Gravitational lensing of cosmological 21 cm emission

A. Pourtsidou[1,2]★ and R. Benton Metcalf[1]

[1]*Dipartimento di Fisica e Astronomia, Universitá di Bologna, viale B. Pichat 6/2, I-40127 Bologna, Italy*
[2]*Institute of Cosmology & Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK*

ABSTRACT

We investigate the feasibility of measuring weak gravitational lensing using 21-cm intensity mapping with special emphasis on the performance of the planned Square Kilometre Array (SKA). We find that the current design for SKA_Mid should be able to measure the evolution of the lensing power spectrum at $z \sim 2$–3 using this technique. This will be a probe of the expansion history of the Universe and gravity at a unique range in redshift. The signal to noise is found to be highly dependent on evolution of the neutral hydrogen fraction in the Universe with a higher H I density resulting in stronger signal. With realistic models for this, SKA Phase 1 should be capable of measuring the lensing power spectrum and its evolution. The dependence of signal to noise on the area and diameter of the array is quantified. We further demonstrate the applications of this technique by applying it to two specific coupled dark energy models that would be difficult to observationally distinguish without information from this range of redshift. We also investigate measuring the lensing signal with 21 cm emission from the Epoch of Reionization (EoR) using SKA and find that it is unlikely to constrain cosmological parameters because of the small survey size, but could provide a map of the dark matter within a small region of the sky.

Key words: gravitational lensing: weak – cosmology: theory – dark energy.

1 INTRODUCTION

21 cm cosmology is a new and exciting area of research with a great deal of potential. Future radio telescopes like the SKA[1] (Square Kilometre Array) will give us access to previously unexplored epochs of the Universe, such as the Epoch of Reionization (EoR), the end of the Dark Ages and the beginning of the Cosmic Dawn (Dewdney 2013).

Another important development in the field has been the H I intensity mapping technique in which the distribution of galaxies can be measured without detecting individual galaxies. Instead, the 21 cm emission is treated as a continuous (unresolved) background, much like the cosmic microwave background (CMB), but extended in the frequency (redshift) dimension. Several groups are planning to use this technique to measure the baryon acoustic oscillations scale at redshifts of order of unity (Chang et al. 2008, 2010; Seo et al. 2010; Masui, McDonald & Pen 2010; Ansari et al. 2012; Battye et al. 2012; Chen 2012; Pober et al. 2013; Smoot & Debono 2014; Bull et al. 2014), an important probe of dark energy (DE).

By combining results from different kinds of surveys (CMB, galaxies, 21 cm), which probe complimentary redshift ranges, we will be able to span a very large observational volume and more

precisely investigate the acceleration of the Universe. Of particular relevance to this paper is the possibility of extending the cosmological probes to redshifts between those accessible with galaxy redshift surveys and the high redshift of the CMB. By studying the evolution of expansion and structure formation across a wide range of redshift we can investigate whether DE or modified gravity effects are present at a higher redshift than the standard cosmological model predicts (Copeland, Sami & Tsujikawa 2006; Clifton et al. 2012).

The 21 cm radiation provides an excellent source for gravitational lensing studies. Earlier work (Zahn & Zaldarriaga 2005; Metcalf & White 2009) had shown that if the EoR is at redshift $z \sim 8$ or later, a large radio array such as SKA could measure the lensing convergence power spectrum and constrain the standard cosmological parameters – however, these studies assumed a much larger survey area than is currently planned for the largest such experiment, SKA_Low (see Section 3 for details). The authors extended the Fourier-space quadratic estimator technique, which was first developed by Hu (2001) for CMB lensing observations to three-dimensional observables, i.e. the 21-cm intensity field $I(\theta, z)$. These studies did not consider 21 cm observations from redshifts after reionization when the average H I density in the Universe is much smaller.

In Pourtsidou & Metcalf (2014), we extended the aforementioned studies to redshifts after reionization, but before those probed by galaxy surveys in the visible bands and showed that lensing can

★ E-mail: alkistispou@gmail.com
[1] www.skatelescope.org

be measured using the H I intensity mapping technique. Here, we develop this concept further. In Section 2, we present a preliminary overview of our study, which spans two distinct regimes (EoR and $z \sim 2$–5), and give the formal form of the lensing estimator and lensing reconstruction noise. In Section 3, we study how well the current SKA design will be able to map the lensing convergence at typical EoR redshifts. In Section 4, we investigate weak lensing intensity mapping with an SKA-like interferometer array at redshifts $z \sim 2$–3. Different models for the H I mass function and different array configurations are considered. We concentrate on the ability of different SKA phases to probe the lensing signal and calculate the corresponding signal-to-noise (S/N) predictions. In Section 5, we show that these measurements can be used to differentiate between some specific and novel interacting DE models. We conclude in Section 6. Throughout the paper, we adopt a flat Λ cold dark matter (ΛCDM) Universe with the *Planck* cosmological parameter values (Planck Collaboration XVI 2014), unless otherwise stated.

2 PRELIMINARIES

We will investigate the lensing of 21 cm emission in two distinct regimes. The first is from the EoR and the second is from $z \sim 2$ to 5. In the first case, the neutral hydrogen fraction is high and H I gas is not restricted to individual galaxies. In this case, the H I distribution will be approximated as a Gaussian random field that roughly follows the distribution of matter. More sophisticated models are under investigation, but will not be considered here. The biggest source of noise for 21 cm observations at these high redshifts comes from foreground contamination. SKA_Low is being planned to probe this regime. Current plans call for it to survey a $5° \times 5°$ sky area with a frequency range of 40–250 MHz corresponding to redshifts $z \sim 5$–35 (Dewdney 2013).

At lower redshift, the H I fraction is much lower and essentially all the H I gas is in discrete galaxies. In this regime, we treat the H I distribution as consisting of discrete sources that are clustered according to the standard CDM paradigm, but also exhibit random Poisson fluctuations. This is the standard way of modelling the distribution of galaxies in redshift surveys. We find that this Poisson or shot noise contribution to the galaxy clustering is important for measuring lensing from 21 cm after (and perhaps during) reionization. The evolution in the H I fraction as a function of redshift is a matter of some debate and speculation and has an important effect on the expected lensing S/N. We will address this uncertainty in Section 4 by adopting several different models.

The quadratic lensing estimator for the CMB or a single redshift slice of the 21 cm emission works by essentially measuring differences in the power spectrum in different regions of the sky which results in correlations in Fourier (or spherical harmonic) modes that would not exist otherwise. For the CMB, the local power spectrum becomes anisotropic because of shearing and the acoustic peaks are slightly shifted to larger (smaller) angular scales by gravitational magnification (demagnification).

The effect of lensing can be divided into a shearing which can make the local power spectrum anisotropic and magnification which isotropically scales the local power spectrum. It can be shown that an isotropic magnification cannot be measured (with a quadratic estimator) if the power spectrum is scale free ($C_\ell \propto \ell^{-2}$). It can also be shown that the shear cannot be measured if the power spectrum is constant (see Bucher et al. 2012 for a nice demonstration of this). As we will see, the projected matter power spectrum in the CDM model is approximately a constant at large scales (small ℓ) and $C_\ell \propto \ell^{-2}$ at small scales so to the extent that the H I distribution

follows the dark matter (DM) distribution, the quadratic estimator picks up the magnification on large scales and the shear on small scales and some combinations in between.

After reionization the remaining H I resides in discrete galaxies. The galaxies are clustered in the same way as matter on large scales (modulo a bias factor), but on small scales their discreteness enters into the power spectrum as shot noise or Poisson noise. The power spectrum of this component is flat (for equal luminosity sources, $C_\ell = 1/\eta$ where η is the number density on the sky) so only the magnification can be measured. Magnifying (demagnifying) a region of the sky reduces (increases) the number density of sources. The result is that the power spectrum will have features on the scale of the lensing.

Unlike the CMB, the 21 cm emission will be observable at many different redshifts. The 3D lensing estimator effectively stacks redshift slices. Since the lensing is coherent for different slices at the same angular position, but the 21 cm emission is statistically independent, the residual effect can be attributed to lensing. This will be made more explicit in the next section.

2.1 The quadratic lensing estimator

Here, we will give a more explicit description of the lensing estimator and noise. The advantage of 21 cm lensing is that one is able to combine information from multiple redshift slices, and that the 21 cm signal extends to far smaller angular scales than CMB fluctuations (Loeb 2004), which are suppressed by Silk damping at $\ell > 3000$. The intensity fluctuations, $I(\theta, z)$, are expressed in discrete Fourier space in the radial direction owing to the finite band width (wave vector $k_\parallel = \frac{2\pi}{\mathcal{L}} j$, where $\mathcal{L}$ is the depth of the observed volume) and in continuous Fourier space perpendicular to the line of sight (wave vector $\boldsymbol{k}_\perp = \boldsymbol{l}/\mathcal{D}$ where $\mathcal{D}$ the angular diameter distance to the source redshift and $\boldsymbol{l}$ is the dual of the angular coordinate on the sky). See Appendix A for the connection between estimators in discrete and continuous Fourier space. Considering modes with different j independent, an optimal estimator can be found by combining the individual estimators for different j modes without mixing them.

The quadratic estimator for the lensing potential Ψ is thus of the form:

$$\hat{\Psi}(\boldsymbol{L}) = \sum_j \sum_{\boldsymbol{\ell}} g(\boldsymbol{\ell}, \boldsymbol{L}, j)\, \tilde{I}_{\boldsymbol{\ell}, j} \tilde{I}^*_{\boldsymbol{\ell}-\boldsymbol{L}, j}, \tag{1}$$

where $\tilde{I}_{\boldsymbol{\ell}, j}$ is the discrete Fourier transform (DFT) of the lensed intensity field $\tilde{I}(\boldsymbol{\theta}, x_\parallel)$. The form of the filter $g(\boldsymbol{\ell}, \boldsymbol{L}, j)$ and the expected noise in the estimator depends on whether the source is discrete or continuous. In Appendices A through C, we derive the cases that are used in the following sections.

The mean observed brightness temperature at redshift z due to the average H I density is $\bar{T}(z) = 180\, \Omega_{H_I}(z)\, h\, \frac{(1+z)^2}{H(z)/H_0}$ mK (Battye et al. 2012). The angular power spectrum of the displacement field (the displacement of points on the sky caused by gravitational lensing) for sources at a redshift z_s can be calculated with

$$C_L^{\delta\theta\delta\theta} = \frac{9\Omega_m^2 H_0^3}{L(L+1)c^3} \int_0^{z_s} \mathrm{d}z\, P(k = L/\mathcal{D}(z), z) \frac{[W(z)]^2}{a^2 E(z)}, \tag{2}$$

(Kaiser 1992) in the weak field limit where $W(z) = (\mathcal{D}(z_s) - \mathcal{D}(z))/\mathcal{D}(z_s)$, $E(z) = H(z)/H_0$ and $P(k)$ is the power spectrum of matter. The convergence power spectrum is simply $C_L^{\kappa\kappa} = (L^2/4)\, C_L^{\delta\theta\delta\theta}$. Note that the Limber approximation (Limber 1953) has been used in deriving equation (2). This approximation is quite accurate even for large scales $\ell \sim 10$, as the resulting formula involves an integral

over the line of sight which effectively averages out the effect of the approximation, see Schmidt et al. ([2008](#)) for a detailed discussion.

The lensing reconstruction noise includes the thermal noise of the array which – assuming a uniform telescope distribution – is calculated using the formula

$$C_\ell^N = \frac{(2\pi)^3 T_{\rm sys}^2}{B t_{\rm obs} f_{\rm cover}^2 \ell_{\rm max}(\nu)^2} , \tag{3}$$

where $T_{\rm sys}$ is the system temperature, B is the chosen frequency window, $t_{\rm obs}$ the total observation time, $D_{\rm tel}$ the diameter (maximum baseline) of the core array, $\ell_{\rm max}(\lambda) = 2\pi D_{\rm tel}/\lambda$ is the highest multipole that can be measured by the array at frequency ν (wavelength λ) and $f_{\rm cover}$ is the total collecting area of the core array $A_{\rm coll}$ divided by $\pi(D_{\rm tel}/2)^2$ (Furlanetto, Oh & Briggs [2006](#)).

Assuming that the temperature (i.e. the 21 cm emissivity) is Gaussian distributed, we find that the lensing reconstruction noise $N(L)$ is (see Zahn & Zaldarriaga [2005](#) and Appendix A)

$$N(L, \nu) =$$
$$\left[\sum_{j=1}^{j_{\rm max}} \frac{1}{L^4} \int \frac{{\rm d}^2\ell}{(2\pi)^2} \frac{[\boldsymbol{l} \cdot \boldsymbol{L} C_{\ell,j} + \boldsymbol{L} \cdot (\boldsymbol{L} - \boldsymbol{l}) C_{|\ell - L|,j}]^2}{2 C_{\ell,j}^{\rm tot} C_{|L-l|,j}^{\rm tot}} \right]^{-1} , \tag{4}$$

where $C_{\ell,j}^{\rm tot} = C_{\ell,j} + C_\ell^N$ and $C_{\ell,j} = [\bar{T}(z)]^2 P_{\ell,j}$ with $P_{\ell,j}$ the underlying DM power spectrum. One can also consider stacking multiple bands ν and then the noise is reduced as $N_L = 1/\sum_\nu [N(L, \nu)]^{-1}$. The Gaussian approximation formalism will be used in the next section to investigate the possibility of measuring the lensing signal with 21 cm emission from the EoR.

3 LENSING STUDIES USING 21 CM RADIATION FROM THE EoR

Observing the EoR through 21 cm has been one of the primary science goals for several large radio telescope projects including the SKA. The collecting area and resolution of LOFAR[2] or MWA[3] are unlikely to be high enough to observe gravitational lensing as we will see. For this reason, we concentrate on the SKA_Low instrument here which is the part of SKA designed for observing the EoR and the Cosmic Dawn.

The current SKA plans call for a 25 square degree survey[4] with SKA_Low (Dewdney [2013](#)). The convergence power spectrum measured in this size field will be dominated by sample variance. This will make it difficult to measure the cosmological parameters through linear growth in the matter power spectrum at a competitive level with these observations even if the S/N in the lensing map is very high. This is not true of the SKA_Mid at lower redshift where the survey area will be much larger – see Section 4.

However, it might be possible to map the lensing convergence within the 25 square degree EoR survey area with high fidelity.

This would allow us to actually 'see' the distribution of DM in a typical region of the sky, something that is only possible with galaxy lensing around very atypical, large galaxy clusters.

For the EoR, the convergence (or, equivalently, the displacement field) estimator and the corresponding lensing reconstruction noise are calculated assuming that the temperature (brightness) distribution is Gaussian, and consider the reionization fraction $f_{\rm H{\small I}} = 1$ until the Universe is rapidly and uniformly reionized at redshift $z_{\rm s}$. Before reionization, this is probably a very good approximation. Reionization is not expected to be homogeneous and may extend over a significant redshift range so for some period there will be ionized regions that grow and intersect until they fill almost all of space when reionization is complete. Because the contrast in the brightness temperature is larger during this period, detecting and mapping the 21 cm emission will be easier, but it will change the power spectrum and increase the fourth order moments of the brightness temperature (Metcalf & White [2009](#)). It is not yet clear how this will affect the detectability of lensing. To assess this will require numerical simulations which we plan to present in a future publication. Note that a semi-analytic model used to describe the patchy regime in Zahn & Zaldarriaga ([2005](#)) showed an increase in noise due to a decrease of the fluctuation level of the 21 cm signal on the smallest resolved scales and the contribution of the connected four-point function which acts as a sample variance term correlating different k-modes. Of particular importance is the non-linear structure effect during the EoR. We will be able to account for these effects once more details about the reionization process are known, but under the assumption that these features appear in higher resolution than that of our considered interferometers there is no need for the connected four-point contribution to the variance of the quadratic lensing estimator to be taken into account (Kovetz & Kamionkowski [2013](#)).

At redshift $z_{\rm s} \sim 8$, we assume the SKA1 Baseline Design (Dewdney [2013](#)) parameters of $A_{\rm coll} \simeq 0.3\,{\rm km}^2$ with maximum baseline $D_{\rm tel} = 4\,{\rm km}$, while for SKA_Mid Phase 2 (SKA2), we can consider $A_{\rm coll} \simeq 1.2\,{\rm km}^2$. The estimated lensing noise is shown in Fig. [1](#) along with the estimated signal. Here, $C_L^{\delta\theta\delta\theta}$ is the convergence field power spectrum at $z_{\rm s} = 8$ and N_L the lensing reconstruction noise assuming a reionization fraction $f_{\rm H{\small I}} = 1$. Note
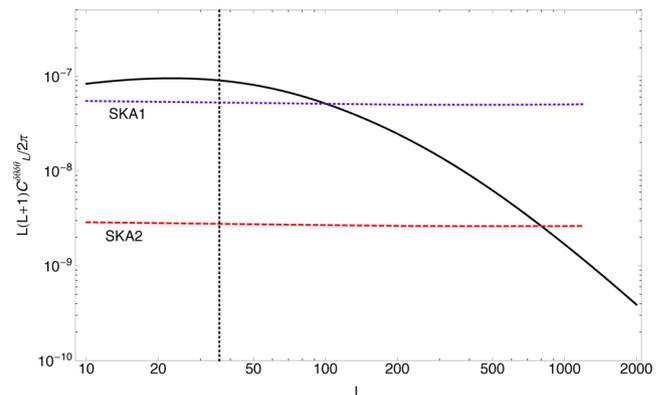

Figure 1. The lensing displacement field power spectrum, $C_L^{\delta\theta\delta\theta}$, for sources at $z = 8$ is shown as a solid black line and lensing reconstruction noise N_L as dashed/dotted lines. The blue dotted curve is for the SKA1 Baseline Design with 108 MHz frequency bins around $z = 8$ spanning the redshift range $z \simeq 6.5$–11. The red dashed line is for SKA2 and the same frequency bins. The vertical line is approximately the lowest L accessible with a** 5-by-5 degree field. Where the noise curves are below C_L, typical fluctuations in the lensing deflection should be recoverable in a map.

that at such high redshifts, the system temperature T_{sys} is dominated by galactic synchrotron radiation and can be approximated by (Dewdney 2013)

$$T_{\text{sys}} = 60 \times \left(\frac{\nu}{300\,\text{MHz}} \right)^{-2.55}\text{K}. \tag{5}$$

For SKA_Low, we can consider a 1000-h observation time and multiple stacked bands ν with $B = 8\,\text{MHz}$.[5] Here, we note that the noise $N(L, \nu)$ converges quickly with j (as going to higher values of j the signal decays quickly below the thermal noise level), so that $j_{\text{max}} \sim 20$ at $z \sim 8$.

It can be seen in Fig. 1 that the noise for at least SKA2 should be well below the power spectrum over a large range in angular scales. This means that typical modes will be measurable with high S/N and a high fidelity map is possible. The noise crosses the power spectrum at $l \sim 800$ for SKA2 which corresponds to an angular resolution of ~ 10 arcmin. This measurement greatly benefits from the larger collecting area that will come with Phase 2 of the SKA. Such a high fidelity map could be cross-correlated with many other observables from lower redshift to determine how gas and stars follow DM as a function of redshift or to test the predictions of general relativity.

4 WEAK LENSING WITH H I INTENSITY MAPPING AFTER REIONIZATION

After reionization, H I gas exists only in discrete galaxies. This introduces some differences in the calculation of the lensing estimator. As will be seen, the noise per mode will increase from the EoR case, but because planned 21-cm intensity mapping surveys will cover a much larger fraction of the sky the sample variance is much lower and a very good measurement of the convergence power spectrum and its redshift evolution should be possible. The S/N is critically dependent on the array parameters, the area surveyed and the evolution of the H I mass fraction with redshift.

4.1 Measuring lensing with the SKA

In previous work (Pourtsidou & Metcalf 2014), we extended the 21-cm intensity mapping lensing method further, taking into account the discreteness and clustering of galaxies, and investigated the possibility of measuring lensing at intermediate redshifts without resolving (in angular resolution) or even identifying individual sources. Here, we perform an improved analysis of the S/N expected from an SKA-like interferometer, and demonstrate the performance of SKA_Mid Phases 1 and 2. A formal treatment of how a quadratic lensing estimator can be derived for discrete sources is included in Appendices B and C.

The expected measurement error in the power spectrum $C_L^{\delta\theta\delta\theta}$ in a band of width ΔL is given by

$$\Delta C_L^{\delta\theta\delta\theta} = \sqrt{\frac{2}{(2L+1)\Delta L f_{\text{sky}}}} \left(C_L^{\delta\theta\delta\theta} + N_L \right). \tag{6}$$

Here, N_L is the derived $\delta\theta$ estimate reconstruction noise (see Appendix C for the detailed derivation and formulas for this case) which involves the underlying DM power spectrum, the H I density

$\Omega_{\text{HI}}(z)$ as well as the H I mass moments up to fourth order and, of course, the thermal noise of the array C_ℓ^N, equation (3). The system temperature in this case is the sum of the receiver and sky noise (Dewdney 2013)

$$T_{\text{sys}} = \left[40 + 1.1 \times 60 \times \left(\frac{\nu}{300\,\text{MHz}} \right)^{-2.55} \right]\text{K}. \tag{7}$$

The moments of the H I mass function (or 21-cm luminosity function) enter into the estimator's noise expressions (C15) through (C20). Assuming that the H I mass function follows the Schechter function (which is an excellent fit to the local Universe data – see for example, the HIPASS survey results in Zwaan et al. 2003), we can write

$$\frac{\mathrm{d}n(M, z)}{\mathrm{d}M}\mathrm{d}M = \phi^\star(z) \left(\frac{M}{M^\star(z)} \right)^\alpha \exp\left[-\frac{M}{M^\star(z)} \right] \frac{\mathrm{d}M}{M^\star(z)}. \tag{8}$$

Damped Lyman α systems observations are used to measure $\Omega_{\text{HI}}(z)$ up to $z \sim 5$, while no measurements of $\phi^\star(z)$ and $M^\star(z)$ are available other than in the local Universe. However, for this form of the mass function we have the relation

$$\Omega_{\text{HI}}(z) = 4.9 \times 10^{-4} \frac{\phi^\star(z)}{\phi^\star(z=0)} \frac{M^\star(z)}{M^\star(z=0)}. \tag{9}$$

At redshift $z = 0$, we use the values $\alpha = -1.3$, $M^\star = 3.47\,h^{-2}10^9\,\text{M}_\odot$, $\phi^\star = 0.0204\,h^3\,\text{Mpc}^{-3}$ reported from the HIPASS survey (Zwaan et al. 2003). We will initially consider the most conservative scenario, i.e. a no evolution model that we will call Model A. Model A was also used in Pourtsidou & Metcalf (2014). Here, we will also show results using a few H I evolution models.

Let us first concentrate on source redshift $z_s \sim 2$. In the recently published SKA1 Baseline Design (Dewdney 2013), SKA_Mid operates in the frequency range ~ 350–3050 MHz divided in three bands. Our chosen z_s corresponds to frequency $\nu_s = 473$ MHz (Band 1). Following the proposed parameters for SKA_Mid, we consider a 2-yr observation time, $f_{\text{sky}} \sim 0.7$, and we further choose $B = 20\,\text{MHz}$ and $\Delta L = 36$. From equation (B6), we see that the measurement errors increase as $1/\sqrt{f_{\text{sky}}}$ and the fact that SKA_Mid will cover a large fraction of the sky contributes to the high S/N significance we predict. An increase in the observation time would reduce the thermal noise of the interferometer and would also lead to an increase of the overall S/N levels. A similar effect comes from increasing the bandwidth B of the observation – however, note that we avoid choosing very wide bands since in that case there would be non-negligible correlations within the band. It is not clear whether these correlations would just lead to an increase of the noise only or if they would also increase the lensing signal. We plan to investigate this issue in a future publication.

Since the galaxies are being treated as point sources, we wish to include only frequency modes that do not resolve the internal structure of the galaxies. Assuming a typical velocity dispersion for a galaxy at $z = 2$ to be around $200\,\text{km s}^{-1}$, we find $j_{\text{max}} = 63$. However, the noise has converged at $j \sim 40$, which is the value we use in our numerical simulations (see Fig. C1 for a demonstration of the convergence with j). Keeping these values constant, we present a S/N contour plot (Fig. 2) at multipole $L = 100$ for $z_s = 2$ on the $(A_{\text{coll}}, D_{\text{tell}})$ parameter space, and show three SKA_Mid performance cases: SKA_Mid Phase 1 (SKA 1) with 50 per cent sensitivity (SKA0), SKA1 and SKA2. We also present the displacement field power spectrum and measurement errors for $z_s = 2$ and 3 using the SKA2 specifications in Fig. 3. These results are not sensitive to the exact value of ℓ_{min}. For example, the S/N values remain practically unchanged if we use a very small $\ell_{\text{min}} = 10$ or

---

[5] As discussed in Metcalf & White (2009), the quadratic lensing estimator is optimized for the case where the statistical properties of the 21-cm radiation signal and noise are constant within a band, and an observation bandwidth of a few MHz is small enough so that this assumption is justified.
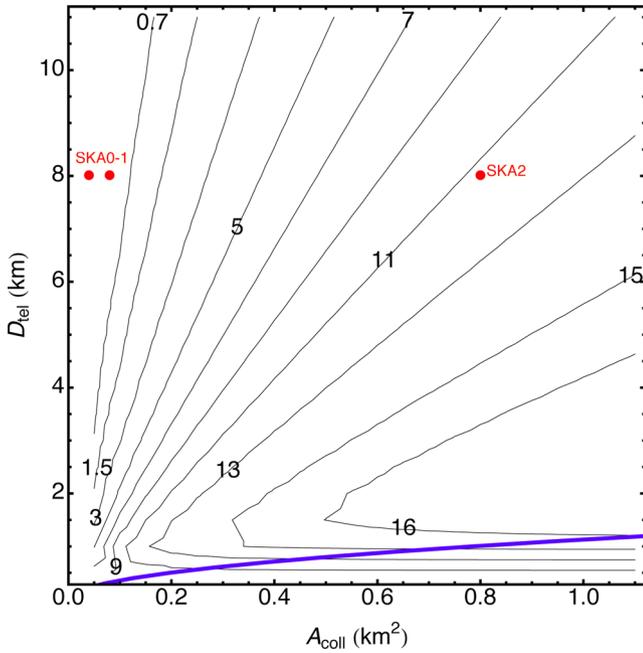
Figure 2. The S/N at $L = 100$ in a $\Delta L = 36$ for various array configurations. Sources are at $z_s = 2$. The contour lines are labelled with the (S/N) values. The area under the thick blue solid line is excluded, since it corresponds to $f_{cover} > 1$. SKA0, SKA1 and SKA2 are shown. The no-evolution in Ω_{H_I} (Model A) scenario is used.
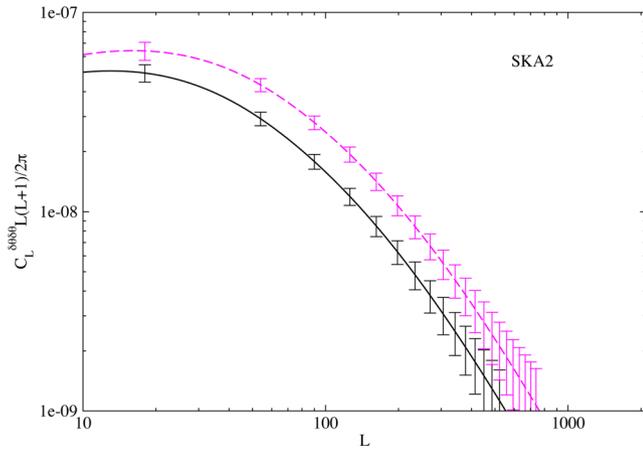

Figure 3. Displacement field power spectrum for $z_s = 2$ (solid black line) and $z_s = 3$ (dashed magenta line) and the corresponding measurement errors using the SKA2 specifications and Model A for the H I mass function.

the one corresponding to the current SKA1 field of view ($\sim$1 deg), $\ell_{min} = 180$.

This technique should enable us to measure the lensing power spectrum at source redshifts well beyond those accessible with more traditional weak lensing surveys based on the shearing of individual galaxy images. The noise is expected to be small enough that evolution in the lensing power spectrum between $z_s = 2$ and 3 will be clearly detectable as illustrated in Fig. 3.

4.2 Importance of the redshift evolution in the H I mass function

It is useful to illustrate how our results depend on a possible evolution of the H I mass function with redshift. We will follow Zhang &
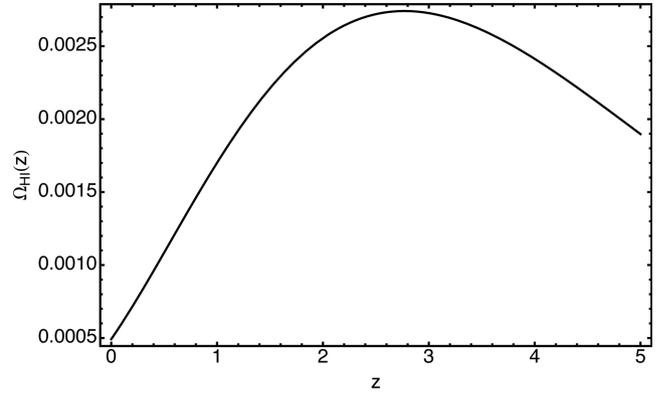

Figure 4. The evolution of $\Omega_{H_I}(z)$ with redshift in Zhang & Pen (2006).

Pen (2006). The authors construct three different evolution models for the H I mass function. Observations of damped Lyman α systems and Lyman α limit systems measure Ω_{H_I} from $z = 0$ to $z \sim 5$ (see for example, Peroux et al. 2003).These observations found that Ω_{H_I} increases by a factor of 5 towards $z \sim 3$ and then decreases towards higher redshift. Combining equation (9) and these observations, one can put constraints on the product of $\phi^\star$ and $M^\star$. Either an increase in $\phi^\star$ or $M^\star$ increases the detectability of lensing, so the no evolution model is considered a conservative choice.

Following (Zhang & Pen 2006), the evolution of Ω_{H_I} will be approximated as

$$g(z) \equiv \frac{\Omega_{H_I}(z)}{\Omega_{H_I}(z = 0)} = (1 + z)^{2.9} \exp(-z/1.3). \tag{10}$$

This function is plotted in Fig. 4 and the constraint for the redshift evolution of the product $\phi^\star M^\star$ is

$$\phi^\star(z)M^\star(z) = \phi^\star(z = 0)M^\star(z = 0)g(z). \tag{11}$$

Here, we will investigate the following evolution scenarios (Zhang & Pen 2006):

  (i) *Model (B)*: No evolution in $M^\star(z)$. $\phi^\star(z) = \phi^\star(z = 0)g(z)$.
  (ii) *Model (C)*: No evolution in $\phi^\star(z)$. $M^\star(z) = M^\star(z = 0)g(z)$.
  (iii) *Model (D)*: $\phi^\star(z)/\phi^\star(z = 0) = M^\star(z)/M^\star(z = 0) = g(z)^{1/2}$.

We also note that H I mass function modelling has been performed by other authors (Abdalla, Blake & Rawlings 2010), where the evolution of the $\Omega_{H_I}(z)$ function agrees with Zhang & Pen (2006) until $z \sim 3$ and so would make very little difference to our predictions.

4.2.1 Model B

Model B assumes no evolution in $M^\star$. The shot-noise terms in the estimator noise are changed in this case. The displacement field power spectrum and measurement errors for Model B are shown in Fig. 5, and the S/N contour at $L = 100$ is shown in Fig. 6 as a function of array parameters. Notice the improvement of the S/N across all of the parameter space from the no-evolution model – the lensing signal can be detected even with SKA0.

4.2.2 Model C

Model C assumes no evolution in $\phi^\star$. Ω_{H_I} will increase, but the shot-noise terms will stay the same as the no-evolution ones (Model A), because the integrals do not get affected by $M^\star$. Hence, the S/N is higher than Model A by a factor of $\sim$1.5 at $L = 100$ using the SKA2
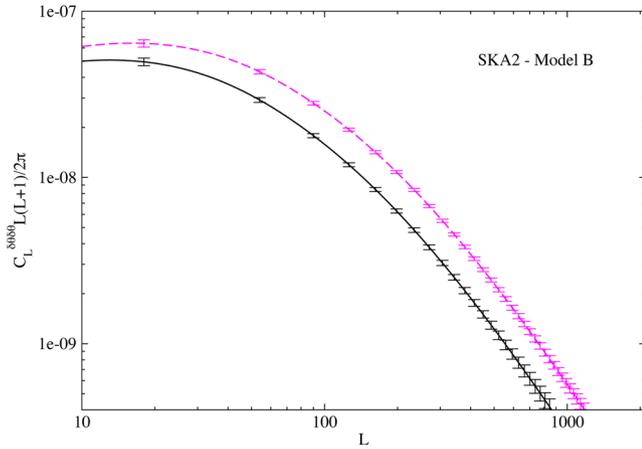
Figure 5. Displacement field power spectrum for $z_s = 2$ (solid black line) and $z_s = 3$ (dashed magenta line) and the corresponding measurement errors using the SKA2 specifications and Model B for the H I mass function.
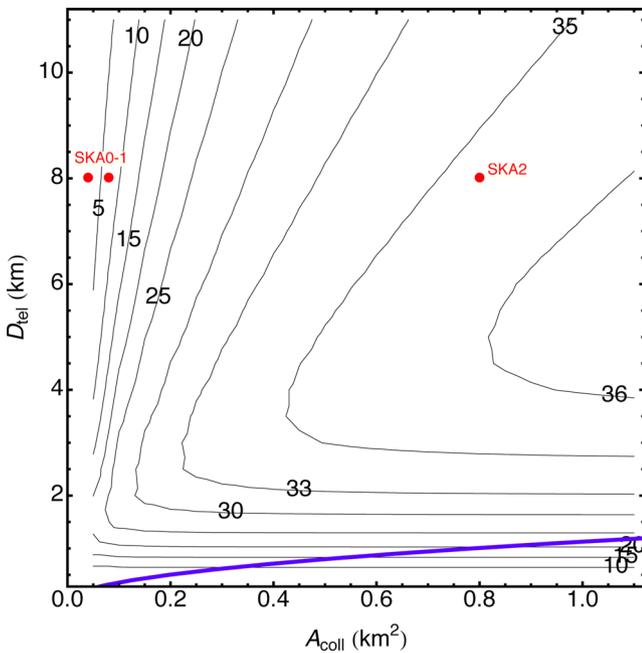

Figure 6. The S/N at $L = 100$ in a $\Delta L = 36$ for various array configurations. Model B for the evolution of the H I mass function is used. Sources are at $z_s = 2$. The contour lines are labelled with the (S/N) values. The area under the thick blue solid line is excluded, since it corresponds to $f_{cover} > 1$. SKA0, SKA1 and SKA2 are shown.

specifications, but lower than Model B by a factor of ~ 0.5 at $L = 100$ using the SKA2 specifications.

4.2.3 Model D

Model D assumes that both $\phi^\star$ and $M^\star$ are evolving as $\phi^\star(z)/\phi^\star(z = 0) = M^\star(z)/M^\star(z = 0) = g(z)^{1/2}$. The S/N is higher than Model C by a factor of ~ 1.5 at $L = 100$ using the SKA2 specifica-tions, but lower than Model B by a factor of ~ 0.7 at $L = 100$ using the SKA2 specifications.

We see that for all models where the H I density increases with redshift the lensing S/N is greater than in the no-evolution model. Furthermore, we learned that there is a stronger dependence on the

evolution of $\phi^\star$ than $M^\star$. This is also expected, since an increase in $\phi^\star$ increases the detectability of lensing – at this point we should stress that in our case the contributions of the Poisson moments contribute both to the signal as well as the noise of our estimator, so the dependence on $\phi^\star$ is crucial. Model B, in which $\phi^\star$ is solely responsible for the increase in $\Omega_{\rm H I}$ is the most optimistic scenario, but the most conservative no-evolution Model A is probably less realistic.

Previously, we used the S/N calculations at a single multipole number ($L = 100$) to demonstrate the capabilities of various arrays (see Figs 2 and 6). A more comprehensive expression would involve contributions from many multipoles. We therefore define a new (S/N) for the detection of the lensing signal as

$$\text{S/N} = \left(\sum_{L_b} \frac{C_{L_b}^2}{\Delta C_{L_b}^2} \right)^{1/2}, \qquad (12)$$

where the sum is over deflection field power spectrum band powers (using $\Delta L = 36$ as before) and ΔC_{L_b} is the forecasted error on each band power. We calculate (S/N) as a function of collecting area and maximum baseline, with $L_{b_{\min}} = 18$ and $L_{b_{\max}} = 990$. We show the S/N contour plots for Models A and B for the H I mass functions in Figs 7 and 8, respectively.

As we have already stressed, Model A is the most conservative, no-evolution scenario one can consider for modelling the H I mass function. From Figs 3 and 7, we can deduce that in order to get a good measurement of the lensing signal up to multipole number $L \sim 1000$ we need a S/N of about 25. Then from Fig. 8, we can see that SKA1 assuming Model B does almost as well as SKA2 assuming no evolution (sources at $z_s = 2$). This shows how crucial the H I mass function evolution is for our method and the SKA's science goals. In Fig. 9, we show the results of our calculations using the SKA1 specifications and Model B – as expected from the

Figure 7. The S/N from equation (12) for various array configurations. Sources are at $z_s = 2$ and we use Model A for the H I mass function. The contour lines are labelled with the (S/N) values. The area under the thick blue solid line is excluded, since it corresponds to $f_{cover} > 1$. SKA0, SKA1 and SKA2 are shown.
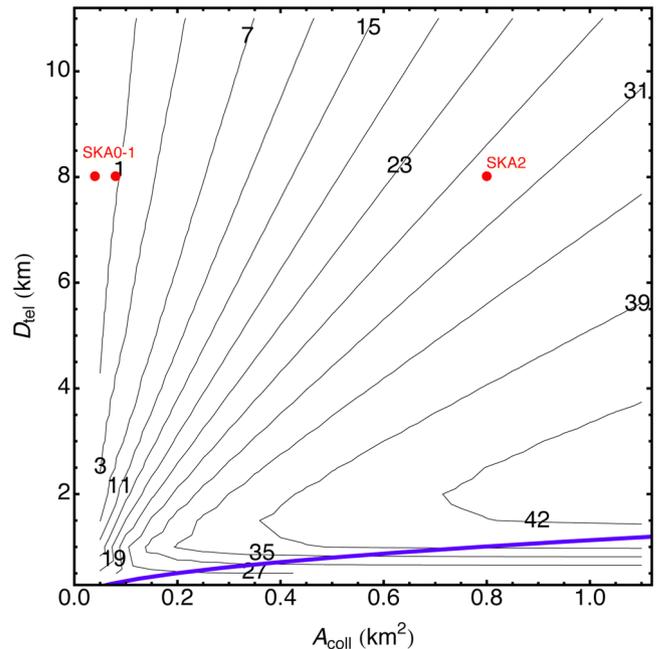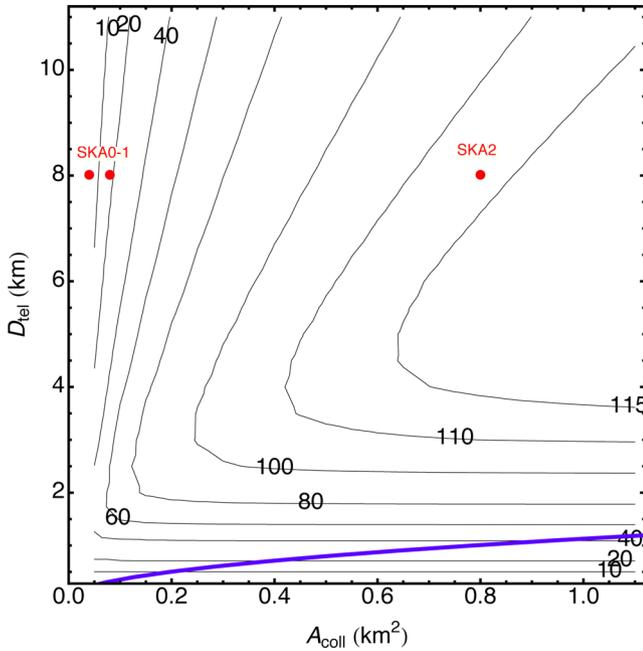
Figure 8. The S/N from equation (12) for various array configurations. Sources are at $z_s = 2$ and we use Model B for the H I mass function. The contour lines are labelled with the (S/N) values. The area under the thick blue solid line is excluded, since it corresponds to $f_{cover} > 1$. SKA0, SKA1 and SKA2 are shown.
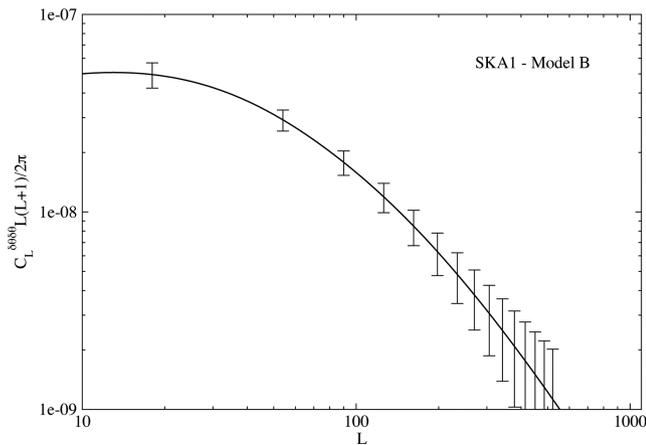

Figure 9. Displacement field power spectrum for $z_s = 2$ and the corresponding measurement errors using the SKA1 specifications and Model B for the H I mass function.

above discussion, the measurement errors are similar to the SKA2 no-evolution case.

5 CONSTRAINING INTERACTING DARK ENERGY MODELS

In this section, we demonstrate the constraining power of 21-cm lensing measurements using the intensity mapping method with an instrument like the SKA. The problems of the nature and evolution of DE and DM are arguably the most important ones in modern cos-

mology, and ambitious future missions like *EUCLID*[6] are dedicated to the exploration and mapping of the dark Universe by investigating the evolution of cosmic structures out to redshifts ∼1.5–2, a period over which DE is thought to dominate the Universe's expansion.

An important question is whether DE is indeed a cosmological constant, like all the available data from CMB measurements (Planck Collaboration XVI 2014), the Hubble constant (Riess et al. 2009) luminosity and distance at high redshift with supernovae Ia (Kowalski et al. 2008) and BAO surveys (Lampeitl et al. 2009) seem to suggest (however, this success comes with the price of the cosmological constant and coincidence problems), an evolving scalar field like quintessence, or a modification of gravity. To differentiate between the plethora of available models, we might have to go even deeper in redshift space $z > 1.5$, and this is where 21 cm radiation becomes a unique probe of cosmology.

Interactions in the dark sector, such as a non-gravitational coupling between DM and DE (for example a coupling between a quintessence field ϕ playing the role of DE and the matter sector, see Amendola 2000), can cause modifications to the background evolution of the matter density and the Hubble parameter, as well as changes in the evolution of structure growth with respect to ΛCDM. However, as we will demonstrate below, the background effects are very difficult to probe as alternative DE and modified gravity models can be very successful in mimicking the ΛCDM background. Therefore, we need to study linear perturbations and quantify the effects of the various models in the CMB and matter power spectra. Modifications in the evolution of structure growth in interacting DE models leave distinctive signatures in the weak lensing signal, and measurements such as those presented in Fig. 3, could be used to constrain such models.

In a recent study, Pourtsidou, Skordis & Copeland (2013) constructed three general classes (types) of models of DE in the form of a scalar field ϕ coupled to CDM. The first class, Type 1, is a generalization of the coupled quintessence (CQ) model suggested by Amendola (2000). In such models, the CDM energy density is not conserved separately (like in the uncoupled case) but there is a non-zero coupling current which represents the decay of DM to DE or vice versa. In the specific Amendola (2000) model, the evolution of the background CDM density $\bar{\rho}_c$ is found to be

$$\bar{\rho}_c = \bar{\rho}_{c,0} a^{-3} e^{\alpha\phi}. \tag{13}$$

We therefore see that the CDM density evolves differently than the uncoupled quintessence case and ΛCDM, which means that for positive coupling parameter α (decay of DM to DE) the CDM density has to be higher in the past in order to reach the same value today.

This model and its variants have been extensively studied in the literature, and constraints on the strength of the coupling parameter (roughly <0.1 assuming it is constant) have been derived using its effects on the CMB and matter power spectra, the growth of structure and the weak lensing signal (e.g. Caldera-Cabral, Maartens & Schaefer 2009; Amendola et al. 2011; De Bernardis et al. 2011; Tarrant et al. 2012; Xia 2013, and references therein).

Here, we will investigate the weak lensing convergence signatures of models belonging to the new classes of coupled DE theories, namely Type 2 and Type 3, constructed in Pourtsidou et al. (2013). Type 2 models involve both energy and momentum transfer between the two components of the dark sector (i.e. like Type 1 models but with a distinctively different coupling mechanism), while Type 3

[6] http://sci.esa.int/euclid/

is a pure momentum transfer theory. The displacement field power spectrum for these models is calculated using (Kaiser 1992)

$$C_L^{\delta\theta\delta\theta} = \frac{9}{L(L+1)c^3} \int_0^{z_s} dz \, [\Omega_m(z)]^2 P(k = L/\mathcal{D}, z)[W(z)]^2$$

$$\times (1+z)^{-4}[H(z)]^3, \tag{14}$$

which simplifies to equation (2) for ΛCDM by setting $\frac{\Omega_m(z)}{\Omega_m} = \frac{H_0^2}{a^3[H(z)]^2}$.

5.1 Type 2 models of CDM coupled to DE

Type 2 models are classified via the Lagrangian function (Pourtsidou et al. 2013)

$$L(n, Y, Z, \phi) = F(Y, \phi) + f(n, Z), \tag{15}$$

where n is the particle number density, $Y = \frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi$ is used to construct a kinetic term for the quintessence field ϕ and $Z = u^\mu\nabla_\mu\phi$ plays the role of a direct coupling of the fluid velocity to the gradient of the scalar field.

We can proceed by making further assumptions. First of all, we want to consider coupled quintessence models, hence we can write $F = Y + V(\phi)$. Secondly, since we choose the scalar field to be coupled to CDM, the function f is separable, i.e. $f = n \, h(Z)$.

The details of the field and fluid equations for the Type 2 class of theories under consideration can be found in Pourtsidou et al. (2013). An important quantity that appears in the cosmological equations is the function $K(Z) = h_z/(h - Zh_z)$, with $h_z = dh/dZ$. We therefore see that we cannot simply choose h to be proportional to Z, i.e. $h(Z) = \beta Z$ with β the (constant) coupling parameter, as in that case $K(Z)$ diverges.

Choosing $h(Z) = Z^\beta$, we find that the background CDM density $\bar{\rho}_c$ evolves as

$$\bar{\rho}_c = \bar{\rho}_{c,0} a^{-3} \bar{Z}^\beta, \tag{16}$$

where $\bar{Z} = -\dot{\bar{\phi}}/a$ (note that the dot denotes derivatives with respect to conformal time). The fact that the CDM density depends on the time derivative $\dot{\bar{\phi}}$ instead of $\bar{\phi}$ itself is a notable difference from the Type 1 (Amendola) case. Note also that the specific form of the coupling function Z^β means that in order to have meaningful solutions $\bar{Z}$ has to be positive throughout the cosmological evolution. In Pourtsidou et al. (2013), the authors derive the background and perturbed Klein–Gordon equations for the evolution of the quintessence field, as well as the perturbed CDM equations for the evolution of the density contrasts and velocities. With these equations at hand, and considering a single exponential potential for the quintessence field, we use a modified version of the CAMB code (Lewis, Challinor & Lasenby 2000) to study the observational signatures of the chosen coupled model. In Figs 10 and 11, we show the Hubble parameter evolution and the CMB TT power spectra for these models together with ΛCDM. We also construct the displacement field power spectrum (Fig. 12) and compare it with the ΛCDM prediction in Fig. 3. Note that the cosmology of each model evolves to the *Planck* cosmological parameter values at $z = 0$ (Ade et al. 2013) and the matter power spectra $P(k)$ have been normalized to CMB fluctuations.

From Fig. 12, we see that the Type 2 model with a coupling parameter $\beta = 0.02$ would be excluded by the proposed lensing/intensity mapping observations. That is because in this case there is energy transfer from DM to DE making the DM density larger in the past compared to the non-interacting case for fixed Ω_m today, the growth is increased, the gravitational potential is higher
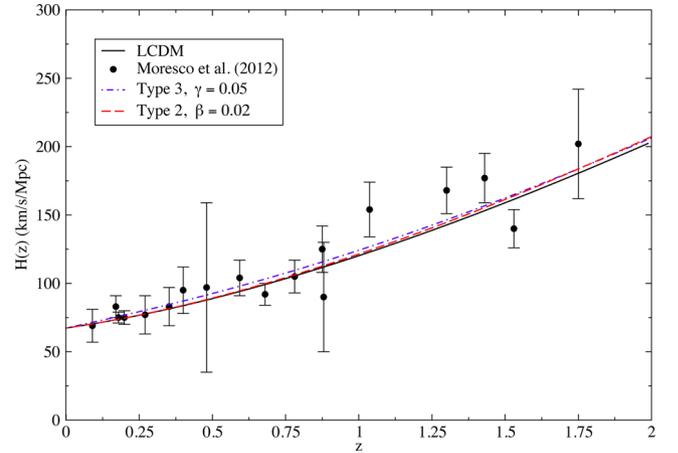

Figure 10. The Hubble parameter versus redshift for ΛCDM (solid black line) and the Types 2 and 3 models under consideration (red dashed line for Type 2, blue dotted dashed line for Type 3), together with expansion history measurements from Moresco et al. (2012a), which combine the observational constraints on the Hubble parameter from Simon, Verde & Jimenez (2005), Stern et al. (2010) and Moresco et al. (2012b).
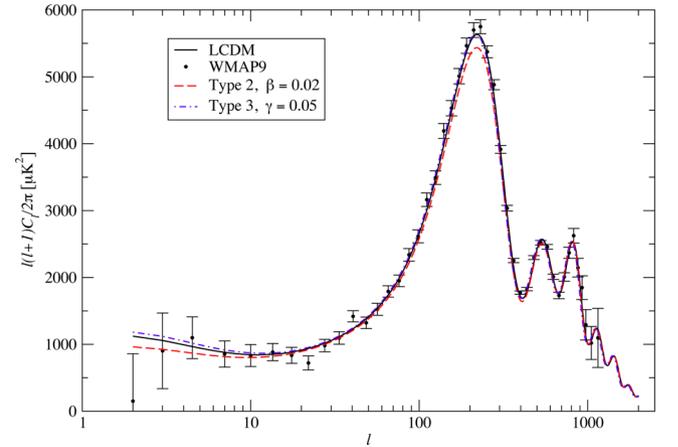

Figure 11. The CMB TT spectra for ΛCDM (solid black line) and the Types 2 and 3 models under consideration (red dashed line for Type 2, blue dotted dashed line for Type 3), together with *WMAP9* measurements.

and the convergence power spectrum is enhanced with respect to ΛCDM. Another important point demonstrated in Fig. 13 is that the Type 2 model linear growth factor is strongly scale dependent, and the difference between the growth of the two scales persists up to the source redshift $z = 2$. Note that there is also an effect on the Hubble parameter evolution (see Fig. 10) and the CMB power spectrum (see Fig. 11), but subtler than the effect on the lensing signal.

5.2 Type 3 models of CDM coupled to DE

Type 3 models are classified via the Lagrangian function (Pourtsidou et al. 2013)

$$L(n, Y, Z, \phi) = F(Y, Z, \phi) + f(n). \tag{17}$$

We can consider a CQ function of the form $F = Y + V(\phi) + h(Z)$, and choose $h(Z) = \gamma Z^2$. Note that this model is physically acceptable for $\gamma < 1/2$, otherwise it suffers from ghost and strong coupling problems (Pourtsidou et al. 2013).
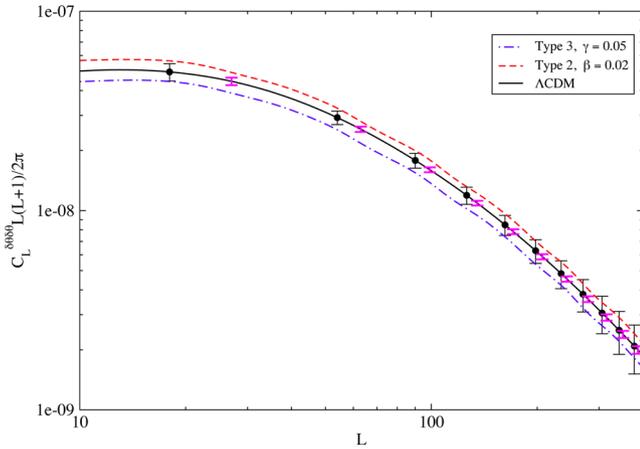
Figure 12. Displacement field power spectrum for ΛCDM (solid black line) compared with Type 2 (red dashed line) and Type 3 (blue dotted dashed line) interacting DE models. Sources are at $z_s = 2$. The circle points with black error bars are the same as in Fig. 3, i.e. using the most conservative scenario (no evolution in the H I mass function), Model A. Assuming a more optimistic scenario like Model B (magenta error bars) would further improve the constraints.
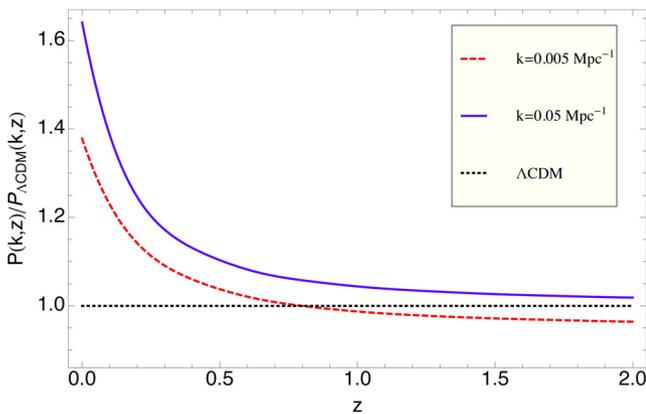

Figure 13. The ratio of the linear matter power spectrum in the Type 2 coupled-DE model to the one in ΛCDM for two characteristic scales as a function of redshift.

As we have already mentioned, Type 3 is a pure momentum transfer theory. The background CDM density $\bar{\rho}_c$ evolves as in the uncoupled case and the energy conservation equation remains uncoupled even at the linear level. Following the same procedure, using the equations derived in Pourtsidou et al. (2013) and the modified version of CAMB, we construct the displacement field power spectrum and compare it with the ΛCDM prediction in Fig. 3.

From Fig. 12, we see that the Type 3 model with a coupling parameter $\gamma = 0.05$ would be excluded. The signature of this model is growth suppression and hence a decrease in the lensing signal. Again, we note that there is also an effect on the Hubble parameter evolution (see Fig. 10) and the CMB power spectrum (see Fig. 11), but subtler than the effect on the lensing signal.

6 CONCLUSIONS

H I intensity mapping is an innovative technique which can be utilized to map the large-scale structure of the Universe in three dimensions. In this paper, we investigated the potential of an SKA-like interferometer to perform accurate measurements of the lensing

signal over a wide range of redshifts using the intensity mapping method. Extending the work in Pourtsidou & Metcalf (2014), we presented our technique for measuring gravitational lensing in H I observations after reionization. Other methods have been proposed based on counting the number of detected galaxies in 21 cm emission (Zhang & Pen 2005, 2006; Zhang & Yang 2011). In such methods, the clustering of galaxies and the shot noise from Poisson fluctuations in the number counts contribute purely to the noise in the lensing estimator. The corresponding errors can be controlled to ∼10–20 per cent level for intermediate source redshifts $z \sim 1$, but for $z_s \gtrsim 2$, the reconstruction is not successful (Zhang & Yang 2011). Our estimator takes into account both the clustering and discreetness of galaxies, which contribute to both the noise and to improving the signal resulting in a significant improvement in performance.

Our calculations showed that the lensing convergence can be measured with high fidelity at redshifts $z \sim 2$–3, hence it can be used for tomographic studies along many redshift (or frequency) bins in order to map, for example, the evolution of the growth function at intermediate redshifts, i.e. higher than those of galaxy shear surveys. We also extended our calculations to include a possible evolution of the H I mass function, showing that the less likely no-evolution scenario is the most conservative one and our claim that a high-S/N measurement can be achieved using this method is robust. To quantify the power of this technique, we performed an optimization study using an SKA-like interferometer and showed that SKA will be able to deliver very good measurements of the lensing signal and its redshift evolution. We also demonstrated the constraining power of our technique by using it to distinguish specific interacting DE models. Our results confirm that the 21 cm technique can be used to compete (and complement) with future galaxy surveys, and SKA in particular can be used to investigate the evolution of DE and DM in order to test the standard ΛCDM paradigm against exotic models with possible observational signatures in higher redshifts than the ones probed by galaxy lensing surveys.

Another exciting prospect is the study of the reionization epoch – we showed that the low-frequency SKA instrument can map the lensing convergence and allow us to 'see' the distribution of DM in a typical region of the sky, something that is only possible with galaxy lensing around very atypical, large galaxy clusters. This would provide a great opportunity to correlate visible objects with mass and test the DM paradigm.

REFERENCES

Abdalla F. B., Blake C., Rawlings S., 2010, MNRAS, 401, 743
Amendola L., 2000, Phys. Rev. D, 62, 043511
Amendola L., Pettorino V., Quercellini C., Vollmer A., 2012, Phys. Rev. D, 85, 103008
Ansari R., Campagne J. E., Colom P., Magneville C., Martin J. M., Moniez M., Rich J., Yeche C., 2012, C. R. Phys., 13, 46
Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, MNRAS, 434, 1239
Bucher M., Carvalho C. S., Moodley K., Remazeilles M., 2012, Phys. Rev. D, 85, 043016

Bull P., Ferreira P. G., Patel P., Santos M. G., 2014, preprint (arXiv:1405.1452)

Caldera-Cabral G., Maartens R., Schaefer B. M., 2009, J. Cosmol. Astropart. Phys., 0907, 027

Chang T., Pen U.-L., Bandura K., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett., 100, 091303

Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463

Chen X., 2012, Int. J. Mod. Phys., 12, 256

Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1

Copeland E. J., Sami M., Tsujikawa S., 2006, Int. J. Mod. Phys. D, 15, 1753

De Bernardis F., Martinelli M., Melchiorri A., Mena O., Cooray A., 2011, Phys. Rev. D, 84, 023504

Dewdney P., 2013, SKA1 System Baseline Design, SKA Project Documents. p. 1

Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181

Hu W., 2001, ApJ, 557, L79

Kaiser N., 1992, ApJ, 522, 272

Kovetz E. D., Kamionkowski M., 2013, Phys. Rev. D, 87, 063516

Kowalski M. et al. (Supernova Cosmology Project Collaboration), 2008, ApJ, 686, 749

Lampeitl H. et al., 2009, MNRAS, 401, 2331

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Limber D. N., 1953, ApJ, 117, 134

Loeb A., Zaldarriaga M., 2004, Phys. Rev. Lett., 92, 211301

McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815

Masui K. W., McDonald P., Pen U.-L., 2010, Phys. Rev. D, 81, 103527

Metcalf R. B., White S. D. M., 2009, MNRAS, 394, 704

Moresco M., Verde L., Pozzetti L., Jimenez R., Cimatti A., 2012a, J. Cosmol. Astropart. Phys., 1207, 053

Moresco M. et al., 2012b, J. Cosmol. Astropart. Phys., 1208, 006

Peroux C., McMahon R. G., Storrie-Lombardi L. J., Irwin M. J., 2003, MNRAS, 346, 1103

Planck Collaboration XVI, 2014, A&A, 571, A16

Pober J. C. et al., 2013, AJ, 145, 65

Pourtsidou A., Metcalf R. B., 2014, MNRAS, 439, L36

Pourtsidou A., Skordis C., Copeland E. J., 2013, Phys. Rev. D, 88, 083505

Riess A. G. et al., 2009, ApJ, 699, 539

Schmidt F., Vallinotto A., Sefusatti E., Dodelson S., 2008, Phys. Rev. D, 78, 043513

Seo H. J., Dodelson S., Marriner J., Mcginnis D., Stebbins A., Stoughton C., Vallinoto A., 2010, ApJ, 721, 164

Simon J., Verde L., Jimenez R., 2005, Phys. Rev. D, 71, 123001

Smoot G. F., Debono I., 2014, preprint (arXiv:1407.3583)

Stern D., Jimenez R., Verde L., Kamionkowski M., Stanford S. A., 2010, J. Cosmol. Astropart. Phys., 1002, 008

Tarrant E. R. M., van de Bruck C., Copeland E. J., Green A. M., 2012, Phys. Rev. D, 85, 023503

Xia J.-Q., 2013, J. Cosmol. Astropart. Phys., 1311, 022

Zahn O., Zaldarriaga M., 2006, ApJ, 653, 922

Zhang P., Pen U. L., 2005, Phys. Rev. Lett., 95, 241302

Zhang P., Pen U. L., 2006, MNRAS, 367, 169

Zhang P., Yang X., 2011, MNRAS, 415, 3485

Zwaan M. A. et al., 2003, AJ, 125, 2842

APPENDIX A: LENSING ESTIMATOR FOR A CONTINUOUS SOURCE

Here, we derive the least squares, or minimum variance, quadratic estimator on a discrete grid for a source that can be approximated as a Gaussian random field. This will be the discrete version of the estimator found by Zahn & Zaldarriaga (2005). The discrete version is necessary for calculations on a finite grid and useful in the deviation of the discrete source case (see Appendix B). This appendix will also serve to introduce some notation.

We write the DFT of the intensity field $I(\boldsymbol{x})$ as

$$I_{\boldsymbol{k}} = \frac{\Omega_{\mathrm{s}}}{N_\perp N_\parallel} \sum_{\boldsymbol{x}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} I(\boldsymbol{x}), \tag{A1}$$

where $\boldsymbol{k} = (\boldsymbol{\ell}, j)$, $\boldsymbol{x} = (\boldsymbol{\theta}, z)$ and $\Omega_{\mathrm{s}} = \Theta_{\mathrm{s}} \times \Theta_{\mathrm{s}}$ for a square survey geometry. We also have

$$I(\boldsymbol{x}) = \frac{1}{\Omega_{\mathrm{s}}} \sum_{\boldsymbol{k}} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} I_{\boldsymbol{k}}. \tag{A2}$$

Here, $N_\parallel$ is the number of cells in the direction parallel to the line of sight and $N_\perp$ is the number of cells on a plane perpendicular to the line of sight. We use a flat sky or small angle approximation.

The correlation between discrete modes is

$$\langle I_{\boldsymbol{k}} I_{\boldsymbol{k}'}^* \rangle = \Omega_{\mathrm{s}}^2 \delta_{\boldsymbol{k}\boldsymbol{k}'}^K P_{\boldsymbol{k}}, \tag{A3}$$

where $P_{\boldsymbol{k}}$ is the discrete power spectrum. The Kronecker delta implies statistical homogeneity. The discrete power spectrum is related to the continuous one, $P(k)$, by

$$P_{\boldsymbol{k}} = \frac{P(k)}{V_{\mathrm{s}}} = \frac{P(k)}{\Omega_{\mathrm{s}}\mathcal{D}^2\mathcal{L}} = \frac{1}{\Omega_{\mathrm{s}}} C_{\ell,j}. \tag{A4}$$

Here, $\mathcal{D}$ is the comoving angular size distance to the source volume from the observer. The angular Fourier coordinate is $\boldsymbol{\ell} = \mathcal{D}\boldsymbol{k}_\perp$ and j denotes the DFT coordinate in the radial direction. The comoving length of the source volume in the radial direction is $\mathcal{L}$. This serves to define the angular power spectrum in angle and radial coordinates $C_{\ell,j}$. So in terms of the angular power spectrum

$$\langle I_{\boldsymbol{k}} I_{\boldsymbol{k}'}^* \rangle = \Omega_{\mathrm{s}}\, C_{\ell,j}\, \delta_{\ell,\ell'}^K \delta_{jj'}^K. \tag{A5}$$

Gravitational lensing causes the observed emission field to be inhomogeneous over a region of the sky with coherent deflection. The observed intensity after lensing is

$$\tilde{I}(\boldsymbol{\theta}, x_\parallel) = I\left(\boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), x_\parallel\right) \tag{A6}$$

$$\simeq I\left(\boldsymbol{\theta}, x_\parallel\right) - \boldsymbol{\alpha}(\boldsymbol{\theta}) \cdot \nabla_\theta I\left(\boldsymbol{\theta}, x_\parallel\right) \tag{A7}$$

$$\simeq I\left(\boldsymbol{\theta}, x_\parallel\right) + \nabla_\theta \Psi(\boldsymbol{\theta}) \cdot \nabla_\theta I\left(\boldsymbol{\theta}, x_\parallel\right), \tag{A8}$$

where $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is the deflection caused by lensing and $\Psi(\boldsymbol{\theta})$ is the lensing potential. The deflection field is a potential field to very good approximation. In Fourier space, this becomes

$$\tilde{I}_{\boldsymbol{\ell},j} \simeq I_{\boldsymbol{\ell},j} + \frac{1}{\Omega_{\mathrm{s}}} \sum_{\boldsymbol{\ell}'} \boldsymbol{\ell}' \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}') I_{\boldsymbol{\ell}',j} \Psi_{\boldsymbol{\ell}-\boldsymbol{\ell}'}. \tag{A9}$$

From this, we can find the correlation between modes to first order

$$\langle \tilde{I}_{\boldsymbol{\ell},j} \tilde{I}^*_{\boldsymbol{\ell}-\boldsymbol{L},j'} \rangle = \Omega_{\mathrm{s}} C_{\boldsymbol{\ell},j} \delta^K_{\boldsymbol{L},0} \delta^K_{jj'} + \delta^K_{jj'} [\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\boldsymbol{\ell},j} + \boldsymbol{L} \cdot (\boldsymbol{L}-\boldsymbol{\ell}) C_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}] \Psi_{\boldsymbol{L}}. \tag{A10}$$

We seek to construct a quadratic estimator of the form

$$\hat{\Psi}(\boldsymbol{L}) = \sum_j \sum_{\boldsymbol{\ell}} g(\boldsymbol{\ell}, \boldsymbol{L}, j)\, \tilde{I}_{\boldsymbol{\ell},j} \tilde{I}^*_{\boldsymbol{\ell}-\boldsymbol{L},j}. \tag{A11}$$

Putting equation (A10) into (A11) and requiring the estimator to be unbiased gives the constraint

$$\sum_j \sum_{\boldsymbol{\ell}} g(\boldsymbol{\ell}, \boldsymbol{L}, j)[\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\boldsymbol{\ell},j} + \boldsymbol{L} \cdot (\boldsymbol{L}-\boldsymbol{L}) C_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}] = 1. \tag{A12}$$

The variance of this estimator is

$$\mathcal{V} = \sum_j \sum_{j'} \sum_{\boldsymbol{\ell}} \sum_{\boldsymbol{\ell}'} g(\boldsymbol{\ell}, \boldsymbol{L}, j) g^*(\boldsymbol{\ell}', \boldsymbol{L}, j') \langle I_{\boldsymbol{\ell},j} I^*_{|\boldsymbol{\ell}-\boldsymbol{L}|,j} I^*_{\boldsymbol{\ell}',j'} I_{|\boldsymbol{\ell}'-\boldsymbol{L}|,j'} \rangle$$

$$= 2\Omega_{\mathrm{s}}^2 \sum_j \sum_{\boldsymbol{\ell}} g^2(\boldsymbol{\ell}, \boldsymbol{L}, j) C^{\mathrm{tot}}_{\boldsymbol{\ell},j} C^{\mathrm{tot}}_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}, \tag{A13}$$

where in calculating the fourth-order correlations the result for a Gaussian random field has been used and

$$C^{\mathrm{tot}}_{\boldsymbol{\ell},j} = C_{\boldsymbol{\ell},j} + C^{\mathrm{N}}_{\boldsymbol{\ell}}, \tag{A14}$$

with $C^{\mathrm{N}}_{\boldsymbol{\ell}}$ the thermal noise of the array.

With the standard Lagrangian multiplier technique $\mathcal{V}$ can be minimized subject to the constraint (A12). The result is

$$g(\tilde{\boldsymbol{\ell}}, \boldsymbol{L}, \tilde{j}) = A_{\mathrm{R}} \frac{[\tilde{\boldsymbol{\ell}} \cdot \boldsymbol{L} C_{\tilde{\boldsymbol{\ell}},\tilde{j}} + \boldsymbol{L} \cdot (\boldsymbol{L}-\tilde{\boldsymbol{\ell}}) C_{|\tilde{\boldsymbol{\ell}}-\boldsymbol{L}|,\tilde{j}}]}{\Omega_{\mathrm{s}}^2 C^{\mathrm{tot}}_{\tilde{\boldsymbol{\ell}},\tilde{j}} C^{\mathrm{tot}}_{|\tilde{\boldsymbol{\ell}}-\boldsymbol{L}|,\tilde{j}}}, \quad A_{\mathrm{R}} = \left[\sum_j \sum_{\boldsymbol{\ell}} \frac{[\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\boldsymbol{\ell},j} + \boldsymbol{L} \cdot (\boldsymbol{L}-\boldsymbol{\ell}) C_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}]^2}{\Omega_{\mathrm{s}}^2 C^{\mathrm{tot}}_{\boldsymbol{\ell},j} C^{\mathrm{tot}}_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}} \right]^{-1}. \tag{A15}$$

The variance of this estimator is

$$\mathcal{V} = 2 \left[\sum_j \sum_{\boldsymbol{\ell}} \frac{[\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\boldsymbol{\ell},j} + \boldsymbol{L} \cdot (\boldsymbol{L}-\boldsymbol{\ell}) C_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}]^2}{\Omega_{\mathrm{s}}^2 C^{\mathrm{tot}}_{\boldsymbol{\ell},j} C^{\mathrm{tot}}_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}} \right]^{-1}. \tag{A16}$$

Now, we can verify the agreement of this estimator with previous work by going to continuous Fourier space using

$$\sum_{\boldsymbol{\ell}} \to \Omega_{\mathrm{s}} \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} \tag{A17}$$

that gives

$$\mathcal{V} = \Omega_{\mathrm{s}} \left[\sum_j \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} \frac{[\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\boldsymbol{\ell},j} + \boldsymbol{L} \cdot (\boldsymbol{L}-\boldsymbol{\ell}) C_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}]^2}{2 C^{\mathrm{tot}}_{\boldsymbol{\ell},j} C^{\mathrm{tot}}_{|\boldsymbol{\ell}-\boldsymbol{L}|,j}} \right]^{-1}, \tag{A18}$$

which with the identification $\Omega_{\mathrm{s}} \to (2\pi)^2 \delta(0)$ is exactly the Zahn & Zaldarriaga (2005) result.

The behaviour of the lensing estimator and noise in the Gaussian approximation has been analysed in Zahn & Zaldarriaga (2005). Higher values of the discretized parallel wave vector $k_{\parallel}$ (i.e. higher values of j) mean higher values of the three-dimensional power spectrum $P(k)$ for each C_{ℓ} value. $P(k)$ is monotonically decreasing on all scales of interest and therefore going to higher values of j the effect of $C_{\ell,j}$ becomes negligible and the signal decays quickly below the thermal noise level. Hence, only a few modes contribute to the estimator (Zahn & Zaldarriaga 2005).

APPENDIX B: LENSING ESTIMATOR FOR UNCLUSTERED DISCRETE SOURCES

Here, we consider a collection of sources with random positions and derive a lensing estimator. The expansion used in the previous section to describe the action of lensing on the surface brightness (equation A7) is not formally valid here because the deflection angle may be large compared to the size of individual sources. For this reason, we derive the estimator by a different method in this case which also gives some insight into how the estimator works.

The fluctuation in the number of sources with luminosity L in a cell i in real space will be

$$\delta n_i^L = n_i^L - \bar{n}_i, \tag{B1}$$

where the average cell occupation is

$$\bar{n}_i = \bar{\eta}\delta V, \tag{B2}$$

$\bar{\eta}$ being the average number density and δV the volume of a cell.

The surface brightness fluctuation $\delta S(i)$ can be written as

$$\delta S(i) = \sum_L \delta n_i^L L. \tag{B3}$$

The continuous limit in luminosity will be taken later. The DFT of the above expression considering the full 3D case with cells $i = (i_\perp, i_\parallel)$ ($i_\perp$ represents the two dimensions perpendicular to the line of sight) is

$$\delta\tilde{S}(j) = \frac{\Omega_s}{N_\perp N_\parallel} \sum_{i_\parallel} \sum_{i_\perp} \delta S(i) e^{i\frac{2\pi i_\perp j_\perp}{N_\perp}} e^{i\frac{2\pi i_\parallel j_\parallel}{N_\parallel}} \tag{B4}$$

$$= \frac{\Omega_s}{N_\perp N_\parallel} \sum_{i_\parallel} \sum_{i_\perp} \sum_L L\delta n_i^L e^{i\frac{2\pi i_\perp j_\perp}{N_\perp}} e^{i\frac{2\pi i_\parallel j_\parallel}{N_\parallel}}. \tag{B5}$$

For the average flux in a cell $\bar{S}$, we will have

$$\bar{S} = \bar{n}_i \langle L \rangle = \bar{\eta}\delta V \langle L \rangle. \tag{B6}$$

Now let us look at correlations between modes:

$$\langle \delta\tilde{S}(j)\delta\tilde{S}^*(j-m) \rangle = \frac{\Omega_s^2}{(N_\perp N_\parallel)^2} \left\langle \sum_{i_\parallel} \sum_{i_\perp} \sum_L L\delta n_i^L e^{i\frac{2\pi i_\perp j_\perp}{N_\perp}} e^{i\frac{2\pi i_\parallel j_\parallel}{N_\parallel}} \sum_{i_\parallel'} \sum_{i_\perp'} \sum_{L'} L'\delta n_{i'}^{L'} e^{-i\frac{2\pi i_\perp'(j_\perp-m_\perp)}{N_\perp}} e^{-i\frac{2\pi i_\parallel'(j_\parallel-m_\parallel)}{N_\parallel}} \right\rangle.$$

Since $\langle \delta n_i^L \rangle = 0$, only the $i = i'$, $L = L'$ terms contribute:

$$\langle \delta\tilde{S}(j)\delta\tilde{S}^*(j-m) \rangle = \frac{\Omega_s^2}{(N_\perp N_\parallel)^2} \left\langle \sum_{i_\parallel} \sum_{i_\perp} \sum_L L^2(\delta n_i^L)^2 e^{i\frac{2\pi i_\perp m_\perp}{N_\perp}} e^{i\frac{2\pi i_\parallel m_\parallel}{N_\parallel}} \right\rangle. \tag{B7}$$

The second moment of the number counts in a cell is given by a Poisson distribution $\langle(\delta n_i)^2\rangle = \bar{n}_i$. If a cell of fixed angular size is magnified by a factor μ_i, the galaxies within it will be a factor μ_i brighter. At the same time, the true volume of that cell will be a factor of $1/\mu_i$ smaller and so the average number of galaxies will go down by the same factor. The result is

$$\langle \delta\tilde{S}(j)\delta\tilde{S}^*(j-m) \rangle = \frac{\Omega_s^2}{(N_\perp N_\parallel)^2} \left\langle \sum_{i_\parallel} \sum_{i_\perp} \sum_L \mu_{i_\perp} L^2 \bar{n}_i e^{i\frac{2\pi i_\perp m_\perp}{N_\perp}} e^{i\frac{2\pi i_\parallel m_\parallel}{N_\parallel}} \right\rangle \tag{B8}$$

$$= \frac{\Omega_s^2}{(N_\perp N_\parallel)^2} \bar{\eta}\delta V \langle L^2 \rangle \left\langle \sum_{i_\perp} \mu_{i_\perp} e^{i\frac{2\pi i_\perp m_\perp}{N_\perp}} \right\rangle \left\langle \sum_{i_\parallel} e^{i\frac{2\pi i_\parallel m_\parallel}{N_\parallel}} \right\rangle \tag{B9}$$

$$= \frac{\Omega_s^2}{N_\perp N_\parallel} \bar{\eta}\delta V \langle L^2 \rangle \tilde{\mu}(m_\perp)\delta_{m_\parallel}^K, \tag{B10}$$

where $\tilde{\mu}(m_\perp)$ is the DFT of the magnification.

Dividing by the average flux $\bar{S}$ from equation (B6) in order to obtain the dimensionless fluctuations, we find

$$\langle \Delta\tilde{S}(j)\Delta\tilde{S}^*(j-m_\perp) \rangle = \frac{\langle \delta\tilde{S}(j)\delta\tilde{S}^*(j-m_\perp) \rangle}{\bar{S}^2}$$

$$= \frac{\Omega_s^2}{N_\perp N_\parallel} \frac{1}{\bar{\eta}\delta V} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \tilde{\mu}(m_\perp) \tag{B11}$$

$$= \frac{\Omega_s^2}{\bar{N}_g} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \tilde{\mu}(m_\perp) \tag{B12}$$

$$= \Omega_s C^{\text{shot}} \tilde{\mu}(m_\perp), \tag{B13}$$

where we have used $\delta V = \frac{V_{\text{tot}}}{N_\perp N_\parallel}$ and $\bar{N}_g = \bar{\eta}V_{\text{tot}}$. Equation (B13) defines C^{shot}. The moments of the luminosity function can now be calculated for a continuous distribution of luminosities.
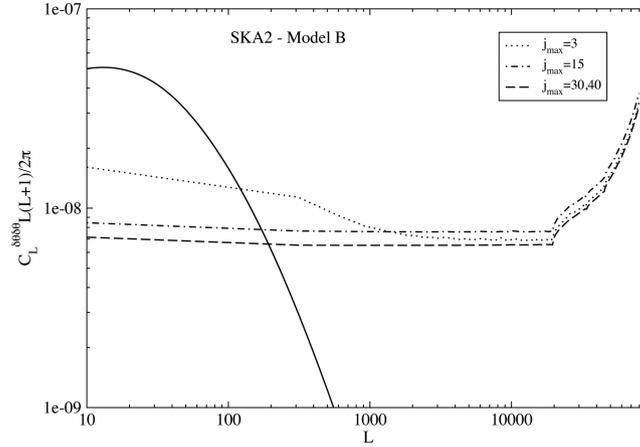
Figure C1. Displacement field power spectrum for $z_s = 2$ and the corresponding lensing reconstruction noise $N(L)$ for different values of j_{max} using the SKA2 specifications and Model B for the H I mass function.

We construct the quadratic estimator $\hat{\mu}(m_\perp)$ as

$$\hat{\mu}(m_\perp) = \frac{1}{\Omega_s C^{shot}} \frac{1}{N_\perp N_\parallel} \sum_{j_\perp} \sum_{j_\parallel} \Delta\tilde{S}(j)\Delta\tilde{S}^*(j - m_\perp). \tag{B14}$$

In this case (with no thermal noise or clustering), the optimal filter ($g(\boldsymbol{\ell}, \boldsymbol{L}, j)$ in the previous section) is a function of only $\boldsymbol{m}_\perp$ and its value is given by the requirement that $\langle\hat{\mu}(m_\perp)\rangle = \tilde{\mu}(m_\perp)$. The variance of this estimator is

$$\mathcal{V} = \langle|\mu(\hat{m}_\perp)|^2\rangle = \langle\hat{\mu}(m_\perp)\hat{\mu}^*(m_\perp)\rangle$$

$$= \frac{1}{(N_\perp N_\parallel)^2} \frac{1}{(\Omega_s C^{shot})^2} \sum_{j_\perp} \sum_{j_\parallel} \sum_{j'_\perp} \sum_{j'_\parallel} \langle\Delta\tilde{S}(j)\Delta\tilde{S}^*(j - m_\perp)\Delta\tilde{S}^*(j')\Delta\tilde{S}(j' - m_\perp)\rangle \tag{B15}$$

$$= \frac{\Omega_s^2}{\bar{N}_g} \frac{\langle L^4\rangle}{\langle L^2\rangle^2} \left(1 + 3\frac{\bar{N}_g}{N_\perp N_\parallel}\right) + 2\Omega_s^2 \frac{N_\perp N_\parallel - 1}{(N_\perp N_\parallel)^2}. \tag{B16}$$

In deriving equation (B16), we have used the higher moments of the Poisson number counts in cells. Note that in the limit $N_\perp, N_\parallel \to \infty$ only the first term survives. In the notation of the previous section $\bar{N}_g = \bar{\eta}\Omega_s\mathcal{D}^2\mathcal{L}$, where $\bar{\eta}$ is the density of galaxies. Note that for 21 cm emission the luminosity is proportional to the H I density so the luminosity moments are also the moments of the H I mass function.

In the weak lensing limit $\mu \simeq 1 + 2\kappa = 1 - \nabla^2\Psi$. So the estimator for the lensing potential is related to the estimator for magnification by

$$\hat{\Psi}(\boldsymbol{L}) = |\boldsymbol{L}|^{-2}\hat{\mu}(\boldsymbol{L}). \tag{B17}$$

As it turns out, this filter is the same one we would have found by plugging a constant power spectrum into the optimal Gaussian filter (A15), but the noise is different.

APPENDIX C: LENSING ESTIMATOR FOR CLUSTERED DISCRETE SOURCES

In this appendix, the results of the two previous appendices are combined to find a lensing estimator and its expected noise in the case of clustered point sources. For the two-point correlation function, we have

$$\langle I(\boldsymbol{x})I(\boldsymbol{x}')\rangle = \frac{1}{\bar{\eta}\delta V} \frac{\langle L^2\rangle}{\langle L\rangle^2}\delta^K_{\boldsymbol{x}\boldsymbol{x}'} + \xi_{\boldsymbol{x}\boldsymbol{x}'}, \tag{C1}$$

where the first part comes from the Poisson fluctuations in the number counts and the second from the clustering. In Fourier space

$$\langle I_{\boldsymbol{k}}I_{\boldsymbol{k}'}^*\rangle = \Omega_s (C_{\ell,j} + C^{shot}) \,\delta^K_{\ell,\ell'}\delta^K_{jj'}. \tag{C2}$$

After lensing the observed correlations will be

$$\langle\tilde{I}_{\ell,j}\tilde{I}_{\ell-L,j'}^*\rangle = \delta^K_{jj'}[\boldsymbol{\ell}\cdot\boldsymbol{L}C_{\ell,j} + \boldsymbol{L}\cdot(\boldsymbol{L}-\boldsymbol{\ell})C_{|\ell-L|,j} + L^2 \, C^{shot}] \,\Psi(\boldsymbol{L})$$

to first order. We will not attempt to find an optimal estimator here. Instead, we use an estimator of the form

$$\hat{\Psi}(\boldsymbol{L}) = f(\boldsymbol{L})\sum_j \sum_\ell \tilde{I}_{\ell,j}\tilde{I}_{\ell-L,j}^*, \tag{C3}$$

where $f(L)$ is a normalization. This would be optimal were the sources unclustered. For an unbiased estimator, we require that $\langle \hat{\Psi}(L) \rangle = \Psi(L)$ so we find

$$f(L) = \left\{ \sum_j \sum_\ell [\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\ell,j} + \boldsymbol{L} \cdot (\boldsymbol{L} - \boldsymbol{\ell}) C_{|\boldsymbol{\ell}-L|,j} + L^2 \, C^{\text{shot}}] \right\}^{-1}$$

$$= \left\{ \sum_j \sum_\ell [\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\ell,j} + \boldsymbol{L} \cdot (\boldsymbol{L} - \boldsymbol{\ell}) C_{|\boldsymbol{\ell}-L|,j}] + (N_\perp N_\parallel) L^2 \, C^{\text{shot}} \right\}^{-1}. \tag{C4}$$

The variance of the estimator $\mathcal{V} = \langle \hat{\Psi}(L) \hat{\Psi}^*(L) \rangle$ is given by

$$\mathcal{V} = [f(L)]^2 \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \langle I_{\ell,j} I^*_{\ell-L,j} I^*_{\ell',j'} I_{\ell'-L,j'} \rangle, \tag{C5}$$

with

$$\langle I_{\ell,j} I^*_{\ell-L,j} I^*_{\ell',j'} I_{\ell'-L,j'} \rangle = \Omega_s \frac{1}{\bar{\eta}^3} \frac{1}{(D^2 \mathcal{L})^3} \frac{\langle L^4 \rangle}{\langle L \rangle^4} \left(1 + 3 \frac{\bar{N}_g}{N_\perp N_\parallel} \right)$$

$$+ \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^3 \rangle}{\langle L \rangle^3} \left(\frac{N_\perp N_\parallel - 1}{N_\perp N_\parallel} \right) [C_{|\ell'-L|,j'} + C_{\ell',j'} + C_{|\ell-L|,j} + C_{\ell,j}]$$

$$+ \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} \left(\frac{N_\perp N_\parallel - 1}{N_\perp N_\parallel} \right) [C_{L,0} + C_{|\ell-\ell'|,|j-j'|} + C_{|\ell+\ell'-L|,j+j'}]$$

$$+ \Omega_s^2 \frac{1}{\bar{\eta}} \frac{1}{(D^2 \mathcal{L})} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \frac{(N_\perp N_\parallel - 1)(N_\perp N_\parallel - 2)}{(N_\perp N_\parallel)^2}$$

$$\times [C_{\ell',j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'} + C_{L-\ell',-j'} \delta^K_{\ell,L-\ell'} \delta^K_{j,-j'} \delta^K_{\ell,L-\ell'} \delta^K_{j,-j'}$$

$$+ C_{\ell',j'} \delta^K_{L-\ell,\ell'} \delta^K_{j',-j} \delta^K_{L-\ell,\ell'} \delta^K_{j',-j} + C_{\ell'-L,j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'}]$$

$$+ \Omega_s^2 \frac{(N_\perp N_\parallel - 1)(N_\perp N_\parallel - 2)(N_\perp N_\parallel - 3)}{(N_\perp N_\parallel)^3}$$

$$\times [C_{\ell,j} C_{|\ell'-L|,j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'} \delta^K_{\ell,\ell'} \delta^K_{j,j'} + C_{\ell,j} C_{\ell',j'} \delta^K_{\ell,L-\ell'} \delta^K_{j,-j'} \delta^K_{\ell,L-\ell'} \delta^K_{j,-j'}]. \tag{C6}$$

Here, we have assumed that the clustering and the noise are Gaussian.

We will now calculate the sum over ℓ, ℓ', j, j' in equation (C5) separately for each of the terms of the fourth moment (C6). We will also consider the limit $N_\parallel \to \infty$ to simplify the constant factors. Note that, to include the thermal noise of the array in our calculation, we just send $C_{\ell,j} \to C^{\text{tot}}_{\ell,j}$, with

$$C^{\text{tot}}_{\ell,j} = C_{\ell,j} + C^N \tag{C7}$$

with C^N the power spectrum of the thermal noise. We have

$$\mathcal{V} = [f(L)]^2 (\mathcal{I}_0 + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4), \tag{C8}$$

with

$$\mathcal{I}_0 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s \frac{1}{\bar{\eta}^3} \frac{1}{(D^2 \mathcal{L})^3} \frac{\langle L^4 \rangle}{\langle L \rangle^4}$$

$$= (N_\perp N_\parallel)^2 \, \Omega_s \frac{1}{\bar{\eta}^3} \frac{1}{(D^2 \mathcal{L})^3} \frac{\langle L^4 \rangle}{\langle L \rangle^4}. \tag{C9}$$

$$\mathcal{I}_1 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^3 \rangle}{\langle L \rangle^3} [C^{\text{tot}}_{|\ell'-L|,j'} + C^{\text{tot}}_{\ell',j'} + C^{\text{tot}}_{|\ell-L|,j} + C^{\text{tot}}_{\ell,j}]$$

$$= \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^3 \rangle}{\langle L \rangle^3} \times (N_\perp N_\parallel) \sum_j \sum_\ell [2 C^{\text{tot}}_{\ell,j} + 2 C^{\text{tot}}_{|\ell-L|,j}]. \tag{C10}$$

$$\mathcal{I}_2 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} [C^{\text{tot}}_{L,0} + C^{\text{tot}}_{|\ell-\ell'|,|j-j'|} + C^{\text{tot}}_{|\ell+\ell'-L|,j+j'}]$$

$$= \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} (N_\perp N_\parallel)^2 C^{\text{tot}}_{L,0}$$

$$+ \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2 \mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} [C^{\text{tot}}_{|\ell-\ell'|,|j-j'|} + C^{\text{tot}}_{|\ell+\ell'-L|,j+j'}]. \tag{C11}$$

$$\mathcal{I}_3 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s^2 \frac{1}{\bar{\eta}} \frac{1}{(D^2\mathcal{L})} \frac{\langle L^2 \rangle}{\langle L \rangle^2}$$

$$\times [C_{\ell',j'}^{\text{tot}} \delta_{\ell,\ell'}^K \delta_{j,j'}^K \delta_{\ell,\ell'}^K \delta_{j,j'}^K + C_{L-\ell',-j'}^{\text{tot}} \delta_{\ell,L-\ell'}^K \delta_{j,-j'}^K \delta_{\ell,L-\ell'}^K \delta_{j,-j'}^K$$

$$+ C_{\ell',j'}^{\text{tot}} \delta_{L-\ell,\ell'}^K \delta_{j',-j}^K \delta_{L-\ell,\ell'}^K \delta_{j',-j}^K + C_{\ell'-L,j}^{\text{tot}} \delta_{\ell,\ell'}^K \delta_{j,j'}^K \delta_{\ell,\ell'}^K \delta_{j,j'}^K]$$

$$= \Omega_s^2 \frac{1}{\bar{\eta}} \frac{1}{(D^2\mathcal{L})} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \sum_j \sum_\ell [2C_{\ell,j}^{\text{tot}} + 2C_{|\ell-L|,j}^{\text{tot}}]. \tag{C12}$$

$$\mathcal{I}_4 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s^2 [C_{\ell,j}^{\text{tot}} C_{|\ell'-L|,j'}^{\text{tot}} \delta_{\ell,\ell'}^K \delta_{j,j'}^K \delta_{\ell,\ell'}^K \delta_{j,j'}^K$$

$$+ C_{\ell,j}^{\text{tot}} C_{\ell',j'}^{\text{tot}} \delta_{\ell,L-\ell'}^K \delta_{j,-j'}^K \delta_{\ell,L-\ell'}^K \delta_{j,-j'}^K]$$

$$= \Omega_s^2 \sum_j \sum_\ell 2C_{\ell,j}^{\text{tot}} C_{|\ell-L|,j}^{\text{tot}}. \tag{C13}$$

From appendices A and B, we can recognize equation (C9) as a pure Poisson clustering term and equation (C13) as a pure Gaussian clustering term. The other terms are from the interplay of these contributions.

To calculate these terms, it is convenient to go into continuous ℓ-space where the integrals can be done numerically. We have

$$f(\boldsymbol{L}) = \left\{ \sum_j \sum_\ell [\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\ell,j} + \boldsymbol{L} \cdot (\boldsymbol{L} - \boldsymbol{\ell}) C_{|\ell-L|,j} + L^2 C^{\text{shot}}] \right\}^{-1}$$

$$= \frac{1}{\Omega_s} \left\{ \sum_j \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} [\boldsymbol{\ell} \cdot \boldsymbol{L} C_{\ell,j} + \boldsymbol{L} \cdot (\boldsymbol{L} - \boldsymbol{\ell}) C_{|\ell-L|,j} + L^2 C^{\text{shot}}] \right\}^{-1}. \tag{C14}$$

$$\mathcal{I}_0 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s \frac{1}{\bar{\eta}^3} \frac{1}{(D^2\mathcal{L})^3} \frac{\langle L^4 \rangle}{\langle L \rangle^4}$$

$$= (N_\parallel)^2 \Omega_s^3 \frac{1}{\bar{\eta}^3} \frac{1}{(D^2\mathcal{L})^3} \frac{\langle L^4 \rangle}{\langle L \rangle^4} \left(\int \frac{\mathrm{d}^2\ell}{(2\pi)^2} \right)^2, \tag{C15}$$

where

$$\int \mathrm{d}^2\ell = \pi(\ell_{\max}^2 - \ell_{\min}^2). \tag{C16}$$

$$\mathcal{I}_1 = \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2\mathcal{L})^2} \frac{\langle L^3 \rangle}{\langle L \rangle^3} \times (N_\perp N_\parallel) \sum_j \sum_\ell [2C_{\ell,j}^{\text{tot}} + 2C_{|\ell-L|,j}^{\text{tot}}]$$

$$= \Omega_s^3 \frac{1}{\bar{\eta}^2} \frac{1}{(D^2\mathcal{L})^2} \frac{\langle L^3 \rangle}{\langle L \rangle^3} N_\parallel \left(\int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} \right) \sum_j \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} [2C_{\ell,j}^{\text{tot}} + 2C_{|\ell-L|,j}^{\text{tot}}]. \tag{C17}$$

$$\mathcal{I}_2 = \sum_j \sum_{j'} \sum_\ell \sum_{\ell'} \Omega_s \frac{1}{\bar{\eta}^2} \frac{1}{(D^2\mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} [C_{L,0}^{\text{tot}} + C_{|\ell-\ell'|,|j-j'|}^{\text{tot}} + C_{|\ell+\ell'-L|,j+j'}^{\text{tot}}]$$

$$= \Omega_s^3 \frac{1}{\bar{\eta}^2} \frac{1}{(D^2\mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} (N_\parallel)^2 \left(\int \frac{\mathrm{d}^2\ell}{(2\pi)^2} \right)^2 C_{L,0}^{\text{tot}} + \Omega_s^3 \frac{1}{\bar{\eta}^2} \frac{1}{(D^2\mathcal{L})^2} \frac{\langle L^2 \rangle^2}{\langle L \rangle^4} \sum_j \sum_{j'} \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} \int \frac{\mathrm{d}^2\ell'}{(2\pi)^2} \times [C_{|\ell-\ell'|,|j-j'|}^{\text{tot}} + C_{|\ell+\ell'-L|,j+j'}^{\text{tot}}].$$

$$\tag{C18}$$

We are not able to probe the $k_\parallel = 0$ mode (its signal will be removed with foreground cleaning), so we ignore the $C_{L,0}$ contribution to the $\mathcal{I}_2$ part of the noise and all our plots are with $j_{\min} = 1$.

$$\mathcal{I}_3 = \Omega_s^2 \frac{1}{\bar{\eta}} \frac{1}{(D^2\mathcal{L})} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \sum_j \sum_\ell [2C_{\ell,j}^{\text{tot}} + 2C_{|\ell-L|,j}^{\text{tot}}]$$

$$= \Omega_s^3 \frac{1}{\bar{\eta}} \frac{1}{(D^2\mathcal{L})} \frac{\langle L^2 \rangle}{\langle L \rangle^2} \sum_j \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} [2C_{\ell,j}^{\text{tot}} + 2C_{|\ell-L|,j}^{\text{tot}}]. \tag{C19}$$

$$\mathcal{I}_4 = \Omega_s^2 \sum_j \sum_\ell 2C_{\ell,j}^{\text{tot}} C_{|\ell-L|,j}^{\text{tot}}$$

$$= \Omega_s^3 \sum_j \int \frac{\mathrm{d}^2\ell}{(2\pi)^2} 2C_{\ell,j}^{\text{tot}} C_{|\ell-L|,j}^{\text{tot}}. \tag{C20}$$

Here, we should note that foreground subtraction techniques will remove the first few $k_\parallel$ modes (Zahn & Zaldarriaga 2005; McQuinn et al. 2006), meaning that we will need to use some $j_{min} > 1$ for the lensing reconstruction. Removing only the first few modes would correspond to a relatively simple foreground contamination, while a large number would denote a much more complex problem. Here, we should note that in our case the S/N is not greatly affected by removing the few first modes, and we plan to investigate foreground contamination and subtraction techniques in future work.

The behaviour of the lensing estimator and noise for clustered discrete sources is more complicated than in the Gaussian approximation due to the complexity of the various contributing terms. As we have stressed in the main text, in our method the Poisson fluctuations contribute both to the signal and the noise of the estimator. More specifically, the contribution of the second moment C^{shot} (see equation C14) is crucial for obtaining a low lensing reconstruction noise level. Of the noise terms the dominant contribution comes from the fourth moment Poisson term $\mathcal{I}_0$, while the $\mathcal{I}_2$ term has the smallest contribution. It is also useful to note that, in a similar manner with the Gaussian estimator, the lensing reconstruction noise converges with increasing j, so that a relatively small number of modes contributes to the final estimator. For example, our calculations for the source redshift $z = 2$ allow us to use a maximum number of parallel modes $j_{max} = 63$, but the noise has already converged at $j \sim 40$. To illustrate that, we show a plot demonstrating the convergence of the lensing reconstruction noise $N(L)$ with j using the SKA2 specifications and Model B for the H I mass function. We see that for this case the noise converges fast (this naturally depends on the interferometer specifications, which determine the contribution of the instrument's thermal noise, as well as the evolution of the H I mass function which affects the signal and the Poisson moments). Also note that the shape of $L^2 N(L)$ is nearly flat up to a scale where noise becomes important – this behaviour is qualitatively similar to the one of the Gaussian case (Zahn & Zaldarriaga 2005).

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