and then to reconstruct the 3D distribution of sources in redshift-space. Assuming that the source population traces the

underlying matter density field, and samples it sufficiently well, one can then constrain the statistical properties of the cosmological matter distribution as a function of scale and redshift. Because the positions and redshifts of the sources can both be measured with high precision, information about the matter distribution is retained in the angular *and* radial directions, unlike with more lossy techniques like intensity mapping and photometric redshifts. The downside is that taking high-resolution spectra for a large number of sources is extremely time-consuming.

For ultra-large scales, we are primarily interested in beating-down sample variance by increasing the effective survey volume. This requires a wide survey area, broad redshift coverage (preferentially at higher redshifts, where longer wavelength Fourier modes are accessible), and a sufficiently high signal intensity/source density that noise will be sub-dominant. Of the methods considered in this paper, spectroscopic galaxy surveys appear the least well-suited to the task of surveying extremely large volumes. Spectroscopy is time-consuming, and while precision redshift information is extremely useful for studying baryon acoustic oscillations and other smaller-scale phenomena, it is less necessary for the largest scales, apart from allowing the survey to be split into a larger number of redshift bins (compare intensity mapping). Nevertheless, spectroscopic galaxy surveys are the most developed of the methods, and so have comparatively well-understood systematic effects. The additional small-scale information also allows them to measure basic cosmological parameters more accurately, which helps to break parameter degeneracies.

We base the specifications of our reference spectroscopic survey on a large emission-line galaxy (ELG) survey along the same lines as *Euclid* (Laureijs et al. 2011), a satellite mission with a near-infrared spectrograph that will detect  $\sim 6 \times 10^7$  H $\alpha$ -emitting galaxies over 15,000 deg<sup>2</sup> in the redshift range  $0.65 \lesssim z \lesssim 2.05$ . A similar ELG survey will be performed by DESI (formerly BigBOSS), which will target [O II] galaxies out to  $z = 1.7$  over 14,000 deg<sup>2</sup> (Schlegel et al. 2011). These are the largest planned spectroscopic surveys,<sup>9</sup> but have the smallest area of the experiments considered here, and the lowest maximum redshift apart from LSST-red (see Section 4.4).

#### 4.3.1. Constraints on Relativistic Effects

The most optimistic forecast for  $\epsilon_{\text{GR}}$ , marginalizing over  $f_{\text{NL}}$  only, yields  $\sigma(\epsilon_{\text{GR}}) = 2.6$  for the H $\alpha$  survey's narrowest redshift binning ( $\Delta z = 0.025$ ). This result is insensitive to bin width, increasing only slightly to 2.7 for the widest binning ( $\Delta z = 0.1$ ); the additional information gained by decreasing the bin width is mostly confined to small scales, where the relativistic effects are essentially negligible. The correlation between  $\epsilon_{\text{GR}}$  and  $f_{\text{NL}}$  is very weak, and there is no change in the constraint whether  $f_{\text{NL}}$  is marginalized or fixed. The survey is also quite close to its ideal (sample variance-limited) performance, with  $\sigma(\epsilon_{\text{GR}})$  improving only slightly to 2.2 in the limit  $N_\ell \rightarrow 0$  for  $\Delta z = 0.025$ . These results, together with the constraints corresponding to a cosmic variance-limited

results (assuming  $N_\ell = 0$  and  $f_{\text{sky}} = 1$ ) are summarized in Table 3, and have a qualitatively similar behavior.

As with the two previous surveys, then, the relativistic effects are undetectable. This is despite the relatively high magnification bias of the H $\alpha$  galaxies, which boosts the size of some of the relativistic correction terms. The H $\alpha$  survey's sky coverage and maximum redshift are smaller than for the other surveys though, which weakens its constraining power. An experiment with the same specifications as the H $\alpha$  survey but covering twice the area (30,000 deg<sup>2</sup>) would give  $\sigma(\epsilon_{\text{GR}}) = 1.8$  (compared with 1.4 in the CV-limited case), which is still not enough to gain a detection.

Nevertheless, the H $\alpha$  surveys's forecast constraint of  $\sigma(\epsilon_{\text{GR}}) = 2.6$  is the best so far, and the CV-limited figure of 1.36 is markedly better than SKA1-MID's value of 1.97, despite the IM survey having a significantly wider redshift range. The enhanced performance of the spectroscopic survey over intensity mapping is primarily due to the different behavior of the bias functions, particularly  $s(z)$ , which caused many of the relativistic effects to cancel for the IM survey.

As before, we also forecast for the detectability of the large-scale lensing effect, parameterized by  $\epsilon_{\text{WL}}$ . After restricting to modes  $\ell \leq 100$  and marginalizing only over  $f_{\text{NL}}$ , we find  $\sigma(\epsilon_{\text{WL}}) = 0.19$  for all three choices of redshift bin width—a strong detection (see Table 3).

#### 4.3.2. Constraints on Primordial Non-Gaussianity

As with the previous two probes, our  $f_{\text{NL}}$  forecasts include a Planck CMB prior and priors on the bias functions ( $\Delta b/b = 0.1$ ,  $\Delta s = 1$ ,  $\Delta f_{\text{evo}} = 1$ ). The results are shown in Table 3 for the  $\Delta z = 0.025$  redshift binning.

The forecast constraint from the H $\alpha$  survey is  $\sigma(f_{\text{NL}}) = 6.8$ , which is worse than the intensity mapping survey by a factor of  $\sim 2$ . While the H $\alpha$  survey has a consistently higher bias (which enhances the non-Gaussian bias signal,  $\propto b - 1$ ), it covers a narrower redshift range and smaller area than the IM survey, so ultimately loses out when the higher-redshift bins of the IM survey are taken into account (see Figure 16, below). The difference in performance remains in the CV-limited case, again mostly due to the wider redshift range of the IM survey.

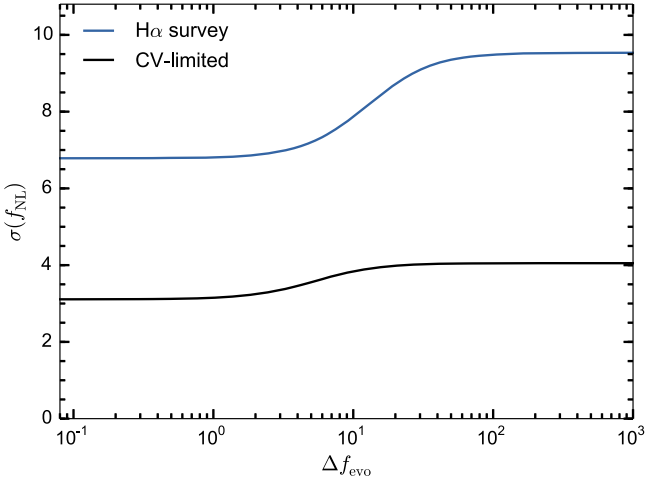
The H $\alpha$  constraint degrades only slightly to  $\sigma(f_{\text{NL}}) = 7.2$  for the widest redshift binning,  $\Delta z = 0.1$ . Similarly, it is only weakly sensitive to  $\ell_{\text{max}}$ , improving from  $\sigma(f_{\text{NL}}) = 7.2$  for  $\ell_{\text{max}} = 200$  to 6.6 for  $\ell_{\text{max}} = 1000$  (both for  $\Delta z = 0.025$ ). The addition of significantly more small-scale information in both the radial and transverse directions is therefore only mildly beneficial.

As with the IM survey, there is a reasonably strong degeneracy between  $f_{\text{NL}}$  and  $f_{\text{evo}}$ , predominantly for the highest-redshift nuisance parameter bin. Figure 10 shows the effect of changing the prior on  $f_{\text{evo}}$ —an  $\mathcal{O}(1)$  prior is sufficient to completely break the degeneracy. The results are insensitive to the prior on the magnification bias, and there is no gain to be had from tightening the bias prior until a very low level of  $\Delta b/b \lesssim 1\%$  is reached.

#### 4.3.3. Systematic Uncertainties

Spectroscopic surveys are prone to systematic effects on large angular scales. Redshift surveys commonly consist of samples from several non-contiguous fields, surveyed during different observing seasons and possibly even with different

<sup>9</sup> A proposed H I galaxy survey with Phase 2 of the SKA would detect  $\sim 10^9$  galaxies over  $\sim 30,000$  deg<sup>2</sup> for redshifts  $0 \leq z \lesssim 2$  (Yahya et al. 2015), but is considerably more futuristic.



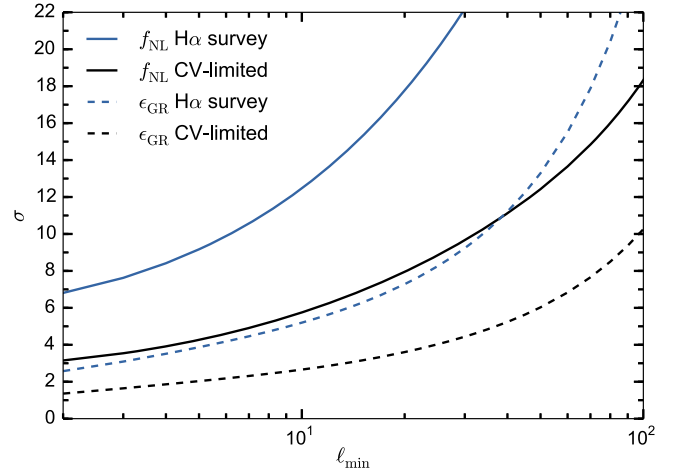
**Figure 10.** 1D marginal errors on  $f_{\text{NL}}$  for a  $\text{H}\alpha$  spectroscopic survey as a function of the prior on  $f_{\text{evo}}$ . A Planck prior on the cosmological parameters is also included.

instruments, which can make it tricky to patch them together into a single coherent survey volume. They may also suffer from the problem of not having a homogeneous magnitude limit, i.e., the magnitude cuts vary, and cannot easily be mapped onto a substantial and complete 3D volume of the sky.

Galactic extinction is a dominant source of systematic error on large angular scales. Dust in our galaxy changes the overall true flux cut of the survey, and introduces number density fluctuations that vary with the shape of the Galaxy. This must be corrected for (e.g., by fitting an extinction template) to avoid biasing the inferred large-scale power. Variations in airmass and seeing also affect the number of photons reaching the detector in a way that is correlated with the elevation of the telescope. This induces additional dispersion in the magnitude of the measured galaxies, and the ability to distinguish them from stars.

Stars themselves are problematic. For example, a small fraction of the observed sources may in fact be misidentified stars that contaminate the galaxy sample. Stars also obscure regions of the sky of the order of the size of the point-spread function, which reduces the observed density of galaxies nearby. While this should be a small effect for an individual star, which will mask an area of  $\sim 10^{-6}$  of a degree, the total obscured area can be significant given the large stellar density (which grows toward the galactic plane). In fact, stellar contamination was found to be a dominant source of systematic error in the recent analysis of the BOSS data (Ross et al. 2011), where it introduced a significant bias in the measurement of the correlation function on large angular scales if left uncorrected. This bias was well above the statistical uncertainty, to the extent that the correlation functions measured in the northern and southern Galactic hemispheres of the survey were inconsistent with one other.

While many of these effects can lead to fluctuations in the number density of galaxies as a function of redshift, inducing systematics along the radial direction, the dominant effect is on large angular scales, reducing the effective area of the survey and hampering accurate recovery of the lowest  $\ell$  modes. Figure 11 shows how the forecast constraints on  $\epsilon_{\text{GR}}$  and  $f_{\text{NL}}$  depend on the minimum recoverable  $\ell$  mode of the survey. There is a rapid loss of information on both parameters as  $\ell_{\text{min}}$



**Figure 11.** 1D marginal errors on  $f_{\text{NL}}$  and  $\epsilon_{\text{GR}}$  as a function of the minimum usable spherical harmonic mode,  $\ell_{\text{min}}$ , for the  $\text{H}\alpha$  survey ( $\Delta z = 0.025$ ). For the  $f_{\text{NL}}$  results, we have added a Planck prior to break degeneracies with other cosmological parameters, as well as priors on the bias functions ( $\Delta b/b = 0.1$ ,  $\Delta s = 1$ ,  $\Delta f_{\text{evo}} = 1$ ).

increases, with  $\sigma(f_{\text{NL}})$  doubling from 6.8 for  $\ell_{\text{min}} = 2$  to around 13 at  $\ell_{\text{min}} = 10$ . The degradation is similar for  $\sigma(\epsilon_{\text{GR}})$ , which also doubles in the same range. As such, future spectroscopic surveys will likely need excellent control over large-scale systematics if they are to be used to constrain  $f_{\text{NL}}$ .

#### 4.4. Photometric Redshift Surveys

One of the main drawbacks of spectroscopic surveys is the long integration times needed to resolve galaxy spectra sufficiently well to yield a good redshift estimate. Because of this, the number of targets selected for spectroscopic follow-up is usually much smaller than the total imaged sample, which significantly limits the survey depth and number density that can be achieved.

In a photometric redshift survey, each galaxy that is detected with a sufficiently high signal-to-noise ratio is imaged in a small number of wide frequency bands. This provides a very coarse measurement of the galaxy's spectrum, convolved with the bandpass of each band, which can be used to statistically infer its redshift. These photometric redshifts (usually abbreviated “photo- $z$ ’s”) have much larger uncertainties than their spectroscopic counterparts, and most of the information about gravitational clustering on radial scales is lost. Photo- $z$ s can be recovered for much fainter galaxies than spectroscopic redshifts, however, and so photometric surveys have the ability to cover significantly larger volumes. This potentially makes them more suitable for constraining cosmological observables on ultra-large scales.

The first wide-area, deep photometric surveys are already underway (Kaiser et al. 2002; Dark Energy Survey Collaboration 2005). Their results will pave the way for the Large Synoptic Survey Telescope (LSST, LSST Collaboration et al. 2009), which will surpass them in terms of area, depth, and angular resolution (although the latter is not a critical factor for this work). We have thus chosen to produce Fisher forecasts for LSST as the best photometric survey that we will have access to in the foreseeable future.

We have assumed that the LSST will observe two separate galaxy populations: early-type (“red”) galaxies and late-type (“blue”) galaxies. Although this is a simplistic picture, it allows



us to study the effect that different properties of the sample will have on the final constraints:

1. Early-type galaxies form preferentially in high-density regions, and are associated with high-mass haloes. They are therefore more highly biased than blue galaxies, which is desirable for measuring  $f_{\text{NL}}$ .

The number density of red galaxies decays very fast beyond redshift  $z \sim 1$ . On the one hand, the drop should be associated with a larger evolution bias, which could enhance the amplitude of the relativistic terms. On the other, it limits the largest scales that the red sample can probe. Finally, the spectra of red galaxies show prominent features, most importantly the 4000 Å Balmer break. These features are easy to locate, even using only photometric information, and therefore photo-zs for red galaxies are more accurate on average.

2. Blue galaxies are found in lower density regions and correspond to lower-mass haloes. They are therefore more faithful tracers of the total matter density field, and have a lower bias, which impacts their usefulness for measuring  $f_{\text{NL}}$ . LSST should be able to observe a significant number of blue galaxies up to much higher redshifts than the red population ( $z \sim 3$ ), however, so their large-scale clustering properties can be studied more accurately. Finally, photometric redshifts for blue galaxies will be more uncertain than those for red galaxies, as discussed above.

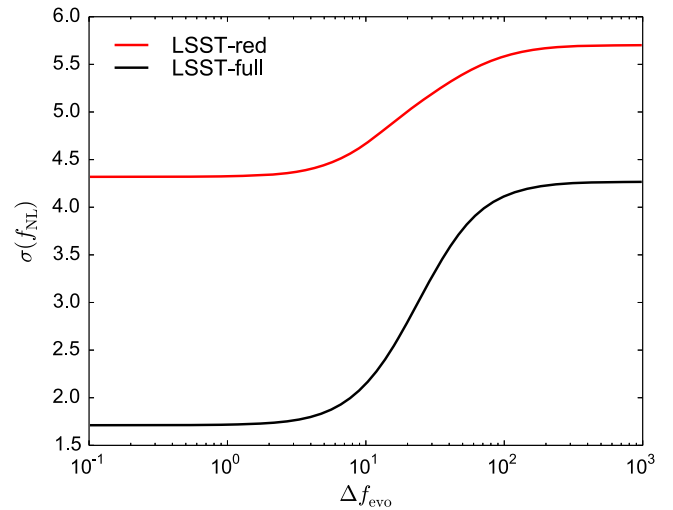
We will provide forecasts for two samples: a sample of red galaxies only, labeled “red,” and a sample containing all of the galaxies observed by LSST (red + blue), which we will call the “full” sample. Although LSST will be able to detect galaxies down to a magnitude limit of 27.5 in the  $r$ -band, it is not clear that the photometric redshift requirements will be satisfied for this survey depth. We have therefore adopted a more stringent magnitude cut of  $i < 25.3$ , corresponding to the so-called LSST “gold” sample (LSST Collaboration et al. 2009). We have further assumed that LSST will cover the whole southern hemisphere ( $f_{\text{sky}} = 0.5$ ).

#### 4.4.1. Constraints on Relativistic Effects

As with the previous experiments, we start by exploring the possibility of detecting the contribution of the relativistic terms to the clustering of LSST galaxies in the best-case scenario, by marginalizing over only  $f_{\text{NL}}$  while keeping all other cosmological parameters fixed to their fiducial values. Two main differences with respect to the previous tracers give some hope for detecting  $\epsilon_{\text{GR}}$  with LSST. First of all, the sharp decay in the number density of red galaxies can enhance the amplitude of the relativistic terms thanks to the large value of  $f_{\text{evo}}$ . Also, LSST covers a wider survey area and redshift range than spectroscopic surveys, so has access to larger scales.

The results are summarized in Table 3. Even though the higher value of  $f_{\text{evo}}$  for red galaxies helps to decrease the forecast uncertainty on  $\epsilon_{\text{GR}}$ , it is still impossible to detect relativistic effects using either sample;  $\sigma(\epsilon_{\text{GR}}) = 1.4$  and 2.3 for the red and full samples respectively. The red sample nevertheless produces the best constraint on  $\epsilon_{\text{GR}}$  of any of the surveys considered above.

As before, we predict the detectability of the large-scale magnification lensing by marginalizing only over  $f_{\text{NL}}$  and using a small-scale cutoff  $\ell_{\text{max}} = 100$ . Our results (see Table 3) show



**Figure 12.** Uncertainty on  $f_{\text{NL}}$  as a function of the size of the prior assumed for  $f_{\text{evo}}$  for the “red” and “full” samples.

that this effect should be clearly detectable (well above  $5\sigma$ ) for both the “red” and “full” samples.

#### 4.4.2. Constraints on Primordial Non-Gaussianity

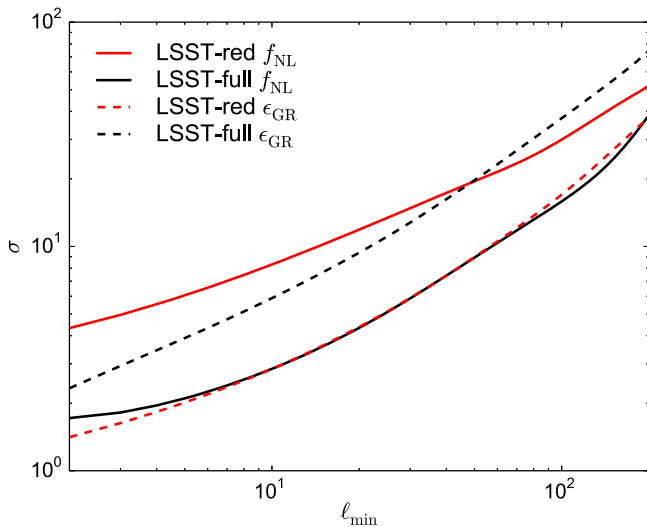
Even though photometric redshifts erase most of the clustering signal on all but the largest radial scales, they are still sufficient to enable a tomographic analysis of galaxy clustering to be performed. This helps immensely in breaking many of the degeneracies reported for continuum surveys (which retain essentially no radial information).

We studied the importance of breaking these degeneracies by again calculating  $\sigma(f_{\text{NL}})$  as a function of the priors on the bias parameters, finding that the forecasted uncertainty is almost completely insensitive to any priors on the clustering and magnification biases,  $b(z)$  and  $s(z)$ . For the reasons outlined in Section 4.1.2, this behavior does not follow for the evolution bias, so it is useful to explore the prior constraints on  $f_{\text{evo}}$  that are required in order to optimize the measurement of  $f_{\text{NL}}$ . Figure 12 shows the dependence of  $\sigma(f_{\text{NL}})$  on a constant Gaussian prior imposed on  $f_{\text{evo}}$ . The degeneracy between the parameters can be largely mitigated by measuring the evolution bias with an accuracy of  $\Delta f_{\text{evo}} \lesssim 1$ .

As in the previous cases, we produced our final forecasts for  $f_{\text{NL}}$  by assuming Planck CMB priors for the cosmological parameters, a 10% uncertainty on the clustering bias, and priors of  $\Delta s = 1$  and  $\Delta f_{\text{evo}} = 1$ . The final results are summarized in Table 3. LSST should be able to impose very tight constraints of  $\sigma(f_{\text{NL}}) \simeq 1.7$  using galaxy clustering autocorrelations (single-tracer) alone.

#### 4.4.3. Systematic Uncertainties

Most of the sources of systematics that affect photometric redshift surveys are exactly the same as for their spectroscopic counterparts: galactic extinction, variations in sky brightness, seeing, and stellar contamination (due to both stars affecting the local observed number density of galaxies and stars erroneously being included in the galaxy sample). All of these effects can potentially contaminate the signal measured on large angular scales. Figure 13 shows the degradation in the constraints on  $f_{\text{NL}}$  when the largest scales are omitted in the



**Figure 13.** Dependence of the uncertainty on  $f_{\text{NL}}$  (solid lines) and  $\epsilon_{\text{GR}}$  (dashed lines) on the minimum scale probed by LSST for the two samples considered here.

analysis. Limiting ourselves to scales  $\ell \geq 10$  would increase our best-case error bars by over 50%, to  $\sigma(f_{\text{NL}}) \sim 2.8$ .

The use of photometric redshifts also carries its own systematic effects. In order to obtain a reliable estimate of the power spectrum of the galaxy density field that we can use to constrain large scales, it is necessary to have a sufficiently accurate model of the window function for every redshift bin. Unless a spectroscopic subsample drawn from the same distribution as the photometric one is available, which is rarely the case, deriving a correct model for the true redshift distribution  $\bar{N}(z)$  is a challenging task, although it has been noted that this issue could be ameliorated by cross-correlating the photometric sample with any spectroscopic survey (Newman 2008). The presence of photometric redshift outliers can also modify the tails of the photo- $z$  distribution, which affects the shape of the redshift window functions. The level to which this effect is problematic will depend on how accurately the photo- $z$  probability density function (pdf) can be characterized.

## 5. DISCUSSION

It has been argued that general-relativistic corrections to the number density of galaxies should be observable with future cosmological surveys. In particular, ultra-large-scale features in the power spectrum of density fluctuations could in principle be detectable with deep and wide surveys such as those that will be carried out by the next generation of experiments. In this paper we set out to systematically examine this claim for the most relevant surveys planned for the next decade or so. At the same time, we performed forecasts for the expected constraints on the scale-dependent bias that arises from primordial non-Gaussianity, another key effect on ultra-large scales. Our analysis uses a more rigorous formalism than is usually followed, based on computing all possible angular cross-correlations between different redshift bins instead of using an approximate 3D power spectrum approach, and avoiding the flat-sky and Limber approximations.

Our main conclusion is that, if we restrict ourselves to the single-tracer power spectrum of the density fluctuations, in either two or three dimensions, all previously undetected contributions to the power spectrum of source number counts

are completely unobservable. Note that we have labelled these terms as “GR effects” in this work, thus excluding the lensing magnification term, which we have treated separately due to its very different properties (see below).

In hindsight, this result is not surprising.<sup>10</sup> A number of papers have previously attempted to forecast the optimal constraints on  $f_{\text{NL}}$  from a variety of surveys, with the general conclusion being that, at best, one can detect a value of  $f_{\text{NL}} \sim 1$ – $2$  at  $1\sigma$ . Our work has confirmed these results. Given that  $\epsilon_{\text{GR}}$  has a similar (although not identical) effect on the power spectrum as  $f_{\text{NL}} \sim 1$ , we expect the same level of sensitivity to relativistic effects. If we are to aspire to a statistically significant detection of the relativistic effects, we would need a sensitivity of  $\sigma(\epsilon_{\text{GR}}) \sim 0.1$ – $0.2$ , which is clearly unachievable with any of the single-tracer survey techniques and strategies analyzed here.

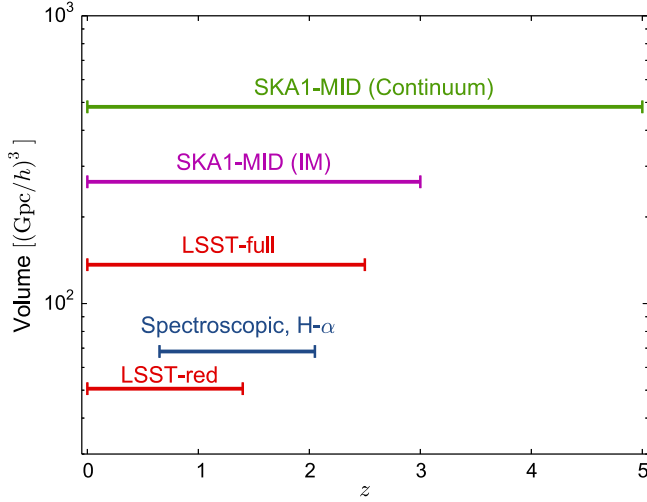
It is interesting to look at each of the survey techniques in turn to see why they fall short of our desired target:

1. On the face of it, *intensity mapping* is a particularly promising approach to efficiently surveying large volumes of the universe—it can simultaneously produce very deep surveys and cover large areas. One would expect this to be ideal for constraining both  $\epsilon_{\text{GR}}$  and  $f_{\text{NL}}$ . Relative to the other techniques, intensity mapping suffers from the fact that one of the substantial corrections on large scales—the perturbation to angular distances—is absent. As was shown in Section 2.1, because one is measuring an intensity rather than source number density, there is an exact cancellation of the lensing contributions to both the number density and angular diameter distance corrections. This has a significant impact on the size of the relativistic correction signal and thus the detectability of  $\epsilon_{\text{GR}}$ . Intensity mapping can be used to obtain reasonably tight constraints on  $f_{\text{NL}}$ , although there too it is placed at a disadvantage by the “Gaussian” clustering bias,  $b^G$ , for neutral hydrogen being smaller than for other types of probe. Note that while most halo-based models predict a low H I bias, measurements of the clustering of damped Ly $\alpha$  systems carried out by Font-Ribera et al. (2012) point toward a significantly higher value at  $z \gtrsim 2$ . If we multiply our fiducial bias function by a factor of  $1.5\times$  (so that  $b \approx 2$  at  $z = 2.2$ ), we obtain a significantly better constraint of  $\sigma(f_{\text{NL}}) = 0.9$  for an IM survey going out to  $z = 3.5$ . Our forecasts for  $f_{\text{NL}}$  are very sensitive to the fiducial bias model, and therefore could change significantly with better empirical measurements.

We must also point out that the survey specifications we assumed are such that the measurements of the power spectra are only cosmic variance-limited on the very largest scales. It might be possible to improve the constraints on  $f_{\text{NL}}$  (by up to 30% in the most optimistic case) by reducing the noise (e.g., by increasing the survey time).

2. *Radio continuum surveys*, while efficient at accessing large volumes, are remarkably poor at constraining both  $\epsilon_{\text{GR}}$  and  $f_{\text{NL}}$  in a single-tracer context. The loss of all radial information through projection over a wide redshift range has a significant toll on their ability to discriminate

<sup>10</sup> See, e.g., Challinor & Lewis (2011), Jeong et al. (2012), and Yoo et al. (2012) for similar statements using less quantitative analyses.



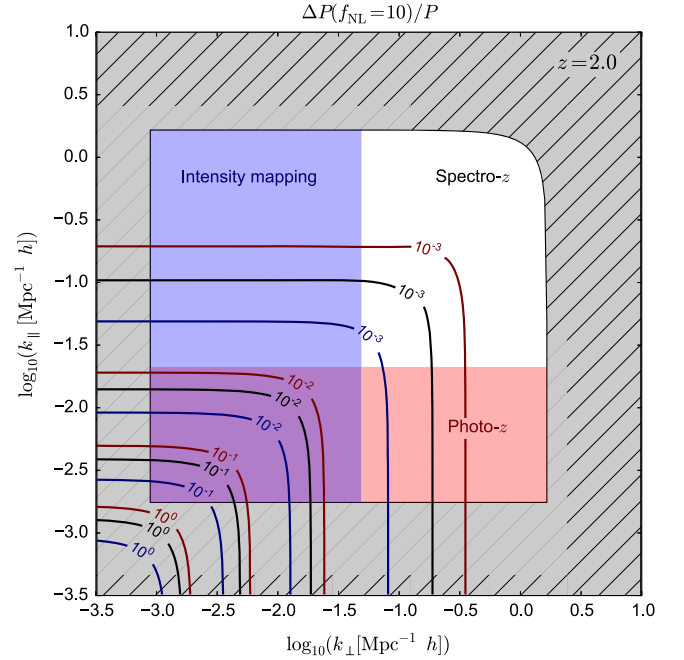
**Figure 14.** Comoving volume and redshift range covered by the experiments considered in this work.

between different effects in the power spectrum. If we are to benefit from continuum surveys, a more sophisticated multi-tracer approach is needed, as we will discuss below.

3. We also looked at the ability of a *spectroscopic survey* to constrain large-scale features in the power spectra. The advantage of this type of survey is ostensibly their ability to obtain high-resolution measurements in both the angular and radial directions—more so than in any other type of survey considered here, although in practice this comes at a price. Spectroscopic surveys are time-intensive, and so surveys are limited to smaller areas of the sky and shallower depths. The average number density of (usable) galaxies is also lower than for other surveys, and hence the effective noise on large angular scales can be more substantial. As was the case with intensity mapping, our forecasts for  $f_{\text{NL}}$  depend crucially on the fiducial clustering bias, which for H $\alpha$  emitters is also relatively low.
4. Deep and wide *photometric surveys* seem to be the most effective method for probing large scales, for the parameters we considered here. The loss of resolution along the radial direction (as compared to a spectroscopic survey) is compensated by the significantly wider redshift coverage and larger survey area. Photometric surveys are also a more rapid way of counting galaxies, so source number densities are higher than for contemporary spectroscopic surveys. Finally, the clustering bias for the sources that LSST will observe is significantly higher than for H I and H $\alpha$  galaxies, boosting its ability to detect  $f_{\text{NL}}$  substantially.

Figure 14 compares the redshift range and comoving volume probed by the various experiments studied here. Measuring ultra-large-scale observables depends critically on the ability to cover very large volumes, but we have seen here that this is far from the only factor. Even though an SKA1 continuum survey should be able to access the largest volume of any of the surveys, its inability to use radial information prevents it from achieving a competitive measurement of  $f_{\text{NL}}$  without the use of the multi-tracer technique.

It is therefore also relevant to compare the range of radial and angular scales covered by each experiment. We do this in

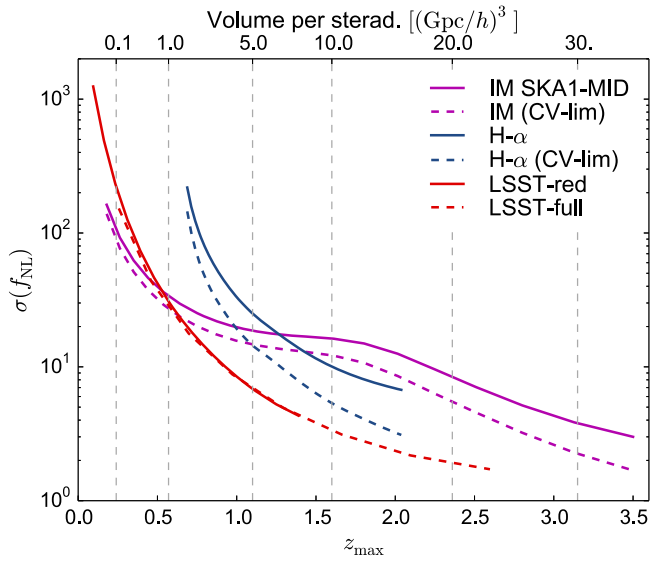


**Figure 15.** Regions in the space of transverse and parallel wavenumbers ( $k_{\perp}$ ,  $k_{\parallel}$ ) accessible for various surveys. Red: region accessible to a photometric redshift survey with  $\sigma_0 = 0.05$ . The large  $k_{\parallel}$ -regime is lost due to the inaccuracy of photo-zs. Blue: region accessible by an SKA1-MID IM survey in single-dish mode ( $D_{\text{dish}} = 15$  m). Small transverse scales are inaccessible due to the beam width. White (containing red and blue regions): region accessible to a full-sky spectroscopic survey covering up to  $z = 2$ . Contours: the relative amplitude of the contribution to the three-dimensional power spectrum due to primordial non-Gaussianity, with  $f_{\text{NL}} = 10$ , for these three experiments, in black (spectroscopic), red (photometric), and blue (intensity mapping). Note that this amplitude depends on the clustering bias of each tracer.

Figure 15 for intensity mapping, photometric, and spectroscopic surveys covering redshifts  $z < 2$ . The gray hashed region corresponds to inaccessible scales, either larger than the survey volume or smaller than the nonlinear scale. While spectroscopic surveys have essentially complete access to the  $k_{\parallel} - k_{\perp}$  plane, photometric surveys and intensity mapping are limited to only large radial or angular scales respectively.

This is not necessarily an important inconvenience for ultra-large-scale observables. The colored contours show the relative amplitude of the  $f_{\text{NL}}$  signal for each of these surveys (spectroscopic in black, photometric in red, and IM in blue) for  $f_{\text{NL}} = 10$ . In the large- $k$  region that only spectroscopic surveys have access to, the amplitude of the signal is relatively low. Note also that radial and angular scales should not be treated equally in this plot; simply by a dimensionality argument (two angular dimensions versus one radial), better angular coverage will usually be more advantageous.

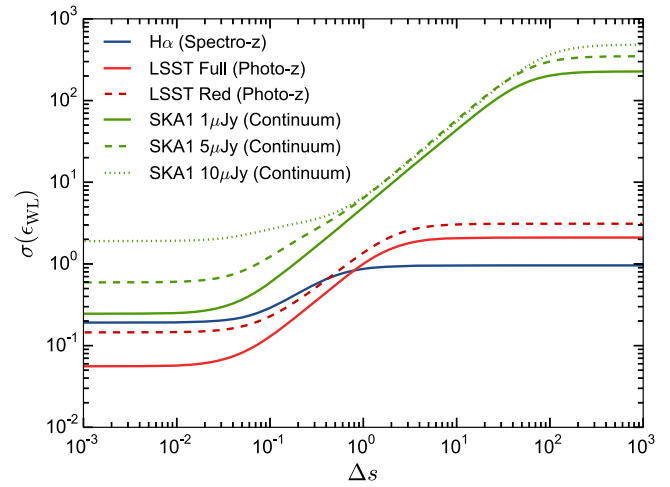
Finally, our ability to measure any signal depends critically on its amplitude. In the case of primordial non-Gaussianity, this depends on the sample's clustering bias, since  $\Delta b_{\text{NG}} \propto (b - 1)$ . This dependence can be observed in Figure 16, where we show the forecasted uncertainty on  $f_{\text{NL}}$  as a function of the maximum redshift covered by each survey. Even though intensity mapping has the potential to cover the largest volume while also preserving radial information, the improvement of  $\sigma(f_{\text{NL}})$  with  $z$  is significantly slower, in particular in the range  $z \sim 1$ – $1.5$ , where the H I bias is very close to unity.



**Figure 16.** Constraints on  $f_{\text{NL}}$  as a function of the maximum redshift covered by each of the surveys considered here. The comoving volume per steradian at each  $z$  is indicated on the upper x-axis.

We emphasize that all of our forecasts depend on a variety of astrophysical model assumptions. By this we mean that we have had to predict the number densities and biases for the surveys as a function of redshift, based on existing data and simulations. Getting these assumptions correct is key for accurate forecasting, and substantial work will have to be done—for all types of surveys—to better estimate these quantities. Nevertheless, we are confident that our models are sufficiently representative that our broad point is correct, and that the numbers we present here give a fair representation of what to expect from future surveys. On this point it is also worth noting that it is in principle possible to further constrain the level of primordial non-Gaussianity from measurements of higher-order correlations. Such measurements are extremely challenging at present due to the large theoretical uncertainties (e.g., in the form of the bias in the presence of primordial non-Gaussianity, the form of the bispectrum and its covariance in redshift space, the value of the nonlinear clustering bias etc.). Further studies of the three-point function might eventually make such measurements possible, possibly superseding the forecasts presented here.

Conservatively, we did not include the lensing magnification contribution to the power spectrum as one of the effects parameterized by  $\epsilon_{\text{GR}}$ , even though it is a fully relativistic effect. This is because magnification has a significant amplitude on sub-horizon scales, and has in fact already been detected by cross-correlating pairs of distant tracers. This is qualitatively different to the situation for the other relativistic terms, which are significant only on ultra-large scales, and which have not been measured yet in large-scale structure. Figure 1 indicates that the lensing magnification contribution to the power spectrum can dominate the GR terms parameterized by  $\epsilon_{\text{GR}}$  at ultra-large scales, however. This large-scale contribution has not been detected in current small-volume surveys, but previous work has forecasted its detectability (Yoo 2009; Namikawa et al. 2011; Yang & Zhang 2011; Montanari & Durrer 2015; Yang et al. 2015). By defining a parameter,  $\epsilon_{\text{WL}}$ , corresponding to the amplitude of the magnification lensing term alone, we have quantitatively verified its detectability,



**Figure 17.** Constraints on  $\epsilon_{\text{WL}}$  as a function of the prior on the magnification bias. The same parameters, and priors for the other bias functions, were used as in the  $f_{\text{NL}}$  analysis, except for the continuum survey, where the clustering bias was not marginalized.

showing that the large-scale lensing contribution should be detectable above  $5\sigma$  in all relevant experiments. It has also been shown by Lopez-Honorez et al. (2012), Namikawa et al. (2011), and Camera et al. (2015a) that omission of the lensing + GR terms leads to a bias in the recovered value of  $f_{\text{NL}}$ . This was predicted to be at the  $\sim 3\sigma$  level for an SKA Phase 2 H I spectroscopic survey (Camera et al. 2015a). Our analysis shows that  $\epsilon_{\text{GR}}$  cannot be responsible for this bias, leading to the conclusion that it must be almost exclusively due to the ultra-large-scale lensing term.

We also tested the consistency of our results with the recent work by Montanari & Durrer (2015), who find  $\sim 1\%$  constraints on  $\epsilon_{\text{WL}}$  (their parameter  $\beta$ ) for a *Euclid* photometric survey. Marginalizing over the same set of cosmological parameters as them (except the neutrino mass,  $m_\nu$ ), and choosing  $\ell_{\text{max}} = 1000$ , we obtain  $\sigma(\epsilon_{\text{WL}}) = 0.06$  for the H $\alpha$  spectroscopic survey with  $\Delta z = 0.1$  bins. This is consistent with their results to within a factor of a few, which is reasonable given the differences in survey specifications, fiducial magnification bias functions, and maximum  $\ell$  used.

As a further test we then repeated this analysis, but restricted ourselves to only large-scale modes,  $\ell \leq 100$ , and marginalized over the bias nuisance parameters as well. While one would likely fix the bias functions to their fiducial forms when attempting a first detection of the large-scale lensing effect, the bias uncertainties must be taken into account for precision measurements of  $\epsilon_{\text{WL}}$  (e.g., if used to test GR, as discussed in Montanari & Durrer 2015). We find a strong correlation between  $\epsilon_{\text{WL}}$  and the magnification bias, especially at higher redshifts, which is to be expected given that both factors multiply the lensing term in the expression for the number count. Figure 17 shows how the  $\epsilon_{\text{WL}}$  constraint depends on the magnification bias prior for the various surveys; the spectroscopic and photometric surveys both require  $\Delta s \lesssim 1$  to measure  $\epsilon_{\text{WL}}$  to better than 100%, and  $\Delta s$  of order a few  $\times 10^{-2}$  to reach their optimal constraints. Continuum surveys require a stronger prior of  $\sim 0.1$  just to reach a 100% constraint on  $\epsilon_{\text{WL}}$  (for 1 and 5  $\mu\text{Jy}$  flux limits), and are also subject to strong correlations with other bias parameters, as discussed in Section 4.2.2. The continuum constraints shown in Figure 17 were derived for fixed clustering bias, but if this is marginalized



over (with a 10% relative prior on  $b$ ),  $\sigma(\epsilon_{\text{WL}})$  increases by a factor of  $\sim 2$ .

The fact that we are unable to detect  $\epsilon_{\text{GR}}$  from the single-tracer power spectrum should not at all lead us to give up hope of seeing the GR corrections, however. Indeed, this is just the first step in identifying the most effective observables for teasing out the ultra-large-scale effects. In order to beat down the cosmic variance that is a fundamental barrier to single-tracer detectability, one must use multi-tracer techniques. In some sense these divide out the stochastic part of the perturbation field, avoiding the effects of cosmic variance for certain (non-stochastic) quantities. By cross-correlating different tracers, with different bias functions, it is possible to isolate a number of terms from the scale dependence due to non-Gaussianity, relativistic effects, and growth of structure, in such a way as to obtain much tighter constraints than those from the overall power spectrum alone. It has been shown that multi-tracer techniques applied to continuum surveys can lead to an improvement by almost an order of magnitude in the detection of  $f_{\text{NL}}$ , for example. We expect that the GR effects will also be detectable via this approach (Yoo et al. 2012). A systematic analysis of multi-tracer techniques in the surveys described here is, therefore, an obvious next step in trying to identify robust methods for measuring ultra-large-scale effects.

*Note added.* While this paper was being finalized, Raccanelli et al. (2015b) appeared, which discusses some related topics.

We would like to thank Rachel Bean, Elisa Chisari, Enea Di Dio, Ruth Durrer, Matt Jarvis, Lance Miller, Francesco Montanari, and Eva-Maria Müller for their very valuable input on this paper. D.A. is supported by ERC grant 259505. P.B. is supported by ERC grant StG2010-257080. P.G.F. acknowledges support from STFC, BIPAC, and the Oxford Martin School. R.M. and M.S. are supported by the South African Square Kilometre Array Project and the South African National Research Foundation. R.M. is also supported by the UK Science & Technology Facilities Council, Grant No. ST/K0090X/1.

## APPENDIX A MODIFICATIONS TO CLASS

In order to compute the power spectra used for the Fisher forecasts, we used the public code CLASS (Lesgourgues 2011). In its current version, CLASS encompasses the extension CLASSgal (Di Dio et al. 2013), which can be used to compute the transfer functions  $\Delta^i$  in Equations (39)–(44). The public version of the code is easy to install, run, and modify, and we encourage its use; however, a number of modifications had to be implemented in order to make it usable for our work. We have made our modified version of the code publicly available at <http://intensitymapping.physics.ox.ac.uk/codes.html>, and document the main changes here:

1. The terms  $\Delta^L$ ,  $\Delta^{P4}$ , and  $\Delta^{\text{ISW}}$  in Equations (39)–(44) can be computationally very demanding to compute. The main reason for this is the wide redshift range covered by their window functions—these terms correspond to integrated effects along the photon trajectory from the source to the observer. In the current version of CLASS, including these terms in the calculation of  $C_{\ell}^{ij}$  for a single redshift bin at  $z = 0.5$  takes about 10 minutes running on a modern four-core computer. This makes including these

terms for the large number of bins used in this project (e.g., 100 bins for intensity mapping) prohibitively expensive, so we invested some time in speeding up the calculation of these terms.

Two modifications were implemented. First of all, the code was parallelized for distributed memory machines using MPI, so that each node computes the transfer functions of a different set of redshift bins. Second, some redundant calculations were circumvented by precomputing the window functions  $W^L(\eta)$ ,  $W^{P4}(\eta)$ , and  $W^{\text{ISW}}(\eta)$  in Equation (45) and storing them in memory. This last modification speeds up the computation of the integrated terms by a factor of  $\sim 4$ , although it requires more memory.

2. We implemented the effect of primordial non-Gaussianity by including the scale-dependent contribution to the bias in Equation (47).
3. We modified the input/output system for the bias parameters  $b(z)$ ,  $s(z)$ , and  $f_{\text{evo}}(z)$ . These can now be supplied as tabulated  $z$ -dependent functions. Furthermore,  $f_{\text{evo}}$  must now be provided separately from the redshift distribution,  $\tilde{N}(z)$ .
4. We implemented the possibility of adding an extra parameter for each redshift bin that corresponds to the photo- $z$  uncertainty, so that the window function in each bin can be computed as in Equation (82). In doing this, we also modified the input/output system for defining the redshift bins. The bin properties must now be supplied in different columns in a separate text file. We believe this system is better suited for a large number of redshift bins.
5. Finally, we included the extra parameter  $\epsilon_{\text{GR}}$ , used in this paper to parameterize the amplitude of the relativistic corrections. This is not a general-purpose modification.

## APPENDIX B SURVEY SPECIFICATIONS

In this appendix we provide the detailed specifications for all four of our reference surveys. The codes used to generate the fiducial redshift distributions and bias functions have been made available online.<sup>11</sup>

### B.1. Intensity Mapping

#### B.1.1. Noise Model and Redshift Binning

For an IM autocorrelation experiment, the simplest case is to assume that the noise is uncorrelated between different frequency channels, and has a white noise power spectrum:

$$N_{\ell}^{ij} = \delta_{ij} \sigma_{\text{sr}}^2. \quad (57)$$

Here  $\sigma_{\text{sr}}^2$  is the noise variance per steradian, and can be calculated as follows: the noise per pointing can be estimated as the rms temperature fluctuation of the system,  $T_{\text{sys}}$ , scaled by the number of independent samples measured (given by  $\delta\nu t_p$ , where  $\delta\nu$  is the frequency channel width and  $t_p$  is the integration time per pointing).  $t_p$  can be approximated by  $t_{\text{tot}} \Delta\Omega / (4\pi f_{\text{sky}})$ , where  $f_{\text{sky}}$  is the surveyed fraction of the sky,  $\Delta\Omega$  is the solid angle covered in each pointing, and  $t_{\text{tot}}$  is the total survey time. Finally, scaling this by the total number of

<sup>11</sup> <http://intensitymapping.physics.ox.ac.uk/codes.html>

dishes in the experiment, we obtain the power spectrum

$$N_{\ell}^{ij} = \delta_{ij} \frac{T_{\text{sys}}^2(\nu_i) 4\pi f_{\text{sky}}}{\delta\nu t_{\text{tot}} N_{\text{dish}}}. \quad (58)$$

Note that the variance per pointing has been multiplied by  $\Delta\Omega$  to obtain the variance per steradian, which cancels the dependence on  $\Delta\Omega$ . The system temperature receives two contributions,  $T_{\text{sys}} = T_{\text{sky}} + T_{\text{inst}}$ , due to atmospheric and background radio emission ( $T_{\text{sky}} \simeq 60 \text{ K} \times (\nu/300 \text{ MHz})^{-2.5}$ ) and instrumental noise ( $T_{\text{inst}}$ ).

Finally, it is worth noting that two different conventions have been adopted in the literature regarding the effect of the beam that defines the angular resolution of the experiment. The difference is in interpreting the beam as smoothing the signal on scales beyond the resolution, or as enhancing the noise at those same scales. We use the former, so that the model for the total observed power spectrum is

$$C_{\ell}^{ij} = C_{\ell}^{S,ij} B_{\ell}^i B_{\ell}^j + N_{\ell}^{ij}, \quad (59)$$

where  $B_{\ell}^i$  is the harmonic transform of the instrumental beam in the  $i$ th frequency bin. We have assumed that the beams are Gaussian,  $B_{\ell}^i = \exp(-\ell(\ell+1)\theta_B^2/2)$ , where  $\theta_B$  is related to the beam FWHM through  $\theta_{\text{FWHM}} = 2\sqrt{2 \ln 2} \theta_B$ . The beam width can be related to the dish diameter approximately as  $\theta_{\text{FWHM}} \simeq c/(\nu D_{\text{dish}})$ .

In our forecasts for SKA1-MID, we used the instrumental parameters  $T_{\text{inst}} = 25 \text{ K}$ ,  $f_{\text{sky}} = 0.75$ ,  $t_{\text{total}} = 10^4 \text{ h}$ ,  $D_{\text{dish}} = 15 \text{ m}$ , and  $N_{\text{dish}} = 254$ , and assumed a minimum frequency of 350 MHz, corresponding to  $z_{\text{max}} \simeq 3$ .

Modern radio receivers have very high frequency resolution (e.g.,  $\delta\nu \sim 0.1 \text{ MHz}$ ), so H I intensity mapping experiments should be able to resolve radial structures on scales much smaller than those relevant for cosmology. We are therefore free to choose the width and shape of the redshift bins used for the cosmological analysis. In order to avoid inhomogeneous coverage of radial scales, we divide the total frequency band into frequency bins of varying width  $\Delta\nu(\nu)$  such that the corresponding comoving size  $\Delta\chi$  is held constant. We estimated the minimum number of frequency bins needed for the constraints on  $f_{\text{NL}}$  and  $\epsilon_{\text{GR}}$  to converge, finding that at least 100 bins were necessary. This corresponds to a radial width of  $\Delta\chi \simeq 44 \text{ Mpc}/h$ .

### B.1.2. Nuisance Parameters and Redshift Evolution

We model the clustering, magnification, and evolution biases for H I using an approach based on the halo model. We first assume that a one-to-one relationship exists between halo mass and H I mass,  $M_{\text{HI}} = M_{\text{HI}}(M, z)$ . The density and clustering bias can then be computed as

$$\rho_{\text{HI}}(z) = \int_{M_{\text{min}}}^{M_{\text{max}}} dM n(M, z) M_{\text{HI}}(M, z), \quad (60)$$

$$b_{\text{HI}}(z) = \int_{M_{\text{min}}}^{M_{\text{max}}} dM n(M, z) b(M, z) \frac{M_{\text{HI}}(M, z)}{\rho_{\text{HI}}(z)}, \quad (61)$$

where  $n(M, z)$  is the halo mass function (comoving number density per unit mass) and  $b(M, z)$  is the halo bias. The background brightness temperature can then be computed in terms of  $\rho_{\text{HI}}$  using Equation (26).

As described in Section 2.1.2, the transverse distance perturbations cancel out for intensity mapping, so that  $s_{\text{HI}}(z) = 2/5$  exactly. Furthermore, since we observe the emission from all of the H I in each patch of the sky, the evolution bias can be computed directly by differentiating  $\rho_{\text{HI}}$  with respect to  $z$ ,

$$f_{\text{evo}} = -\frac{d \log[\rho_{\text{HI}}(z)]}{d \log(1+z)}. \quad (62)$$

All that remains is to specify the function  $M_{\text{HI}}(M, z)$ . As in Bull et al. (2015), we assume a power-law relation  $M_{\text{HI}}(M, z) \propto M^{\alpha}$  with an exponent  $\alpha \simeq 0.6$ , and with the normalization set by constraints on  $\Omega_{\text{HI}}$  at  $z = 0.8$  from Switzer et al. (2013).

## B.2. Radio Continuum Surveys

### B.2.1. Number Counts, Bias Functions, and Noise Model

The models for the signal and noise power spectra are very simple compared with the other probes, as the galaxy sample is distributed in a single redshift bin, with a window function given by the redshift distribution,

$$W(z) \propto \bar{N}(z). \quad (63)$$

We estimate  $\bar{N}(z)$  for the radio sources from empirical estimates of the luminosity functions of the main radio populations, since these also contain the necessary information to estimate the magnification and evolution biases. We consider four main radio galaxy types: star-forming galaxies (SF), starbursts (SB), radio-quiet quasars (RQQ), and Faranoff–Riley type I active galactic nuclei (AGNs) (FRI).<sup>12</sup> The luminosity functions for each population were computed following the prescriptions of Wilman et al. (2008).

We will now outline the procedure used to calculate the redshift distribution,  $s$ , and  $f_{\text{evo}}$  in all cases, and refer the reader to Wilman et al. (2008) and the references Yun et al. (2001) (SF and SB), Ueda et al. (2003) (RQQ), and Willott et al. (2001) (FRI) for details on the observations that the luminosity functions are based on. These details are also summarized in Appendix C. As with optical and IR surveys (see Sections 4.3 and 4.4), the  $k$ -correction to the flux measured in a given band is also needed in order to accurately estimate the observed number counts. This can be done for radio sources by assuming a particular SED for each population. For this, we again used the models from Wilman et al. (2008).

In its rest frame, a radio source has a luminosity per unit frequency given by

$$L_{\nu} \equiv \frac{dE_e}{dt_e d\nu_e} = L_{\nu_*} \frac{\varphi(\nu)}{\varphi(\nu_*)}, \quad (64)$$

where  $\nu_*$  is a pivot frequency and  $\varphi(\nu)$  is the source SED. This is related to the flux per unit frequency measured by the observer by

$$S_{\nu} \equiv \frac{dE_o}{dt_o d\nu_o dA_o} = \frac{L_{\nu(1+z)}}{4\pi\chi^2(z)(1+z)} (1-2\delta_{\perp}). \quad (65)$$

<sup>12</sup> We also considered FRII galaxies, but their number density is so low that they contribute negligibly to the total number counts.

Consider a radio survey in a given frequency band ( $\nu \in [\nu_0, \nu_f]$ ). The average flux density measured for one source is defined as

$$\bar{S}(\nu_0, \nu_f) = \int_{\nu_0}^{\nu_f} S_\nu \frac{d\nu}{\nu_f - \nu_0}. \quad (66)$$

A given source will be detected if its average flux is above the detection limit  $S_{\text{cut}}$ , and therefore all sources with a pivot luminosity above a minimum value  $L_{\nu_*}^{\text{cut}} \equiv \bar{L}_{\nu_*}^{\text{cut}}(1 + 2\delta_\perp)$  will be included in the sample, where the average threshold luminosity is

$$\bar{L}_{\nu_*}^{\text{cut}}(z, S_{\text{cut}}) = \frac{4\pi S_{\text{cut}} \chi^2(z)(1+z)\varphi(\nu_*)}{\bar{\varphi}(\nu_0(1+z), \nu_f(1+z))}, \quad (67)$$

and the average SED in the observed band is

$$\bar{\varphi}(\nu_1, \nu_2) \equiv \int_{\nu_1}^{\nu_2} \varphi(\nu) \frac{d\nu}{\nu_2 - \nu_1}. \quad (68)$$

Given a pivot frequency  $\nu_*$ , the luminosity function  $\bar{n}_s(z, \ln L_{\nu_*})$  at that frequency, and a characteristic SED  $\varphi(\nu)$ , the redshift distribution of sources can be computed as

$$\bar{N}(z) = \frac{c \chi^2(z)}{(1+z)^3 H(z)} \bar{N}(z, > \bar{L}_{\nu_*}^{\text{cut}}), \quad (69)$$

where  $\bar{N}$  is defined as in Equation (8).  $s(z)$  and  $f_{\text{evo}}(z)$  are then calculated from  $\bar{N}$  using Equations (28) and (29).

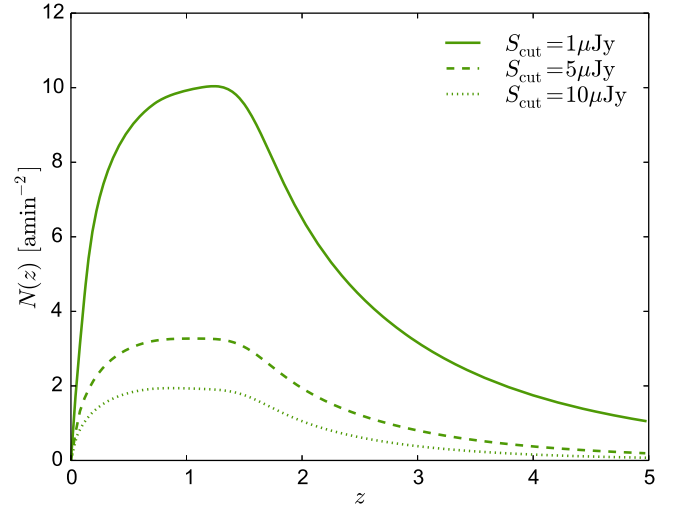
For the clustering bias, we follow the same approach used in Wilman et al. (2008) and assign a fixed halo mass to each population. The corresponding bias is then found as the halo-model bias for that mass as a function of redshift. For this we parameterize the halo bias as in Sheth & Tormen (1999). We have also explored the approach followed in Ferramacho et al. (2014), where each population is given a distribution of halo masses rather than a fixed one, and the bias is found by averaging over that distribution. No significant differences were found between the approaches, so we use the first, simpler one.

Once the redshift distribution and bias functions have been calculated for each population, we compute them for the combined sample as a weighed average of the individual ones,

$$\bar{N}_{\text{tot}}(z) = \sum_a \bar{N}_a(z), \quad (70)$$

$$(b, s, f_{\text{evo}})_{\text{tot}}(z) = \sum_a (b, s, f_{\text{evo}})_a(z) \frac{\bar{N}_a(z)}{\bar{N}_{\text{tot}}(z)}, \quad (71)$$

where  $a$  labels the population. Separating the different populations is a very costly observational task. For our purposes, the main benefit of doing this is to allow the use of the multi-tracer technique to circumvent cosmic variance. Since we have postponed the multi-tracer analysis for future work, we will only report our forecasts here for the combined sample of radio sources. Figure 18 shows the redshift distribution for the combined sample for the three different detection limits considered here. The luminosity functions, redshift distributions, and bias can be obtained using a code that we have made publicly available.<sup>13</sup> This provides an easy



**Figure 18.** Angular number density of radio galaxies as a function of redshift for the combined continuum sample considered here, and for different flux cuts.

way to obtain number counts and power spectra without needing to query the full simulation of Wilman et al. (2008). Our results are consistent with the simulation except for the total number counts of star-forming galaxies, where our numbers are a factor of  $2.5 \times$  higher. This is consistent with what is described in Jarvis et al. (2015) and references therein, however.

The most relevant source of noise in a clustering analysis of discrete sources is shot noise, determined by the number density of sources. The angular number density of radio sources in the sample,  $\bar{N}_\Omega$ , is determined by integrating the redshift distribution

$$\bar{N}_\Omega \equiv \int_0^\infty dz \bar{N}(z), \quad (72)$$

and the noise power spectrum is given by

$$N_\ell = \bar{N}_\Omega^{-1}. \quad (73)$$

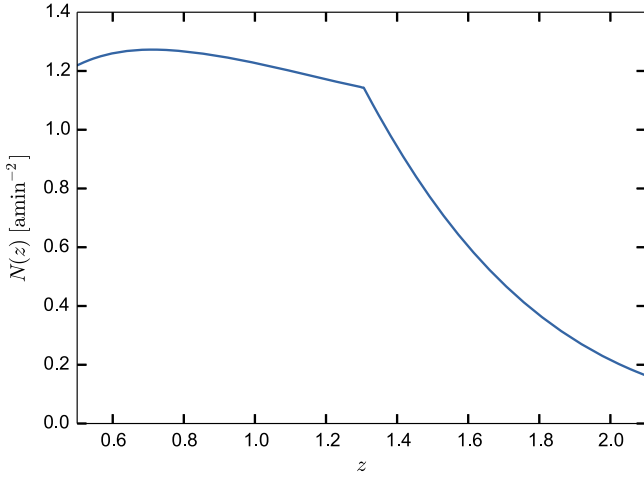
### B.3. Spectroscopic Redshift Surveys

#### B.3.1. Noise and Signal Model

Spectroscopic galaxy surveys measure Equation (11), the perturbation to flux-limited number counts  $\Delta_N(z, \hat{n}, > \log F_{\text{cut}})$ . This depends on the clustering, magnification, and evolution bias functions for the source population. The latter two can be derived from the background luminosity function of the sources,  $n_s(z, \log L)$ , given the flux limit and efficiency of the survey. Number counts are subject to a Poisson shot noise term that depends on the number density of sources, which can also be obtained from the luminosity function.

We follow the current set of public specifications for a realistic H $\alpha$  spectroscopic survey, described in Amendola et al. (2013). The number counts presented in Table 1.3 of that paper can be approximately reproduced by using the best-fit H $\alpha$

<sup>13</sup> <http://intensitymapping.physics.ox.ac.uk/codes.html>



**Figure 19.** Predicted number density of galaxies as a function of redshift for the H $\alpha$  spectroscopic survey.

Schechter luminosity function found by Geach et al. (2010),

$$n(z, L > L_{\text{cut}}) = \int_{x_{\text{cut}}}^{\infty} \epsilon \varphi^* x^{\alpha} e^{-x} dx;$$

$$x \equiv L / L^*(z),$$

$$L^*(z) = \begin{cases} 5.1 \times 10^{41} (1+z)^{3.1} & (z < 1.3) \\ 6.8 \times 10^{42} & (1.3 < z < 2.2) \end{cases} \text{erg s}^{-1},$$

where  $L$  is the H $\alpha$  line luminosity and we have assumed a flux limit of  $F_{\text{cut}} = 3 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ , an efficiency of  $\epsilon = 0.45$ , a faint-end slope of  $\alpha = -1.35$ , and comoving number density normalization  $\phi^* = 1.37 \times 10^{-3} \text{ Mpc}^{-3}$ . The predicted galaxy number density as a function of redshift is shown in Figure 19. There are significant uncertainties in this model, which we account for by marginalizing over a set of bias function nuisance parameters, as explained in Section 3. Finally, for the clustering bias we use the simplified prescription from Amendola et al. (2013),  $b(z) = \sqrt{1+z}$ , which we also subject to the nuisance parameterization.

The assumed bias functions for the H $\alpha$  survey were shown in Figure 2. The evolution bias is large and negative, and grows more negative with redshift until  $z = 1.3$ , where there is a discontinuity in the  $L^*(z)$  model, beyond which  $L^*(z) = \text{const.}$  and so  $f_{\text{evo}} = 0$ . This is not particularly realistic, but is the best that can be done until updated constraints on the high-redshift H $\alpha$  luminosity function become available. The magnification bias grows rapidly with redshift, and deviates significantly from  $s = 0.4$  over most of the range, meaning that there is little or no cancellation of the transverse scale perturbation as there was in the intensity mapping survey (although this is contingent on the uncertain behavior of the luminosity function model for  $z > 1.3$ ). The clustering bias is relatively close to unity across the entire redshift range, slowly evolving from a minimum of  $b \approx 1.3$  at  $z = 0.65$  to a maximum of  $b \approx 1.8$  at  $z \approx 2$ .

We consider three constant-width redshift binnings for the selection function over the interval  $0.65 \leq z \leq 2.05$ :  $\Delta z = (0.025, 0.05, 0.10)$ , yielding (56, 28, 14) bins respectively. The target redshift uncertainty for, e.g., *Euclid* is  $\sigma_z \leq 0.001 (1+z)$  (Laureijs et al. 2011), which is always significantly smaller than the narrowest bin width. As such, we assume a uniform (top-hat)

selection function, weighted by the source redshift distribution,  $\bar{N}(z)$ .

Our noise model assumes that only shot noise is relevant, and that effects such as spectroscopy failures and point source masking have been taken into account in the survey efficiency,  $\epsilon$ . The shot noise angular power spectrum is

$$N_{\ell}^{ij} = \delta_{ij} / n_i; \quad n_i \equiv \int_{z_i} \bar{N}(z) dz. \quad (74)$$

with  $n_i$  measured in units of  $\text{sr}^{-1}$ . In our forecasts, we assumed that a wide range of multipoles can be recovered,  $2 \leq \ell \leq 1000$ , with no cuts at high or low  $\ell$  due to systematics, nonlinear effects, and so on (this assumption was relaxed in Section 4.3.3).

## B.4. Photometric Redshift Surveys

### B.4.1. Redshift Distribution and Bias Parameters

In order to compute  $\bar{N}(z)$ ,  $s(z)$ , and  $f_{\text{evo}}(z)$  for our two samples, we need an estimate of the luminosity function for both red and blue galaxies, preferably in the  $r$ -band, for which the LSST specifications are provided. We describe the method used for this task here.

For the red sample we follow a method similar to that used by Joachimi et al. (2011). First, an estimate of the  $B$ -band luminosity function for red galaxies is obtained from Faber et al. (2007) as a Schechter function with constant slope  $\alpha = -0.5$ , and  $z$ -dependent  $\phi_*$  and  $M_*$ , measured in a number of redshift bins in the interval  $z \in (0.2, 1.2)$ . We extrapolate the luminosity function to higher/lower redshifts by fitting the values of these parameters, measured by Faber et al. (2007), to the models

$$M_*(z) = M_0 + M_1 z, \quad (75)$$

$$\phi_*(z) = \frac{\phi_0}{1 + (z/z_0)^a} [10^{-3} \text{ Mpc}^{-3}] \quad (76)$$

with  $M_0 = -20.6$ ,  $M_1 = -0.49$ ,  $\phi_0 = 1.82$ ,  $z_0 = 1.04$ , and  $a = 7.17$ . In order to translate this into an  $r$ -band luminosity function, we use  $B - r = 1.32$  (Fukugita et al. 1995) and assume that  $B - r$  does not evolve significantly for the red sample in the redshift range under study.

For the full sample, we use the  $r'$ -band luminosity function found by Gabasch et al. (2006) and approximate  $r' \simeq r$ . This is again given as a Schechter function with constant slope  $\alpha = -1.33$  and  $z$ -dependent  $M_*$  and  $\phi_*$ , for which we have used the following parameterizations:

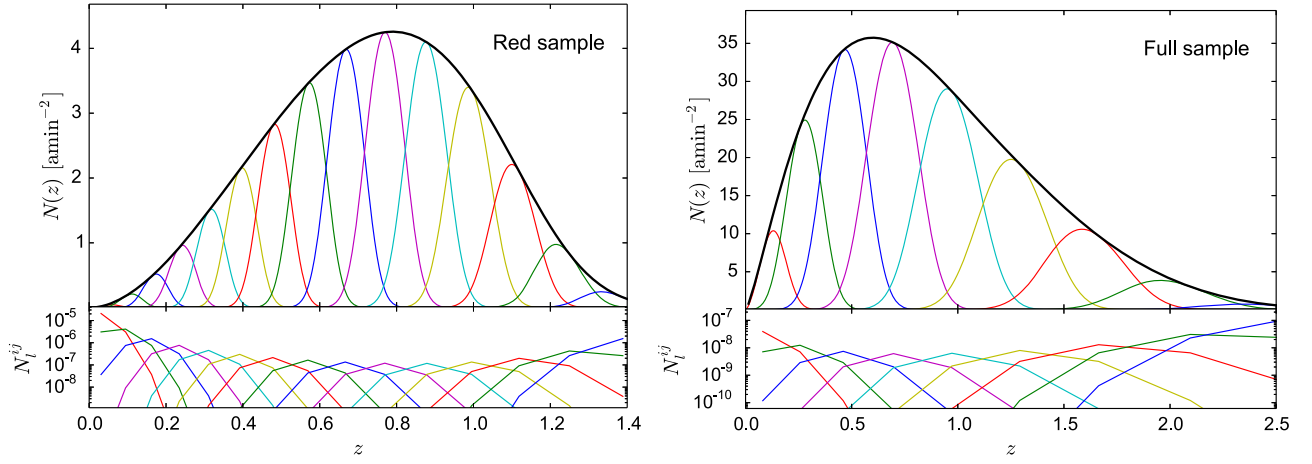
$$M_*(z) = M_0 + a \ln(1+z), \quad (77)$$

$$\phi_*(z) = (\phi_0 + \phi_1 z + \phi_2 z^2) [10^{-3} \text{ Mpc}^{-3}] \quad (78)$$

with  $M_0 = -21.49$ ,  $a = -1.25$ ,  $\phi_0 = 2.59$ ,  $\phi_1 = -0.136$ , and  $\phi_2 = -0.081$ . The luminosity function for blue galaxies is then estimated as the difference between those of the full and red samples.

An absolute magnitude,  $M$ , measured in a given rest-frame band for a galaxy at redshift  $z$  is related to the apparent





**Figure 20.** Angular number density of galaxies as a function of redshift for the LSST “red” (left plot) and “full” (right plot) samples. The colored lines show the window functions of the different redshift bins considered here.

magnitude in the observer-frame band,  $m$ , by

$$M = m - 25 - 5 \log_{10} \left[ \frac{d_L(z)}{1 \text{ Mpc } h^{-1}} \right] + \log_{10} h - k(z), \quad (79)$$

where  $d_L(z)$  is the luminosity distance and  $k(z)$  is the  $k$ -correction corresponding to that galaxy’s SED redshifted to  $z$ . We estimated  $k(z)$  for red and blue galaxies by running the code `kcorrect` (Blanton & Roweis 2007) on the spectra of an elliptical galaxy and a barred spiral galaxy (Sbc) respectively, as measured by Coleman et al. (1980). We approximate and extrapolate these  $k$ -corrections as  $k_{\text{red}}(z) \sim 2.5z$  and  $k_{\text{blue}} \sim 1.5z$ . We verified that these parameterizations are compatible with the  $k$ -corrections shown in Fukugita et al. (1995) for both types, and also that the final redshift distributions did not vary significantly when other functional forms are used. For a given magnitude limit  $m_{\text{cut}}$ , we use Equation (79) to compute the corresponding luminosity cut  $\bar{L}_{\text{cut}}$  as a function of redshift. The redshift distribution, magnification bias, and evolution bias for each population are then estimated using Equations (28)–(30), respectively, and with the luminosity functions described above.

Regarding the clustering bias, for the full sample we use the parameterization  $b_{\text{full}}(z) = 1 + 0.84z$ , based on the simulations of Weinberg et al. (2004) and quoted in the LSST science book (LSST Collaboration et al. 2009). Red galaxies should have a larger bias, which we parameterize as  $b_{\text{red}}(z) \simeq 1 + z$ . This parameterization is compatible with bias measurements at redshifts  $z < 1$  (e.g., Coil et al. 2008). Since the red population dies off at  $z \sim 1.4$ , this extrapolation should not significantly influence our final result, especially as we ultimately marginalize over  $b(z)$ .

The redshift distributions for the two samples considered here are shown in Figure 20. For this figure, as well as in our forecasts, we assume a magnitude limit of  $i = 25.3$ , corresponding to  $r \sim 26$  for typical galaxy colors, as quoted in LSST Collaboration et al. (2009). According to the models used here, the total number of galaxies observed by LSST should be  $\sim 40$  per  $\text{arcmin}^2$ , in qualitative agreement with previous results (Ilbert et al. 2006; LSST Collaboration et al. 2009).

#### B.4.2. Photometric Redshifts, Binning, and Shot Noise

The quality of a given photo- $z$  algorithm is normally quoted in terms of its rms error,  $\sigma_z^2 \equiv \langle (z_{\text{photo}} - z_{\text{true}})^2 \rangle$ , which is typically parameterized as

$$\sigma_z(z) = \sigma_0(1 + z). \quad (80)$$

The photometric redshift requirement for the LSST gold sample, as quoted in LSST Collaboration et al. (2009), is  $\sigma_0 < 0.05$ , with a goal of 0.02. As described above, since the spectral properties of red galaxies make their photometric redshifts more accurate than those of the blue population, we have assumed the following photometric redshift uncertainties for the two samples:

$$\sigma_0^{\text{red}} = 0.02, \sigma_0^{\text{full}} = 0.05. \quad (81)$$

We have also assumed that the data will be analyzed by dividing the sample into a number of photo- $z$  bins. Let  $z_0^i$  and  $z_f^i$  be the limits of the  $i$ th bin. The window function in this bin must trace the true  $z$  distribution of galaxies within it, and is therefore given by the product of the overall redshift distribution and the photo- $z$  probability distribution, integrated over the bin:

$$W^i(z) \propto \bar{N}(z) w^i(z), \quad (82)$$

$$w^i(z) = \int_{z_0^i}^{z_f^i} dz_p p(z_p|z), \quad (83)$$

where  $p(z_p|z)$  is the photo- $z$  pdf. Assuming a Gaussian photo- $z$  distribution, we can write the window functions  $w^i(z)$  analytically as

$$w^i(z) = \frac{1}{2} \left[ \text{erf} \left( \frac{z - z_0^i}{\sqrt{2} \sigma_z} \right) - \text{erf} \left( \frac{z - z_f^i}{\sqrt{2} \sigma_z} \right) \right]. \quad (84)$$

The tails of  $w^i(z)$  correlate different redshift bins to a much larger degree than the intrinsic correlations due to gravitational clustering do, which is an expression of the loss of information on radial scales. In order to reduce this correlation and avoid redundant calculations, the widths of the redshift bins are usually defined to be of the order of  $\sigma_z$ . In this work we have chosen to define the width of our bins to be three times the photo- $z$  dispersion at the bin center. For the values of  $\sigma_z$

assumed here, this results in 15 bins for the red sample and 9 bins for the full sample. The window functions for the bins are shown in Figure 20.

Finally, as in the case of spectroscopic and continuum surveys, the main source of statistical noise in the measurement of clustering anisotropies is shot noise. The noise bias term for photometric surveys is thus given by

$$N_{\ell}^{ij} = \frac{\delta_{ij}}{n^i n^j}, \quad (85)$$

where

$$n^i \equiv \int_0^\infty dz \bar{N}(z) w^i(z), \quad (86)$$

Note that this reduces to Equation (74) in the case of top-hat windows (i.e.,  $\sigma_z \rightarrow 0$ ).

## APPENDIX C

### LUMINOSITY FUNCTION OF RADIO SOURCES

We will summarize here the steps that were followed to compute the luminosity functions of the different radio sources that would be observable with a continuum survey. We follow almost exactly what was done by Wilman et al. (2008) to simulate the distribution of radio galaxies, but we would like to describe the details of the calculation here for the benefit of the potential users of the codes used in this paper.

We will discuss five main galaxy populations: star-forming galaxies, starbursts, radio-quiet AGNs, and radio-loud AGNs of types FRI and FRII. We note that we use the convention for the luminosity function used in Section 2.1, where  $n_s(z, L)$  is the physical (not comoving) number density of objects per unit  $\ln L$ .

#### C.1. Star-forming Galaxies and Starbursts

The luminosity functions of normal star-forming galaxies and starbursts used here are based on the luminosity function derived from the IRAS 2  $\mu$ Jy sample by Yun et al. (2001) at 1.4 GHz. This is modeled as a sum of two Schechter functions, which we identify with two populations: normal and starburst galaxies respectively. Thus, for both normal star-forming galaxies and starbursts, the  $z = 0, 1.4$  GHz luminosity function is parameterized as

$$n_s(z = 0, L) = n_* \left( \frac{L}{L_*} \right)^\alpha \exp \left[ -\frac{L}{L_*} \right], \quad (87)$$

where in each case

$$n_* = \begin{cases} 3.2 \times 10^{-4} \text{ Mpc}^{-3} & \text{for normal galaxies} \\ 8.3 \times 10^{-6} \text{ Mpc}^{-3} & \text{for starbursts} \end{cases} \quad (88)$$

$$L_* = \begin{cases} 2.10 \times 10^{22} [\text{W Hz}^{-1} \text{ sr}^{-1}] & \text{for normal galaxies} \\ 1.44 \times 10^{23} [\text{W Hz}^{-1} \text{ sr}^{-1}] & \text{for starbursts} \end{cases} \quad (89)$$

and  $\alpha = -0.633$  in both cases. This parameterization is valid for luminosities above  $\log_{10} L / [\text{W Hz}^{-1} \text{ sr}^{-1}] = 19.6$ , and it is assumed to be constant for lower values of  $L$ , in agreement with Mauch & Sadler (2007).

In agreement with Rowan-Robinson et al. (1993), we assume a pure luminosity evolution of this luminosity function,

$$n_s(z, L) = (1 + z)^3 n_s(z = 0, L f(z)), \quad (90)$$

$$f(z) = \begin{cases} (1 + z)^{3.1} & \text{for } z < z_0 \\ (1 + z_0)^{3.1} & \text{for } z \geq z_0, \end{cases} \quad (91)$$

with  $z_0 = 1.5$ . This parameterization was obtained assuming a cosmological model  $(\Omega_M, \Omega_\Lambda) = (1, 0)$ . We adapt it to our fiducial cosmology by scaling the number densities and luminosities by the ratios of comoving volumes and luminosity distances in both models,

$$n_s^{(1)}(z, L | \Omega_1) = n_s^{(2)} \left( z, L \frac{\chi_{(2)}^2(z)}{\chi_{(1)}^2(z)} \right) \frac{\chi_{(2)}^2(z) H_{(1)}(z)}{\chi_{(1)}^2(z) H_{(2)}(z)}, \quad (92)$$

where the indices (1) and (2) indicate quantities computed in two different cosmological models.

The SED for both normal and starburst galaxies was assumed to be

$$\varphi(\nu) \propto \nu_{\text{GHz}}^2 (1 - e^{-\tau}) (1 + 10 \nu_{\text{GHz}}^{-0.65}) \quad (93)$$

with  $\nu_{\text{GHz}} \equiv \nu / (1 \text{ GHz})$  and  $\tau = (\nu_*/\nu_{\text{GHz}})^{2.1}$ , with  $\nu_* = 0.005$  for normal galaxies and  $\nu_* = 1$  for starbursts. This corresponds to a combination of thermal free-free emission and non-thermal synchrotron from supernovae. We did not include a dust component in this SED, as mentioned in Wilman et al. (2008), which should be irrelevant for the range of redshifts and frequencies studied here.

#### C.2. Radio-loud AGNs

The luminosity function for FRI and FRII radio-loud AGNs was based on model ‘‘C’’ of the luminosity function at 151 MHz derived by Willott et al. (2001). This luminosity function consists of low-luminosity and high-luminosity components, which we identify with FRI and FRII sources respectively.

For FRI galaxies, the  $z = 0$  luminosity function takes the form of a Schechter function

$$n_s(z = 0, L) = \frac{n_*}{\ln 10} \left( \frac{L}{L_*} \right)^\alpha \exp \left[ -\frac{L}{L_*} \right], \quad (94)$$

with  $n_* = 10^{-7.12} \text{ Mpc}^{-3}$ ,  $\log_{10} L_*/(1 \text{ W Hz}^{-1} \text{ sr}^{-1}) = 26.1$ , and  $\alpha = -0.539$ . The model also assumes a pure density evolution,

$$n_s(z, L) = (1 + z)^3 f(z) n_s(z = 0, L) \quad (95)$$

$$f(z) = \begin{cases} (1 + z)^{4.3} & \text{for } z < z_0 \\ (1 + z_0)^{4.3} & \text{for } z \geq z_0, \end{cases} \quad (96)$$

where  $z_0 = 0.706$ .

For FRII galaxies, the luminosity function takes the form of a Schechter function with an inverted exponential term,

$$n_s(z = 0, L) = \frac{n_*}{\ln 10} \left( \frac{L}{L_*} \right)^\alpha \exp \left[ -\frac{L_*}{L} \right], \quad (97)$$

with  $n_* = 10^{-6.196} \text{ Mpc}^{-3}$ ,  $\log_{10} L_*/(1 \text{ W Hz}^{-1} \text{ sr}^{-1}) = 26.95$ , and  $\alpha = -2.27$ . As for FRI, a pure density evolution is

assumed, with

$$f(z) = \exp\left(-\frac{(z - z_0)^2}{2\sigma_*^2(z)}\right) \quad (98)$$

$$\sigma_*(z) = \begin{cases} 0.559 & \text{for } z < z_0 \\ 1.378 & \text{for } z \geq z_0, \end{cases} \quad (99)$$

where  $z_0 = 1.91$ . As in the case of star-forming galaxies, these luminosity functions were derived for an Einstein–de Sitter background, so had to be adapted to our fiducial cosmology using Equation (92).

A power-law SED with  $\varphi(\nu) \propto \nu^{-0.75}$  was assumed for both types of radio-loud AGNs.

### C.3. Radio-quiet AGN

Radio-quiet AGNs make up the majority of the total AGN population, as is observed from the hard X-ray luminosity function. This can be combined with the relation between hard X-ray and radio luminosities (Brinkmann et al. 2000),

$$\begin{aligned} \log_{10}(L_{2-10 \text{ keV}}/(\text{erg s}^{-1})) \\ = 1.012 \log_{10}(L_{1.4 \text{ GHz}}/(\text{W Hz}^{-1} \text{ sr}^{-1})) + 21.3, \end{aligned} \quad (100)$$

to derive the 1.4 GHz luminosity function.

For this we use the X-ray luminosity function of Ueda et al. (2003) in the 2–10 keV band, parameterized at  $z = 0$  as

$$n_s(z = 0, L_X) = \frac{A}{\ln 10} \left[ (L_X/L_*)^{\gamma_1} + (L_X/L_*)^{\gamma_2} \right]^{-1}, \quad (101)$$

where  $L_X$  is the X-ray luminosity,  $A = 5.04 \times 10^{-6} \text{ Mpc}^{-3}$ ,  $\gamma_1 = 0.86$ ,  $\gamma_2 = 2.23$ , and  $\log_{10}(L_*/(\text{erg s}^{-1})) = 43.94$ .

The evolution with redshift is parameterized as a luminosity-dependent density evolution,

$$n_s(z, L_X) = (1 + z)^3 f(z, L_X) n_s(z = 0, L_X), \quad (102)$$

with

$$f(z, L_X) = \begin{cases} (1 + z)^{4.23} & \text{for } z < z_0 \\ (1 + z_0)^{4.23} \left( \frac{1 + z}{1 + z_0} \right)^{-1.5} & \text{for } z \geq z_0, \end{cases} \quad (103)$$

where  $z_0$  is a luminosity-dependent function

$$z_0(L_X) = \begin{cases} 1.9(L_X/L_1)^{0.335} & \text{for } L_X < L_1 \\ 1.9 & \text{for } L_X \geq L_1, \end{cases} \quad (104)$$

with  $\log_{10}(L_1/(\text{erg s}^{-1})) = 44.6$ .

A power-law SED with  $\varphi(\nu) \propto \nu^{-0.7}$  was assumed for radio-quiet AGNs.

## REFERENCES

- Alonso, D., Bull, P., Ferreira, P. G., & Santos, M. G. 2015, *MNRAS*, **447**, 400
- Amendola, L., Appleby, S., Bacon, D., et al. 2013, *LRR*, **16**, 6
- Baldauf, T., Seljak, U., Senatore, L., & Zaldarriaga, M. 2011, *JCAP*, **10**, 31
- Battye, R. A., Davies, R. D., & Weller, J. 2004, *MNRAS*, **355**, 1339
- Baumann, D. 2009, arXiv:0907.5424
- Bertacca, D., Maartens, R., Raccanelli, A., & Clarkson, C. 2012, *JCAP*, **1210**, 025
- Blanton, M. R., & Roweis, S. 2007, *AJ*, **133**, 734
- Bonvin, C., & Durrer, R. 2011, *PhRvD*, **84**, 063505
- Bonvin, C., Durrer, R., & Gasparini, M. A. 2006, *PhRvD*, **73**, 023523
- Braun, R., Bourke, T. L., Green, J. G., Keane, E. F., & Wagg, J. 2015, in *Advancing Astrophysics with the Square Kilometre Array*, PoS, PoS (AASKA14)174
- Brinkmann, W., Laurent-Muehleisen, S. A., Voges, W., et al. 2000, *A&A*, **356**, 445
- Bruni, M., Crittenden, R., Koyama, K., et al. 2012, *PhRvD*, **85**, 041301
- Bull, P., Ferreira, P. G., Patel, P., & Santos, M. G. 2015, *ApJ*, **803**, 21
- Camera, S., Maartens, R., & Santos, M. G. 2015a, *MNRAS*, **451**, L80
- Camera, S., Santos, M. G., Ferreira, P. G., & Ferramacho, L. 2013, *PhRvL*, **111**, 171302
- Camera, S., Santos, M. G., & Maartens, R. 2015b, *MNRAS*, **448**, 1035
- Challinor, A., & Lewis, A. 2011, *PhRvD*, **84**, 043516
- Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, *PhRvL*, **100**, 091303
- Chen, S., & Schwarz, D. J. 2015, *PhRvD*, **91**, 043507
- Cheung, C., Creminelli, P., Fitzpatrick, A. L., Kaplan, J., & Senatore, L. 2008, *JHEP*, **0803**, 014
- Coil, A. L., Newman, J. A., Croton, D., et al. 2008, *ApJ*, **672**, 153
- Cole, S., Helly, J., Frenk, C. S., & Parkinson, H. 2008, *MNRAS*, **383**, 546
- Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, *ApJS*, **43**, 393
- Dalal, N., Doré, O., Huterer, D., & Shirokov, A. 2008, *PhRvD*, **77**, 123514
- Dark Energy Survey Collaboration 2005, arXiv:astro-ph/0510346
- Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, *IEEEP*, **97**, 1482
- Di Dio, E., Montanari, F., Lesgourgues, J., & Durrer, R. 2013, *JCAP*, **11**, 44
- Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, *ApJ*, **665**, 265
- Fergusson, J., & Shellard, E. 2009, *PhRvD*, **80**, 043510
- Ferramacho, L. D., Santos, M. G., Jarvis, M. J., & Camera, S. 2014, *MNRAS*, **442**, 2511
- Font-Ribera, A., Miralda-Escudé, J., Arnau, E., et al. 2012, *JCAP*, **11**, 59
- Fukugita, M., Shimasaku, K., & Ichikawa, T. 1995, *PASP*, **107**, 945
- Gabash, A., Hopp, U., Feulner, G., et al. 2006, *A&A*, **448**, 101
- Geach, J. E., Cimatti, A., Percival, W., et al. 2010, *MNRAS*, **402**, 1330
- Giannantonio, T., Porciani, C., Carron, J., Amara, A., & Pillepich, A. 2012, *MNRAS*, **422**, 2854
- Giannantonio, T., Ross, A. J., Percival, W. J., et al. 2014, *PhRvD*, **89**, 023511
- Giannantonio, T., Scranton, R., Crittenden, R. G., et al. 2008, *PhRvD*, **77**, 123520
- Hall, A., Bonvin, C., & Challinor, A. 2013, *PhRvD*, **87**, 064026
- Hildebrandt, H., van Waerbeke, L., & Erben, T. 2009, *A&A*, **507**, 683
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, **457**, 841
- Jarvis, M. J., Bacon, D., Blake, C., et al. 2015, arXiv:1501.03825
- Jeong, D., Schmidt, F., & Hirata, C. M. 2012, *PhRvD*, **85**, 023504
- Joachim, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011, *A&A*, **527**, A26
- Kaiser, N., Aussel, H., Burke, B. E., et al. 2002, *Proc. SPIE*, **4836**, 154
- Komatsu, E., & Spergel, D. N. 2001, *PhRvD*, **63**, 063002
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv:1110.3193
- Lesgourgues, J. 2011, arXiv:1104.2932
- Liguori, M., Hansen, F., Komatsu, E., Matarrese, S., & Riotto, A. 2006, *PhRvD*, **73**, 043505
- Lindsay, S. N., Jarvis, M. J., & McAlpine, K. 2014, *MNRAS*, **440**, 2322
- Lombriser, L., Yoo, J., & Koyama, K. 2013, *PhRvD*, **87**, 104019
- Lopez-Honorez, L., Mena, O., & Rigolin, S. 2012, *PhRvD*, **85**, 023511
- LSST Collaboration 2009, arXiv:0912.0201
- Maartens, R., Zhao, G.-B., Bacon, D., Koyama, K., & Raccanelli, A. 2013, *JCAP*, **2**, 44
- Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, *ApJL*, **763**, L20
- Matarrese, S., & Verde, L. 2008, *ApJL*, **677**, L77
- Mauch, T., & Sadler, E. M. 2007, *MNRAS*, **375**, 931
- Mo, H. J., & White, S. D. M. 1996, *MNRAS*, **282**, 347
- Montanari, F., & Durrer, R. 2015, *JCAP*, **10**, 70
- Namikawa, T., Okamura, T., & Taruya, A. 2011, *PhRvD*, **83**, 123514
- Newman, J. A. 2008, *ApJ*, **684**, 88
- Peacock, J. A., & Smith, R. E. 2000, *MNRAS*, **318**, 1144
- Percival, W. J. 2014, *Proc. Int. Sch. Phys. Fermi*, **186**, 101
- Planck Collaboration 2014a, *A&A*, **571**, A15
- Planck Collaboration 2014b, *A&A*, **571**, A16
- Planck Collaboration 2014c, *A&A*, **571**, A24
- Planck Collaboration 2015a, arXiv:1502.01589
- Planck Collaboration 2015b, arXiv:1502.01592
- Raccanelli, A., Bertacca, D., Doré, O., & Maartens, R. 2014, *JCAP*, **8**, 22
- Raccanelli, A., Doré, O., Bacon, D. J., et al. 2015a, *JCAP*, **1**, 42
- Raccanelli, A., Montanari, F., Bertacca, D., Doré, O., & Durrer, R. 2015b, arXiv:1505.06179
- Ross, A. J., Ho, S., Cuesta, A. J., et al. 2011, *MNRAS*, **417**, 1350

- Rowan-Robinson, M., Benn, C. R., Lawrence, A., McMahon, R. G., & Broadhurst, T. J. 1993, [MNRAS](#), **263**, 123
- Sachs, R. K., & Wolfe, A. M. 1967, [ApJ](#), **147**, 73
- Santos, M. G., Bull, P., Alonso, D., et al. 2015, arXiv:1501.03989
- Schlegel, D., Abdalla, F., Abraham, T., et al. 2011, arXiv:1106.1706
- Scranton, R., Ménard, B., Richards, G. T., et al. 2005, [ApJ](#), **633**, 589
- Seljak, U. 2009, [PhRvL](#), **102**, 021302
- Sheth, R. K., & Tormen, G. 1999, [MNRAS](#), **308**, 119
- Smith, K. M., & Zaldarriaga, M. 2011, [MNRAS](#), **417**, 2
- Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, [MNRAS](#), **434**, L46
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, [ApJ](#), **598**, 886
- Verde, L., Wang, L.-M., Heavens, A., & Kamionkowski, M. 2000, [MNRAS](#), **313**, L141
- Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2004, [ApJ](#), **601**, 1
- Weinberg, S. 2008, [PhRvD](#), **77**, 123541
- Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, [MNRAS](#), **322**, 536
- Wilman, R. J., Miller, L., Jarvis, M. J., et al. 2008, [MNRAS](#), **388**, 1335
- Wyithe, J. S. B., & Loeb, A. 2008, [MNRAS](#), **383**, 606
- Yadav, A. P., & Wandelt, B. D. 2008, [PhRvL](#), **100**, 181301
- Yahya, S., Bull, P., Santos, M., et al. 2015, [MNRAS](#), **450**, 2251
- Yang, X., & Zhang, P. 2011, [MNRAS](#), **415**, 3485
- Yang, X., Zhang, P., Zhang, J., & Yu, Y. 2015, [MNRAS](#), **447**, 345
- Yoo, J. 2009, [PhRvD](#), **79**, 023517
- Yoo, J. 2010, [PhRvD](#), **82**, 083508
- Yoo, J., & Desjacques, V. 2013, [PhRvD](#), **88**, 023502
- Yoo, J., Fitzpatrick, A. L., & Zaldarriaga, M. 2009, [PhRvD](#), **80**, 083514
- Yoo, J., Hamaus, N., Seljak, U., & Zaldarriaga, M. 2012, [PhRvD](#), **86**, 063514
- Yoo, J., & Seljak, U. 2015, [MNRAS](#), **447**, 1789
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, [ApJ](#), **554**, 803