



Kinematic Sunyaev-Zel'dovich effect with projected fields. II. Prospects, challenges, and comparison with simulations

Simone Ferraro,^{1,2,*} J. Colin Hill,³ Nick Battaglia,⁴ Jia Liu,³ and David N. Spergel^{4,5}

¹*Berkeley Center for Cosmological Physics and Department of Astronomy, University of California, Berkeley, California 94720, USA*

²*Miller Institute for Basic Research in Science, University of California, Berkeley, California 94720, USA*

³*Department of Astronomy, Pupin Hall, Columbia University, New York, New York 10027, USA*

⁴*Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, New Jersey 08544, USA*

⁵*Center for Computational Astrophysics, 162 5th Ave, New York, NY 10003, USA*

(Received 18 May 2016; published 28 December 2016)

The kinematic Sunyaev-Zel'dovich (kSZ) signal is a powerful probe of the cosmic baryon distribution. The kSZ signal is proportional to the integrated free electron momentum rather than the electron pressure (which sources the thermal SZ signal). Since velocities should be unbiased on large scales, the kSZ signal is an unbiased tracer of the large-scale electron distribution, and thus can be used to detect the “missing baryons” that evade most observational techniques. While most current methods for kSZ extraction rely on the availability of very accurate redshifts, we revisit a method that allows measurements even in the absence of redshift information for individual objects. It involves cross-correlating the *square* of an appropriately filtered cosmic microwave background (CMB) temperature map with a projected density map constructed from a sample of large-scale structure tracers. We show that this method will achieve high signal-to-noise when applied to the next generation of high-resolution CMB experiments, provided that component separation is sufficiently effective at removing foreground contamination. Considering statistical errors only, we forecast that this estimator can yield $S/N \approx 3$, 120 and over 150 for Planck, Advanced ACTPol, and a hypothetical Stage IV CMB experiment, respectively, in combination with a galaxy catalog from WISE, and about 20% larger S/N for a galaxy catalog from the proposed SPHEREx experiment. We show that the basic estimator receives a contribution due to leakage from CMB lensing, but that this term can be effectively removed by either direct measurement or marginalization, with little effect on the kSZ significance. We discuss possible sources of systematic contamination and propose mitigation strategies for future surveys. We compare the theoretical predictions to numerical simulations and validate the approximations in our analytic approach. This work serves as a companion paper to the first kSZ measurement with this method, where we used CMB temperature maps constructed from Planck and WMAP data, together with galaxies from the WISE survey, to obtain a $3.8\text{--}4.5\sigma$ detection of the kSZ^2 amplitude.

DOI: [10.1103/PhysRevD.94.123526](https://doi.org/10.1103/PhysRevD.94.123526)

I. INTRODUCTION

The amount of baryonic matter in the Universe is tightly constrained at high redshift by measurements of the primordial cosmic microwave background (CMB) anisotropies [1,2] and of the abundance of light elements formed through the process of big bang nucleosynthesis (BBN) [3]. The baryonic abundance of the present-day Universe must satisfy these primordial constraints, assuming the absence of unknown, exotic physics. However, the cosmic baryon census at low redshifts has long fallen short of the expected value (e.g., [4,5]), especially for halos smaller than galaxy clusters, such as individual galaxies or groups of galaxies. One hypothesis is that these “missing baryons” reside in an ionized, diffuse component known as the warm-hot

intergalactic medium [6], which has been difficult to detect in x -ray emission due to its relatively low density and temperature. Observations of highly ionized gas in quasar absorption lines provide some evidence and constraints on its properties [7,8].

The kinematic Sunyaev-Zel'dovich (kSZ) effect is caused by Compton scattering of CMB photons off of free electrons moving with a nonzero line-of-sight (LOS) velocity [9–11]. The corresponding shift in the observed CMB temperature is proportional to both the total number of electrons (or optical depth) and their LOS velocity, which is equally likely to be positive or negative. Moreover, the kSZ signal should be unbiased, in the sense that halos of different masses move in the same large-scale cosmic velocity field, and therefore it is a direct probe of the electron density. Thus it can be used to measure the ionized gas abundance and distribution in galaxies and clusters.

*sferraro@berkeley.edu

These measurements can be performed as a function of mass and redshift (and other galaxy properties of interest), informing us about the extent and nature of feedback processes.

If the cluster optical depth can be determined through other methods, the kSZ effect can be used to measure statistics of LOS velocities, which are sensitive to the rate of growth of structure and are hence a powerful probe of dark energy or modified gravity [12].

The kSZ effect was first detected in Atacama Cosmology Telescope (ACT) data by studying the pairwise momenta of luminous galaxies in the Baryon Oscillation Spectroscopic Survey (BOSS) DR9 catalog [13]. Recent analyses of the Planck, ACTPol and South Pole Telescope (SPT-SZ) data sets have found additional evidence for the signal, using large-scale structure catalogs from the Sloan Digital Sky Survey (SDSS) and Dark Energy Survey (DES) [14–17]. A high-resolution analysis of a particular galaxy cluster also found evidence for the kSZ effect in that system [18,19].

Most kSZ estimators in the literature [12,20–23] require spectroscopic redshifts. The use of photometric redshifts leads to a large degradation in the statistical significance of the kSZ detection [24,25]. In this paper, we revisit a method that only makes use of projected fields and therefore does not require individual redshifts for each object, but only a statistical redshift distribution for the low-redshift tracers used in the analysis. Such a distribution could be constructed from photometric redshift data, but even photometric redshifts are not necessarily required—a well-understood subsample cross-matched to existing redshift catalogs would suffice. The main motivation of this estimator is that photometric or imaging surveys are much cheaper than their spectroscopic counterparts and are able to map larger volumes of the Universe. An excellent example is the Wide-Field Infrared Survey Explorer (WISE) data set [26], which covers the full sky in the midinfrared. Moreover, the kSZ technique described here will have comparable statistical power to and yield information independent of the traditional methods when applied to future high-resolution CMB experiments, if component separation allows an effective removal of frequency-dependent foregrounds.

The basic idea behind this estimator is that because of the equal likelihood of positive and negative kSZ signals, an appropriately filtered version of the CMB temperature map must be squared in real space before cross-correlating with tracers (e.g., galaxies, quasars, or gravitational lensing convergence); we thus refer to this as the kSZ²-tracer cross-correlation. Crucially, the CMB temperature map must be cleaned of foreground (non-kSZ) emission associated with the tracer objects, and thus a multifrequency analysis is necessary. First suggested in [27] and studied further in [28], the kSZ²-tracer cross-correlation probes the mass and LOS velocity of the ionized gas associated with the tracer objects in the large-scale structure sample. In

other words, the CMB temperature itself contains kSZ information, and this is just the lowest-order nonzero estimator that allows one to extract the signal from a given tracer population without requiring 3D information. This is in essence a measurement of the *squeezed* limit of the bispectrum of two powers of the CMB temperature and one power of the projected tracer field, and it can be shown to be the configuration containing most of the information (we leave a full treatment of optimality to future work).

Because the estimator is quadratic in temperature, it is affected by leakage from weak lensing of the CMB, and this lensing contribution—which can be larger than the signal in some instances—must be appropriately removed or marginalized over. Fortunately, the multipole dependence of the lensing leakage is quite different than the kSZ² signal, and thus it can be marginalized with very little effect on the statistical significance of the kSZ² signal.

We have recently presented the first measurement of the baryon abundance with this technique in a companion paper [29] (hereafter H16). We used a galaxy catalog constructed from WISE data [26] and CMB temperature maps cleaned via “local-generalized morphological component analysis” (LGMCA) [30] constructed from the Planck full mission [2] and Wilkinson Microwave Anisotropy Probe nine-year survey (WMAP9) data [31]. We detected the kSZ² signal with signal-to-noise (S/N) \approx 3.8–4.5, depending on the use of external CMB lensing information, and thus obtained a 13% measurement of the baryon abundance at $z \approx 0.4$.

Except where explicitly stated otherwise, we use cosmological parameters from the 2015 Planck data release [32].

The remainder of this paper is organized as follows. In Sec. II we review the theory, including the approximations in our analytic approach. In Sec. III we present forecasts for current and future experiments, while Sec. IV discusses the lensing contribution and ways to remove it. In Sec. V, we present a comparison of the theory with numerical simulations to check the accuracy of our approximations. We discuss our recent measurement using this method in Sec. VI. Foreground contamination poses a serious challenge for this type of measurement, which we discuss in Sec. VII. We conclude in Sec. VIII.

II. THEORY

The kSZ effect produces a CMB temperature change, $\Theta^{\text{kSZ}}(\hat{\mathbf{n}}) = \Delta T^{\text{kSZ}}/T_{\text{CMB}}(\hat{\mathbf{n}})$, in a direction $\hat{\mathbf{n}}$ on the sky (in units with $c = 1$):

$$\Theta^{\text{kSZ}}(\hat{\mathbf{n}}) = - \int d\eta g(\eta) \mathbf{p}_e \cdot \hat{\mathbf{n}} \quad (1)$$

$$= -\sigma_T \int \frac{d\eta}{1+z} e^{-\tau} n_e(\hat{\mathbf{n}}, \eta) \mathbf{v}_e \cdot \hat{\mathbf{n}}, \quad (2)$$

where σ_T is the Thomson scattering cross section, $\eta(z)$ is the comoving distance to redshift z , τ is the optical depth to Thomson scattering, $g(\eta) = (d\tau/d\eta)e^{-\tau}$ is the visibility function, n_e is the physical free electron number density, \mathbf{v}_e is the peculiar velocity of the electrons, and we have defined the electron momentum $\mathbf{p}_e = (1 + \delta_e)\mathbf{v}_e$.

For concreteness, we consider galaxies as tracers in the following, but the formalism extends straightforwardly to any other tracer of the late-time density field (such as quasars, lensing convergence, or 21 cm fluctuations).

The projected galaxy overdensity δ_g is given by

$$\delta_g(\hat{\mathbf{n}}) = \int_0^{\eta_{\max}} d\eta W^g(\eta) \delta_m(\eta \hat{\mathbf{n}}, \eta), \quad (3)$$

where η_{\max} is the maximum source distance, δ_m is the (three-dimensional) matter overdensity, and $W^g(\eta)$ is the projection kernel:

$$W^g(\eta) = b_g p_s(\eta). \quad (4)$$

Here $p_s(\eta) \propto dn/d\eta$ is the redshift distribution of the galaxies (normalized to have unit integral) and b_g is the linear galaxy bias.

As explained in the Introduction, the cross-correlation between the kSZ signal and low-redshift tracers is expected to vanish on small scales [where the contribution from the integrated Sachs-Wolfe (ISW) effect is expected to be negligible] because of the $\mathbf{v}_e \rightarrow -\mathbf{v}_e$ symmetry. We therefore square the CMB temperature fluctuation map in real space before cross-correlating it with a tracer density map.

In order to downweight angular scales dominated by noise (in our case primary CMB fluctuations and detector noise), we filter the temperature map in harmonic space with a Wiener filter F before squaring in real space:

$$F(\ell) = \frac{C_\ell^{\text{kSZ}}}{C_\ell^{\text{tot}}}, \quad (5)$$

where C_ℓ^{kSZ} is the (theoretical) kSZ power spectrum and C_ℓ^{tot} is the total fluctuation power, which includes primary CMB, kSZ, ISW, noise, and any residual foregrounds. Our template for C_ℓ^{kSZ} in Eq. (5) is derived from cosmological hydrodynamics simulations [33].

Moreover, the CMB is observed through a finite beam $b(\ell)$, so that the total filtered map Θ_f is related to the underlying (true) CMB anisotropy Θ by

$$\Theta_f(\ell) = F(\ell)b(\ell)\Theta(\ell) \equiv f(\ell)\Theta(\ell) \quad (6)$$

where we have defined $f(\ell) = F(\ell)b(\ell)$.

In this work, we are interested in the cross-correlation $C_\ell^{\text{kSZ}^2 \times \delta_g}$ between the square of the filtered CMB map and tracers:

$$\langle \Theta_f^2(\ell) \delta_g(\ell') \rangle = (2\pi)^2 \delta_D(\ell + \ell') C_\ell^{\text{kSZ}^2 \times \delta_g}. \quad (7)$$

Following [27,28] we can write the angular power spectrum of the kSZ²-galaxy cross-correlation as

$$C_\ell^{\text{kSZ}^2 \times \delta_g} = \int_0^{\eta_{\max}} \frac{d\eta}{\eta^2} W^g(\eta) g^2(\eta) \mathcal{T}(k = \ell/\eta, \eta), \quad (8)$$

where we have used the Limber approximation [34], and the *triangle power spectrum* \mathcal{T}

$$\mathcal{T}(k, \eta) = \int \frac{d^2\mathbf{q}}{(2\pi)^2} f(q\eta) f(|\mathbf{k} + \mathbf{q}| \eta) B_{\delta p_{\hat{\mathbf{n}}} p_{\hat{\mathbf{n}}}}(\mathbf{k}, \mathbf{q}, -\mathbf{k} - \mathbf{q}). \quad (9)$$

Here, the hybrid bispectrum $B_{\delta p_{\hat{\mathbf{n}}} p_{\hat{\mathbf{n}}}}$ is the three-point function of one density contrast and two LOS electron momenta, $p_{\hat{\mathbf{n}}}$. The triangle power spectrum \mathcal{T} is the integral over all triangles with sides \mathbf{k} , \mathbf{q} , and $-\mathbf{k} - \mathbf{q}$, lying on planes of constant redshift. Since the momentum field is $\mathbf{p} \sim \mathbf{v}\delta$ on small scales, the hybrid bispectrum $B_{\delta p_{\hat{\mathbf{n}}} p_{\hat{\mathbf{n}}}}$ is the sum of terms of the form $\langle vv \rangle \langle \delta\delta\delta \rangle$, $\langle v\delta \rangle \langle \delta\delta v \rangle$, etc., and a connected part $\langle vv\delta\delta\delta \rangle_c$. DeDeo *et al.* [28] argued that the former term $\langle vv \rangle \langle \delta\delta\delta \rangle$ dominates on small scales ($k \gg k_{\text{eq}}$) and we will assume that the non-Gaussianity is weak enough that the connected part can be neglected.

On small scales we can therefore approximate the hybrid bispectrum in terms of the 3D velocity dispersion v_{rms}^2 and the nonlinear matter bispectrum B_m^{NL} [27,28]:

$$B_{\delta p_{\hat{\mathbf{n}}} p_{\hat{\mathbf{n}}}} \approx \frac{1}{3} v_{\text{rms}}^2 B_m^{\text{NL}}. \quad (10)$$

We use fitting functions from [35] for the nonlinear matter bispectrum B_m^{NL} and the velocity dispersion v_{rms}^2 is computed in linear theory, which should be an excellent approximation.¹ We test the validity of the approximations made here by comparison to numerical simulations in Sec. V, and we find that these are excellent on the scales relevant for the analysis of a Planck-like experiment.

At late times, some fraction of the cosmological abundance of electrons lies in stars or neutral media and therefore does not take part in the Thomson scattering that produces the kSZ signal. We define f_{free} as the fraction of free electrons, and note that in general this quantity will be redshift dependent. The visibility function $g(\eta)$ in Eq. (1) is proportional to f_{free} , so that $C_\ell^{\text{kSZ}^2 \times \delta_g}$ scales like f_{free}^2 and

¹Numerical simulations [36] show that linear theory is a very good approximation to the velocity power spectrum up to $k \approx 0.5h/\text{Mpc}$, and therefore the velocity dispersion, which receives most of its contribution from larger scales, should be well approximated by linear theory.

hence can be used to measure the free electron fraction. In H16 we note that the signal is also proportional to the (square of the) baryon fraction $f_b = \rho_b/\rho_m$, so that if we allow f_b to vary, the amplitude of $C_\ell^{\text{kSZ}^2 \times \delta_g}$ provides a measurement of the product $f_{\text{free}} f_b$. For convenience in what follows we will fix $f_b = 0.155$, the fiducial value in our assumed cosmology.

Technically, the bispectrum in Eq. (10) is the three-point function of one matter and two electron overdensities, but for the purpose of forecasts, we will assume that the free electrons trace the dark matter down to the scales of interest. While this is expected to be true for an experiment with the resolution of Planck, this assumption will not hold as experiments proceed to higher resolution. The overall amplitude of the signal is set by f_{free} , but the shape of the cross-correlation on small scales is directly related to the baryon profiles around galaxies and clusters, which are expected to be heavily influenced by feedback processes (for a measurement of the kSZ signal as a function of scale for group-size tracers see [16]).

III. FORECASTS

In this section we present forecasts for detection of the kSZ^2 signal. As discussed above, the amplitude of $C_\ell^{\text{kSZ}^2 \times \delta_g}$ is proportional to the galaxy bias b_g so that we can define

$$(C_\ell^{\text{kSZ}^2 \times \delta_g})_{\text{measured}} = b_g \mathcal{A}_{\text{kSZ}^2} (C_\ell^{\text{kSZ}^2 \times \delta_g})_{\text{fiducial}} \quad (11)$$

where the fiducial prediction assumes unit galaxy bias and full ionization, such that $\mathcal{A}_{\text{kSZ}^2} \propto f_{\text{free}}^2$. It is often the case that the galaxy bias is either known externally to high accuracy (for example from the autocorrelation function or in cross-correlation with CMB lensing maps) or absent (for example if our tracer were lensing convergence). Therefore, in this section we will assume that we have an external sharp prior on the bias, so that the fractional error on $(b_g \mathcal{A}_{\text{kSZ}^2})$ is the same as on $\mathcal{A}_{\text{kSZ}^2}$. If this is not the case, we will show that the bias can be jointly fit together with $\mathcal{A}_{\text{kSZ}^2}$, thanks to the fact that there is a lensing contribution to the measured $C_\ell^{\text{kSZ}^2 \times \delta_g}$ which is proportional to b_g , but independent of the kSZ amplitude, as explained in Sec. IV. If the galaxy bias is obtained by a joint fit, there will be some (generally small) degradation in significance that depends on the experimental configuration,² but this can also serve as a very useful consistency check, since the bias obtained must agree with that determined from external data (e.g., the galaxy autocorrelation).

The maximum S/N ratio can be estimated by using Fisher's formula,

²For an experiment with Planck resolution and noise, the degradation in S/N when jointly fitting $\mathcal{A}_{\text{kSZ}^2}$ and b_g is about 15% (see H16).

$$\left(\frac{\Delta \mathcal{A}_{\text{kSZ}^2}}{\mathcal{A}_{\text{kSZ}^2}} \right)^{-2} = f_{\text{sky}} \sum_{\ell} \frac{(2\ell + 1) (C_\ell^{\text{kSZ}^2 \times \delta_g})^2}{C_\ell^{\bar{T}\bar{T}^2, f} C_\ell^{\delta_g \delta_g} + (C_\ell^{\text{kSZ}^2 \times \delta_g})^2}, \quad (12)$$

where f_{sky} is the observed sky fraction, $C_\ell^{\delta_g \delta_g}$ is the tracer density power spectrum (including shot noise), and for $C_\ell^{\bar{T}\bar{T}^2, f}$ we use the Gaussian approximation:

$$C_\ell^{\bar{T}\bar{T}^2, f} \approx 2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} C_L^{\bar{T}\bar{T}, f} C_{|\ell - \mathbf{L}|}^{\bar{T}\bar{T}, f}. \quad (13)$$

Here $C_\ell^{\bar{T}\bar{T}, f} = F^2(\ell) b^2(\ell) (C_\ell^{TT} + C_\ell^{\text{kSZ}} + N_\ell)$ and C_ℓ^{TT} is the lensed primary CMB temperature power spectrum. The noise power spectrum N_ℓ , here assumed to be white noise,³ is given by

$$N_\ell = \Delta_T^2 b^{-2}(\ell) \approx \Delta_T^2 \exp\left(\frac{\theta_{\text{FWHM}}^2 \ell^2}{8 \ln 2}\right) \quad (14)$$

where Δ_T^2 is the pixel noise level of the experiment (usually quoted in $\mu\text{K-arcmin}$) and θ_{FWHM} is the beam full width at half maximum (FWHM).

Since $\mathcal{A}_{\text{kSZ}^2} \propto f_{\text{free}}^2$, if we are interested in a measurement of the free electron fraction f_{free} , the fractional error is given by

$$\left(\frac{\Delta f_{\text{free}}}{f_{\text{free}}} \right) \approx \frac{1}{2} \left(\frac{\Delta \mathcal{A}_{\text{kSZ}^2}}{\mathcal{A}_{\text{kSZ}^2}} \right). \quad (15)$$

Table II shows the expected results for a selection of CMB experiments and large-scale structure probes. For concreteness we have picked the WISE galaxy catalog and a catalog from the proposed SPHEREx [37] space-based experiment as our large-scale structure surveys of choice, but we note that the next decade will see a large number of galaxy surveys, both ground and space based.

The assumed redshift distributions for WISE and SPHEREx galaxies are derived from Refs. [38] and [37], respectively, and are shown in Fig. 1. For WISE we have approximately 50 million galaxies over half of the sky, while on the same footprint, the full SPHEREx galaxy catalog is predicted to have about 290 million objects.

The galaxy bias is assumed constant for WISE, while for SPHEREx we use the (redshift-dependent) bias model from [37].

The effective noise level for Advanced ACTPol [39] is determined by assuming that the component separation procedure yields a multiplicative increase over the proposed 150 GHz channel noise equal to that found for the 2015 Planck + WMAP9 LGMCA map compared to the Planck 143 GHz channel noise (a factor of $47/33 = 1.4$).

³This should be a reasonable approximation on the scales selected by the filter function.

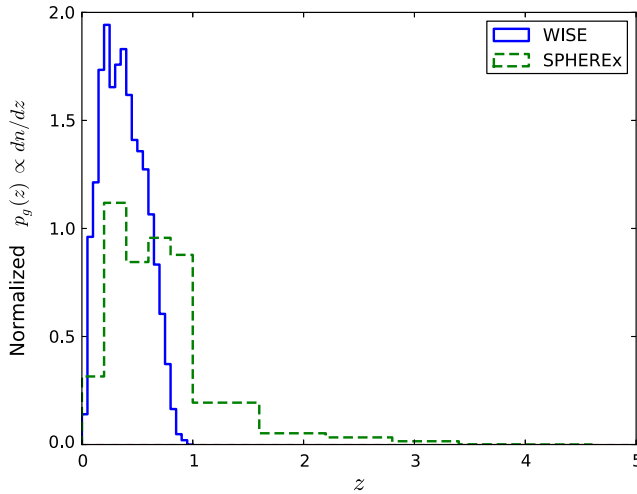


FIG. 1. Normalized redshift distributions for WISE and SPHEREx galaxies. The weight $p_g(z)$ is related to $p_g(\eta)$ in Eq. (4) by $p_g(\eta) = H(z)p_g(z)$.

The filters used in these forecasts are shown in Fig. 2, $f(\ell) = F(\ell)b(\ell)$, where $F(\ell)$ is constructed from Eq. (5) and $b(\ell)$ is the beam.

As seen in Table II, the statistical S/N for future CMB experiments is enormous, and thus the actual results are likely to be limited by systematics such as foreground component separation or theoretical modeling uncertainties. These and other challenges are discussed in Sec. VII.

For a CMB experiment with the angular resolution of Planck, this method is suboptimal (in terms of S/N per object) when 3D information is available and should only be used in the absence of reliable spectroscopic redshifts. This is easy to understand: our method uses the observed CMB temperature as a proxy for the cluster peculiar

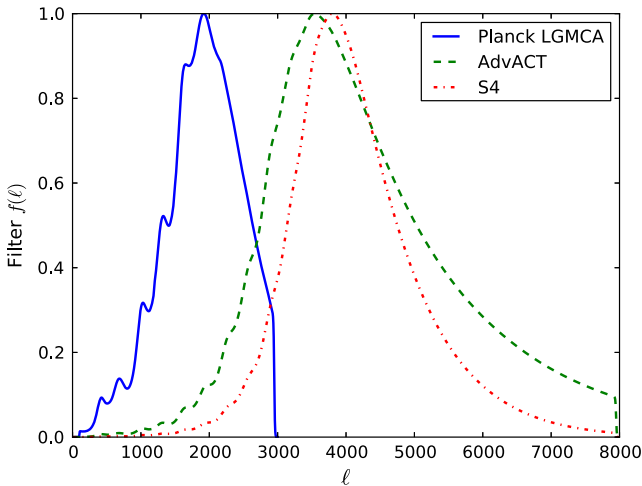


FIG. 2. Filters $f(\ell) = F(\ell)b(\ell)$ for the three CMB experiments considered here. The normalization is arbitrary and the results are independent of the normalization. The filter for a hypothetical CMB-S4 experiment is matched to the configuration of case 1 in Table I.

TABLE I. Specifications for the CMB experiments assumed in the forecasts.

| CMB experiment | Beam FWHM (arcmin) | Effective noise ^a Δ_T ($\mu\text{K}\cdot\text{arcmin}$) |
|------------------------------|--------------------|---|
| Planck (2015 LGMCA map) | 5 | 47 |
| Advanced ACTPol | 1.4 | 10 |
| CMB-S4 (case 1) ^b | 3 | 3 |
| CMB-S4 (case 2) | 1 | 3 |
| CMB-S4 (case 3) | 3 | 1 |
| CMB-S4 (case 4) | 1 | 1 |

^aHere by “effective noise” we mean the residual cleaned CMB map noise after component separation.

^bSpecifications for a future S4 experiment are not yet set; therefore here we consider a few cases for illustration purposes. Actual properties may be different.

velocity, rather than the 3D position of the tracers. On large angular scales ($\ell \lesssim 3000$), the primary anisotropy is much larger than the kSZ amplitude and the signal-to-noise per mode is very small. As high-resolution CMB experiments allow us to access smaller scales, we expect very high- S/N detections with Advanced ACTPol and CMB-S4. In fact, at $\ell \gtrsim 4000$, the fluctuation field is dominated by kSZ and not by the primary anisotropy. This point is illustrated in Fig. 3, where we show the correlation coefficient between the total temperature field that has a blackbody spectrum (i.e., lensed primary CMB and kSZ on the scales of interest) and the kSZ field. While this cross-correlation is small at low ℓ (including all of the ℓ range probed by Planck and WMAP), it grows to order unity at high ℓ . This means that in the absence of other frequency-dependent foregrounds and noise, high-resolution CMB maps are a direct probe of the integrated electron momentum.

TABLE II. Forecasts for determining f_{free} from the kSZ²-galaxy cross-correlation. The baryon profile on small scales is very uncertain and in order to minimize the theoretical uncertainties, we have fixed the filter for all S4 cases to the lower resolution case 1. In all cases the noise and the resolution are treated self-consistently.

| | f_{sky} | ℓ range | $(\frac{\Delta f_{\text{free}}}{f_{\text{free}}})^{-1}$ |
|----------------------------------|------------------|--------------|---|
| Planck \times WISE | 0.7 | 100–3000 | 5.2 |
| Planck \times SPHEREx | 0.7 | 100–3000 | 5.4 |
| Advanced ACTPol \times WISE | 0.5 | 100–8000 | 232 |
| Advanced ACTPol \times SPHEREx | 0.5 | 100–8000 | 280 |
| CMB-S4 (case 1) \times WISE | 0.5 | 100–8000 | 296 |
| CMB-S4 (case 1) \times SPHEREx | 0.5 | 100–8000 | 356 |
| CMB-S4 (case 2) \times WISE | 0.5 | 100–8000 | 704 |
| CMB-S4 (case 2) \times SPHEREx | 0.5 | 100–8000 | 866 |
| CMB-S4 (case 3) \times WISE | 0.5 | 100–8000 | 702 |
| CMB-S4 (case 3) \times SPHEREx | 0.5 | 100–8000 | 858 |
| CMB-S4 (case 4) \times WISE | 0.5 | 100–8000 | 822 |
| CMB-S4 (case 4) \times SPHEREx | 0.5 | 100–8000 | 1014 |



FIG. 3. Correlation coefficient between the kSZ “field” and the total temperature fluctuation field that has a blackbody frequency spectrum. Noise is not included here since it is experiment dependent. The kSZ field includes both late-time and reionization contributions. Note that while the exact shape and amplitude of the kSZ signal are still uncertain, the qualitative features should be correct. Here the C_{ℓ}^{TT} power includes lensing.

Also note that even at Planck resolution, this method allows us to use much larger photometric catalogs such as WISE, instead of smaller spectroscopic samples. As seen in Table II, we expect the combination of Planck and WISE to yield constraints that are comparable to recent analyses that use the full 3D (spectroscopic) information in the galaxy density field, thanks to the fact that we can use a much larger sample of tracer objects ($\sim 10^8$ in our work with WISE, compared to $\sim 10^4$ – 10^5 for previous works [13,14,16]).

IV. CMB LENSING CONTRIBUTION

Since our kSZ^2 estimator is quadratic in the CMB temperature, it can potentially receive a contribution from weak lensing of the CMB, due to matter inhomogeneities between us and the surface of last scattering (see Ref. [40] for a review on CMB lensing). In this section, we define $\Theta = \Delta T/T$ to be the unlensed (primary) CMB temperature fluctuation and $\tilde{\Theta}$ to be the corresponding lensed fluctuation.

We first note that if we could observe the CMB with an infinitesimally small beam and did not apply any filter, then the lensing contribution to our estimator would vanish. This is because CMB lensing preserves the total variance, since

the lensing amounts to a remapping of perturbations on the last scattering surface to a slightly different point in the sky [40].

This argument no longer applies when we observe the CMB through a finite resolution experiment and the map is filtered as described above; in this case the weak lensing contribution can be large. As before, we define the lensed $\tilde{\Theta}_f(\ell) = f(\ell)\tilde{\Theta}(\ell)$, where $f(\ell) = F(\ell)b(\ell)$ is the product of a filter F and the beam function b . We would like to compute the Fourier transform of $\langle \tilde{\Theta}_f^2(\mathbf{x})\delta_g(\mathbf{y}) \rangle$ and we shall assume the flat sky approximation:

$$\begin{aligned} \langle \tilde{\Theta}_f^2(\ell_1)\delta_g(\ell_2) \rangle &= \int \frac{d^2\mathbf{L}}{(2\pi)^2} \langle \tilde{\Theta}_f(\mathbf{L})\tilde{\Theta}_f(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle \\ &= \int \frac{d^2\mathbf{L}}{(2\pi)^2} f(L)f(|\ell_1 - \mathbf{L}|) \langle \tilde{\Theta}(\mathbf{L})\tilde{\Theta}(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle. \end{aligned} \quad (16)$$

The lensed fluctuation field can be expanded in terms of the unlensed field [40]:

$$\tilde{\Theta}(\mathbf{x}) = \Theta(\mathbf{x}) + \nabla\psi \cdot \nabla\Theta(\mathbf{x}) + \dots \quad (17)$$

where ψ is the lensing potential, so that we can express

$$[\nabla\psi \cdot \nabla\Theta](\mathbf{L}) = - \int \frac{d^2\mathbf{L}'}{(2\pi)^2} \mathbf{L}' \cdot (\mathbf{L} - \mathbf{L}')\psi(\mathbf{L}')\Theta(\mathbf{L} - \mathbf{L}'). \quad (18)$$

Up to first order in the lensing potential we have

$$\begin{aligned} \langle \tilde{\Theta}(\mathbf{L})\tilde{\Theta}(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle &= \langle \Theta(\mathbf{L})\Theta(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle \\ &+ \langle [\nabla\psi \cdot \nabla\Theta](\mathbf{L})\Theta(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle \\ &+ (\mathbf{L} \rightarrow \ell_1 - \mathbf{L}) + \dots \end{aligned} \quad (19)$$

The first term is simply the fiducial $C_{\ell}^{\text{kSZ}^2 \times \delta_g}$ for the kSZ²-galaxy cross-correlation that was computed in Sec. II, while the second and third terms are the lowest order CMB lensing contribution and are equal in magnitude by symmetry. Plugging Eq. (19) into (16) we find

$$\begin{aligned} \langle \tilde{\Theta}_f^2(\ell_1)\delta_g(\ell_2) \rangle &= \langle \Theta_f^2(\ell_1)\delta_g(\ell_2) \rangle \\ &- 2 \int \frac{d^2\mathbf{L}}{(2\pi)^2} f(L)f(|\ell_1 - \mathbf{L}|) \int \frac{d^2\mathbf{L}'}{(2\pi)^2} \mathbf{L}' \cdot (\mathbf{L} - \mathbf{L}') \langle \psi(\mathbf{L}')\Theta(\mathbf{L} - \mathbf{L}')\Theta(\ell_1 - \mathbf{L})\delta_g(\ell_2) \rangle + \dots \end{aligned} \quad (20)$$

The four-point function of the form $\langle \psi \Theta \Theta \delta_g \rangle$ on the right-hand side of Eq. (20) can be decomposed into a connected four-point function (technically nonvanishing because of ISW, but subdominant to the other terms in the range of scales considered here), and two nonzero contractions $\langle \psi \delta_g \rangle \langle \Theta \Theta \rangle$ and $\langle \psi \Theta \rangle \langle \Theta \delta_g \rangle$, the latter again nonzero due to ISW. Consider the first one and write

$$\begin{aligned} \langle \psi(\mathbf{L}') \delta_g(\ell_2) \rangle &= (2\pi)^2 C_{\ell_2}^{\psi \delta_g} \delta_D(\mathbf{L}' + \ell_2) \\ \langle \Theta(\mathbf{L} - \mathbf{L}') \Theta(\ell_1 - \mathbf{L}') \rangle &= (2\pi)^2 C_{|\ell_1 - \mathbf{L}'|}^{TT} \delta_D(\ell_1 - \mathbf{L}'). \end{aligned}$$

Then the main correction due to lensing⁴ is [from the right-hand side of Eq. (20)]

$$-2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} f(L) f(|\mathbf{L} - \ell_1|) \ell_1 \cdot (\mathbf{L} - \ell_1) C_{\ell_1}^{\psi \delta_g} C_{|\mathbf{L} - \ell_1|}^{TT}. \quad (21)$$

Similarly, the other contraction gives rise to

$$-2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} f(L) f(|\mathbf{L} - \ell_1|) \ell_1 \cdot (\mathbf{L} - \ell_1) C_{\ell_1}^{\psi T} C_{|\mathbf{L} - \ell_1|}^{T \delta_g} \quad (22)$$

which is due to ISW and numerically is found to be factor of $\sim 10^4$ – 10^5 smaller than the former contribution on the scales considered here. Thus it will be neglected in the following.

Changing variables in Eq. (21) to $\mathbf{L}' = \mathbf{L} - \ell_1$, we can rewrite the leading-order lensing contribution as

$$\begin{aligned} \Delta C_{\ell}^{T^2 \times \delta_g} &\approx -2 \frac{\ell C_{\ell}^{\psi \delta_g}}{(2\pi)^2} \int_0^{\infty} dL' L'^2 f(L') C_{L'}^{TT} \\ &\quad \times \int_0^{2\pi} d\phi f(|\mathbf{L}' + \ell|) \cos \phi. \end{aligned} \quad (23)$$

Finally, we see that in the absence of a filter and beam [i.e., $f(\ell) = \text{constant}$], the lensing correction vanishes as expected. Examples of the lensing contribution are shown in Figs. 6 and 7. It displays a characteristic oscillatory behavior that makes it nearly orthogonal to the $k\text{SZ}^2$ signal.

Heuristically, the shape of the lensing contribution is determined by the interplay of multiple effects with opposite sign, such as power transfer from large to small scales, magnification of lensed regions, and the use of a bandpass filter in harmonic space. The exact magnitude and shape of the lensing contribution depends on the CMB experiment, ℓ -space filter, and large-scale structure survey used in the analysis.

⁴Here C_{ℓ}^{TT} denotes the *unlensed* primary anisotropy power spectrum.

V. COMPARISON TO NUMERICAL SIMULATIONS

In this section we compare our theoretical predictions for the $k\text{SZ}^2$ signal and lensing contribution to two different sets of numerical simulations. The first is a cosmological hydrodynamics simulation [33], while the second is constructed from a dark-matter-only tree-particle-mesh simulation, in which halos are populated with gas in postprocessing using a polytropic equation of state and hydrostatic equilibrium [41]. In this section only, the cosmological parameters for the theory curves are chosen to match the respective simulations and will in general differ from the fiducial cosmology assumed in the rest of the paper.

As a first test, we set the filter $f(\ell)$ to a constant and compare our theoretical prediction to the simulations from [33]. These are hydrodynamic simulations of cosmological volumes (box side-length $L = 165 \text{ Mpc}/h$) using a modified version of the GADGET-2 code [42]. Included in these simulations are subgrid physics models for active galactic nuclei (AGN) feedback [33], cosmic ray physics [43–45], radiative cooling, star formation, galactic winds, and supernova feedback [46]. The halo catalogs from these simulations are incomplete below masses of $\approx 5 \times 10^{13} M_{\odot}$ [47], and thus we cannot construct simulated galaxy density maps to mock the WISE or SPHEREx samples. Instead, we consider weak gravitational lensing convergence (κ_{CFHT}) as the large-scale structure tracer of choice in this analysis. For the present comparison, we construct mock lensing convergence maps using mass shells extracted from the simulations and a source galaxy redshift distribution matching that of the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) [48].

The $k\text{SZ}$ and the CFHTLenS-like lensing convergence maps are made at each redshift snapshot following the methods described in [47] and [49], respectively. We compute the cross-power spectrum for each redshift output and then average the cross-power spectra over ten initial condition realizations. We sum these average spectra over the redshift outputs to compute the final spectrum. The results of this comparison are shown in Fig. 4, with error bars computed from the scatter amongst the ten realizations.

One important caveat when comparing theory to simulations is that the velocity field is coherent on very large scales and thus finite-box simulations can underpredict the expected signal, since they lack contributions from velocity modes with wavelength larger than the box size [50]. To be more quantitative, from Eq. (10), the signal is proportional to v_{rms}^2 , and we find that about half of the contribution to v_{rms}^2 comes from $k < 0.06 \text{ h}/\text{Mpc}$. As seen in Fig. 4, the agreement between theory and simulations is excellent when using the same k_{min} and k_{max} as the simulations, but there is a large discrepancy if we neglect the effect of the finite box size.

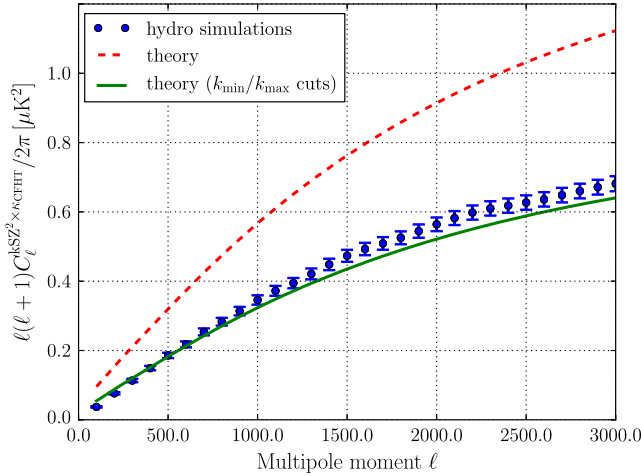


FIG. 4. Theoretical cross-correlation between the unfiltered kSZ^2 signal and CFHTLenS-like weak lensing convergence maps. The blue points with error bars show the result measured from cosmological hydrodynamics simulations [33]. The dashed red curve shows the fiducial theory computation, while the solid green curve shows the theory computation with wave number cutoffs matching those of the simulation ($k_{\text{min}} = 0.038 h/\text{Mpc}$ and $k_{\text{max}} = 76 h/\text{Mpc}$). This comparison shows that the simulation results are biased low due to the lack of superbox long-wavelength velocity modes. This effect can be large, as seen here, but if properly accounted for, the theory and simulations agree to $\lesssim 5\%$ over the whole range considered. Here we have used $f_{\text{free}} = 0.85$, independent of redshift. The small difference between theory and simulations might be explained by the redshift evolution of f_{free} or by the intrinsic uncertainty on our theory curve due to the fitting function for the nonlinear bispectrum, which is of order 5%–10% [35].

Next we compare our predictions to the full-sky simulation of Ref. [41], with a nontrivial filter $f(\ell)$ that includes the weighting and beam appropriate for the Planck experiment (in particular, as constructed from the 2013 LGMCA map [51]). Note that the simulation box is 1 Gpc/ h on a side, so effects related to the low- k cutoff discussed above are substantially reduced here. For this analysis, we consider CMB lensing convergence (κ_{CMB}) as our large-scale structure tracer, since ray-traced maps of this quantity have already been computed from this simulation. The results are shown in Fig. 5. In this case, since only one simulation is available, we estimate error bars from the scatter within each multipole bin.

We find agreement between theory and simulation to better than 10% at $\ell \gtrsim 500$. There is a minor discrepancy at very low ℓ , but this might be explained by the filtering applied to the simulations: because of the way that the light cone was constructed, the kSZ signal from the intergalactic medium was overpredicted on large scales and therefore a filter of the form $w(\ell) = 1 - e^{-(\ell/500)^2}$ was applied to the kSZ map to suppress the large-scale excess. The authors of [41] caution that “since the simple filtering modifies the signal at $\ell < 1000$, the maps should not be used to predict

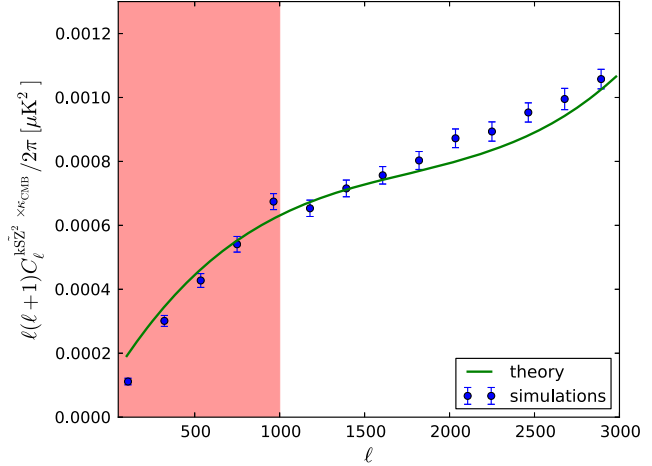


FIG. 5. Cross-correlation between the filtered kSZ^2 signal and CMB lensing convergence maps. The blue points with error bars show the result measured from the simulation of Ref. [41], using a filter $f(\ell)$ appropriate for Planck data. The agreement is better than 10% at high ℓ and the difference at very low ℓ is likely due to the fact that these simulations do not accurately predict the kSZ power spectrum on large scales (see Sec. 2.4 of [41]). The shaded region at $\ell < 1000$ represents the scales that may be unreliable in the simulation.

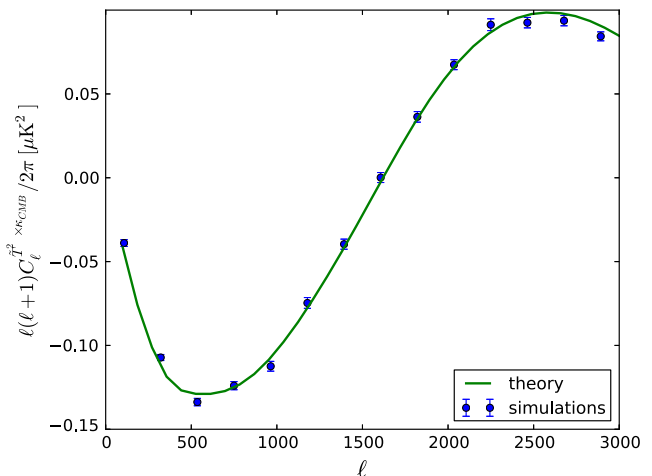


FIG. 6. Lensing leakage in the kSZ^2 estimator, inferred from simulated lensed CMB temperature maps cross-correlated with CMB lensing convergence maps, i.e., $\langle \tilde{T}_f^2 \kappa_{\text{CMB}} \rangle$. (To be clear, κ_{CMB} is simply the large-scale structure tracer considered in this test; the lensing leakage calculation can be applied to any tracer.) The blue points show the cross-correlation measured from the simulation of Ref. [41], while the solid line is the analytic calculation presented in this paper [see Eq. (23)], and is obtained using fitting functions for the nonlinear matter power spectrum of [52]. The agreement is better than 3% over the entire multipole range probed here, with the exception of the highest ℓ point (where the nonlinear corrections are largest and which deviates from the theory curve by $\approx 6\%$).

the kSZ signal at these scales” and thus the slight low- ℓ discrepancy is not a significant cause for concern.

Finally, we test our lensing leakage prediction from Eq. (23), using κ_{CMB} rather than δ_g as the large-scale structure tracer of choice. For this comparison, we calculate the cross-correlation between the square of the lensed, filtered CMB temperature map (with no other secondary anisotropy) and the CMB weak lensing convergence map. The result is shown in Fig. 6, indicating an agreement to better than 6% on all scales. Therefore we conclude that higher order corrections are subleading and can be neglected at the current level of precision.

VI. EXAMPLE: MEASUREMENT USING WMAP, PLANCK, AND WISE

In H16, we recently presented the first measurement of the kSZ signal using this method. Here we briefly summarize the analysis as an example of an application to real data. Some specific technical details are found in H16. We also discuss several of the challenges of this measurement in Sec. VII.

We use a cleaned CMB temperature map constructed from a joint analysis of the nine-year WMAP [31] and Planck full-mission [2] full-sky temperature maps [30].⁵

The CMB is separated from other components in the microwave sky using the LGMCA technique, relying on the sparse distribution of non-CMB foregrounds in the wavelet domain. We refer the reader to [30,53] for a thorough description of this component separation technique and characterization of the resulting maps. The method reconstructs a full-sky CMB map with minimal dust contamination and essentially zero contamination from the thermal SZ (tSZ) effect, which is explicitly projected out in the map construction (unlike in, e.g., the official Planck SEVEM, NILC, or SMICA component-separated CMB maps, which all possess significant tSZ residuals). Since the kSZ signal preserves the CMB black-body spectrum, it is not removed by the component separation algorithm. We further clean the LGMCA map to explicitly deproject any residual emission associated with the WISE galaxies (e.g., from dust)—see H16 for details.

As discussed in Sec. II, a filter is applied to the CMB map before squaring in real space to downweight scales that are dominated by the primary CMB or noise. The filter used in H16 is shown in Fig. 2 (including multiplication by the FWHM = 5 arcmin beam of the LGMCA map).

The WISE [26] source catalog contains more than 500 million objects, roughly 70% of which are star-forming galaxies [38]. Color cuts can be used to separate galaxies from stars and other objects. We use the same selection criteria as Ref. [54] to select a sample of galaxies, originally

based on previous work [55], and we refer the reader to these papers for a detailed explanation.

The redshift distribution of WISE-selected galaxies has been shown to be fairly broad, with a peak at $z \approx 0.3$ and extending to $z = 1$ [38]. Here we note that the galaxy selection is imperfect and that there is some residual stellar contamination, especially close to the Galactic plane. However, Galactic stars are expected to be uncorrelated with the kSZ signal, and any contamination will only lead to larger noise (which is taken into account in our analysis), but not a bias. For this reason, we apply a mask that leaves $f_{\text{sky}} = 0.447$ and 46.2 million galaxies.

Our model for the expected total signal is the sum of the theoretical kSZ² and lensing templates, the amplitude of each being $\mathcal{A}_{\text{kSZ}^2} b_g$ and b_g , respectively (where we have defined $\mathcal{A}_{\text{kSZ}^2} \propto f_{\text{free}}^2$ as the amplitude of the kSZ² signal, with a fiducial expectation of unity):

$$\mathbf{t}(\mathcal{A}_{\text{kSZ}^2}, b_g) = \mathcal{A}_{\text{kSZ}^2} b_g \mathbf{t}_{\text{kSZ}^2} + b_g \mathbf{t}_{\text{lens}}. \quad (24)$$

The best-fit amplitude is found by minimizing the function

$$\chi^2(\mathcal{A}_{\text{kSZ}^2}, b_g) = (\mathbf{d} - \mathbf{t}(\mathcal{A}_{\text{kSZ}^2}, b_g))^T C^{-1} (\mathbf{d} - \mathbf{t}(\mathcal{A}_{\text{kSZ}^2}, b_g)). \quad (25)$$

\mathbf{d} is the data vector (from the measured cross-correlation) and C^{-1} is the inverse of the noise covariance matrix estimated from the data itself, sourced by primary CMB fluctuations and other sources of noise. For the best fit we find $\chi^2_{\text{b.f.}}/\text{dof} = 13.1/11$, indicating a good fit.

Figure 7 shows the total best fit to the data, as well as the individual contributions from the kSZ² and lensing templates (matching Fig. 1 of H16). In our fiducial analysis, we

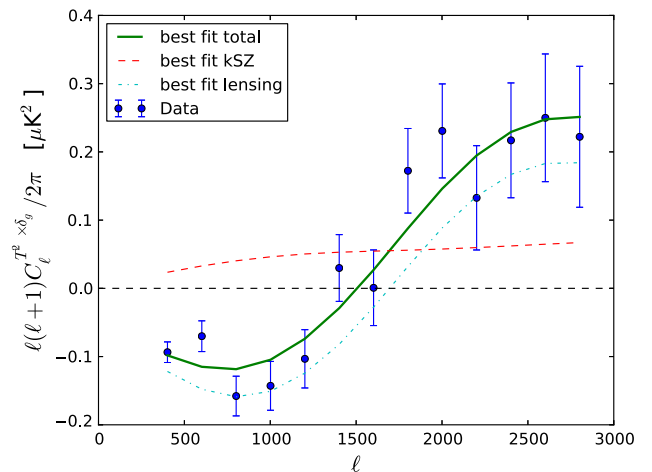


FIG. 7. Results for the $C_{\ell}^{\text{kSZ}^2 \times \delta_g}$ analysis of H16, shown in blue. The dashed red curve is the best-fit kSZ² template, the dashed-dotted cyan curve is the best-fit lensing template, and the solid green curve is the sum of the two. No external prior on the galaxy bias is used in the fit shown in this plot.

⁵http://www.cosmostat.org/research/cmb/planck_wpr2.

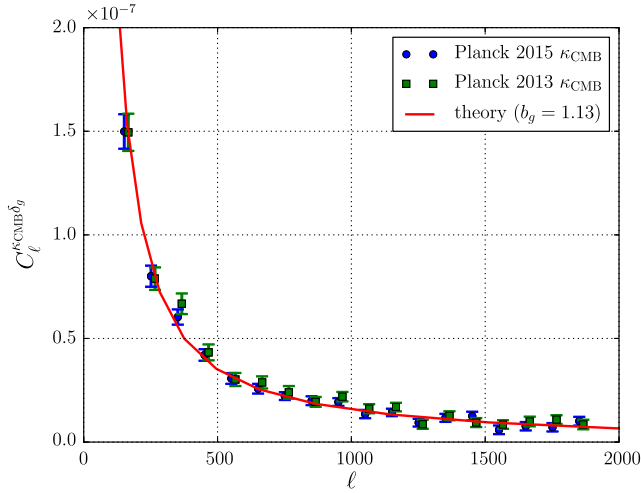


FIG. 8. Cross-correlation between Planck lensing convergence maps and WISE galaxies, shown for both the 2013 and 2015 version of the lensing maps. The best-fit galaxy bias from the 2015 map is $b_g = 1.13 \pm 0.02$.

marginalize over the lensing contribution, but as a check we also obtain the galaxy bias by cross-correlating the WISE sample with Planck CMB lensing maps [56,57]. This cross-correlation is shown in Fig. 8.

The posteriors for $\mathcal{A}_{\text{kSZ}^2}$ with and without the prior on b_g from the external CMB lensing data are shown in Fig. 9.

The best-fit kSZ^2 amplitude and galaxy bias are presented in Table III. The results indicate that marginalization over the lensing contribution leads to a degradation of $\approx 15\%$ in the error bar on $\mathcal{A}_{\text{kSZ}^2}$. The corresponding posterior for f_{free} for our fiducial case (where both the kSZ^2 amplitude and galaxy bias are obtained without using external CMB lensing data) is shown in Fig. 10. Since $f_{\text{free}} \sim \sqrt{\mathcal{A}_{\text{kSZ}^2}}$, the posterior is fairly non-Gaussian and

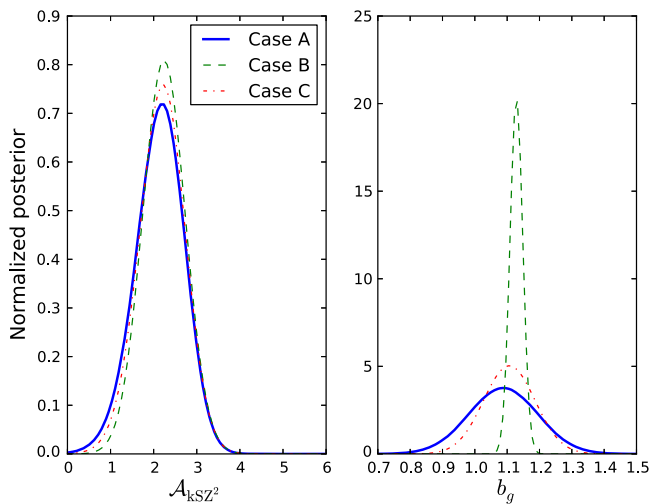


FIG. 9. Posterior distribution for $\mathcal{A}_{\text{kSZ}^2}$ and b_g for the three analysis cases given in Table III.

TABLE III. Best-fit parameters for the kSZ^2 -WISE galaxies cross-correlation from H16. We include three analysis scenarios:

(A) using only the $C_\ell^{T_{\text{clean}}^2 \times \delta_g}$ data and marginalizing over the lensing contribution amplitude (i.e., the galaxy bias); (B) including an external prior on the WISE galaxy bias from our measurement of $C_\ell^{\text{kCMB} \delta_g}$; (C) same as (B), but including an additional 10% theoretical systematic error on the b_g constraint from $C_\ell^{\text{kCMB} \delta_g}$, due to uncertainties from nonlinear evolution and baryonic physics.

| Case | $\mathcal{A}_{\text{kSZ}^2}$ | b_g |
|--|------------------------------|-----------------|
| (A): $C_\ell^{T_{\text{clean}}^2 \times \delta_g}$ only | 2.18 ± 0.57 | 1.10 ± 0.11 |
| (B): $C_\ell^{T_{\text{clean}}^2 \times \delta_g}$ and $C_\ell^{\text{kCMB} \delta_g}$ | 2.24 ± 0.50 | 1.13 ± 0.02 |
| (C): +10% error on $C_\ell^{\text{kCMB} \delta_g}$ | 2.21 ± 0.53 | 1.11 ± 0.08 |

shows a considerable negative skewness. For this reason, our best-fit measurement $f_{\text{free}} = 1.48$ is only in mild tension with the fiducial value of $f_{\text{free}} = 1$, and from the posterior we estimate that the probability of $f_{\text{free}} \leq 1$ is 5.4%, so that if the posterior were Gaussian, this would correspond to a 1.6σ upward fluctuation.

VII. CHALLENGES

A. Foregrounds

As in most cross-correlation analyses, there are a number of possible contaminants that have to be carefully scrutinized. In particular, any emission or imprint of the tracer

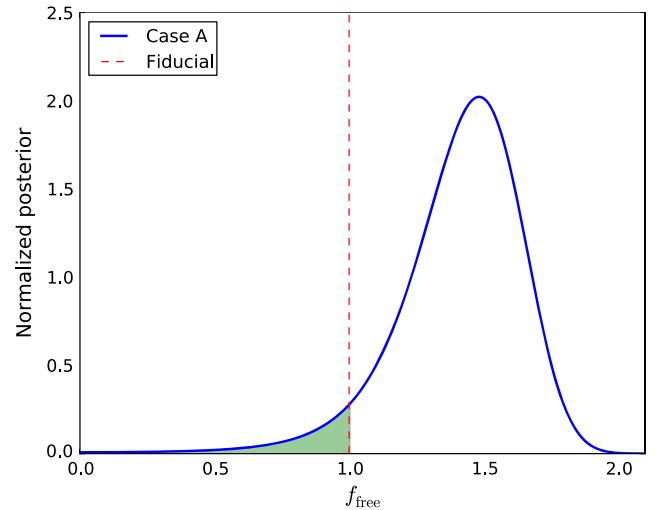


FIG. 10. Posterior probability for f_{free} , obtained by replacing $\mathcal{A}_{\text{kSZ}^2}$ with f_{free}^2 in Eq. (24) and marginalizing over b_g with no external prior (i.e., Case A from Table III). The maximum-posterior value is $f_{\text{free}} = 1.48$, slightly larger than the fiducial $f_{\text{free}} = 1$, but we find that the probability of $f_{\text{free}} \leq 1$ (the area shaded in green) is 5.4%. If the posterior were Gaussian, this would correspond to a 1.6σ upward fluctuation.

galaxies⁶ that leaks into the CMB maps will contribute to $C_\ell^{\text{kSZ}^2 \times \delta_g}$ and could be mistaken for the kSZ^2 signal. Since the kSZ signal arises from a Doppler shift in photon energy, it preserves the blackbody spectrum of the CMB, simply producing a small shift in the effective temperature. On the contrary, most other foregrounds give rise to emission that differs considerably from a blackbody at T_{CMB} and can therefore (at least in principle) be separated using multifrequency analysis (see for example [58]). In particular, while most of the kSZ estimators that require spectroscopic redshifts can be applied to single-frequency CMB temperature maps, our method explicitly requires multifrequency analysis for foreground subtraction.

To ensure that the foreground separation is effective a number of null tests can be performed; here we comment on some, but this is far from an exhaustive list. We have noted before that because of the $\mathbf{v}_e \rightarrow -\mathbf{v}_e$ symmetry, the kSZ contribution to the cross-correlation between tracers and the CMB temperature vanishes, i.e., $\langle \Theta_f^{\text{kSZ}} \delta_g \rangle = 0$. Checking that this quantity is consistent with zero is therefore a powerful test for the absence of contamination by foregrounds, such as Galactic or extra-Galactic dust, tSZ, or radio emission.⁷

Another useful test for dust or radio contamination is to replace one power of the cleaned temperature map in the standard analysis with a tracer of foregrounds (for example the 545 GHz Planck map is an excellent tracer of dust emission, while the 30 GHz map traces radio emission). Schematically, we can look at $\langle \Theta_f^{\text{clean}} \Theta_f^{\text{foreground}} \delta_g \rangle$, where any potential contamination will be greatly enhanced over its contribution to $\langle (\Theta_f^{\text{clean}})^2 \delta_g \rangle$.

The two null tests just described ensure that foregrounds are subtracted correctly *on average*. Spatially varying source properties (such as fluctuations in the spectral index) can lead to a situation where a subset of the sources have been oversubtracted and the others have been under-subtracted; the previous null tests, being linear in Θ^{clean} , are not guaranteed to be sensitive to such a contamination,⁸ but our estimator is, since it is quadratic in Θ^{clean} . One way to test for the latter scenario is to generate mock catalogs in which galaxies are associated with spatially varying emission (with the relevant parameters drawn from a random distribution with scatter matching known source properties). These mock catalogs can then be subjected to the foreground separation pipeline and used in place of Θ^{clean} in the kSZ^2 cross-correlation to estimate the expected

amplitude of the effect. All of these null tests were performed in H16.

Note that our method only requires the removal of foregrounds that are correlated with the large-scale structure tracers under consideration. For example, it is well known that at high ℓ , the cosmic infrared background (CIB) is a major contributor to the measured CMB power spectrum, but the bulk emission of the CIB originates from unresolved galaxies at $z \sim 1-3$ [59,60].

We have shown in H16 that component separation can be used to detect the kSZ^2 signal with $S/N \approx 4$ on angular scales up to $\ell \approx 3000$. We have estimated that the residual contamination is a small fraction of the current statistical uncertainty. It is not yet known how well multifrequency cleaning techniques will perform at higher ℓ and with lower noise levels. This could potentially be the limiting factor in the future performance of this method, and will be the subject of future analysis.

B. Gravitational secondary anisotropies

There are other secondary CMB anisotropies that preserve the blackbody spectrum of the CMB and therefore cannot be removed by multifrequency component separation: the contribution from weak lensing and the ISW effect [61], as well as its nonlinear generalization known as the Rees-Sciama effect [62]. As we have noted in Sec. IV, the weak lensing contribution can be large and must be accounted for, but its characteristic ℓ dependence and the possibility of using external priors allow it to be cleanly separated from the kSZ^2 signal.

Regarding ISW, we should distinguish between the linear and nonlinear contributions. The linear part is due to the decay of the gravitational potential on large scales because of the late-time cosmic acceleration. This is a very large-scale effect and detectable at $\ell < 100$ (for a measurement of ISW with WISE galaxies, see [54,63]). For this reason any analysis of the kSZ^2 signal should explicitly filter out scales with ℓ less than a few hundred. The nonlinear contribution is expected to be subdominant to kSZ on all scales with $\ell >$ few hundred. Perturbation theory and halo model calculations indicate that it is at least 2 orders of magnitude smaller than kSZ on the scales of interest [64–66]. If nonperturbative effects are large or the kSZ S/N is large enough (e.g., $\gtrsim 100$), then this contribution will need to be modeled and accounted for.

C. Theoretical uncertainties

Finally, we note that the approximations presented here, while more than adequate for the analysis in H16, may need to be improved for the high S/N regime. In particular, we have used fitting functions for the nonlinear matter power spectrum and bispectrum, which have a calibration uncertainty $\approx 5\%-10\%$ [35,52], consistent with the level of agreement found when comparing to simulations in Sec. V.

⁶Or emission from other objects that are correlated with the tracer population.

⁷There is a small contribution to this correlation due to the ISW effect, which is detectable on large scales and is discussed later.

⁸They can be sensitive to it if the specific intensity of emission of the sources is correlated with the spectral index or other properties.

In addition, the $C_{\ell}^{\text{kSZ}^2 \times \delta_g}$ signal depends steeply on the cosmological parameters, for example scaling as σ_8^{6-7} [27]. Moreover, for the purpose of this work we have assumed that the baryons follow the dark matter exactly on the scales of interest. This should be a good approximation on the scales probed by H16, but it is known not to be the case on small scales. However, the baryon profile in the outskirts of galaxies and clusters is still very uncertain. In fact, the small-scale shape of $C_{\ell}^{\text{kSZ}^2 \times \delta_g}$ can be used as a probe of the free electron profile, which is sensitive to the effects of feedback and energy injection into the intracluster and intergalactic media (for a measurement of the baryon profile with kSZ and comparison to dark matter, see [16]).

For this analysis, we have used a scale- and redshift-independent galaxy bias, and moreover we have assumed that the shape of the lensing contribution is known exactly, up to a multiplicative constant (that is, the galaxy bias). While marginalizing over the galaxy bias can mitigate some of the theoretical uncertainties on the amplitude of the lensing term, scale-dependent bias or baryonic effects can introduce systematic effects in the high S/N regime, which may require appropriate treatment in the future.

D. Photometric redshift uncertainties

Errors in the assumed redshift distribution dn/dz of the tracer sample can also bias the measured kSZ^2 amplitude. To quantify the effect on the analysis in the previous section, we shift the assumed peak of the redshift distribution of WISE sources up and down in redshift by 20%, while preserving the shape. When repeating the analysis, we find that the best-fit amplitude changes by 5.5% when shifting up and 3.7% when shifting down, with the main difference coming from the shape of the lensing template. While the redshift distribution of WISE sources can be subject to considerable errors, future photometric surveys are expected to have good control over redshift uncertainty [$\sigma(z)/(1+z) \lesssim 0.01$]. If a large number of outliers is present, that would need to be modeled in the analysis.

E. Future directions

As discussed in the previous section, unmodeled scale-dependent effects in the lensing contribution can potentially mimic the kSZ^2 signal and bias the results in the high S/N regime. It is possible to write down estimators that use temperature and/or polarization that are insensitive to the lensing signal, regardless of its amplitude and shape. In particular, CMB polarization is lensed by the same gravitational potential as the CMB temperature, while it receives a negligible contribution from the kSZ effect. Therefore an appropriate combination of temperature and polarization can cancel the lensing signal, while preserving the correct kSZ^2 amplitude.

Another improvement that can be implemented in future analyses is optimal redshift weighting of the projected

tracer field in Eq. (3), which has not been considered in this work. Reference [27] shows that the peak differential contribution to the kSZ^2 signal comes from $z \sim 0.5$, and the WISE galaxy distribution is fairly well matched to the signal redshift distribution. Optimal weighting should especially benefit surveys for which the source distribution is peaked at higher redshift or with very extended tails. For example, the SPHEREx experiment might benefit from downweighting the high-redshift population tail and it may be possible to obtain higher statistical significance than that predicted in Table II.

These points will be explored in future work.

VIII. CONCLUSIONS

We have revisited a kSZ estimator based on projected fields, which does not require expensive spectroscopic data. This will allow the use of large, full-sky imaging catalogs for kSZ measurements, yielding accurate determinations of the low-redshift baryon abundance and the free electron distribution associated with galaxies and clusters. In a companion paper (H16), we have shown that this method is already competitive with other kSZ approaches when applied to current data, allowing a detection of the kSZ signal with $S/N \approx 4$ by combining Planck and WMAP microwave temperature maps with a WISE galaxy catalog. If foreground cleaning methods in future experiments are effective at separating the CMB blackbody component from other microwave sky signals, we forecast kSZ measurements with $S/N > 100$ for Advanced ACTPol and CMB-S4. This will allow precision measurements of both the abundance and profile of the baryons associated with the tracer sample. Since both of these properties are expected to vary with mass and redshift, the tracer population can be split into multiple samples that can be compared to high precision. In addition, other properties such as color, star formation rate, or AGN activity are expected to influence the gas distribution, and comparing the kSZ^2 signal from multiple different tracer populations will shed light on galaxy evolution and feedback processes. When combined with tSZ measurements of the same objects, the gas temperature, density, and pressure of the intergalactic medium can be simultaneously inferred, providing information about the amount of energy injection.

It is also important to point out that while for concreteness we have shown forecasts for “galaxy overdensity” as our tracer, any tracer of the late-time density can be used in this approach. In particular we expect interesting measurements when using galaxy lensing as a tracer (our measurement will then probe the matter-gas correlation), or 21 cm observations (to probe the ionized-neutral gas correlation).

Finally, these measurements will soon complement kSZ measurements obtained from the small-scale CMB power spectrum [67,68], and will be useful to disentangle the

contributions due to late-time structure from those produced during “patchy” cosmic reionization.

ACKNOWLEDGMENTS

We are grateful to Olivier Doré, Zoltan Haiman, Emmanuel Schaan, Blake Sherwin, Kendrick Smith and the anonymous referees for very useful conversations and comments. We also thank the local-generalized morphological component analysis (LGMCA) team for publicly releasing their CMB maps. S. F., J. C. H. and D. N. S. acknowledge support from NASA Theory Grant No. NNX12AG72G and NSF AST-1311756. S. F. thanks

the Miller Institute for Basic Research in Science at the University of California, Berkeley, for support. This work was partially supported by a Junior Fellow award from the Simons Foundation to J. C. H. N. B. acknowledges support from the Lyman Spitzer Fellowship. J. L. is supported by NSF Grant No. AST-1210877. Some of the results in this paper have been derived using the HEALPix package [69]. This publication makes use of data products from the Wide-Field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

-
- [1] G. Hinshaw, D. Larson, E. Komatsu *et al.*, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
- [2] R. Adam and P. A. R. Ade *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A1 (2016).
- [3] G. Steigman, *Annu. Rev. Nucl. Part. Sci.* **57**, 463 (2007).
- [4] M. Fukugita, C. J. Hogan, and P. J. E. Peebles, *Astrophys. J.* **503**, 518 (1998).
- [5] J. N. Bregman, *Annu. Rev. Astron. Astrophys.* **45**, 221 (2007).
- [6] R. Cen and J. P. Ostriker, *Astrophys. J.* **650**, 560 (2006).
- [7] J. K. Werk, J. X. Prochaska, J. Tumlinson *et al.*, *Astrophys. J.* **792**, 8 (2014).
- [8] M. Bonamente, J. Nevalainen, E. Tilton *et al.*, *Mon. Not. R. Astron. Soc.* **457**, 4236 (2016).
- [9] R. A. Sunyaev and Y. B. Zel’dovich, *Comments Astrophys. Space Phys.* **4**, 173 (1972).
- [10] R. A. Sunyaev and I. B. Zeldovich, *Annu. Rev. Astron. Astrophys.* **18**, 537 (1980).
- [11] J. P. Ostriker and E. T. Vishniac, *Astrophys. J.* **306**, L51 (1986).
- [12] S. Bhattacharya and A. Kosowsky, *Phys. Rev. D* **77**, 083004 (2008).
- [13] N. Hand, G. E. Addison, E. Aubourg *et al.*, *Phys. Rev. Lett.* **109**, 041101 (2012).
- [14] P. A. R. Ade, N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **586**, A140 (2016).
- [15] C. Hernández-Monteagudo, Y.-Z. Ma, F. S. Kitaura, W. Wang, R. Génova-Santos, J. Macías-Pérez, and D. Herranz, *Phys. Rev. Lett.* **115**, 191301 (2015).
- [16] E. Schaan, S. Ferraro, M. Vargas-Magaña *et al.*, *Phys. Rev. D* **93**, 082002 (2016).
- [17] B. Soergel, S. Flender, K. T. Story *et al.*, *Mon. Not. R. Astron. Soc.* **461**, 3172 (2016).
- [18] T. Mroczkowski, S. Dicker, J. Sayers *et al.*, *Astrophys. J.* **761**, 47 (2012).
- [19] J. Sayers, T. Mroczkowski, M. Zemcov *et al.*, *Astrophys. J.* **778**, 52 (2013).
- [20] S. Ho, S. Dedeo, and D. Spergel, [arXiv:0903.2845](https://arxiv.org/abs/0903.2845).
- [21] J. Shao, P. Zhang, W. Lin, Y. Jing, and J. Pan, *Mon. Not. R. Astron. Soc.* **413**, 628 (2011).
- [22] M. Li, R. E. Angulo, S. D. M. White, and J. Jasche, *Mon. Not. R. Astron. Soc.* **443**, 2311 (2014).
- [23] P. G. Ferreira, R. Juszkiewicz, H. A. Feldman, M. Davis, and A. H. Jaffe, *Astrophys. J.* **515**, L1 (1999).
- [24] R. Keisler and F. Schmidt, *Astrophys. J.* **765**, L32 (2013).
- [25] S. Flender, L. Bleem, H. Finkel, S. Habib, K. Heitmann, and G. Holder, *Astrophys. J.* **823**, 98 (2016).
- [26] E. L. Wright, P. R. M. Eisenhardt, A. K. Mainzer *et al.*, *Astron. J.* **140**, 1868 (2010).
- [27] O. Doré, J. F. Hennawi, and D. N. Spergel, *Astrophys. J.* **606**, 46 (2004).
- [28] S. DeDeo, D. N. Spergel, and H. Trac, [arXiv:astro-ph/0511060](https://arxiv.org/abs/astro-ph/0511060).
- [29] J. C. Hill, S. Ferraro, N. Battaglia, J. Liu, and D. N. Spergel, *Phys. Rev. Lett.* **117**, 051301 (2016).
- [30] J. Bobin, F. Sureau, and J. Starck, *Astron. Astrophys.* **591**, A50 (2016).
- [31] C. L. Bennett, D. Larson, J. L. Weiland *et al.*, *Astrophys. J. Suppl. Ser.* **208**, 20 (2013).
- [32] P. A. R. Ade, N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A13 (2016).
- [33] N. Battaglia, J. R. Bond, C. Pfrommer, J. L. Sievers, and D. Sijacki, *Astrophys. J.* **725**, 91 (2010).
- [34] D. N. Limber, *Astrophys. J.* **117**, 134 (1953).
- [35] H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, and L. Verde, *J. Cosmol. Astropart. Phys.* **2** (2012) 047.
- [36] O. Hahn, R. E. Angulo, and T. Abel, *Mon. Not. R. Astron. Soc.* **454**, 3920 (2015).
- [37] O. Doré, J. Bock, M. Ashby *et al.*, [arXiv:1412.4872](https://arxiv.org/abs/1412.4872).
- [38] L. Yan, E. Donoso, C.-W. Tsai *et al.*, *Astron. J.* **145**, 55 (2013).
- [39] S. W. Henderson, R. Allison, J. Austermann *et al.*, *J. Low Temp. Phys.* **184**, 772 (2016).
- [40] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006).
- [41] N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac, *Astrophys. J.* **709**, 920 (2010).
- [42] V. Springel, *Mon. Not. R. Astron. Soc.* **364**, 1105 (2005).

- [43] C. Pfrommer, V. Springel, T. A. Enßlin, and M. Jubelgas, *Mon. Not. R. Astron. Soc.* **367**, 113 (2006).
- [44] T. A. Enßlin, C. Pfrommer, V. Springel, and M. Jubelgas, *Astron. Astrophys.* **473**, 41 (2007).
- [45] M. Jubelgas, V. Springel, T. Enßlin, and C. Pfrommer, *Astron. Astrophys.* **481**, 33 (2008).
- [46] V. Springel and L. Hernquist, *Mon. Not. R. Astron. Soc.* **339**, 289 (2003).
- [47] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, *Astrophys. J.* **758**, 75 (2012).
- [48] C. Heymans, L. Van Waerbeke, L. Miller *et al.*, *Mon. Not. R. Astron. Soc.* **427**, 146 (2012).
- [49] N. Battaglia, J. C. Hill, and N. Murray, *Astrophys. J.* **812**, 154 (2015).
- [50] H. Park, P. R. Shapiro, E. Komatsu, I. T. Iliev, K. Ahn, and G. Mellema, *Astrophys. J.* **769**, 93 (2013).
- [51] J. Bobin, F. Sureau, J.-L. Starck, A. Rassat, and P. Paykari, *Astron. Astrophys.* **563**, A105 (2014).
- [52] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, *Astrophys. J.* **761**, 152 (2012).
- [53] J. Bobin, J.-L. Starck, F. Sureau, and S. Basak, *Astron. Astrophys.* **550**, A73 (2013).
- [54] S. Ferraro, B. D. Sherwin, and D. N. Spergel, *Phys. Rev. D* **91**, 083533 (2015).
- [55] T. H. Jarrett, M. Cohen, F. Masci *et al.*, *Astrophys. J.* **735**, 112 (2011).
- [56] P. A. R. Ade, N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **571**, A17 (2014).
- [57] P. A. R. Ade, N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A15 (2016).
- [58] R. Adam, P. A. R. Ade *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A9 (2016).
- [59] J. C. Hill and D. N. Spergel, *J. Cosmol. Astropart. Phys.* **2** (2014) 30.
- [60] G. E. Addison, J. Dunkley, and J. R. Bond, *Mon. Not. R. Astron. Soc.* **436**, 1896 (2013).
- [61] R. K. Sachs and A. M. Wolfe, *Astrophys. J.* **147**, 73 (1967).
- [62] M. J. Rees and D. W. Sciama, *Nature (London)* **217**, 511 (1968).
- [63] A. J. Shajib and E. L. Wright, *Astrophys. J.* **827**, 116 (2016).
- [64] P. M. Merkel and B. M. Schäfer, *Mon. Not. R. Astron. Soc.* **431**, 2433 (2013).
- [65] R. E. Smith, C. Hernández-Monteagudo, and U. Seljak, *Phys. Rev. D* **80**, 063528 (2009).
- [66] A. Cooray, *Phys. Rev. D* **65**, 083518 (2002).
- [67] O. Zahn, C. L. Reichardt, L. Shaw *et al.*, *Astrophys. J.* **756**, 65 (2012).
- [68] E. Calabrese, R. Hlozek, N. Battaglia *et al.*, *J. Cosmol. Astropart. Phys.* **8** (2014) 010.
- [69] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005).