Magnification–Temperature Correlation: the Dark Side of ISW Measurements

Marilena LoVerde[1,2], Lam Hui[1,2], Enrique Gaztañaga[3,4]
[1]*Institute for Strings, Cosmology and Astro-particle Physics (ISCAP)*
[2]*Department of Physics, Columbia University, New York, NY 10027*
[3]*Institut de Ciències de l'Espai, CSIC/IEEC, Campus UAB, F. de Ciències, Torre C5 par-2, Barcelona 08193, Spain*
[4]*INAOE, Astrofísica, Tonantzintla, Puebla 7200, Mexico*
marilena@phys.columbia.edu, lhui@astro.columbia.edu, gazta@aliga.ieec.uab.es
(Dated: February 5, 2008)

Integrated Sachs-Wolfe (ISW) measurements, which involve cross-correlating the microwave background anisotropies with the foreground large-scale structure (e.g. traced by galaxies/quasars), have proven to be an interesting probe of dark energy. We show that magnification bias, which is the inevitable modulation of the foreground number counts by gravitational lensing, alters both the scale dependence and amplitude of the observed ISW signal. This is true especially at high redshifts because (1) the intrinsic galaxy-temperature signal diminishes greatly back in the matter dominated era, (2) the lensing efficiency increases with redshift and (3) the number count slope generally steepens with redshift in a magnitude limited sample. At $z \gtrsim 2$, the magnification-temperature correlation dominates over the intrinsic galaxy-temperature correlation and causes the observed ISW signal to increase with redshift, despite dark energy subdominance – a result of the fact that magnification probes structures all the way from the observer to the sources. Ignoring magnification bias therefore can lead to (significantly) erroneous conclusions about dark energy. While the lensing modulation opens up an interesting high z window for ISW measurements, high redshift measurements are not expected to add much new information to low redshift ones if dark energy is indeed the cosmological constant. This is because lensing introduces significant covariance across redshifts. The most compelling reasons for pursuing high redshift ISW measurements are to look for potential surprises such as early dark energy domination or signatures of modified gravity. We conclude with a discussion of existing measurements, the highest redshift of which is at the margin of being sensitive to the magnification effect. We also develop a formalism which might be of more general interest: to predict biases in estimating parameters when certain physical effects are ignored in interpreting observations.

PACS numbers: 98.80.-k, 98.80.Es, 98.70.Vc, 95.36.+x, 98.62.Sb

I. INTRODUCTION

In a universe with accelerating expansion, the gravitational potentials associated with large–scale structures decay. A photon traveling through a decaying potential will experience a net change in energy. This leads to a secondary anisotropy in the cosmic microwave background (CMB) called the integrated Sachs–Wolfe (ISW) effect [1, 2]. The ISW effect has been proposed as a tool to examine dark energy [3]. Recent studies have detected the ISW signature providing an additional confirmation of the presence of dark energy [4, 5, 6, 7, 8, 9, 10, 11, 12].

Since inhomogeneities in the matter distribution trace gravitational potentials, one expects a significant correlation between the CMB temperature anisotropies and the distribution of large–scale structure at redshifts where there is a non–negligible fraction of dark energy. A universe with non–negligible curvature or contributions from anything other than pressureless matter can do the same. Bearing in mind the evidence for spatial flatness [13, 14], we will only consider dark energy.

The ISW anisotropy can be isolated from the primordial anisotropies in the CMB by cross–correlating the temperature and galaxy/quasar fluctuations. (Hereafter, galaxy and quasar can be considered roughly synonymous: essentially all statements about one apply to the other.) Cross-correlating the temperature anisotropy in direction $\hat{\boldsymbol{\theta}}$, $\delta_T(\hat{\boldsymbol{\theta}}) = \delta T(\hat{\boldsymbol{\theta}})/\bar{T}$, with the galaxy fluctuation at redshift z in direction $\hat{\boldsymbol{\theta}}'$, $\delta_n(\hat{\boldsymbol{\theta}}', z) = \delta n(\hat{\boldsymbol{\theta}}', z)/\bar{n}$, gives

$$\left\langle \delta_T(\hat{\boldsymbol{\theta}}) \delta_n(\hat{\boldsymbol{\theta}}', z) \right\rangle = w_{nT}(\theta, z) \tag{1}$$

where $\cos\theta = \hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\theta}}'$. This scale and redshift dependent signal provides information about the growth rates of large–scale structure at the redshift of the galaxy sample.

Gravitational lensing alters the observed galaxy fluctuation in two ways. First, lensing alters the area of the patch of sky being observed thus modifying the apparent number density. Second, lensing can focus light promoting intrinsically faint objects above the magnitude threshold thus increasing the measured number density [15, 16]. Including these effects changes the observed galaxy fluctuation to

$$\delta_n = \delta_g + \delta_\mu \tag{2}$$

where δ_g is the intrinsic galaxy fluctuation and δ_μ is the *magnification bias* correction due to gravitational lensing. The effect of magnification bias on the galaxy–galaxy and galaxy–quasar correlation functions is well

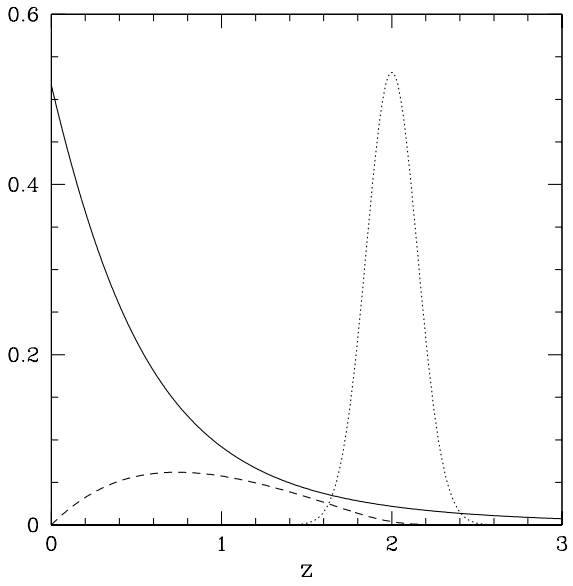

FIG. 1: The derivative $(d/dz)[D(z)(1+z)]$ (solid line) where D is the linear growth factor, a galaxy selection function (dotted line) and the corresponding lensing weight function divided by c/H_0 (dashed line). The galaxy–temperature correlation is proportional to the integral of the product of the solid and dotted lines, while the magnification–temperature correlation is proportional to the integral of the product of the solid and dashed lines.

studied [16, 17, 18, 19, 20, 21]. The most recent measurements of this effect are discussed in [22, 23, 24]. Discussions of earlier measurements can be found in the references therein.

With magnification bias the cross–correlation signal becomes

$$w_{nT}(\theta, z) = w_{gT}(\theta, z) + w_{\mu T}(\theta, z). \tag{3}$$

The galaxy–temperature term, $w_{gT}(\theta, z)$, is significant only when dark energy is non–negligible (assuming a flat universe, as is throughout this paper). Thus for high redshift galaxy samples this term is very small. However the magnification–temperature term, $w_{\mu T}(\theta, z)$, depends on the lower–redshift distribution of lensing objects so at high redshifts it may dominate over the $w_{gT}(\theta, z)$ term. Figure 1 illustrates this.

We examine the effect of the magnification bias term on measurements of the ISW effect from cross–correlation. The magnification–temperature correlation has a different scale and redshift dependence from the galaxy–temperature correlation. We will demonstrate that the magnification term can be large though its magnitude depends on the population of galaxies under consideration.

This paper is organized as follows. We present the basic expressions that govern the anisotropies and correlation functions in §II and §III. The ISW cross–correlation signal, especially in the presence of magnification bias, is sensitive to the sample of objects one is cross-correlating with. We define two illustrative samples in §IV. The cross–correlation signals for these two samples, with and without magnification, are discussed in §V. In §VI, we investigate how erroneous conclusions regarding dark energy would be reached if one ignores magnification bias when interpreting ISW measurements. In §VII, we study how projections for the signal–to–noise and parameter estimation might be altered by the presence of magnification bias. We conclude with a discussion of existing measurements in §VIII. Appendix A is devoted to a technical discussion of how to estimate error and bias. In particular, we develop a formalism for predicting the estimation–bias for parameters of interest when certain physical effects (such as magnification bias) are ignored in interpreting data. We keep the discussion there quite general with a view towards possible applications other than ISW measurements.

Two comments are in order before we proceed. Some of the existing ISW measurements come from cross-correlating the microwave background temperature with other diffuse backgrounds such as the X-ray background e.g. [7]. Magnification bias, an effect that affects number counts, does not affect these measurements. This is not to say that lensing has no effect on diffuse backgrounds, gravitational lensing does have an effect through stochastic deflections. In particular, one might wonder how the lensing of the microwave background itself might impact the observed galaxy–temperature cross–correlation. The effect appears to be small on the large scales where the ISW measurements are generally considered interesting (we consider $\ell = 2 - 100$). The reader is referred to [4] and references therein for more discussions.

II. GALAXY, TEMPERATURE AND MAGNIFICATION BIAS ANISOTROPIES

The temperature anisotropy due to the ISW effect is expressed as an integral over conformal time from 0 to today η_0 [25]

$$\delta_T^{\rm ISW}(\hat{\boldsymbol{\theta}}) = 2\int_0^{\eta_0} d\eta\, \mathrm{e}^{-\tau(\eta)} \frac{\partial \phi}{\partial \eta} \tag{4}$$

where $\tau(\eta)$ is the optical depth between η_0 and η. For perturbations sufficiently within the horizon, the gravitational potential ϕ is related to the mass (or matter) fluctuation $\delta = \delta\rho/\bar{\rho}$ in Fourier space by [25]

$$\phi(\mathbf{k}, z) = -\frac{3}{2}\frac{H_0^2}{c^2}\Omega_m(1+z)\frac{\delta(\mathbf{k}, z)}{k^2} \tag{5}$$

where Ω_m is the ratio of the matter density to the critical density today, H_0 is the Hubble constant today, c is the speed of light, z is the redshift, and k is the comoving wavenumber. The mass fluctuation δ on sufficiently large scales grows according to linear theory: $\delta \propto D(z)$, where $D(z)$ is commonly referred to as the growth factor.

We are interested in cross–correlating the temperature anisotropies, δ_T, with the observed galaxy overdensity,

δ_n. With magnification, the measured galaxy fluctuation is a sum of two terms (eq. [2]). The first is the intrinsic galaxy fluctuation

$$\delta_g(\hat{\boldsymbol{\theta}}, z_i) = \int dz\, b(z)\, W(z, z_i)\delta(\chi(z)\hat{\boldsymbol{\theta}}, z), \qquad (6)$$

where $b(z)$ is an assumed scale–independent bias factor relating the galaxy overdensity to the mass overdensity i.e. $\delta_g = b\,\delta$, $W(z, z_i)$ is a normalized selection function about some mean redshift z_i, and $\chi(z)$ is the comoving distance to redshift z.

The magnification bias term is also expressed as an integral over redshift [16]

$$\begin{aligned}\delta_\mu(\hat{\boldsymbol{\theta}}, z_i) &= 3\Omega_m \frac{H_0^2}{c^2}(2.5s(z_i) - 1) \\ &\times \int dz \frac{c}{H(z)} g(z, z_i)(1+z)\delta(\chi(z)\hat{\boldsymbol{\theta}}, z). \qquad (7)\end{aligned}$$

The lensing weight function, $g(z, z_i)$, can be thought of as being proportional to the probability of sources around z_i to be lensed by intervening matter at z [26].

$$g(z, z_i) = \chi(z) \int_z^\infty dz' \frac{\chi(z') - \chi(z)}{\chi(z')} W(z', z_i) \qquad (8)$$

For sources at a distance χ, the lensing weight function peaks at $\sim \chi/2$. The prefactor of the magnification term depends on the slope of the number count of the source galaxies. This is defined as

$$s \equiv \frac{d \log_{10} N(< m)}{dm} \qquad (9)$$

where m is the limiting magnitude of one's sample, and $N(< m)$ represents the count of objects brighter than m. Magnification bias vanishes in the case that $s = 0.4$. The values of the slope s, as well as the galaxy bias b in eq. [6], depend on the population of galaxies under consideration.

III. CROSS–CORRELATION FUNCTIONS

For the most part, we focus on the spherical harmonic transform of the various angular correlation functions $w_{gT}(\theta)$, $w_{\mu T}(\theta)$ and so on. They can all be neatly described by the following expression:

$$C_\ell^{AB}(z_i, z_j) = \frac{2}{\pi} \int k^2 dk P(k) \delta A_\ell(k, z_i) \delta B_\ell(k, z_j) \qquad (10)$$

where $P(k)$ is the matter power spectrum today as a function of the wavenumber k, and the functions δA_ℓ and δB_ℓ in the integrand can be one of the following:

$$[\delta_g]_\ell(k, z_i) = b_i \int dz\, W(z, z_i) D(z) j_\ell(k\chi(z)) \qquad (11)$$

$$\begin{aligned}[\delta_\mu]_\ell(k, z_i) &= 3\Omega_m \frac{H_0^2}{c^2}(2.5s_i - 1) \qquad (12) \\ &\times \int dz \frac{c}{H(z)} g(z, z_i)(1+z) D(z) j_\ell(k\chi(z))\end{aligned}$$

$$\begin{aligned}\left[\delta_T^{\rm ISW}\right]_\ell(k) &= 3\frac{H_0^2}{c^2}\Omega_m \qquad (13) \\ &\times \int dz \frac{d}{dz}\left[D(z)(1+z)\right] \frac{j_\ell(k\chi(z))}{k^2}\end{aligned}$$

The symbol j_ℓ denotes the spherical Bessel function. Here we have elected to assume a constant galaxy–bias b_i and slope s_i for each redshift bin centered around z_i. In the temperature integral we have neglected the factor containing the optical depth. This leads to an error of at most about 3% at the highest redshift we consider.

Notice that the temperature function $[\delta_T]_\ell(k)$ is independent of redshift, so the cross–correlation functions C_ℓ^{gT} and $C_\ell^{\mu T}$ are functions of only a single redshift z_i. This is in contrast to the other correlation functions which depend on the redshifts of the two samples being correlated, e.g. $C_\ell^{gg}(z_i, z_j)$, $C_\ell^{g\mu}(z_i, z_j)$, $C_\ell^{\mu g}(z_i, z_j)$ and $C_\ell^{\mu\mu}(z_i, z_j)$. Note also that eq. [10] and eq. [13] are not suitable for computing C_ℓ^{TT} since that would account for only the ISW contribution to C_ℓ^{TT}.

Examining the correlation functions, we can see that the relative magnitude of the intrinsic galaxy–temperature correlation, C_ℓ^{gT}, and the magnification–temperature correlation, $C_\ell^{\mu T}$, is redshift and scale dependent. The lensing efficiency increases with the redshift of the source galaxies causing the magnification bias effect to increase with redshift as well. The scale dependence arises because each term depends on the matter distribution at different distances: the galaxy term depends on the matter distribution at the source redshifts while the magnification term depends on the distribution at the lenses.

Additionally, the coefficients of C_ℓ^{gT} and $C_\ell^{\mu T}$ depend on properties of the galaxies via b and s. The magnitude of the magnification term compared with the galaxy term depends on the ratio $(2.5s - 1)/b$. Thus the correction may be negative for $(s < 0.4)$ and in any case increases in magnitude as $|s - 0.4|$. The bias and slope depend on the choice of galaxy sample and are redshift dependent themselves.

The Limber approximation, which is quite accurate when ℓ is not too small ($\ell \gtrsim 10$), can be obtained from eq. [10] by setting $P(k) = P(k = (\ell + 1/2)/\chi(z))$ and using the fact that $(2/\pi)\int k^2 dk j_\ell(k\chi) j_\ell(k\chi') = (1/\chi^2)\delta(\chi - \chi')$. We find that the substitution $k = (\ell + 1/2)/\chi(z)$ is a better approximation to the exact expressions than $k = \ell/\chi(z)$ (see also [9]). More explicitly, C_ℓ^{gT} and $C_\ell^{\mu T}$ under the Limber approximation are given by:

$$\begin{aligned}C_\ell^{gT}(z_i) &= \frac{3\Omega_m H_0^2}{c^2} \frac{b_i}{(\ell + 1/2)^2} \int dz W(z, z_i) \frac{H(z)}{c} D(z) \\ &\times \frac{d}{dz}[D(z)(1+z)] P(k_\perp = (\ell + 1/2)/\chi) \qquad (14)\end{aligned}$$

Sample-I $m < 27$	z-bin 1	z-bin 2	z-bin 3	z-bin 4	z-bin 5	z-bin 6
z_i	0.49	1.14	1.93	2.74	3.54	4.35
b_i	1.08	1.37	2.02	2.90	3.89	4.81
s_i	0.15	0.20	0.31	0.43	0.54	0.63
n_i	34.6	29.0	10.1	3.89	1.68	0.81
Sample-II $m < 25$	z-bin 1	z-bin 2	z-bin 3	z-bin 4	z-bin 5	–
z_i	0.48	1.07	1.85	2.67	3.46	–
b_i	1.13	1.51	2.73	4.57	6.63	–
s_i	0.19	0.35	0.86	1.31	1.75	–
n_i	16.1	8.6	0.87	0.11	0.02	–

TABLE I: The mean redshift z_i, the galaxy–bias b_i, the slope s_i and the number of galaxies per square arcminute n_i are given for samples I and II, and for each corresponding redshift bin.

$$C_\ell^{\mu T}(z_i) = \left(\frac{3\Omega_m H_0^2}{c^2}\right)^2 \frac{2.5s_i - 1}{(\ell+1/2)^2} \int dz\, g(z,z_i)(1+z) \times \quad D(z)\frac{d}{dz}[D(z)(1+z)]P(k_\perp = (\ell+1/2)/\chi) \qquad (15)$$

When displaying the above quantities in figures, we follow the custom of multiplying them by 2.725 K.

Unless otherwise stated, we adopt throughout this paper the following values for the cosmological parameters when making predictions for the various correlations of interest: we assume a flat universe with a matter density of $\Omega_m = 0.27$, a cosmological constant of $\Omega_\Lambda = 0.73$, a Hubble constant of $h = 0.7$, a baryon density of $\Omega_b = 0.046$, a scalar spectral index of $n_s = 0.95$ and a fluctuation amplitude of $\sigma_8 = 0.8$. The matter power spectrum is computed using the transfer function of Eisenstein and Hu [27]. The microwave background temperature power spectrum C_ℓ^{TT} is computed using the publicly available code CAMB [28], with an optical depth of $\tau = 0.09$ and no tensor modes. Throughout this paper, we consider ISW measurements at $\ell = 2 - 100$. Including higher ℓ modes does not significantly change our conclusions. This is because the signal-to-noise of ISW measurements is dominated by $\ell = 10 - 100$ (see e.g. [31, 32], but also [33]). We adopt the Limber approximation for all discussions in §VI and VII, which concern issues related to signal–to–noise, estimation–bias and parameter forecast, since these issues are not significantly affected by the very low ℓ modes. The discussions and figures in §V, however, make use of the exact expressions for the correlation functions (eq. [10]).

IV. SURVEY SAMPLES

The predictions for the galaxy–temperature and magnification–temperature correlation depend on three quantities that are sample dependent: the redshift distribution of galaxies, the galaxy–bias, and the slope of the number count (at the magnitude cut–off). For illustration, we define two semi-realistic samples.

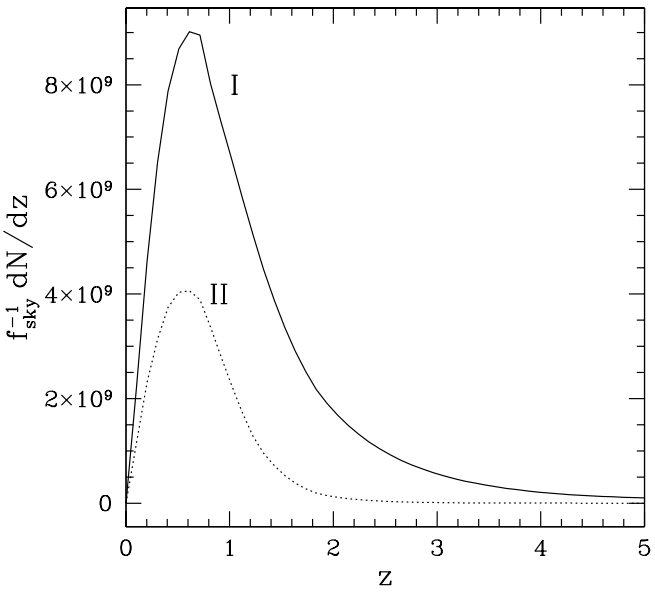

FIG. 2: The number of galaxies per unit redshift over the whole sky as a function of redshift, for two illustrative galaxy samples I and II (see text). The actual dN/dz for a given survey equals the sky coverage $f_{\rm sky}$ times the above.

Sample I is defined by an observed I–band (centered around 7994$\mathring{A}$) magnitude cut of 27 which implies a redshift distribution dN/dz shown as a solid line in Figure 2, adopting the observed redshift-dependent luminosity function given by [29]. This gives a net galaxy angular number density of 80 per square arcminute. We assume a sky coverage of $f_{sky} = 0.5$. These survey specifications are similar to those of the Large Synoptic Survey Telescope [30]. We divide these galaxies into 6 redshift bins, following the procedure of [31]:

$$W(z,z_i) \propto \frac{1}{2}\frac{dN(z)}{dz}\Big[\,\mathrm{erfc}\left(\frac{(i-1)\Delta - z}{\sigma(z)\sqrt{2}}\right) - \mathrm{erfc}\left(\frac{i\Delta - z}{\sigma(z)\sqrt{2}}\right)\Big] \qquad (16)$$

where $\Delta = 0.8$, $\sigma(z) = 0.02(1+z)$, and z_i denotes the mean redshift of the bin i. The complementary error function is defined as $\mathrm{erfc}(x) = (2/\sqrt{\pi})\int_x^\infty \exp(-t^2)dt$. The normalization of W is fixed by demanding that $\int dz W(z,z_i) = 1$. We have checked that none of our conclusions are significantly altered by increasing the number of bins.

Given the observed luminosity function from [29], the slope s (eq. [9]) at the specified apparent magnitude cut can be obtained in a straightforward manner. The

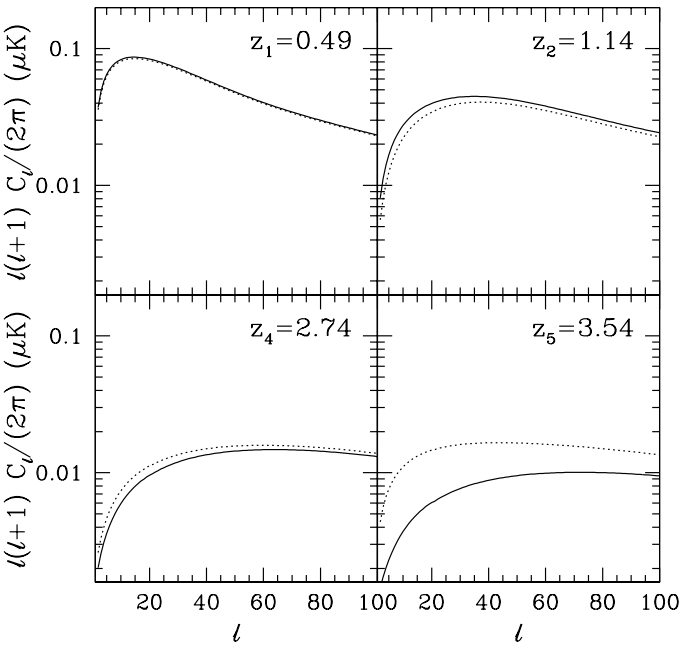

FIG. 3: The galaxy–temperature correlation (solid lines) compared with the net galaxy–temperature plus magnification–temperature correlation (dotted lines) for four different redshift bins. This is for sample I.

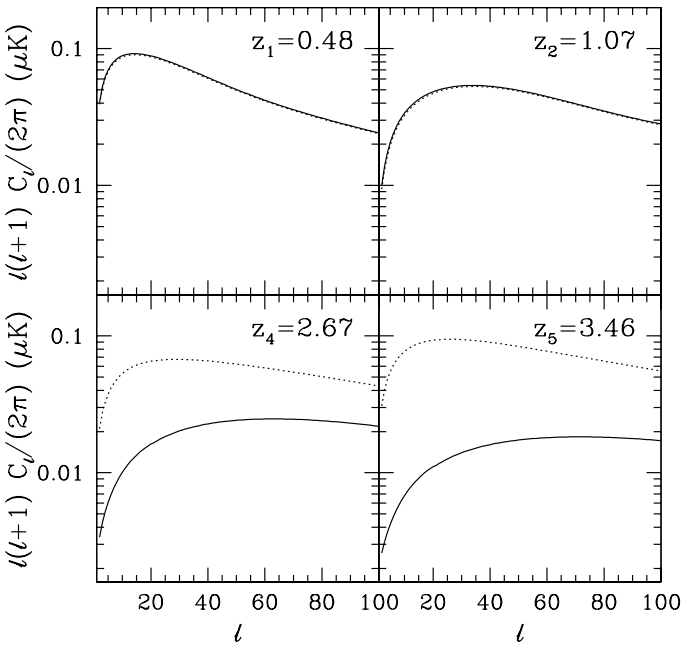

FIG. 4: The analog of Figure 3 for sample II.

galaxy–bias b, on the other hand, requires some theoretical input. Here, we adopt the method of Kravtsov et al. [34]. It works basically by number matching: one matches the the number of galaxies brighter than some magnitude with the number of halos and subhalos above some mass threshold. Once this mass threshold is known, the galaxy–bias can be computed. More specifically, the number of (central plus satellite) galaxies within a halo of a given mass M is

$$\frac{1}{2}\,\mathrm{erfc}\left(\frac{\ln M_0 - \ln M}{0.1\sqrt{2}}\right) \times (1 + \frac{M}{20M_0}) \qquad (17)$$

Here, the complementary error function gives essentially a step function with a modest spread (of 0.1), and M_0 can be thought of as the mass threshold. The quantity $20M_0$ specifies the (parent) halo mass at which the expected number of satellites is unity. This factor is in principle redshift dependent, but we ignore such complications here. We determine M_0 by demanding that the integral of the halo occupation number over the halo mass function equals the number of galaxies for the sample under consideration. The bias of this sample of galaxies can be obtained by integrating the product of the halo occupation number and the halo–bias over the halo mass function, normalized by the total number of galaxies. We adopt the mass function and halo-bias of Sheth and Tormen [36].

Sample II is defined in a similar way to sample I, except that the limiting I–magnitude is taken to be 25 instead of 27. The corresponding dN/dz is shown as a dotted line in Figure 2. The net galaxy angular number density is 26 per square arcminute. We assume the same sky coverage of $f_{sky} = 0.5$. This is a shallower survey, and we divide the sample into only 5 redshift bins i.e. apply eq. [16] up to $i = 5$.

A summary of the mean redshift, the galaxy–bias b, the slope s and the galaxy number density for each redshift bin and each sample is given in Table I. The tabulated dN/dz for each sample is available from [35].

V. THE CROSS–CORRELATION SIGNAL

In Figures 3 and 4 we show the cross–correlations $C_\ell^{gT}(z_i)$ and $C_\ell^{nT}(z_i) = C_\ell^{gT}(z_i) + C_\ell^{\mu T}(z_i)$ at several redshift bins for both sample I and sample II, calculated using equations (10) - (13).

There are several features in Figures 3 and 4 that are worth commenting. As is expected, the galaxy–temperature correlation C_ℓ^{gT} decreases rapidly with redshift: at high z, dark energy becomes subdominant making the time derivative of the gravitational potential quite small (eq. [4]). Note that this decrease of C_ℓ^{gT} occurs despite the overall increase of the galaxy–bias with redshift.

The magnification–temperature correlation $C_\ell^{\mu T}$, on the other hand, generally increases with redshift because of the increase of the lensing efficiency. The magnification–temperature correlation $C^{\mu T}$ can be high even as C^{gT} becomes small, because the rather broad lensing weight function makes $C^{\mu T}$ sensitive to structures at low redshifts where the gravitational potential has a

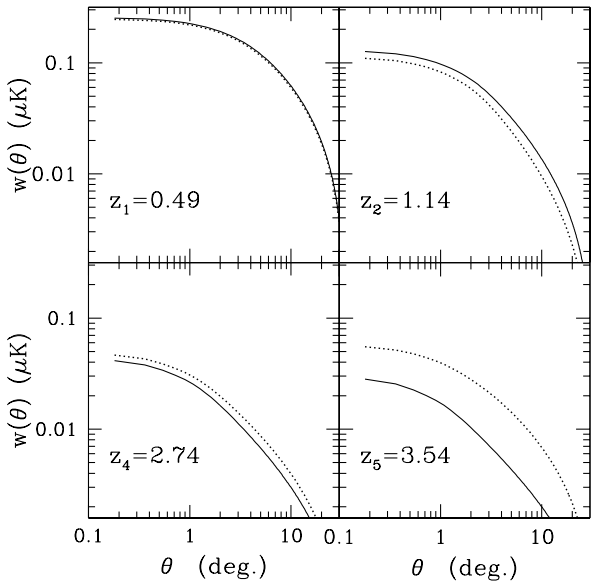

FIG. 5: The 2-point galaxy–temperature correlation $w_{gT}(\theta, z_i)$ (solid lines) compared with the net galaxy–temperature plus magnification–temperature correlation $w_{nT}(\theta, z_i)$ (dotted lines) in several of our redshift bins. This is for sample I.

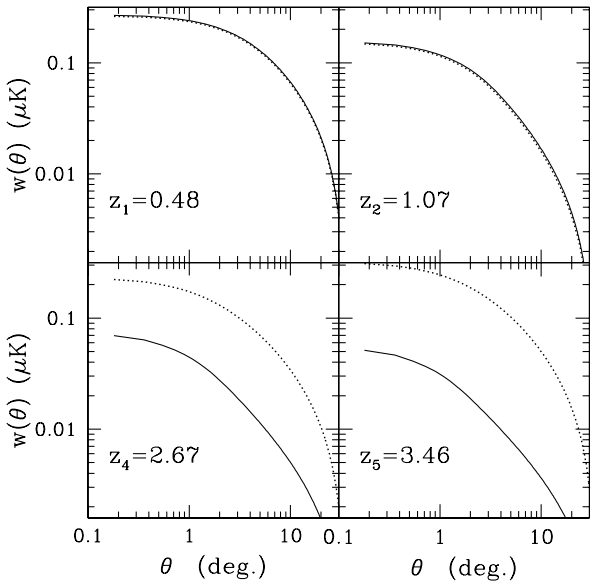

FIG. 6: The analog of Figure 5 for sample II.

significant time derivative. The general increase of $C_\ell^{\mu T}$ with redshift is aided also by the increase of the number count slope s — at a higher z one is generally looking at intrinsically brighter galaxies which reside in the steep part of the luminosity function. The net result is that the total cross-correlation C_ℓ^{nT} starts to climb with redshift, for $z \gtrsim 2$. In fact at a sufficiently high z, C_ℓ^{nT} even becomes comparable with C_ℓ^{nT} or C_ℓ^{gT} at the lowest redshift bin. Comparing sample I and sample II, one can see that a brighter magnitude cut leads to a more pronounced magnification bias effect. This is due to the steeper number count slope.

Another feature visible in Figures 3 and 4 is that magnification changes the shape of C_ℓ^{nT} with ℓ. The magnification term, being sensitive to structures at lower redshift, peaks at a lower ℓ than the galaxy term. For positive magnification contribution, the peak of C_ℓ^{nT} is at a lower ℓ than the peak of C_ℓ^{gT}.

In Figures 5 and 6 we plot the real space correlation functions $w_{gT}(\theta, z_i)$ and $w_{nT}(\theta, z_i)$ without the monopole and dipole contributions. These are calculated from

$$w_X(\theta, z_i) = \sum_{\ell=2}^{500} \frac{2\ell+1}{4\pi} C_\ell^X(z_i) P_\ell(\cos\theta). \qquad (18)$$

where X symbolizes gT or nT and $P_\ell(\cos\theta)$ are the Legendre polynomials. We have extended the range of the sum to $\ell = 500$ to insure the convergence of the sum for small values of θ. The general discussion about C_ℓ^{nT} applies to w_{nT} as well; at $z \gtrsim 2$ magnification bias grows to become an important component of the cross–correlation signal, and the amplitude is preferentially increased at large angular scales.

Magnification bias can of course have either sign. At low redshifts, one is generally looking at intrinsically faint galaxies which leads to a small s that can be less than 0.4. This is why $C_\ell^{nT} < C_\ell^{gT}$ at low redshifts for our samples. The transition redshift from $s < 0.4$ to $s > 0.4$ is sample dependent. For a sample with an apparent limiting magnitude cut, the brighter the cut, the lower the transition redshift. It is conceivable that a sufficiently faint cut can be achieved such that s remains less than 0.4 out to redshifts where $|C_\ell^{\mu T}| > C_\ell^{gT}$. In such a case the net cross-correlation C_ℓ^{nT} would become negative. This, however, requires an exceptionally deep survey: an I-band magnitude limit of 28 or higher, according to the observed luminosity function [29]. Conversely, if a sufficiently bright galaxy cut were used at low redshift it is possible that magnification could become important at lower redshifts than evidenced by our samples.

In the literature the cross–correlation signal is often presented in the normalized form $C_\ell^{gT}(z_i)/\sqrt{C_\ell^{gg}(z_i)}$ or $w_{gT}(\theta, z_i)/\sqrt{w_{gg}(\theta, z_i)}$ to remove the dependence on galaxy bias. However, when magnification bias is taken into account the auto–correlation is modified to be

$$C_\ell^{nn}(z_i, z_j) = \qquad (19)$$
$$C_\ell^{gg}(z_i, z_j) + C_\ell^{g\mu}(z_i, z_j) + C_\ell^{\mu g}(z_i, z_j) + C_\ell^{\mu\mu}(z_i, z_j).$$

Note, then, that the galaxy–bias would be completely removed by the division only if magnification bias were absent. Had we plotted the normalized versions of the cross-correlation signals, the fractional difference between $C_\ell^{gT}/\sqrt{C_\ell^{gg}}$ and $C_\ell^{nT}/\sqrt{C_\ell^{nn}}$ would remain about the same as that between the un-normalized versions. This

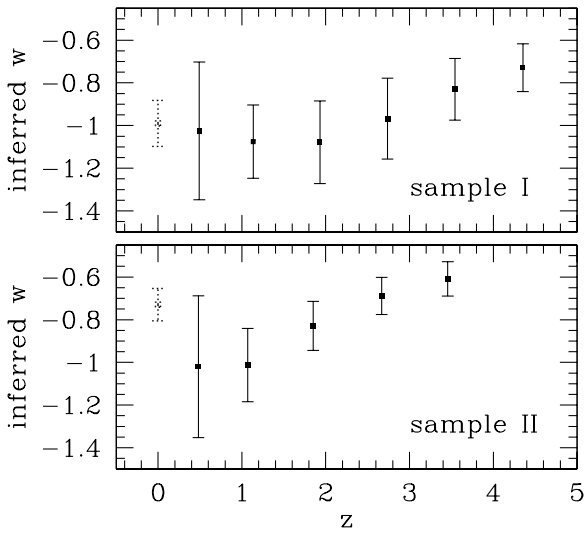

FIG. 7: The average *inferred* dark energy equation of state w from ISW measurements when (erroneously) ignoring magnification. The true w equals -1. The solid symbols show the inferred w for each redshift bin separately. The dotted symbols at $z = 0$ show the inferred w combining all redshift bins. The error-bars are 1 σ — note that they are dependent upon the *inferred* model (see text). The upper panel is for sample I, and the lower panel is for sample II.

is because for the cases considered here $C_\ell^{nn}(z_i, z_i) \sim C_\ell^{gg}(z_i, z_i)$ to within about 20%.

VI. ESTIMATION–BIAS DUE TO IGNORING MAGNIFICATION —- A THOUGHT EXPERIMENT

Clearly, magnification bias should be taken into account when interpreting ISW measurements, especially at high redshifts where the magnification–temperature correlation actually dominates over the usual galaxy–temperature signal. Ignoring it would lead to erroneous conclusions about dark energy. We quantify this by computing the estimation–bias in the dark energy equation of state w if magnification were ignored when interpreting measurements from samples I and II.

A thought experiment is set up as follows. Suppose the true dark energy equation of state is $w = -1$ i.e. a cosmological constant (and other cosmological parameters take the values stated at the end of §III). Suppose further one were to infer w (and other parameters) from measurements of the total cross-correlation C_ℓ^{nT} by fitting them with a model that ignores magnification i.e. C_ℓ^{gT}, and one were to assign errorbars based on the (erroneous) model with no magnification. The fitted parameters include: the dark energy equation of state w (i.e. $w = P/\rho$ where P and ρ are the dark energy pressure and energy density respectively), the matter density Ω_m (we assume a flat universe so that the dark energy density is $1 - \Omega_m$), the amplitude of fluctuations σ_8, the spectral index n_s, the Hubble constant h, the baryon density Ω_b and the galaxy-bias for each redshift bin b_i. Constraints from other data (such as from the temperature anisotropy itself C_ℓ^{TT} and from the galaxy power spectrum) are simulated by the inclusion of priors: the fractional (1 σ) errors on $\Omega_m h^2$, $\Omega_b h^2$, h and σ_8 are assumed to be 5%, the fractional error on b_i is 10%, and n_s has an (absolute) error of 0.02. These are similar to current constraints, depending somewhat on assumptions [13]. A simple analytic expression can be derived for the parameter estimation–bias, which is discussed in detail in Appendix A (eq. [A14]).

The inferred w, together with its marginalized error, for samples I and II are shown in Figure 7. The modes used range from $\ell = 2$ to $\ell = 100$. The solid symbols show the inferred w from each redshift bin separately. The sign of the resulting estimation–bias is determined by the slope s. If $s < 0.4$, as is the case at low redshifts, the inferred w tends to be lower than the true value. This is understandable because the *observed* C^{nT} is suppressed by magnification in that case, and the only way to stick with an erroneous model without magnification is to resort to a lower w to delay dark energy domination. The reverse happens at high redshifts, where $s > 0.4$. In fact, at the highest redshift bin, the bias can become quite large. Ignoring magnification bias is simply unacceptable.

It is worth emphasizing that the errorbars are model dependent i.e. they are computed using the *inferred* model, which is different for each redshift bin. The general trend is for the errorbar to shrink as one goes to higher redshifts, despite the drop in signal–to–noise (see §VII). This is due to the fact that $d\ln C^{gT}/dw$ rises with both z and w. The generally positive bias in w at high redshifts helps to diminish the corresponding errorbars.

The dotted symbols at $z = 0$ show the inferred w and its associated errorbar when measurements from all the redshift bins are combined. The previous remarks on the model dependent nature of the errorbar applies here as well. In the case of sample II, one finds that the overall inferred w is biased high by almost 3 σ.

VII. SIGNAL–TO–NOISE AND PARAMETER FORECAST

The above considerations suggest that magnification could cause the signal–to–noise of ISW measurements to remain favorable even at relatively high redshifts, deep into the matter dominated regime. This is borne out by Figure 8 [37], which shows (with $f_{\rm sky}$ divided out)

$$\left[\frac{S}{N}(z_i)\right]^2 = \sum_\ell \frac{[C_\ell^{nT}(z_i)]^2 f_{\rm sky}(2\ell+1)}{(C_\ell^{nn}(z_i,z_i) + \frac{1}{n_i})C_\ell^{TT} + [C_\ell^{nT}(z_i)]^2} \quad (20)$$

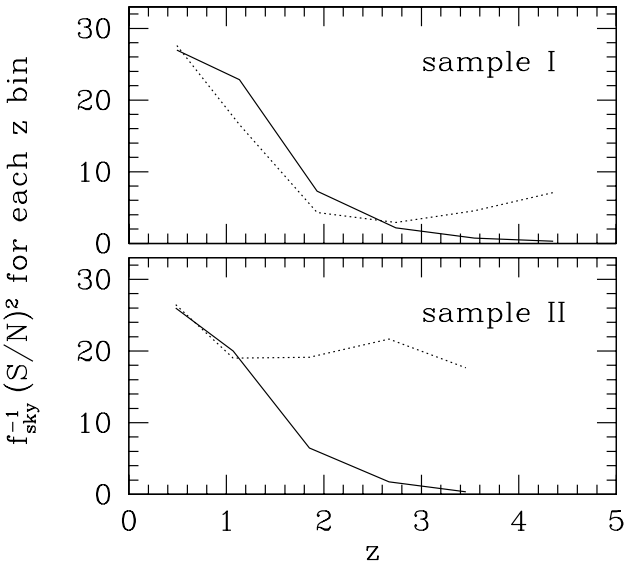

FIG. 8: The signal–to–noise (squared) of ISW measurements from each redshift bin for samples I and II. The solid line ignores magnification (equivalent to setting $s = 0.4$) while the dotted line includes magnification. Note $(S/N)^2 \propto f_{\rm sky}$ i.e. the division by $f_{\rm sky}$ takes out the dependence on sky coverage.

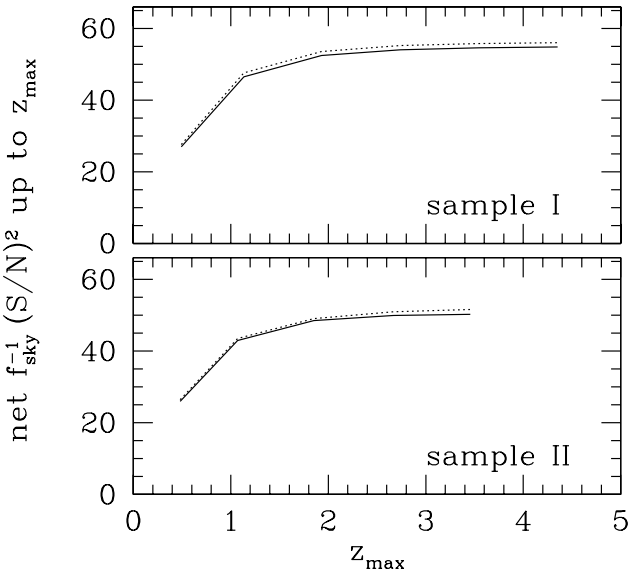

FIG. 9: The net cumulative signal–to–noise (squared) of ISW measurements from all redshift bins up to $z_{\rm max}$, for samples I and II. The solid line ignores magnification (equivalent to setting $s = 0.4$) while the dotted line includes magnification (the two are very close to each other). The division by $f_{\rm sky}$ removes the dependence on sky coverage.

where n_i is the mean surface density of galaxies, and we include modes from $\ell = 2$ to $\ell = 100$ (as in the rest of the paper). The result for sample II is particularly striking: the signal–to–noise remains more or less flat out to high redshifts (dotted line), in sharp contrast with the signal–to–noise if magnification is ignored (solid line; equivalent to putting $s = 0.4$, or replacing the superscript n by g in eq. [20]).

Magnification bias, therefore, opens up a high redshift window for ISW measurements which otherwise would not exist. That is the good news. The bad news, however, is that since the magnification of redshift z_i galaxies probes structures at redshifts $z < z_i$, the C_ℓ^{nT} measurements at high redshifts are in fact quite correlated with those at low redshifts. Taking into account such correlations, we show in Figure 9 the net accumulated signal–to–noise (squared) for measurements from all redshift bins up to a given $z_{\rm max}$, defined as follows:

$$\left[\frac{S}{N}(z_{\rm max})\right]^2 = \sum_{z_i, z_j \le z_{\rm max}} \sum_\ell C_\ell^{nT}(z_i) \left[{\rm Cov}_\ell^{-1}\right]_{ij} C_\ell^{nT}(z_j) \tag{21}$$

where

$$[{\rm Cov}_\ell]_{ij} = \frac{(C_\ell^{nn}(z_i, z_j) + \delta_{ij}/n_i) C_\ell^{TT} + C_\ell^{nT}(z_i) C_\ell^{nT}(z_j)}{f_{\rm sky}(2\ell+1)}. \tag{22}$$

It appears the cumulative signal–to–noise reaches a plateau by $z \sim 2$ whether or not magnification bias is included. As is expected, magnification bias makes much less of a difference in this case: the solid and dotted lines are very close together. This happens despite the significant difference in signal–to–noise on a redshift bin by redshift bin basis, shown in Figure 8. The culprit is the significant correlation between different redshift bins introduced by lensing — the high redshift measurements are in fact only probing large–scale structure at low redshifts.

The reader might also wonder why the two samples have such similar cumulative signal–to–noise. This is because at low redshifts (where most of the cumulative signal–to–noise comes from), neither sample is shot-noise dominated i.e. for the most part, the noise is dominated by the term associated with $C_\ell^{nn} C_\ell^{TT}$ (eq. [22]).

Ultimately, we are interested in cosmological constraints from ISW measurements on, for instance, the dark energy equation of state w. It is therefore useful to show the expected errors for w, both on a redshift bin by redshift bin basis (as in Figure 8) and on a cumulative basis (as in Figure 9) (see also [32]). This is displayed in Figures 10 and 11.

The fiducial cosmological model here, as is in the case of the computation of S/N, is that described at the end of §III. The cosmological parameters that are marginalized

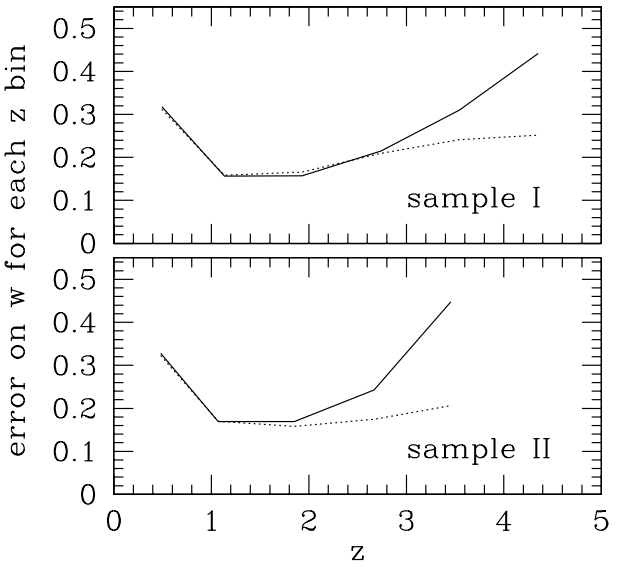

FIG. 10: The marginalized error on w, $\sqrt{\langle \delta w^2 \rangle}$, for ISW measurements from each redshift bin. The solid line ignores magnification (equivalent to setting $s = 0.4$) while the dotted line includes magnification.

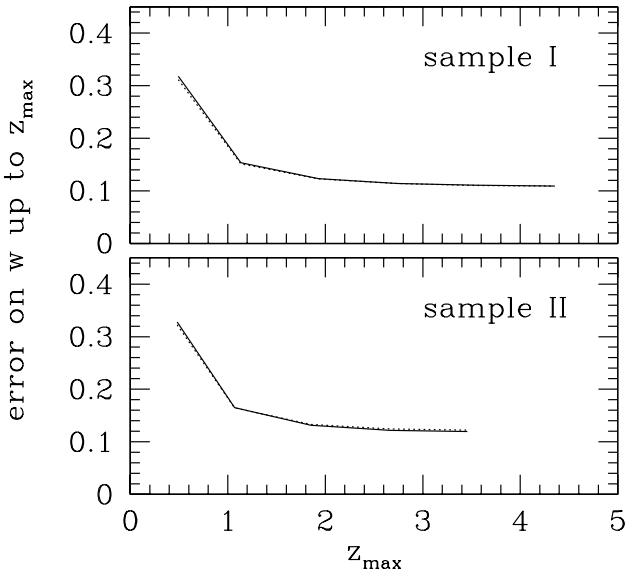

FIG. 11: The marginalized error on w, $\sqrt{\langle \delta w^2 \rangle}$, for ISW measurements from all redshift bins combined up to $z_{\rm max}$. The solid line ignores magnification (equivalent to setting $s = 0.4$) while the dotted line includes magnification (the two are very close together).

over and the set of priors assumed are identical to those described in §VI (with the addition of s_i, the number count slope for each redshift bin, as a new parameter, for which we impose a prior of $\Delta s_i = 0.02$), and the Fisher matrix formalism for doing so is explained in Appendix A, in particular eq. [A17]. Once again, we see that while the presence of magnification bias improves the errors at high redshifts on a redshift bin by redshift bin basis, it does not in fact add much cumulative information if one considers the errors from combining redshift bins. We have experimented with changing the priors and varying sample definitions, and it appears this conclusion is quite robust.

The most compelling reason for doing high redshift ($z \gtrsim 2$) ISW measurements is therefore not that the expected cumulative error on w would continue to improve, but that there might be surprises e.g. dark energy might be quite different from the cosmological constant and actually remain a significant component of the universe even at high redshifts, or gravity may be modified in non-trivial ways on large scales [38, 39, 40, 41, 44]. At the very least, high redshift ISW measurements constitute a consistency check that one should make. As explained in §VI, such high redshift measurements should be interpreted with care: magnification bias must be taken into account.

VIII. DISCUSSION

A natural question: to what extent should we worry about magnification bias when interpreting existing ISW measurements? This is partially addressed in Figure 12, which shows the cross-correlation signal at 6^o separation.

The points with errorbars are measurements taken from a compilation in [42] with the addition of recent results from [10] and [11]. All involve correlating the microwave background with galaxies, except the highest redshift one [11] which involves correlation with quasars. The median redshifts of the measurements are: $\bar{z} \sim 0.1, 0.15, 0.3, 0.5, 0.9, 1.5$. The ISW signal shown is normalized by the galaxy/quasar bias, with the bias values taken from [10, 11, 42]. They are, in the order of increasing redshift: $1.1, 1.0, 1.2, 1.4, 1.7, 2.3$.

The solid lines show the theoretical prediction for w_{gT}/b (i.e. ignoring magnification) adopting the cosmological parameters stated in §III. Because the selection function is somewhat uncertain, we choose to illustrate the prediction with a 'broad' and a 'narrow' selection function, parametrized as follows:

$$W(z) = \frac{\beta}{\Gamma\left[\frac{\alpha+1}{\beta}\right]} \frac{z^\alpha}{z_0^{\alpha+1}} e^{-(z/z_0)^\beta}. \qquad (23)$$

where $\alpha = 2$, $\beta = 1.5$ is broad, and $\alpha = 2\,(1+3\,z_0)$, $\beta = \alpha$ is narrow. For each selection function z_0 is adjusted to

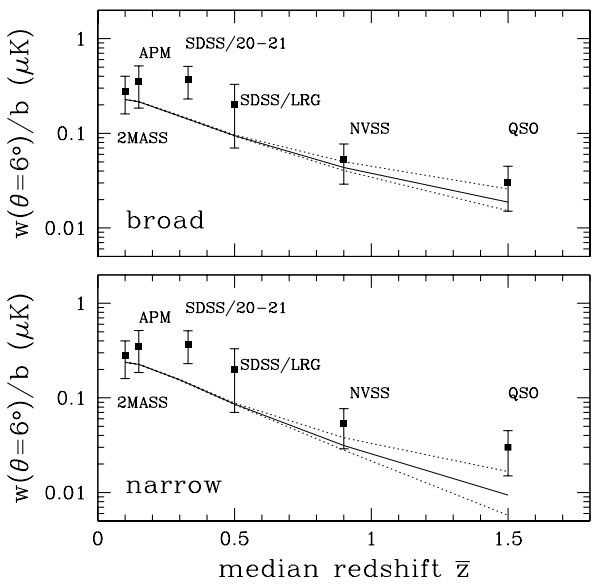

FIG. 12: The two panels show the 2-point correlation function divided by the galaxy bias, $w(\theta,\bar{z})/b(\bar{z})$, at $\theta = 6^o$ as a function of the median redshift of the sources, $\bar{z}$. The points with errorbars are the existing measurements. The solid line shows the theoretical prediction ignoring magnification for a flat cosmological constant model (see §III). The dotted lines include magnification for the cases $s = 0.8$ (upper dotted line) and $s = 0.2$ (lower dotted line). The upper panel uses a broad galaxy selection function while the lower panel uses a narrow selection function (see text).

reproduce the median redshifts of the measurements.

The dotted lines [43] in Figure 12 show the theoretical prediction for w_{nT}/b (i.e. accounting for magnification) for $s = 0.8$ (upper) and $s = 0.2$ (lower). These values for the number count slope seem to span the range observed in quasar samples [23] similar to the one used in the highest redshift ISW measurement, where magnification bias is most relevant.

Considering all angular scales and both selection functions, we find that for the four lowest redshift measurements $|w_{\mu T}/w_{gT}| < 0.1$, unless $s > 1.4$ or $s < -0.6$. For the NVSS sample at $\bar{z} \sim 0.9$, we estimate the slope of the NVSS radio sample to be $s = 0.320 \pm 0.07$ which results in $|w_{\mu T}/w_{gT}| \lesssim 0.1$. This is small compared with the measurement error. It appears that the quasar–temperature correlation at $\bar{z} \sim 1.5$ is the only measurement for which magnification *could* be significant. To determine precisely to what extent this measurement is affected by magnification requires a more detailed analysis of the quasar sample. It is conceivable that a suitably chosen quasar subsample could exhibit a large magnification effect.

Figure 12 also shows that for $z \gtrsim 1$ there is a significant dependence on the width of the selection function. This is because the mass–temperature correlation decreases rapidly with redshift. If a broad selection function is used, the galaxy–temperature correlation will decline less rapidly because low redshift sources will contribute. On the other hand, $w_{\mu T}$ depends on the lensing weight function (eq. [8]) which is broadly distributed regardless of the width of the source distribution. In other words, changing the width of the selection function from broad to narrow causes the galaxy–temperature term to decrease, while leaving the magnification–temperature term largely unchanged. Figure 12 illustrates this for a fixed θ, however, the statement holds for other angles. Source distributions used in the existing ISW measurements are probably closer to the broad selection function.

To conclude, we find that magnification bias alters the observed galaxy/quasar–temperature cross–correlation significantly at high redshifts (Figures 3 and 4). The precise magnitude of the modification depends on the sample of objects under consideration. Three facts more or less guarantee magnification plays a non–negligible or even dominant role at $z \gtrsim 2$: the generally steepening slope of the number count at high redshifts (because one is looking at intrinsically brighter objects), the decay of the intrinsic galaxy–temperature correlation C^{gT} as one enters the matter–dominated era, and the increase of the lensing efficiency with redshift. Ignoring magnification bias when interpreting high redshift ISW measurements would lead to erroneous conclusions about the nature of dark energy. For instance, the estimated equation of state w can differ from the true one by more than 3 σ in some cases (Figure 7).

The boosting of the ISW signal by magnification (assuming $s > 0.4$ which is generally the case at high redshifts when one is looking at intrinsically bright objects) implies that, despite naive expectations, ISW measurements remain viable even at $z \gtrsim 2$. However, because of the correlated nature of the lensing signal (across different redshifts), the cumulative information from low z's to high z's is not significantly enhanced by magnification. The most compelling scientific justification for pursuing high redshift ISW measurements is to look for surprises: for instance, dark energy might remain significant out to high redshifts. At the very least, one would like to perform a consistency check of the very successful cosmological constant model of dark energy. The large–scale nature of the ISW signal also means it is a natural place to look for spatial fluctuations in dark energy [31] and possible signatures of modified gravity [44].

When interpreting future high redshift ISW measurements, accounting for the effect of magnification bias is a must. The net cross–correlation signal depends not only on the cosmological parameters (e.g. Ω_m, Ω_Λ, h, n_s, σ_8) but also on the sample–dependent parameters b and s. The slope s can be estimated from the observed number counts. The bias b can be inferred by comparing the amplitude of the galaxy/quasar auto–correlation with the amplitude of the mass auto–correlation predicted by the microwave background, modulo the fact that the observed galaxy/quasar auto–correlation function is itself affected by magnification bias (though it appears to be at

most a 20% effect at the scales and redshifts we study). Alternatively, the bias can be estimated by examining the galaxy/quasar higher order correlations. The upshot is that separating the galaxy–temperature and the magnification–temperature signals is in principle possible, at least in a model dependent manner. Other possibilities are to exploit the sample–dependent nature of b and s to study the variation of the total ISW signal with different sample cuts, or to use the galaxy–galaxy correlation across different redshift bins to check for consistency with the ISW–magnification bias signal.

An interesting question remains to be explored: given a survey with many different kinds of objects, such as galaxies/quasars of different luminosities and types, how should one go about choosing a weighting scheme that optimizes the total signal and/or its magnification component? The magnification component might be particularly interesting if the galaxy/quasar bias turns out to be uncertain. We leave this for future investigation.

Acknowledgments

We would like to thank Ryan Scranton for useful discussions regarding magnification bias, Asantha Cooray for addressing questions we had about his results on the conditional luminosity function, Elena Zucca for kindly sending us her group's results on the luminosity function for several galaxy types (see [45]), and Ben Johnson for helpful discussions regarding luminosity functions and galaxy samples. Support for this work is provided in part by the DOE, grant number DE-FG02-92-ER40699, and the Initiatives in Science and Engineering Program at Columbia University. EG acknowledges support from Spanish Ministerio de Ciencia y Tecnologia (MEC), project AYA2006-06341 with EC-FEDER funding, research project 2005SGR00728 from Generalitat de Catalunya.

APPENDIX A: ON ESTIMATION BIAS AND ERROR

The technology for making error forecasts is well documented in the literature [25]. What seems to be less often discussed is the issue of predicting the bias: what is the resulting bias in the estimation of a parameter if some particular physical effect is ignored or some approximation is made in modeling the data? In our case, the parameter of interest is w and the physical effect is the magnification bias. We keep our discussion here fairly general since it might be of interest to researchers working on other problems. Much of our discussion is a straightforward generalization of the classic paper by Rybicki and Press [46].

In general, we have the problem of trying to obtain constraints on m parameters, labeled p_α with $\alpha = 1, 2, ..., m$, from an n-dimensional data vector, $\hat{d}_i$ with $i = 1, 2, ..., n$. Assuming Gaussian distributed data, the likelihood is proportional to $\exp\left[-\chi^2/2\right]$, where

$$\chi^2 = \sum_{i,j} (\hat{d}_i - d_i) C_{ij}^{-1} (\hat{d}_j - d_j) \tag{A1}$$

where $C_{ij} \equiv \langle (\hat{d}_i - d_i)(\hat{d}_j - d_j) \rangle$, with $d_i \equiv \langle \hat{d}_i \rangle$. The vector (component) d_i is in general some complicated function of the parameters p_α. The trick is to turn this into a linear estimation problem by supposing that p_α is close to some fiducial value $\bar{p}_\alpha$, and therefore:

$$d_i = \bar{d}_i + \frac{\partial d_i}{\partial p_\alpha} \delta p_\alpha \quad , \quad \delta p_\alpha \equiv p_\alpha - \bar{p}_\alpha \tag{A2}$$

where $\bar{d}_i$ is the expected data average if the parameter values were indeed $\bar{p}_\alpha$, and the partial derivative is evaluated at $p_\alpha = \bar{p}_\alpha$.

Substituting the linear model eq. [A2] into the expression for χ^2, we obtain

$$\chi^2 = \sum_{\alpha,\beta} (\delta p_\alpha - \delta \hat{p}_\alpha^{\max}) F_{\alpha\beta} (\delta p_\beta - \delta \hat{p}_\beta^{\max}) \tag{A3}$$

up to an additive constant independent of δp_α. Here,

$$\delta \hat{p}_\alpha^{\max} \equiv \sum_{\beta,i,j} F_{\alpha\beta}^{-1} \frac{\partial d_i}{\partial p_\beta} C_{ij}^{-1} (\hat{d}_j - \bar{d}_j) \tag{A4}$$

where the $\hat{}$ on top of $\delta \hat{p}_\alpha^{\max}$ is there to remind us that this quantity depends on the data $\hat{d}_j$ (i.e. a stochastic quantity), and $F_{\alpha\beta}$ is the Fisher matrix

$$F_{\alpha\beta} \equiv \sum_{i,j} \frac{\partial d_i}{\partial p_\alpha} C_{ij}^{-1} \frac{\partial d_j}{\partial p_\beta} \tag{A5}$$

Therefore, the maximum likelihood estimate for p_α is

$$\hat{p}_\alpha^{\max} = \bar{p}_\alpha + \delta \hat{p}_\alpha^{\max} \tag{A6}$$

and the error covariance associated with such an estimate is given by the inverse of the Fisher matrix:

$$\langle \hat{p}_\alpha^{\max} \hat{p}_\beta^{\max} \rangle - \langle \hat{p}_\alpha^{\max} \rangle \langle \hat{p}_\beta^{\max} \rangle = F_{\alpha\beta}^{-1} \tag{A7}$$

Combining eq. [A4] and [A6], we see that

$$\langle \hat{p}_\alpha^{\max} \rangle = \bar{p}_\alpha + \sum_j w_{\alpha j} (\langle \hat{d}_j \rangle - \bar{d}_j) \tag{A8}$$

$$w_{\alpha j} \equiv \sum_{\beta,i} F_{\alpha\beta}^{-1} \frac{\partial d_i}{\partial p_\beta} C_{ij}^{-1}$$

The estimation bias of interest is therefore $\langle \hat{p}_\alpha^{\max} \rangle - p_\alpha^{\rm true}$, where $p_\alpha^{\rm true}$ is the true value of the underlying parameter. Note that both the weights $w_{\alpha j}$ and the vector $\bar{d}_j$ are model dependent. For instance, one could choose to ignore certain physical effects, such as magnification bias, in constructing these quantities. The result would be a biased estimate of p_α.

Let us suppose first that one is using the correct or *exact* model. Then, substituting $\langle\hat{d}_j\rangle = \bar{d}_j + [\partial d_j/\partial p_\gamma](p_\gamma^{\rm true} - \bar{p}_\gamma)$ into eq. [A8], we have $\langle\hat{p}_\alpha^{\rm max}\rangle = p_\alpha^{\rm true}$, as expected.

Suppose instead one is using an incorrect or *approximate* model, where a certain physical effect (e.g. magnification bias) has been ignored. This means $\bar{d}_j$ and the weights $w_{\alpha j}$ are all computed by ignoring this effect, whereas $\langle\hat{d}_j\rangle$, which is the expectation value of the observed data, of course has such a physical effect (and all other relevant physical effects) taken into account. The preceding argument (concerning the *exact* model) no longer works and one is left with an estimation bias i.e. $\langle\hat{p}_\alpha^{\rm max}\rangle \neq p_\alpha^{\rm true}$. Note that this bias depends on the assumed fiducial parameter value $\bar{p}_\alpha$ (as opposed to the case of unbiased estimation, where $\langle\hat{p}_\alpha^{\rm max}\rangle$ does not depend on $\bar{p}_\alpha$ at all). Generally, a modeler would choose a fiducial value that is close to the *perceived* (i.e. possibly incorrect, due to the use of an incorrect/approximate model) maximum likelihood value. We therefore obtain the estimation bias by iteration: starting with some fiducial parameter value, compute $\langle\hat{p}_\alpha^{\rm max}\rangle$, then adopt the result as a new fiducial value and iterate. In all cases investigated in this paper, a few iterations are sufficient to guarantee convergence. The modeler would assign a (theoretical) errorbar to his/her biased estimate based on eq. [A7], where the Fisher matrix is, again, that of the *approximate* model.

Lastly, let us address the question of priors. Priors can be thought of as just another kind of data, and they add to the χ^2 in eq. [A1] a term of the form:

$$\sum_{\alpha,\beta}(\hat{p}_\alpha - p_\alpha)M^{-1}_{\alpha\beta}(\hat{p}_\beta - p_\beta)$$

where $\hat{p}_\alpha$ denotes the assumed prior value (e.g. $\Omega_m = 0.27$, etc) and $M_{\alpha\beta}$ is the covariance associated with this set of priors. Completing square as before, one obtains, up to an irrelevant additive constant, a χ^2 slightly different from the one in eq. [A3]:

$$\chi^2 = \sum_{\alpha,\beta}(\delta p_\alpha - \delta\hat{p}_\alpha^{\rm max})F'_{\alpha\beta}(\delta p_\beta - \delta\hat{p}_\beta^{\rm max}) \qquad \text{(A9)}$$

where $F'_{\alpha\beta}$ is a generalized Fisher matrix, which is related to the original Fisher matrix by

$$F'_{\alpha\beta} = F_{\alpha\beta} + M^{-1}_{\alpha\beta} \qquad \text{(A10)}$$

and the maximum likelihood estimate $\hat{p}_\alpha^{\rm max} = \bar{p}_\alpha + \delta\hat{p}_\alpha^{\rm max}$ is given by (superseding eq. [A4]):

$$\delta\hat{p}_\alpha^{\rm max} \equiv \sum_\beta {F'}^{-1}_{\alpha\beta}\Big[\sum_{i,j}\frac{\partial d_i}{\partial p_\beta}C_{ij}^{-1}(\hat{d}_j - \bar{d}_j) + \sum_\gamma M^{-1}_{\beta\gamma}(\hat{p}_\gamma - \bar{p}_\gamma)\Big] \qquad \text{(A11)}$$

Therefore, the expectation value of $\hat{p}_\alpha^{\rm max}$ is given by

$$\langle\hat{p}_\alpha^{\rm max}\rangle = \bar{p}_\alpha + \sum_\beta {F'}^{-1}_{\alpha\beta}\Big[\sum_{i,j}\frac{\partial d_i}{\partial p_\beta}C_{ij}^{-1}(\langle\hat{d}_j\rangle - \bar{d}_j) + \sum_\gamma M^{-1}_{\beta\gamma}(\langle\hat{p}_\gamma\rangle - \bar{p}_\gamma)\Big] \qquad \text{(A12)}$$

A set of priors that are unbiased implies $\langle\hat{p}_\gamma\rangle = p_\gamma^{\rm true}$. The estimation bias is given by $\langle\hat{p}_\alpha^{\rm max}\rangle - p_\alpha^{\rm true}$, and the error covariance is given by the inverse of the generalized Fisher matrix:

$$\langle\hat{p}_\alpha^{\rm max}\hat{p}_\beta^{\rm max}\rangle - \langle\hat{p}_\alpha^{\rm max}\rangle\langle\hat{p}_\beta^{\rm max}\rangle = {F'}^{-1}_{\alpha\beta} \qquad \text{(A13)}$$

Eq. [A12] and [A13] constitute the main results of this Appendix. Let us illustrate their use by applying them to a question of interest in this paper: what is the resulting bias on cosmological parameters if one ignores magnification bias when interpreting ISW measurements? The parameters p_α stand for w, Ω_m and so on. The data $\langle\hat{d}_j\rangle$ are the $C^{gT} + C^{\mu T}$ at various ℓ's and redshifts (collectively labeled by the index j here). The C_{ij} matrix here is the covariance described in eq. [22]. To figure out the estimation bias on parameter p_α (which in our case is mainly the dark energy equation of state w since all other parameters are fairly constrained by our priors), one applies eq. [A12] with $F'_{\alpha\beta}$, $\partial d_i/\partial p_\beta$, C_{ij} and $\bar{d}_j$ all computed ignoring magnification bias. To be completely explicit, eq. [A12] tells us that the average inferred value for parameter p_α (e.g. $p_1 = w$ and so on) is:

$$p_\alpha^{\rm inferred} = \bar{p}_\alpha + \sum_\beta {F'}^{g\,-1}_{\alpha\beta} \Big[\sum_{\ell,i,j}\frac{\partial C_\ell^{gT}(z_i)}{\partial p_\beta}[\,{\rm Cov}_\ell^g]^{-1}_{ij}(C_\ell^{nT}(z_j) - C_\ell^{gT}(z_j)) + \sum_\gamma M^{-1}_{\beta\gamma}(p_\gamma^{\rm true} - \bar{p}_\gamma)\Big] \qquad \text{(A14)}$$

where F'^g is the generalized Fisher matrix (eq. [A10]) (the superscript g is supposed to remind us that magnification is ignored) i.e.

$${F'^g}_{\alpha\beta} = M^{-1}_{\alpha\beta} + \sum_{\ell,i,j}\frac{\partial C_\ell^{gT}(z_i)}{\partial p_\alpha}[\,{\rm Cov}^{\,g}_\ell]^{-1}_{ij}\frac{\partial C_\ell^{gT}(z_j)}{\partial p_\beta} \qquad \text{(A15)}$$

and the covariance (with magnification ignored) is:

$$[\,{\rm Cov}_\ell^g]_{ij} = \frac{(C_\ell^{gg}(z_i,z_j) + \delta_{ij}/n_i)C_\ell^{TT} + C_\ell^{gT}(z_i)C_\ell^{gT}(z_j)}{f_{\rm sky}(2\ell+1)} \qquad \text{(A16)}$$

where n_i is the mean number of galaxies per unit area in redshift bin i. Note that the covariance given above is not exact (nor is the implicit assumption that the different ℓ modes are uncorrelated), but it should be fairly accurate for modes with $\ell \gtrsim 10$ (recall that we use $f_{\rm sky} = 0.5$ in our worked examples), which dominate the signal-to-noise. The same statement can be made about the Fisher matrix.

The fiducial parameter values $\bar{p}_\alpha$'s should be chosen to be close to the *perceived* best–fit values — recall that to a

modeler who ignores magnification, the *perceived* best–fit differs from the true values. All quantities on the right hand side of eq. [A14], such as F'^g, C_ℓ^{gT} and the covariance, should be evaluated with the parameters set at the fiducial values $\bar{p}_\alpha$'s. The only exception is $C_\ell^{nT}(z_j)$ which represents the average *observed* cross-correlation and should of course be computed using the true parameter values (e.g. $w=-1$ and so on in our thought experiment in §VI). In principle, the *perceived* best-fit values can differ quite a bit from the true values. We therefore apply eq. [A14] iteratively: starting with $\bar{p}_\alpha = p_\alpha^{\rm true}$, we use eq. [A14] to compute $p_\alpha^{\rm inferred}$, and then setting $\bar{p}_\alpha = p_\alpha^{\rm inferred}$, we reapply eq. [A14] and so on. Generally, convergence is achieved with only a few iterations.

The estimation-bias of interest is $p_\alpha^{\rm inferred} - p_\alpha^{\rm true}$. The associated (theoretical/model-dependent) error covariance is given by the inverse of the generalized Fisher matrix F'^g (eq. [A15]) that ignores magnification, and is evaluated at the same $\bar{p}_\alpha$'s as above.

If one is instead interested in making error forecasts when magnification bias is taken into account (as in §VII), one should use eq. [A13] with a $F'_{\alpha\beta}$ that is computed without ignoring magnification bias i.e.

$$F'_{\alpha\beta} = M^{-1}_{\alpha\beta} + \sum_{\ell,i,j} \frac{\partial C_\ell^{nT}(z_i)}{\partial p_\alpha} [\,{\rm Cov}\,_\ell]^{-1}_{ij} \frac{\partial C_\ell^{nT}(z_j)}{\partial p_\beta} \quad \text{(A17)}$$

where the covariance is given by eq. [22].

[1] R. Sachs, A. Wolfe, Astrophys. J 147, 73 (1967)
[2] L. A. Kofman, A. A. Starobinskii, Sov. Astron. Lett. 11, 271 (1985)
[3] R. G. Crittenden, N. Turok, Phys. Rev. Lett. 76, 575 (1996)
[4] P. Fosalba, E. Gaztañaga, F. Castanader, Astrophys. J. Lett. 597, L89 (2003)
[5] R. Scranton, et al astro-ph/0307335 (2003)
[6] M. R. Nolta, *et al.*, Astrophys. J. 608, 10 (2004)
[7] S. P. Boughn, R. G. Crittenden, Nature 427 45 (2004)
[8] P. Fosalba, E. Gaztañaga, Mon. Not. R. Astron. Soc. 350 , L37 (2004)
[9] N. Afshordi, Y.-S. Loh, M. Strauss, Phys. Rev. D 69 083524 (2004)
[10] A. Cabré, E. Gaztañaga, M. Manera, P. Fosalba Mon. Not. R. Astron. Soc. Lett. , 372, L23 (2006)
[11] T. Giannantonio *et al.*, Phys. Rev. D 74 063520 (2006)
[12] D. Pietrobon, A. Balbi, D. Marinucci, Phys. Rev. D 74 043524 (2006)
[13] M. Tegmark *et al.*, Phys. Rev. D 74, 123507 (2006)
[14] D. Spergel, *et al.*Astrophys. J. submitted (2006) astro-ph/0603449
[15] R. Narayan, Astrophys. J. Lett. 339, 53 (1989)
[16] T. J. Broadhurst, A. N. Taylor, J. A. Peacock, Astrophys. J. 438, 49 (1996)
[17] J. Villumsen, Mon. Not. R. Astron. Soc. 281 , 369 (1996)
[18] J. Villumsen, W. Freudling, L. N. da Costa, Astrophys. J. 481, 578 (1997)
[19] R. Moessner, B. Jain, J. Villumsen, Mon. Not. R. Astron. Soc. 294, 291 (1998)
[20] R. Moessner, B. Jain, Mon. Not. R. Astron. Soc. 294, L18 (1998)
[21] N. Kaiser, Astrophys. J. 498, 26 (1998)
[22] E. Gaztañaga, Astrophys. J. 589, 82 (2003)
[23] R. Scranton, *et al.*, Astrophys. J. 633, 589 (2005)
[24] B. Jain, R. Scranton, R. K. Sheth, Mon. Not. R. Astron. Soc. 345, 62 (2003)
[25] For a review, see S. Dodelson, Modern Cosmology, Academic Press (2003). Note that our ϕ here equals Ψ there, which also equals $-\Phi$ there in the absence of anisotropic stress.
[26] Strictly speaking, the factor of $2.5s-1$ should be absorbed into g where s is made a function of the integration variable z'. Here, we make the approximation that s is roughly constant within the redshift bin i.
[27] D. J. Eisenstein, W. Hu, Astrophys. J. 496, 605 (1998)
[28] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. 538, 473 (2000), http://camb.info
[29] A. Gabasch *et al.*, Astron. & Astrophys. 421, 41 (2004) [astro-ph/0403535]; A. Gabasch *et al.*, Astron. & Astrophys. in press (2006) [astro-ph/0510339]. To be precise: we make use of Table 4 of the former, and Table 9 (Case 3) of the latter, and linearly interpolate between adjacent rest bands. Our assumed observed I-band blueshifts (going back in time) beyond the bluest rest-frame band (centered at $1500\mathring{A}$) given in the above papers by a redshift of about 4.33. For redshifts higher than this, we simply use the luminosity function centered at rest-frame $1500\mathring{A}$ to obtain dN/dz etc – altering the extrapolation scheme should have only a mild impact on the highest z bin we consider.
[30] H. Zhan, L. Knox, J. A. Tyson, V. Margoniner, Astrophys. J. 640, 8 (2006)
[31] W. Hu, R. Scranton, Phys. Rev. D 70, 123002 (2004)
[32] N. Afshordi, Phys. Rev. D 70 083536 (2004)
[33] There is a slight caveat to this statement. If $s<0.4$, $C^{\mu T}$ will be negative. For a given z_i, $C_\ell^{\mu T}(z_i)$ is peaked at lower ℓ than $C_\ell^{gT}(z_i)$, thus for $s<0.4$ the peak of the sum $C_\ell^{nT}(z_i) = C_\ell^{gT}(z_i) + C_\ell^{\mu T}(z_i)$ is at higher ℓ than it would be if there were no magnification. So, in principle if the magnification term, $C^{\mu T}$ were large and negative the range of largest signal to noise may be shifted beyond $\ell = 100$. However this is unlikely in practice because magnification bias does not become important until high redshift ($z \gtrsim 2$), and at high redshifts one is typically observing bright objects which tend to have $s > 0.4$, unless one has an exceptionally deep survey.
[34] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, S. Gottlober, B. Allgood, J. R. Primack, Astrophys. J. 609, 35 (2004)
[35] http://phys.columbia.edu/∼marilena/research.html
[36] R. K. Sheth, G. Tormen, Mon. Not. R. Astron. Soc. 308, 119 (1999)
[37] Strictly speaking, for Figures 8 and 10, the lines should really be represented as points, since the quantities plotted are fundamentally derived from discrete redshift bins.

We join them up as lines for clarity of presentation. For Figures 9 and 11, there is one good justification for connecting the dots: in the limit the redshift bins become smaller and smaller (and the number of bins increases), the cumulative signal-to-noise or the cumulative error should approach a smooth curve approximated by the one we present.

[38] J. Garriga, L. Pogosian, T. Vachaspati, Phys. Rev. D 69, 063511 (2004)

[39] L. Pogosian, JCAP 0504, 015 (2005)

[40] L. Pogosian, P.-S. Corasaniti, C. Stephan-Otto, R. Crittenden, R. Nichol Phys. Rev. D 72, 103519 (2005)

[41] P. Zhang, Astrophys. J. 647, 55 (2006)

[42] E. Gaztañaga, M. Manera, T. Multamäki, Mon. Not. R. Astron. Soc. 365, 171 (2006)

[43] Strictly speaking, for Figure 12, the dotted lines should really be represented as points, since w_{nT}/b has a residual dependence on the bias b, and the bias is fundamentally sample specific and does not necessarily evolve in a smooth manner. We join up the points as lines for clarity of presentation.

[44] A. Lue, R. Scoccimarro, G. Starkman, Phys. Rev. D 69, 044005 (2004); P. Zhang, Phys. Rev. D 73, 123504 (2006); I. Sawicki, S. M. Carroll, preprint (2005) [astro-ph/0510364], Y.-S. Song, I. Sawicki, W. Hu, submitted to Phys. Rev. D (2006) [astro-ph/0606286]

[45] E. Zucca *et al.*, Astron. & Astrophys. in press (2006) [astro-ph/0506393]

[46] G. B. Rybicki, W. H. Press, Astrophys. J. 398, 169 (1992)